Institute for Strategy and Business Economics

University of Zurich

Working Paper Series

ISSN 1660-1157

Working Paper No. 66

Cooperation in the Cockpit: Evidence of Reciprocity and Trust among Swiss Air Force Pilots

Beat Hedinger und Lorenz Götte

Oktober 2006

Cooperation in the Cockpit: Evidence of Reciprocity and Trust among Swiss Air Force Pilots[1]

Beat Hedinger[2] and Lorenz Goette[3]
University of Zurich
October 23, 2006

Abstract

Cooperation between workers is important for firms. Cooperation can be maintained through positive or negative reciprocity between workers. In an environment where cooperation yields high efficiency gains negative reciprocity may, however, result in high costs for firms. Therefore positive reciprocity should be prevailing in these environments. To test this assumption we conduct experiments with Swiss Air Force pilots and a student reference group.

We find that pilots' cooperation is based on stronger positive reciprocal behaviour. We conclude that Swiss Air Force pilots maintain team-work with high levels of positive reciprocity, regardless of the identity of their partner.

---

1 The support by the commander of the unit of the Swiss Air Force professional pilots (Ueberwachungsgeschwader), Colonel Bruno Morgenthaler is gratefully acknowledged. We thank Uschi Backes-Gellner, Claudio Noto, Jeremy Faux, Hannah Scott, Paul Maugin and Martina Schmid for helpful comments.
2 Institute for Strategy and Business Economics, Plattenstrasse 14, CH - 8032 Zurich,
Corresponding Author: Tel.: + 41 - 44 – 261 98 28, Fax: + 41 - 44 – 261 98 28, Mail: beathedinger@bluewin.ch
3 Institute for Empirical Research in Economics, Bluemlisalpstrasse 10, CH-8006 Zurich

1. Introduction

Voluntary cooperation is important in organisations due to the incomplete nature of contacts. Much attention has been devoted to the cooperation between firms and workers (Prendergast, 1999; Kandel and Lazear, 1992; La Porta et al., 1997; Akerlof and Kranton, 2005). In this case the typical problem involves the difficulty to enforce complex contracts to render a firm efficient. This issue can often be resolved through repeated interactions, so that a worker's incentive to comply is to profit from future interactions.[4] If, in addition to that, a minimal fraction of subjects is not purely selfishly motivated, cooperation is enhanced and sometimes even possible in one-shot interactions (Fehr and Schmidt, 1999; Brown et al., 2002).[5]

In many cases, cooperation between a worker and his fellow worker is just as important for a firm as is its efficiency (La Porta et al. 1997). This interaction differs in two important aspects from the interaction between a worker and a firm:

First, the interaction is harder to observe for the firm: It is unlikely to be enforced by incentives such as those created through repeated interactions between workers and firms (La Porta et al., 1997). Since employees often refrain from reporting an incident to the firm, where fellow workers do not cooperate, a worker cannot be disciplined by the fear of loosing his job. Hence it is difficult for the employer to establish cooperation among workers.

Second, social preferences may be more powerful in interactions between workers, because workers mostly interact with fellow workers. Therefore they might identify more strongly with one another than for example with their principal (Akerlof and Kranton, 2003).

It follows that firms have to rely on social preferences of their employees to ensure the cooperation among their workers. These social preferences can be expressed through positive and negative reciprocal behaviour. If such reciprocity is anticipated by other workers, it can lead to trust and the cooperation of a whole group (Fehr and Gaechter, 2000b; Falk and Fischbacher, 2000; La Porta et al., 1997; Carpenter et al., 2005).

---

[4] The effects of reputation can prove efficient (Fehr and Schmidt, 1999; Fehr and Gaechter, 2000a; Brown et al., 2002; Camerer and Fehr, 2002).

[5] Non-purely selfish individuals are commonly modelled by utility functions that include other individuals' actions and payoffs (see e.g. Levine (1998), Fehr and Schmidt (1999), Falk and Fischbacher (2000), Bolton and Ockenfels (2000) or Charness and Rabin (2002)).

In this study, we measure positive and negative reciprocity in a group of subjects for which cooperation at work is literally a question of life or death: We conducted a moonlighting game with pilots of the Swiss Air Force. While previous studies have shown, that workers cooperate more than students (Fehr and List, 2004; Carpenter et al., 2005), we examine two novel issues: First we study preferences for positive versus negative reciprocity in workers when negative reciprocity can have very high costs to the principal. Second we examine the robustness of reciprocity to social distance when cooperation is indispensable.

Starting with the first issue, we argue, that in some cases negative reciprocity may bear more adverse effects than the efficiency gains achieved by enforced cooperation. In the context of aviation, for instance, there are spectacular examples of how negative reciprocity can be costly to an organisation: On October 28, 2003, an airliner flew from Leipzig to Zurich. After an uneventful smooth flight, suddenly the captain decided to perform the duties of the co-pilot in addition to his own. The co-pilot stopped doing his duties almost altogether, because he felt himself left out. Consequently, he did not use the proper map to monitor the captain's approach nor did he check the visual references the captain claimed to see on final approach. In heavy fog, the aircraft crashed next to the runway due to the captain's bad judgment. A subsequent investigation ruled out any technical failure (AAIB, 2006). Although the aircraft was a total loss, on account of pure luck, no one was hurt.

It can be argued, that negative reciprocity played an important role in the occurrence of the accident: The co-pilot was unwilling to perform his duties after having been left out by the captain, even though it would have been especially important to assist a captain, who is performing the duties of two crewmembers at the same time. Hence it is particularly interesting to study if the vital cooperation among pilots is typically established through higher levels of positive reciprocity rather than higher levels of negative reciprocity.

Focusing on our second issue, it can be argued, that a further feature may have contributed to the detriment in our example: that the captain relied on his rank to relieve the co-pilot of his duties. Any identification process between the pilots, which could have lead to stronger positive and weaker negative reciprocity, would have been destroyed. Therefore social distance, in the form of hierarchy, may further hinder cooperation in such an environment.[6] As cooperation is essential among pilots to remain effective, neither negative reciprocity nor

[6] In addition to that, it has been shown, that perceived social hierarchy may hinder the optimal performance of subjects (Hoff and Pandey, 2006).

social distance should be a critical factor. Hence it is also interesting to study if social distance, particularly in the form of hierarchy, influences levels of reciprocity.

To research the first issue, we compare Air Force pilots and students playing the moonlighting game for real money. We find that positive reciprocity is indeed stronger than negative reciprocity among pilots compared to students. In addition trust was shown to be dependent on the level of expected reciprocity.

To research the second issue, we first compare positive and negative reciprocity of pilots paired with higher-ranking pilots to positive and negative reciprocity of pilots paired with pilots of equal rank. We do not find any effect of perceived social hierarchy on actions and reactions of pilots and no statistically significant difference in the levels of positive and negative reciprocities was detected.

To further stress the second issue, we test amongst the pilots the general assumption, that high levels of cooperation among workers are based on a higher degree of group identity similar to the studies of Gaechter and Fehr (1999) and Goette et al. (2006). The pilots were paired with students to investigate pilots' behaviour towards total outsiders. Surprisingly positive reciprocity does not vanish.

Our results suggest differences in basic social preferences between pilots and students. It seems tempting to conclude, that these differences are achieved through screening, training or socialization of the Swiss Air Force.

The remainder of this paper is organized as follows: In Section 2 we provide some details about Swiss Air Force pilots. Section 3 discusses the experimental design and provides the predictions. Results are discussed in Section 4. Finally section 5 summarizes.

2. Institutional Background

In order to better understand the screening, training and socialization of Swiss Air Force pilots, this section provides background details about their careers.

Every year the Swiss Air Force picks 6 to12 potential pilots out of hundreds of twenty-year-old applicants, who, among other criteria, are chosen for their teamwork capabilities (Noser, 2003; Airforcepilot, 2004).

To further train cooperation, the applicants have to attend 56 lessons and 3 years of practical education in "Crew Resource Management" during pilot school.[7] This particular course has been developed on the basis of management training and aims to optimise the use of human resources in the cockpit (Helmreich et al., 1999).[8] After pilot school, every flying officer is required to attend to refresher courses on a regular basis.

In their work environment, pilots become more socialized because of the heavy dependence on each other: Large helicopters are always crewed by two pilots due to their complexity. Jet fighters never fly alone to increase efficiency and firepower. A single pilot is unlikely to succeed in accomplishing an actual task especially when confronted with an emergency situation. Therefore crew members must fully cooperate in order to minimize the probability of loss of human lives and best accomplish their task.

For all these circumstances, we view Swiss Air Force squadrons as being a work environment where cooperation has high efficiency gains and where screening, training and socialization is used to enhance such cooperation.

3. Experimental Design

3.1. The Game

The subjects participated in several moon lighting games (Abbink et al., 2000). In this particular game, Player A has the choice of being friendly, neutral or unfriendly to Player B. Player B in turn rewards or punishes Player A for his actions. We can learn about B’s positive and negative reciprocity when looking at his rewards and punishments. Behaviour of Player A may carefully be interpreted as trust.

The game is summarized in Figure 1. Player A and Player B are each endowed with 20 points at the beginning of the game. Player A has to decide whether to pass on 10 points to Player B, not to pass on any points, or take 5 points from Player B. The experimenter in any case doubles the transfer. Hence, if Player A passes on 10 points, Player B will receive 20 points. Conversely, if Player A takes 5 points, Player B will loose 10 points.

---

[7] The Commander of Screening & pilot basic training Swiss Air Force provided this information.

[8] It enhances communication, teamwork, situational awareness, decision-making capabilities, leadership and stress management (Helmreich et al., 1999).

Player B can then decide how to respond to Player A's action. Player B can either reward or punish Player A. If Player B spends one point on rewarding Player A, Player A will receive two points. If Player B spends one point on punishing Player A, Player A will lose two points. Player B can use up to ten points for either punishing or rewarding Player A. We apply the strategy method for Player B, i.e., Player B makes a choice for each possible case that can arise (Player A passing on 10 points, passing on 0 points, or taking 5 points).

The sub game-perfect equilibrium with selfish preferences is easy to derive: Player B would never use any points to reward or punish, since Player A has already chosen his action. Consequently, Player A will maximize his payoff by taking 5 points from Player B (who looses 10 points). But behaviour is more interesting when agents have non-selfish preferences. Different social preferences can be examined for different players: First, the behaviour of Player B can be used to measure reciprocity. Player B can display positive reciprocity if he rewards transfers of 10 points relative to a neutral transfer (or if he rewards Player A for not taking any points away). Conversely Player B can display negative reciprocity if he punishes when Player A takes 5 points from him (or if he punishes Player A for not passing on any points).

Second, Player A's actions may be considered as trusting behaviour. However, they must be interpreted more carefully than in a standard trust game due to the ambiguous role of risk preferences. In a standard trust game (e.g. Glaeser et al., 2000; Fehr and List, 2004; Falk and Zehnder, 2006), players exhibit trust by transferring positive amounts to Player B. Risk aversion reduces transfers of first movers, because this reduces the variance in payoffs. In the moonlighting game Player A could also be motivated to pass on 10 points to Player B if he fears substantial punishment otherwise. Risk preferences have a more complicated impact on Player A's behaviour: A "neutral" action may be risky as well, as it may be rewarded or punished. These "outbursts" of reciprocity may create an even larger variance in the payment of Player A when he acts "neutral". Therefore the interpretation of Player A's actions has to be very careful and involve his beliefs concerning Player B's actions.

3.2. Treatments

Figure 2 summarizes the different treatments that have been conducted.

Student-Student Baseline: Our control condition, that is best comparable to other experiments, is the Student-Student baseline treatment (referred to as S-S in the remainder of the paper). 34

students participated as Players A, and 33 as Players B.[9] The students were told that their partners in the experiments were other students, not present now. Immediately after their choices, we elicited beliefs about the behaviour of their counterparts in the experiment, as we did in all treatments.

Student-Pilot (S-P): In this treatment, the choices of another 58 students as Players A were used. They were told that they would be matched with a pilot from the Swiss Air Force as Player B. 112 pilots participated in this treatment. The students were informed that their choices may be used in multiple matchings, and that they would be paid the total amount they earned from all their matchings.[10]

Higher-ranking pilots - lower-ranking pilots (H-L): In this treatment, the pilots were assigned the roles of A players and B players based on their rank in the Air Force. Pilots with the rank of major or above (N = 49) were assigned the role of Player A and were matched with a pilot with the rank of captain or below (N = 67).[11] Some of the choices of Players A were used twice to determine all the payoffs of Players B.

Pilot-Pilot (P-P): Pilots in this treatments where paired with their peer-group. Lower-ranking pilots were paired with lower-ranking pilots, and told so. Similarly, higher-ranking pilots were paired with higher-ranking pilots and informed so.

3.3. Procedures

The experiment for the baseline and the first part of the *S-P* treatment were conducted using students of the University of Zurich, who volunteered to participate as subjects of behavioural research studies. No economic students were allowed to participate in the sessions to avoid economic background bias due to their possible knowledge of game theory (Carter and Irons, 1991; Frank et al., 1993).

The data for the other treatments were collected on the occasion of a compulsory workshop for all Swiss Air Force pilots on the subject of "work satisfaction" and "burn out syndromes" on December 12, 2005. Due to the compulsory nature of the workshop any participation bias seems unlikely.

---

[9] The decision of the 34th player A was randomly assigned to the decision of a player B.

[10] As can be seen, a students' choice was used almost exactly twice.

[11] The separation between captains and majors is due to the fact that every squadron member becomes a captain after a certain amount of time, while squadron leaders and other superiors hold at least the rank of a major. Therefore the working level is represented by the maximum rank of a captain, while superiors are represented by the minimum rank of a major.

The pilots were seated according to their military rank. They were told that they had to make decisions and could earn money in doing so. Additionally they were informed that the data would be treated anonymously.[12] Thereafter the instructions and the decision sheets for the *S-P* treatment were distributed. Treatment *H-L* and *P-P* followed while every effort was made to ensure the subjects would not speak to each other and never hold more than one decision sheet at a time.

Only after the pilots had finished the last treatment were they issued a closed envelope with a copy of the decision sheet of a randomly assigned student to estimate their earnings from the *S-P* treatment. This late distribution of the partners' decision in the first treatment helped preventing a learning effect (Egas and Riedl, 2005). Payments where distributed at the end of the day.

3.4. Subject Characteristics

As former studies have shown, it is very important to obtain social background variables when comparing students to other subjects (Bellmare and Kroeger, 2005). Therefore the question of whether students can be compared to pilots in an efficient manner mainly depends on the overlap in the social background variables of the two populations. Table 1 shows the statistical means, minima, maxima, standard errors and numbers for pilots and students.

The main difference between the two groups can be found in their mean age; the disparity is no less than 24 years. However, the two populations are overlapping, as the youngest pilot is 23 years old and the oldest student 45. Another difference can be found in the distribution of the genders within the two populations. As there are only 6 female pilots in the Swiss Air Force and only two of them participated in the workshop, males were the more dominant gender among pilots than among students. As for education[13], the difference between students and pilots is small: On the one hand, we consider students to hold a high-school diploma, as

---

[12] Special care was taken to ensure anonymity as not to endanger the actual team spirit of the pilots involved (Carpenter et al., 2005). Furthermore, each subject was informed about his right to tell the experimenter not to use his decision-data for behavioural research.

[13] Education was coded as theoretical years of education according the highest degree a subject has accomplished (see for example Bonjour (1997) or Falter and Ferro (2000)).

this is an entry requirement to university.[14] Military pilot applicants, on the other hand, are currently required to hold a high school diploma or reach an equivalent educational level.[15]

Despite the above discrepancies in mean age and gender, we consider the two groups as sufficiently overlapping and therefore assume that statistical tests are not subject to out-of-sample predictions.

3.5. Behavioural Predictions

Our design allows us to examine in detail how the behaviour of Player B (and to some extent of Player A) is affected by the different treatments.

The treatment most closely related to the previous literature is S-S: it provides us with the baseline measures of social preferences. Evidence from other studies suggest that when one controls for social background variables, students become comparable in their social preferences to the general public in their social preferences (Gueth et al., 2002; Gaechter et al., 2004). Bellmare and Kroeger (2005) find, that especially age, gender and education have significant effects on investments and rewards of a representative panel of the Dutch population and that after controlling for social background variables, no difference could be found between the representative panel and students.

Players B

Comparing the behaviour of Players B in treatment S-S to treatment P-P sheds light on our first issue, the preference for positive versus negative reciprocity. It shows the difference in social preferences between students and a group of subjects that interacts on a daily basis in an environment that yields high efficiency gains from cooperation.

There is strong evidence from previous studies that social ties between participants strongly influence behaviour towards more cooperation (see e.g. Burks et al., 2006).[16] Moreover, the scope for cooperation in the work environment of pilots may shape their social preferences.

---

[14] No exact data on completed education was available for students.

[15] There are older pilots who have fewer years of schooling and a small group of pilots who have earned a university degree. These two groups almost offset each other.

[16] Burks et al. (2006) find that bike messengers cooperate more as second movers in a sequential prisoners' dilemma even to the extend that 30 percent cooperate in spite of defection of the first-mover.

This has been suggested by Fehr and List (2004), who show more trustworthiness in managers of coffee planting cooperatives than in students.[17]

In light of possible negative consequences of negative reciprocity, we predict the following:

Cooperation among pilots is mainly based on stronger positive instead of stronger negative reciprocity. Therefore positive reciprocity is stronger and negative reciprocity is weaker among pilots than among students.

Our second focus lies on hierarchy: We compare Player B's behaviour in P-P to H-L treatment, where we match low-ranking pilots with high ranking-pilots.

It can be argued that ranks between pilots matter, as only pilots of similar rank are similar in many dimensions, including age and years of service. Further, similar ranking pilots are more likely to interact with each other while performing similar tasks, since the higher-ranking pilots mostly pursue management duties within the Air Force. Thus, we might expect a higher social distance when pilots of "unequal" rank are matched in an experiment. Further evidence from Hoff and Pandey (2006) and La Porta et al. (1997) show that social hierarchy can have a strong impact on behaviour.

However, as perceived hierarchy might bear negative externalities in such an environment as discussed in the introduction, we predict the following:

Social distance induced by hierarchy does not affect positive and negative reciprocity of pilots.

To finally test whether high levels of cooperation are based on group-specific norms, we increase the social distance to a degree of total outsiders. The comparison of P-P treatment vs. S-P treatment shows whether pilots' degrees of reciprocity are different when interacting with a student rather than with a pilot. There is strong evidence from earlier studies that pro-social behaviour is different when individuals interact with someone they perceive to belong to a different social group. Most closely related, Goette et al. (2006) find evidence consistent with in-group favouritism when officer candidates in the Swiss Army interact with a member of

[17] Holding transfers of the first mover constant, the managers in their experiment return significantly more than the student control group. In contrast to the standard trust game, that has been applied by Fehr and List (2004), our experiment lets us distinguish between positive and negative reciprocity.

their own platoon relative to when they interact with a member of a different platoon.[18] Similarly Ruffle and Sosis (2006) find that members of Kibbutzim are more cooperative in interactions with member of Kibbutz than with subjects living in "normal" cities.[19] Fershtman and Gneezy (2001) finally find that subjects' social preferences strongly depend on demographic differences to partners in experiments. Hence we predict the following:

When pilots are paired with students positive reciprocity decreases and negative reciprocity increases compared to pilots interacting with fellow pilots or superior pilots.

Player A

In the case of Player A, we are not only interested in his actions. Beliefs might be important as well, as they influence the actions of Player A.[20] Trusting behaviour can be explained as follows: If Player A believes in a strong positive reciprocity from Player B, he is prone to express more trust by transferring more money to Player B.[21] Additionally, in the moonlighting game he might be influenced by the fear of negative reciprocity of Player B. If, however, his transfer is high in absence of corresponding expectations about Player B’s back transfer, his action might be motivated by risk preferences. These are ambiguous as discussed in section 3.1.. The fact that military pilots have a risk of death, which is at least 60 times higher than their student counterparts, could play a role in this case.[22] Hence, only if believed positive reciprocity is high, a statement in the favour of trust can be made. But even then caution has to be exercised when interpreting the results because risk preferences might be interacting with the beliefs of Player A.

---

[18] Bernhard et al. (2006) find similar results for self-selected groups: Members of clans in Papua-New Guinea show the same behavioural tendencies when interacting with somebody from their own clan relative to interactions with somebody from a different clan.

[19] Further, when Kibbutz members interact with city members, cooperation is at the same level as that of city members.

[20] This is due to the fact that for player A, the strategy method could not be applied

[21] Fehr and List (2004) find higher transfers of first movers in their trust game with managers. In Burks et al. (2006) bike messengers cooperate more as first movers than students in a sequential prisoners dilemma.

[22] The mean risk of death for a Swiss worker in 2001 amounts to 0.5 deaths per 10’000 workers. In civil aviation it amounts to 1.5 deaths per 10’000 workers (SUVA, 2004). For Swiss Air Force pilots, it amounts to 33 deaths per 10’000 pilots each year for the period of 1981 until 2001. During the period of 1941 until 2001 the Swiss Air Force had even lost 81.5 pilots to crashes per 10’000 pilots each year. As some of the older pilots among our subjects may have based their decision on early statistics, even this number could be claimed representative.

The comparison of S-S vs. P-P is used to establish the differences in stated beliefs and actions between pilots and students.[23] In pursuing the argument for Player B we propose the following:

Pilots exhibit more trust: They anticipate higher levels of positive reciprocity, believe in higher back transfers and thus transfer more than students.

As a final point we test the effect of in-group favouritism by introducing hierarchy when comparing treatment P-P to H-L. One might argue that social distance hinders trust (Hoffmann et al., 1996). However, as cooperation across ranks is essential among pilots in the same crew to remain effective, social distance should not be critical for the actions of pilots. Hence we predict the following:[24]

Beliefs and transfers of Players A remain the same whether pilots are paired with lower ranking pilots or with pilots of the same rank.

4. Results

4.1. Comparison of S-S to Previous Literature

The S-S treatment serves as a benchmark to previous studies with students as subjects. In our experiment, the B player punishes negative transfers with 3.6 points, rewards zero transfers with essentially zero (0.3) points, and rewards a transfer of 10 points with 5.6 points on average. In relative terms, our subjects thus invested 36% of the points they lost into punishment and 28% of the points they gained into rewards. The latter figure is strikingly close to Abbink et al. (2000), who find a corresponding figure of 26%. Evidence in Abbink et al. (2000) points to a slightly stronger negative reciprocity than our study, although this could be related to the details of the game.[25]

---

[23] Although we do not have a treatment that introduces experimental variation in beliefs to identify the beliefs' effects on behaviour the comparison of pilots and students beliefs can be telling.

[24] As we do not have a P-S treatment (see Figure 2), we cannot test for total strangeness for player A.

[25] The version of the moonlighting game in Abbink et al (2000) give the A players a continuous choice to either transfer or extract points from B, hence only a comparison of the relative shares is possible. The continuous nature of the game also makes it possible to extract only a few points from player B, which usually are punished quite severely. This may explain why there is more evidence for negative reciprocity in Abbink et al. (2002). Furthermore for player A positive transfers are tripled, while negative transfers are not altered by the experimenter. In the case of player B, only negative transfers are tripled. The moonlighting game of Cox et al. (2002) shows further differences: Players B can not only reward or punish but even take money from player A.

In our base treatment, 14% of players B choose to take 5 points, 27% choose to transfer 0 points and 59% of players transferred 10 points. In Abbink et al. (2000) 18.8 % choose a negative amount (12.5% more than half of the possible points), 15.6% choose to send 0 points and 65.6% sent a positive amount (43.8% more than half of the possible points).[26]

4.2. ***Behaviour of Player B***

- ***Result 1:*** *Pilots are significantly more positive reciprocal and slightly less negative reciprocal towards pilots than students towards students.*

Evidence for this result can be found in Figure 3 in the appendix, which shows the average back transfers to different actions of players A. A clear difference between back transfers of pilots to back transfers of students can be found. Particularly, kind actions of players A are more rewarded by pilots than students, which points to stronger positive reciprocity. Additionally, unkind actions of players A are less revenged by pilots than by students, which points to weaker negative reciprocity. Table 2 in the appendix presents a more stringent statistical test.[27] An OLS regression is performed with heteroscedastic standard errors. The regression controls for age, squared age, gender and possible differences in education. According to the results in column 1 through 3 of Table 2, variation in gender, age or education between students and pilots do not explain the whole difference in back transfers. Furthermore, it cannot be rejected *(p* < 0.01) that pilots have a more positive reciprocal nature than students. For negative reciprocity on the other hand we do not find a statistically significant difference between pilots and students. We therefore infer that the fact that negative reciprocity of pilots is not stronger compared to those of students cannot be rejected. To further substantiate our findings we added a conservative significance test, which adjusts for the fact that we test multiple hypotheses (Holm, 1979, as cited in Romano and Wolf,

---

Cox et al. (2002) find, that 38% of their players B, who have lost money, take revenge on player A. Players B, who have made money, reinvested 39% of the money they have gained in to rewards.

[26] In Cox et al. (2002) 43% choose a negative amount, 10% send 0 points and 47% a positive amount.

[27] As a paper and pencil method was applied, we were not able to prevent missing actions and beliefs. In addition to that, some individuals stated beliefs, that added up to more than 100% and where therefore disregarded. Finally, few individuals probably misinterpreted the decision sheets. They stated decisions that are exactly the opposite of reciprocal, as they rewarding unfriendly and punished friendly actions. We believe, that we should not consider these data for our analysis. However, additional regression results show, that none of our findings are sensitive to the inclusion of these data. Exact tables of missing and excluded observations and beliefs can be found in Table 7 through Table 10 in the appendix.

2005). The data are robust to this test, as the significance level for positive reciprocity is not affected.[28]

The data also show some demographic differences. Age appears to have an effect on positive reciprocity, which follows a u-shape with a minimum at 46 years. These findings are very close to those of Bellmare and Kroeger (2005). This is important because they use a representative panel of the population. In their trust game, they find a similar relation between age and the propensity to reward investments with a minimum between 35 and 40 years of age. A further similarity in results exists for the dummy variable "male", which has a positive effect on back transfer when treated kindly (column 3).[29] Lastly, education has no statistically significant effect on back transfers in our results, which might be caused by the low variation in our data. Bellmare and Kroeger (2005), however, find a negative educational effect for responders in an investment game. Due to the otherwise high comparability of our results to Bellmare and Kroeger (2005), who do not rely on student data only, we argue that the behaviour of our subjects might be representative of social preferences among the general population in an anonymous setting.

To examine the exact nature of the cooperation differential we introduce social distance, first in the form of hierarchy by pairing high-ranking pilots with low-ranking pilots and finally in the form of total strangeness by pairing students with pilots.

- ***Result 2****: Cooperative behaviour of pilots, mainly based on strong positive reciprocity, is not confined to fellow pilots. There is no difference in the back transfers of pilots to higher-ranking pilots or even students.*

Figure 4 shows virtually no difference between the back transfers of treatments P-P, H-L and S-P.[30] More stringent results are displayed in Table 4 where OLS-regressions with cluster

[28] To further assess the difference in social preferences between pilots and students, we examine the costs the subjects are willing to incur to respond to players A. The OLS regression results with the dependent variable being the absolute amount of money spent are represented in columns 1 through 3 of Table 3. The only significant difference we find between pilots and students is in the response to a positive transfer of 10 points. Comparing column 1 of Table 3 to Table 2, we might infer from the fact, that the point estimate drops from 2 to 0 points, that pilots are not generally less engaged in interactive behaviour, but do it in a less negative way when treated unfriendly. This result gives some support to the assumption that pilots may be less negatively reciprocal.

[29] Equal results can be found in Gueth et al. (2002).

[30] There is no concern for effects of subjects playing both roles in the experiment (see for example Burks et al, 2003), as this is not the case in the same treatment. The procedure, however, does not exclude a possible order effect, which is minimized by not informing the subjects about their earnings before the end of the last treatment.

analysis have been run. Dummy-variables have been added for the S-P and the H-L treatment, the P-P treatment being the base group in this case. According to columns 1-3, we must conclude that the two treatment effects are not statistically significant. No differences between the three treatments where detected. The precision of this result is fairly high, as differences above 0.6 points would have been picked up. The respective F-tests confirm the result.[31]

It is interesting to ask whether pilots perceive any difference at all between pilots and students playing part A.[32] According to Figure 5 pilots have identical beliefs about pilots of the same or higher rank. By contrast, their beliefs about students' behaviour are more negative. They expect a lower percentage of the students to transfer 10 points, and a higher percentage to take 5 points. Again, Table 5 presents statistical tests. OLS-regressions with cluster analysis confirm the results. This shows that pilots are well aware of their partners and all the same do not change their behaviour.

It is tentative to conclude, that the Swiss Air Force is able to screen, train or socialize its pilots in a way that cooperation is primarily sustained through higher levels of positive reciprocity, and the effect is very robust to social hierarchy. It is astonishing, however, that pilots do not change their behaviour when paired with outsiders. This result is surprising, as other studies have found strong effects of social distance (La Porta et al., 1997; Fershtman and Gneezy, 2001; Ruffle and Sosis, 2006; Hoff and Pandey, 2006; Bernhard et al., 2006; Goette et al., 2006). Therefore we conclude that differences in the behaviour of pilots compared to other subjects are not mainly caused by social ties but rather by more cooperative preferences of pilots.

Even though the situation is less clear for Player A, the next subsection analyses his behaviour.

[31] To detect if superiors behave differently from normal pilots, probably due to higher salaries, we introduced an additional variable for ranks starting with major. As no statistical significant difference can be detected we furthermore conclude that there is also no wealth effect (as majors earn more than captains) involved in the equation, which confirm findings of Slonim and Roth (1998) and Bellmare and Kroeger (2005).

[32] This has been suggested by Manski (2002) and applied by Bellmare et al. (2005). However, as in our case the strategy method has been applied we do not consider using beliefs as an independent variable in the regression. Every pilot had the chance to make conditional decisions, which cannot depend on his beliefs of the distribution of the actions of first movers. Hence the actual decision of player A cannot in any logical sense influence the decision of player B.

4.3. Behaviour of Player A

- ***Result 3:*** *Pilots do not seem to transfer more than students.*

The formal statistical test can be found in the second column of Table 6,[33] where the amount transferred to Player B is regressed on the treatment condition. We do not find a statistically significant difference between the transfer of pilots compared to students when controlling for gender, age, age^2 and education. The maximum-likelihood ordered probit estimation in the next column of Table 6, which accounts for the restricted choice characteristic for Player A, confirms the results. If we compare these results to Figure 6, which shows the different distributions in percents of choices of students and pilots in the position of Player A, we are astonished: As we had originally expected, clearly more pilots than students send 10 points to Player B. But the control variables of the regression seem to explain the whole difference.[34] Furthermore the estimators for age point towards an inverted u-shape of the age effect with a maximum at the age of 37. This result is very comparable to those of Bellmare and Kroeger (2005), who find in their representative panel an inverted u shaped age-effect with a maximum at 37 years as well.[35] Hence we conclude that pilots do not seem to transfer more than students as players A, as any difference is already explained by the control variables.
As we expected that pilots transfer more because they believe in higher back-transfers of fellow pilots, it is interesting to study whether actions of players A correlate with different beliefs about the conditional back transfer of players B (Camerer and Fehr, 2006; Goette et al., 2006).[36]

- ***Result 4****: Actions seem to be mainly motivated by players' beliefs in positive and negative reciprocity, i.e. by more optimistic beliefs about back-transfers.*

---

[33] The 4 observations that had no respective beliefs (see Table 7 and Table 10) had to be dropped to keep estimations comparable. This caused no significance difference in the results.

[34] To find out, whether there is a good reason to include the control variables, we run different regressions and found, that finally the squared age effect is responsible for the high p-value of the treatment factor. To show the effect, we included in the first column of Table 6 a regression without this factor, where the statistical effect of the P-P treatment is quantitatively large and significant as originally expected. However, as age^2 has a p-value of 0.088 and age one of 0.105, we may not exclude the factor.

[35] Additionally Alesina and La Ferrara (2002) find similar effects. A possible explanation can be found in Glaeser et al. (2002), who argue that investment in social capital follows the same pattern as investment in human capital.

[36] Fehr et al. (2006) accordingly find in a trust game, that not all the differences in transfers between Americans and Germans can be explained by different beliefs.

An OLS and ordered probit regression incorporating subjects' beliefs as independent variables are displayed in the fourth and fifth column of Table 6. They show, that indeed beliefs are statistically significant factors for subjects' actions.[37] All the more, the r-squared climbs from 11% (in the regression without beliefs) to 27%, which indicates the high relevance of the beliefs. Finally, the relevant combined F-tests confirm the results. It appears as if higher transfers are motivated by more optimistic beliefs about positive, neutral and negative actions.[38]

Finally we study the beliefs of players A.

- ***Result 5****: There is weak evidence that Pilots believe in higher back-transfers than students.*

The formal statistical test can be found in the last columns of Table 6, where the beliefs of players A about back-transfers of Player B after negative, neutral and positive transfer are regressed on the treatment condition. We find a statistically significant difference between the beliefs of pilots and students when controlling for gender, age, and education. The estimated difference in believed back-transfers after a positive action is 3.9 points and is statistically significant with a p-value of 0.005.[39] However, as we include the squared effect of age, only beliefs about back-transfers after positive actions are positive and marginally statistically significant with a p-value of 0.090.[40] This result is astonishing when compared to Figure 7 where clearly pilots believe in higher back-transfers of Players B than do the students. The statistical result seem to be driven by the squared effect of age, which is not statistically significant in this case and does not seem to improve the regression as it renders age insignificant as well.

It seems unclear whether pilots do not transfer more because they do not have more optimistic beliefs about back-transfers than students, or whether the ambiguous effects of risk preferences in the moonlighting game offset the slightly more optimistic beliefs of pilots.

---

[37] The respective p-values for the factors of the two regressions are not very different but climb over the 10% margin for the OLS-regression in the case of A's beliefs after neutral and negative actions.

[38] However, one has to be cautious interpreting these results as beliefs might be affected by risk aversion. Furthermore as their beliefs have been asked after they had made their decision, it seems possible that the subjects acted according their intuition and when asked about their expectations, justified their decisions through appropriate beliefs (Kahnemann, 2003), although the experimenter promised the subjects money for accuracy in beliefs.

[39] When we apply the multiple hypotheses test (Holm, 1979) the p-value climbs to 0.015.

[40] They even become statistically insignificant if the multiple hypothesis-test is applied (Holm, 1979).

To finally test for social distance due to hierarchical effects, we compare the P-P to the H-L treatment.

- ***Result 6:*** *Higher-ranking pilots do not trust less in lower-ranking pilots and therefore do not transfer less to lower ranks, compared to pilots interacting with equal ranks.*

Only graphical evidence for this result can be given in Figure 8 as just four players A changed their action from H-L to P-P treatment. Surprisingly it was all to the worse of their partners. The results complement the previous findings from players B. They suggest that hierarchy in this setting has no disruptive effect on social preferences. This must be one feature of the very special work environment even though it is within the armed forces where hierarchy plays a more influential role than in the civilian world (Akerlof and Kranton, 2005).

5. Conclusions

Cooperation within its workforce is a highly important factor for almost every firm in the market (La Porta et al., 1997). However, it is difficult to enforce cooperation among workers: It is hard to observe and therefore unlikely to be enforced by incentives created by repeated interactions between workers and firms. Hence high levels of cooperation must be established on account of higher grades of positive or negative reciprocity.

Previous studies have shown that negative reciprocity is a strong behavioural force when subjects are mostly college students (Charness and Rabin, 2002). However, negative reciprocity among workers may bear inefficient outcomes for a firm. Furthermore, if workers of different levels have to cooperate, perceived hierarchy may hinder cooperation. Therefore we hypothesized that firms could assure cooperation by appropriate screening, training or socialization of its workers.

To test this hypothesis, we use a subject pool of professional Swiss Air Force pilots, who must arguably cooperate very closely to achieve their mission goals. To detect both positive and negative reciprocity we apply the moonlighting game (Abbink et al., 2000).

We find that pilots are indeed significantly more positive reciprocal and slightly less negative reciprocal towards pilots than students towards students.[41] Furthermore the estimators of our

[41] This Result does not contradict Mas (2006) who finds strong negative reciprocity among police agents. It seems impossible in his set-up to detect positive reciprocity at all, as either the expectations of the policemen are met or not met but never exceeded. Furthermore he researches the interaction of firms and workers, while in our case we examine the interaction between workers and workers.

control variables are very comparable to those of Bellmare and Kroeger (2005), who use data of a representative panel of the Dutch population. Therefore we infer that our results are of general relevance. We can furthermore show that trust may depend on the level of expected reciprocity. However, the higher levels of trust expressed by pilots seem to be mostly explained by control variables. But this unexpected result may be biased due to the ambiguous effects of risk preferences in the moonlighting game.

To test for the effects of perceived hierarchy, we conduct further experiments and find that pilots' stronger positive and weaker negative reciprocity is not confined to the interaction with fellow pilots. Pilots do not treat higher-ranking pilots different than fellow pilots.

We finally study how pilots behave towards strangers. We find that pilots do not change their behaviour, as there is no difference between their back transfers to students and their back transfers to their fellow pilots. Thus our results suggest that these pilots have different preferences and are even outside their environment more cooperative. This is astonishing in light of previous studies which indicate strong effects of group identity (Gaechter and Fehr, 1999; La Porta et al., 1997; Fershtman and Gneezy, 2001; Ruffle and Sosis, 2006; Hoff and Pandey, 2006; Bernhard et al., 2006; Goette et al., 2006).

These findings indicate that the Swiss Air Force is in fact able to screen, train or socialize its pilots in a way that cooperation among its workers is primarily maintained through positive reciprocity, and hierarchy does not compromise teamwork. Future research should address this issue and try to separate these different possible channels.

Not surprisingly there is evidence which points towards the screening thesis: Hedinger (2004) shows that reserve pilots of the Swiss Air Force that work in non-flying jobs (e.g. as engineers or medical doctors) and pass the same screening as the professional pilots statistically earn significantly more than comparable individuals with a equal job.[42] This could be interpreted as an exceptionally high level of productivity, which may be based on high teamwork abilities. Hence, if reserve pilots show the same levels of reciprocity, cooperation could be claimed to screening effects.

The findings of this study imply that an employer who claims having such a workforce does not need to enforce cooperation among its workers. Furthermore, he might even profit from the exceptionally high levels of positive reciprocity, as his employees do not distinguish between superiors or fellow workers. Therefore he can forgo substantial gains from the

[42] After including various control variables, propensity-matching estimates a salary difference of 25%.

positive reciprocity of his workers, if he does not offer them appropriate conditions of employment.

References

Abbink, K. Irlenbusch, B. 2000. The Moonlighting Game. An experimental Study on reciprocity and retribution. Journal of Economic Behavior & Organization (42). 265–277.

Aircraft Accident Investigation Bureau (AAIB), 2006. Report No: u1882, Bern.

Airforcepilot, 2004. We expect, http://www.vbs-ddps.ch/internet/luftwaffe/en/home/airforcepilot/chapter0/chapter0.html from 19.8.2004.

Akerlof, G., Kranton, R., 2005. Identity and the Economics of Organizations. Journal of Economic Perspectives 19(1). 9–32.

Alesina, A., La Ferrara, E., 2000. Who Trust Ohters. Journal of Public Economics 85(2). 207-234.

Bellmare, C., Kroeger, S., 2005. On representative social capital. IZA Discussion Paper No 1145.

Bellmare, C., Kroeger, S., Van Soest, A., 2005. Actions and Beliefs: Estimating Distribution-Based Preferences using a large scale Experiment with Probability Questions on Expectations. IZA Discussion Paper No 1666.

Bernhard, H., Fehr, F., Fischbacher, U., 2006. Group Affiliation and Altruistic Norm Enforcement. American Economic Review 96(2). 217-221.

Bolton, G., Ockenfels, A., 2000. A Theory of Equity, Reciprocity and Competition. The American Economic Review 90 (1). 166-193.

Bonjour, D., 1997. Lohndiskriminierung in der Schweiz. Paul Haupt Verlag, Bern.

Brown, M., Falk, A., Fehr, E., 2002. Contractual Incompleteness and the Nature of Market Interaction. IEW-Working Paper 38.

Burks, S., Carpenter, J., Goette, L., 2006. Performance Pay and the Erosion of Worker Cooperation: Field experimental evidence. IZA Working Paper 2020.

Burks, S., Carpenter, J., Verhoogen, E., 2003. Playing both roles in the trust game. Journal of Economic Behaviour & Organization 51, 195-216.

Camerer, C., Fehr, E. 2002. Measuring Social Norms and Preferences using Experimental Games: A Guide for Social Scientists. Working Paper 97, IEW-Working-Papers…

Camerer, C., Fehr, E., 2006. When Does "Economic Man" dominate Social Behaviour. Science 311, 47-52.

Carpenter, J., Burks, S., Verhoogen, E., 2005. Comparing Students to Workers: The effects of social framing on behaviour in distribution games. Research in Experimental Economics 10, 261-290.

Carter, J., Irons, M., 1991. Are Economics Different, and if so, why? The Journal of Economic Perspectives 5, 171-177.

Charness, G., Rabin, M., 2002. Understanding Social Preferences with Simple Tests. Quarterly Journal of Economics 117(3). 817-870.

Cox, J., Sadiraj, K., Sadiraj, V., 2002. Trust, Fear, Reciprocity and Altruism. University of Arizona discussion paper.

Egas, M., Riedl, A., 2005. The Economic of Altruistic Punishment and the Demise of Cooperation. IZA Discussion Paper No 1646.

Falk, A., Fischbacher, U., 2000. A Theory of Reciprocity. Working Paper 6, IEW-Working Papers.

Falk, A., Zehnder, C., 2006. Trust and Discrimination. A City-Wide Experiment. Manuscript. University of Zurich.

Falter, J., Ferro Luzzi, G., 2000. Public-Privat Sector Wage Differentials in Switzerland. Schweizerische Zeitschrift für Volkswirtschaft und Statistik 136(3), 319-340.

Fehr, E., Gaechter, S., 2000a. Cooperation and Punishment in Public Goods Experiments. The American Economic Review 90, 980-994.

Fehr, E., Gaechter, S. 2000b. Fairness and Retaliation: The Economics of Reciprocity. The Journal of Economic Perspectives 14(3). 159-181.

Fehr, E., List, J., 2004. The hidden costs and returns of incentives: Trust and trustworthiness among CEO's. Journal of European Economic Association 2, 743-771.

Fehr, E., Naef, M., Fischacher, U., Schupp, J., Wagner, G., 2006. Why Americans are more trusting than Germans. Working Paper. University of Zurich.

Fehr, E., Schmidt, K., 1999. A theory of Fairness, Competition and Cooperation. The Quarterly Journal of Economics, 817-868.

Fershtman, C., Gneezy, U., 2001. Discrimination in a segmented society: an experimental Approach. The Quarterly Journal of Economics 116(1), 351-377.

Frank, R., Gilovich, Th., Regan, D., 1993. Does Studying Economics Inhibit Cooperation? The Journal of Economic Perspectives 7, 159-171.

Gaechter, S., Fehr, E., 1999. Collective Action as a Social Exchange. Journal of Economic Behavior & Organisation 39, 341–369.

Gaechter, S., Herrmann, B., Thoni, C., 2004. Trust, voluntary cooperation, and socio-economic background: survey and experimental evidence. Journal of Economic Behavior & Organization 55(4), 505-531.

Glaeser, E., Laibson, D., Scheinkman, J., Soutter, C., 2000. Measuring Trust. The Quarterly Journal of Economics 115 (3), 811-846.

Glaeser, E., Laibson, D., Sacerdote, B., 2002. An Economic Approach to Social Capital. The Economic Journal 112 (483), 437-458.

Goette, L., Huffman, D., Meier, S., 2006. The Impact of Group Membership on Cooperation and Norm Enforcment: Evidence using Random Assignment to Real Social Groups. American Economic Review 96(2), 212-216.

Gueth, Werner, Carsten Schmidt, and Matthias Sutter. 2007. “Bargaining Outside the Lab: A Newspaper Experiment of a Three-Person Ultimatum Game.” Economic Journal, Vol. 117, No. 518 (Mar), pp. 449-469.

Hedinger, B., 2004. Employability of Specialists in situations of structural change. A theoretical and empirical Analysis. Diplomathesis, University of Zurich.

Helmreich, R., Merritt, A., Wilhelm, J., 1999. The evolution of Crew Resource Management training in commercial aviation. International Journal of Aviation Psychology 9, 19-32.

Hoff, K., Pandey, P., 2006. Discrimination, Social Identity, and Durable Inequalities. American Economic Review 96(2), 206-211.

Hoffmann, E., McCabe, K., Smith, V., 1996. Social Distance and Other-Regarding Behavior in Dictator Games. The American Economic Review 86(3), 653-660.

Holm, S., 1979. A Simple Sequentially Rejective Multiple Test Procedure. Scandinavian Journal of Statistic 6, 65-70.

Kahneman, D., 2003. Maps of Bounded Rationality: Psychology for Behavioral Economics. The American Economic Review 93(5). 1449-1475.

Kandel, E., Lazear, E., 1992. Peer Pressure and Partnerships. The Journal of Political Economy 100(4). 801-817.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R., 1997. Trust in Large Organizations. The American Economic Review 87(2). 333-338.

Levine, D., 1998. Modeling Altruism and Spitefulness in Experiments. Review of Economic Dynamics 1. 593-622.

Manski, C., 2002. Identification of decision rules in experiments on simple games of proposal and response. European Economic Review 46, 880-891.

Mas, A., 2006. Pay, Reference Points, and Police Performance. Working Paper, University of California.

Noser, P., 2003. Untersuchung des Auswahlsystems für die Fliegerische Vorschulung des Bundes (FVS) aus dem Jahre 2000. Diplomathesis, University of Zurich.

Prendergast, C., 1999. The Provision of Incentives in Firms. Journal of Economic Literature 37(1). 7-63.

Romano, J., Wolf, M., 2005. Exact and Approximate Stepdown Methods for Multiple Hypothesis Testing. Journal of the American Statistical Association 100 (469). 94-108.

Ruffle, B., Sosis, R., 2006. Cooperation and the In-Group-Out-Group Bias: A Field Test on Israeli Kibbutz Members and City Residents. Journal of Economic Behavior & Organization 60 (2). 147-163.

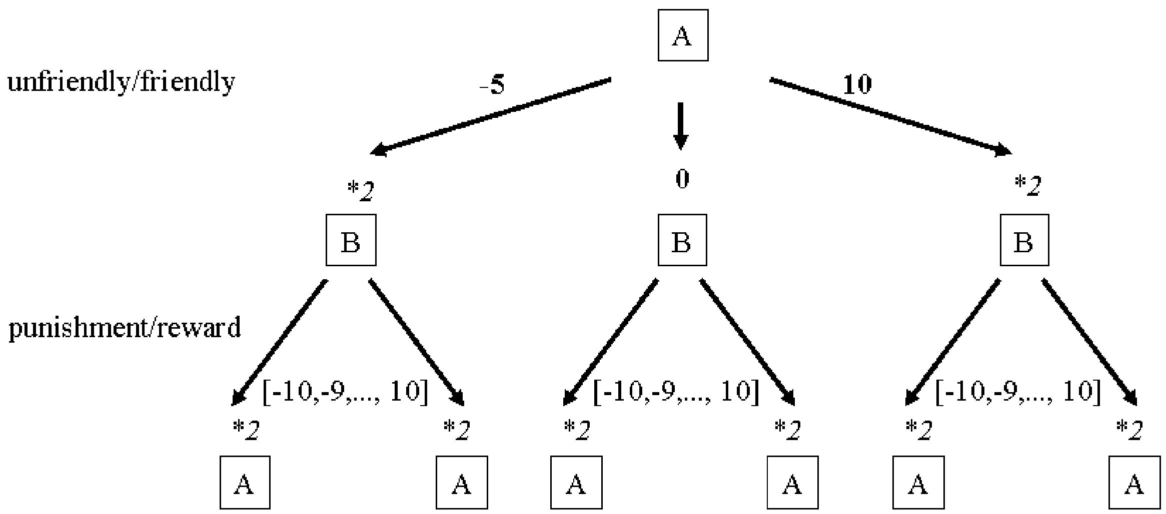
Schweizerische Unfallversicherungsanstalt Suva, 2004. Unfallstatistik UVG 1998-2002. SUVA, Luzern.

Slonim, R., Roth, A., 1998. Learning in High Stakes Ultimatum Games: An Experiment in the Slovak Republic. Econometrica 66(3), 569-596

Figures and tables

Figure 1

Experimental Design: Player A's choices vs Player B's strategic options



Examples: A takes 5, so B looses 10, then B punishes -10, he has to pay 10, so A looses 20.

A gives 10, so B gets 20, then B rewards +10, he has to pay 10, so A wins 20.

Figure 2

Treatments and Subjects

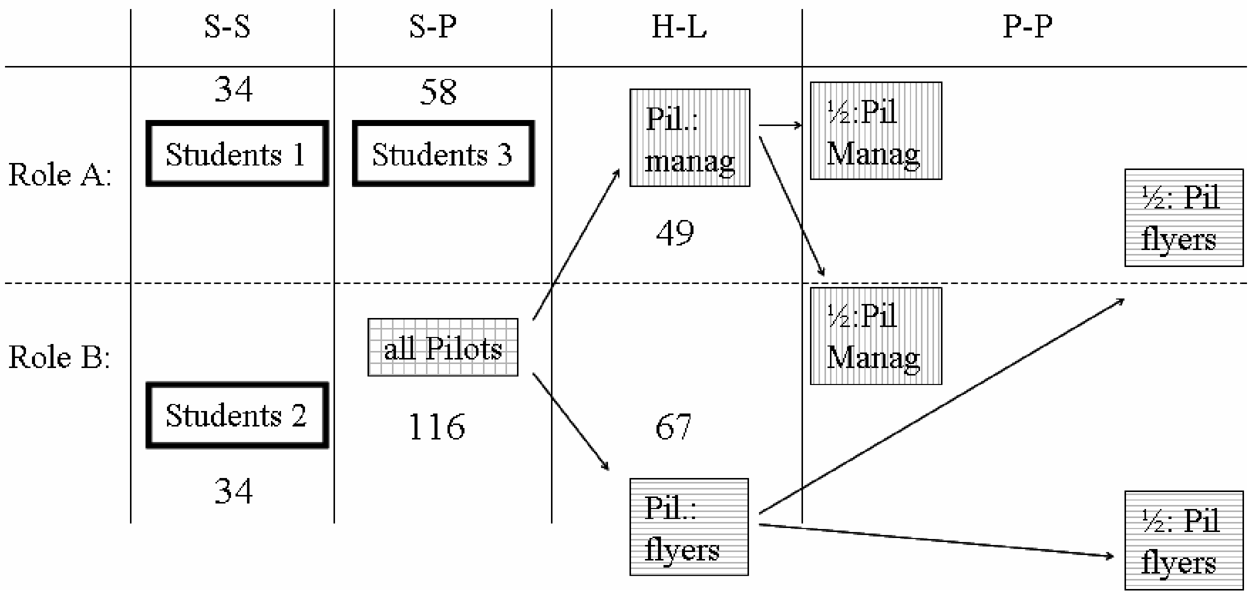


Figure 3

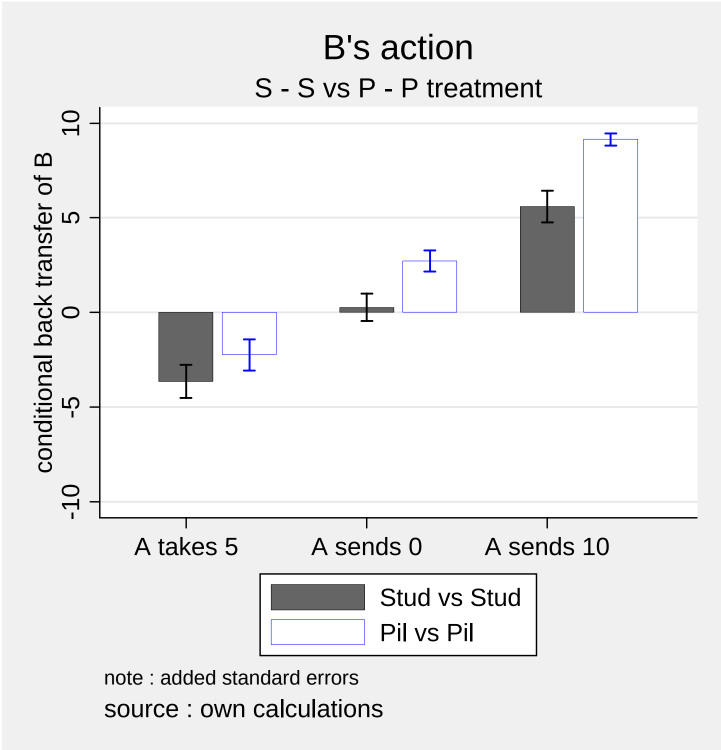


Figure 4

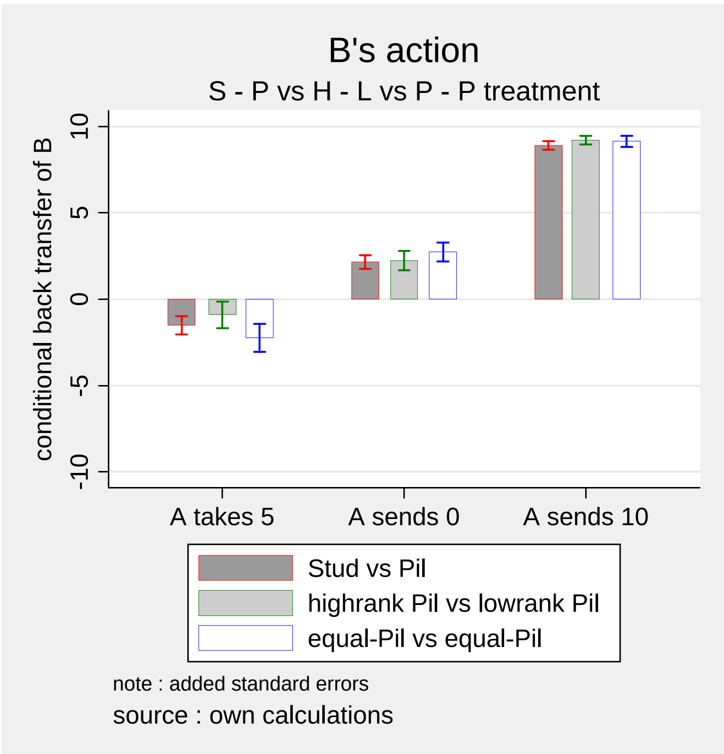


Figure 5:

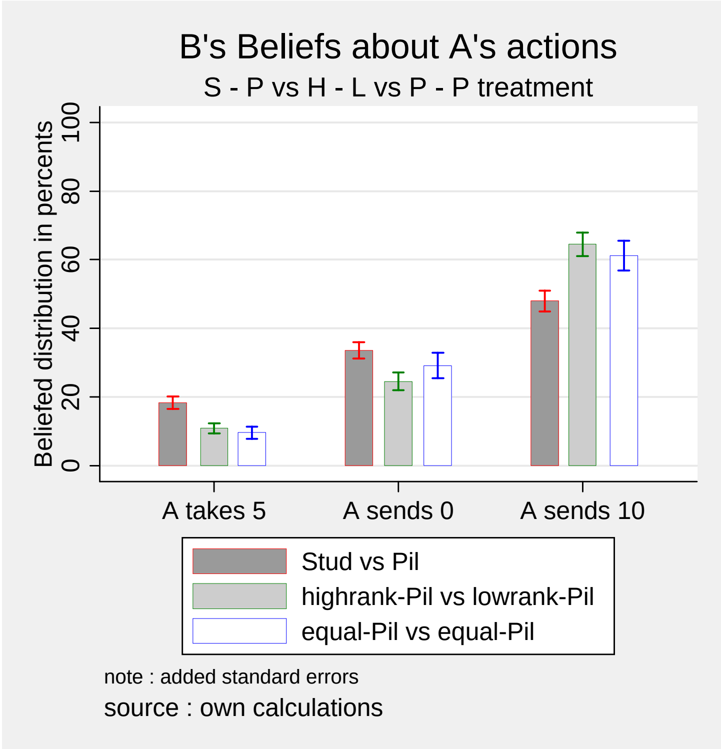


Figure 6

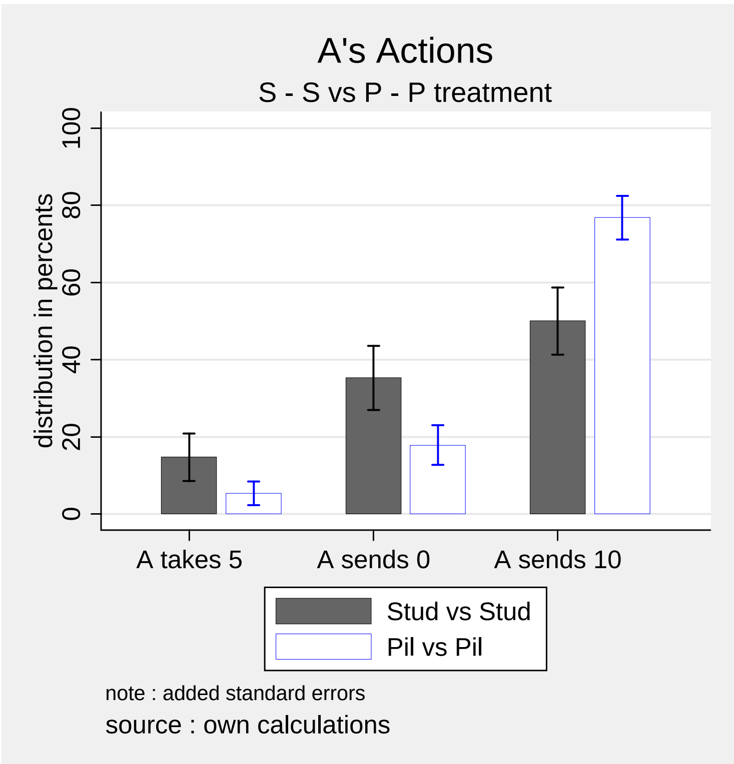


Figure 7

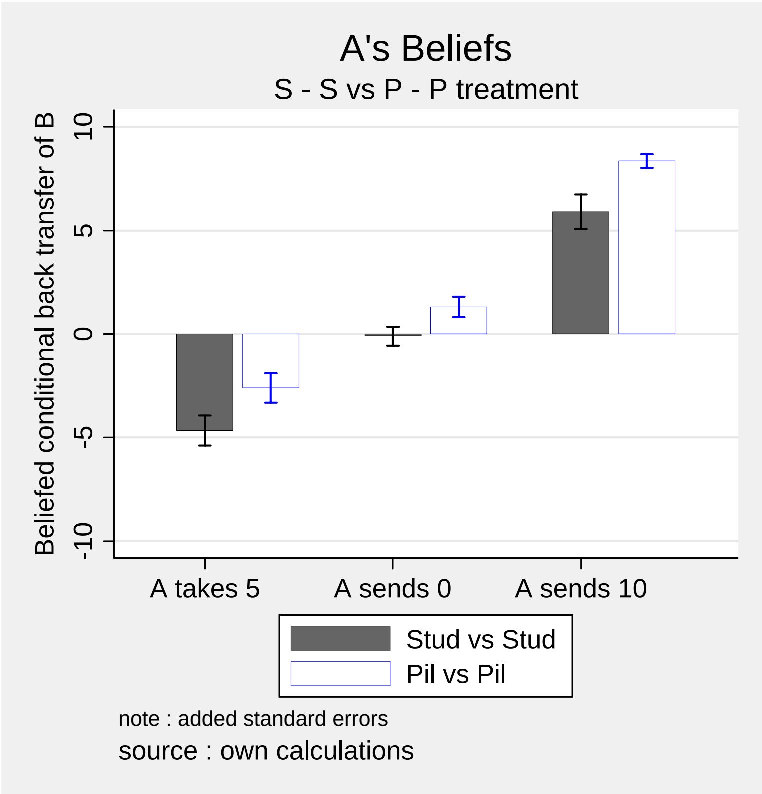


Figure 8

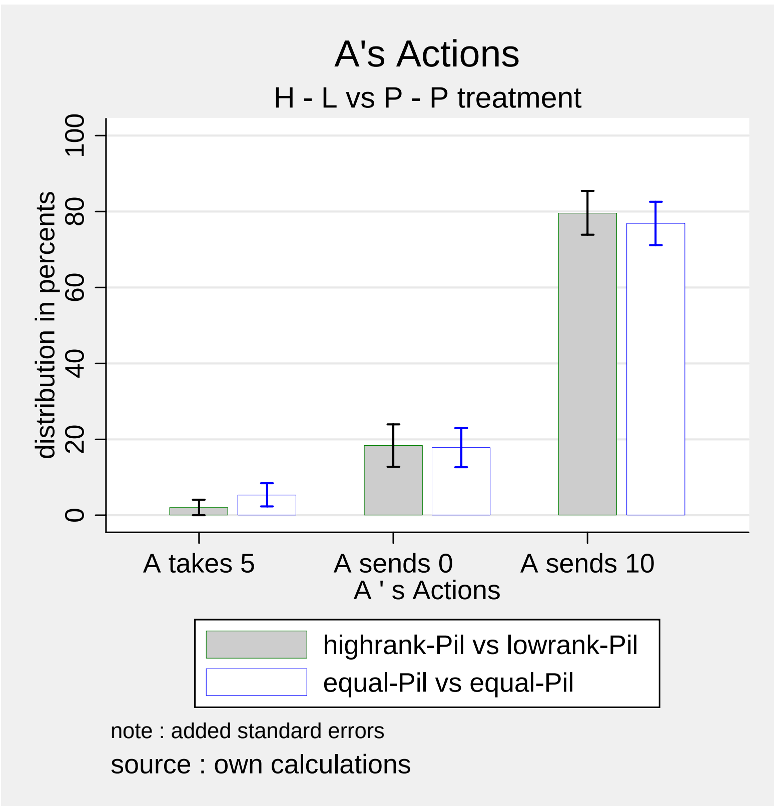


Table 1: Subjects characteristics: Pilots' and students' age, gender and education.

	pilots					students				
Variable	mean	min	max	sd	n	mean	min	max	sd	n
age	36.50	23	57	9.960	116	22.32	18	45	3.568	126
male	0.983	0	1	0.131	116	0.556	0	1	0.499	126
educ	12.71	12	19	1.552	116	12.5	12.5	12.5	0	126

Source: own calculations based on experimental evidence from December 2005.

Table 2: B's actions conditional different actions of A
(Comparison of S-S and P-P treatments)

	B's action conditional a negative action of A	B's action conditional a neutral action of A	B's action conditional a positive action of A
pilot	1.996	2.516	3.974###
	(2.212)	(1.719)	(1.081)***
male	-1.023	0.804	3.132
	(1.643)	(1.414)	(1.428)**
age	-0.476	-0.117	-0.555
	(0.508)	(0.348)	(0.265)**
age^2	-0.007	0.001	0.006
	(0.007)	(0.005)	(0.003)*
educ	-0.243	0.246	0.162
	(0.393)	(0.379)	(0.151)
const.	-6.928	-1.147	11.364
	(9.798)	(7.277)	(4.488)**
R^2	0.056	0.092	0.333
Prob > F	0.464	0.132	0.000
n	88	88	88

***Notes*: Dependent variable: B's transfer after different actions of A. Coefficients of OLS-regression (Robust standard errors in parentheses).**
***Level of significance:* *0.1<p, **0.01<p<0.05, ***p<0.01, corrected levels for multiple hypothesis (Holm, 1979): #0.1<p, ##0.01<p<0.05, ###p<0.01**
***Source:* own calculations based on experimental evidence from December 2005.**

Table 3: B's absolute spendings conditional different actions of A

(Comparison of S-S and P-P treatments)

	B's absolute spending cond. negative action of A	B's absolute spending cond. neutral action of A	B's absolute spending cond. positive action of A
pilot	-0.049	0.084	4.051###
	(1.561)	(1.500)	(1.053)***
male	2.560	0.674	2.550
	(1.373)	(1.166)	(1.294)*
age	-0.366	0.182	-0.596
	(0.357)	(0.345)	(0.261)**
age^2	0.005	-0.003	0.007
	(0.005)	(0.005)	(0.003)**
educ	-.062	0.213	0.168
	(0.381)	(0.387)	(0.150)
const.	9.654	-3.519	12.561
	(7.407)	(7.371)	(4.303)***
R^2	0.059	0.020	0.320
Prob > F	0.307	0.907	0.0000
n	88	88	88

***Notes*: Dependent variable: absolute amount spent by B after different actions of A. Coefficients of OLS-regression (Robust standard errors in parentheses).**
***Level of significance:* *0.1<p, **0.01<p<0.05, ***p<0.01**
corrected levels for multiple hypothesis (Holm, 1979):
#0.1<p, ##0.01<p<0.05, ###p<0.01
***Source:* own calculations based on experimental evidence from December 2005.**

Table 4: B's actions conditional different actions of A

(Comparison of P-P, S-P and H-L treatments)

	B's action conditional a **negative** action of A	B's action conditional a **neutral** action of A	B's action conditional a **positive** action of A
S-P	0.725	-0.482	-0.227
	(0.719)	(0.525)	(0.327)
H-L	0.284	-0.336	-0.073
	(0.386)	(0.253)	(0.148)
≥major	0.593	-0.394	-0.552
	(2.075)	(1.578)	(0.874)
male	-2.505	0.447	0.709
	(3.255)	(2.160)	(0.336)**
age	0.229	0.510	-0.007
	(0.457)	(0.321)	(0.167)
age^2	-0.005	-0.007	-0.000
	(0.006)	(0.004)*	(0.002)
educ	-0.069	0.328	0.231
	(0.289)	(0.288)	(0.088)***
const.	-0.544	-10.476	6.033
	(9.766)	(7.330)	(2.974)**
R^2	0.052	0.045	0.033
Prob > F	0.225	0.286	0.073
n	233	233	233
clusters	113	113	113
Prob >F for S-P and H-L	0.587	0.408	0.784

***Notes*: Dependent variable: B's transfer after different actions of A. Coefficients of OLS-regression (Robust standard errors adjusted for clustering on individuals in parentheses).**
***Level of significance:* *0.1<p, **0.01<p<0.05, ***p<0.01**
corrected levels for multiple hypothesis (Holm, 1979):
$^{\#}$0.1<p, $^{\#\#}$0.01<p<0.05, $^{\#\#\#}$p<0.01
***Source:* own calculations based on experimental evidence from December 2005.**

Table 5: B's beliefs about percentages of players A making different actions. (Comparison of P-P, S-P and H-L treatments)

	B's Beliefs about % of A making a **negative** transfer	B's Beliefs about % of A **not** transferring anything	B's Beliefs about % of A making a **positive** transfer
S-P	8.584###	4.729	-13.314###
	(1.922)***	(3.162)	(3.827)***
H-L	0.606	-0.038	-0.569
	(0.904)	(1.881)	(2.267)
≥major	-8.259	1.670	6.589
	(4.226)*	(8.935)	(10.239)
male	-5.247	10.530	-5.283
	(2.591)**	(3.045)***	(4.048)
age	-0.151	3.577	-3.426
	(1.261)*	(1.609)**	(2.098)
age^2	0.009	-0.039	0.030
	(0.017)*	(0.021)*	(0.028)
educ	0.771	0.112	-0.883
	(1.225)	(1.431)	(1.908)
const.	1.660	-57.558	155.898
	(21.575)	(32.947)*	(44.080)***
R^2	0.089	0.110	0.149
Prob > F	0.000	0.000	0.000
n	204	204	204
clusters	101	101	101

***Notes*: Dependent variable: B's belief about how many percents of A making different actions. Coefficients of OLS-regression (Robust standard errors adjusted for clustering on individuals in parentheses).**
***Level of significance:* *0.1<p, **0.01<p<0.05, ***p<0.01**
corrected levels for multiple hypothesis (Holm, 1979):
#0.1<p, ##0.01<p<0.05, ###p<0.01
***Source:* own calculations based on experimental evidence from December 2005.**

Table 6: A's actions / A's beliefs about B's back-transfer conditional different transfers of player A / A's actions according beliefs (Comparison of P-P and S-S treatments)

	A's action			A's action according Belief		A's Belief about B's back transfer					
	OLS		ordered Probit	OLS	ordered Probit	after a **negative** transfer		after a n**eutral** transfer		after a **positive** transfer	
P-P	3.467	1.298	0.293	-0.706	-0.218	2.550	1.722	1.975	2.008	$3.891^{\#\#}$	3.165
	(1.655)**	(1.896)	(0.445)	(1.875)	(0.476)	(1.539)	(1.743)	(1.032)*	(1.277)	(1.356)***	(1.847)*
male	0.796	0.977	0.228	1.501	0.473	0.784	0.853	-0.139	-0.142	-0.702	-0.641
	(2.223)	(2.110)	(0.402)	(2.027)	(0.419)	(1.334)	(1.336)	(0.714)	(0.733)	(1.456)	(1.487)
age	-0.053	0.809	0.183	0.777	0.222	-0.048	0.281	-0.033	-0.047	-0.082	0.206
	(0.065)	(0.494)	(0.124)	(0.499)	(0.128)*	(0.058)	(0.410)	(0.042)	(0.273)	(0.038)**	(0.344)
age^2		-0.011	-0.003	-0.010	-0.003		-0.004		0.000		-0.004
		(0.065)*	(0.002)	(0.007)	(0.002)*		(0.005)		(0.003)		(0.004)
educ	-0.006	0.021	0.027	-0.119	-0.039	-0.510	-0.499	-0.187	-0.187	0.239	0.248
	(0.500)	(0.383)	(0.089)	(0.520)	(0.117)	(0.338)	(0.299)*	(0.153)	(0.151)	(0.234)	(0.267)
A's belief after				0.490	0.093						
pos action				(0.264)*	(0.055)*						
A's belief after				0.382	0.131						
neut action				(0.263)	(0.077)*						
A's belief after				-0.182	-0.062						
neg action				(0.124)	(0.034)*						
cut1			3.131		3.962						
			(2.239)		(2.435)						
cut2			2.230		2.916						
			(2.206)		(2.432)						
const.	5.244	-8.475		-10.514		2.261	-2.972	3.039	3.252	5.109	0.521
	(7.091)	(9.119)		(9.833)		(4.427)	(7.060)	(1.797)*	(4.073)	(3.042)*	(6.700)*
(pseudo) R^2	0.076	0.106	0.063	0.266	0.171	0.068	0.073	0.055	0.055	0.143	0.150
Prob > F	0.180	0.158		0.010		0.115	0.163	0.173	0.232	0.010	0.004
Prob > Chi^2			0.219		0.045						
n	86	86	86	86	86	86	86	86	86	86	86
Prob >F				0.082							
Prob > Chi^2 for all Beliefs					0.048						

***Notes*: Dependent variable: First Column: A's action (-5, 0, 10), coefficients of OLS-regression. Second column: A's action, coefficient of ordered Probit-regression, third through fifth column: A's belief about B's back-transfer conditional different actions of A. Coefficients of OLS-regressions. Sixth column: A's transfer to B, coefficients of OLS-regression, seventh column: coefficient of ordered Probit-regression (Robust standard errors in parentheses).**

***Level of significance:* *0.1<p, **0.01<p<0.05, ***p<0.01 corr. levels for mult. hypo. (Holm, 1979): $^{\#}$0.1<p, $^{\#\#}$0.01<p<0.05, $^{\#\#\#}$p<0.01**

***Source:* own calculations based on experimental evidence from December 2005.**

Table 7: Useable actions and beliefs in treatment S-S

Student-Student	A	B
participants	34	34
missing actions	-	-
antireciprocical actions	-	1
useable actions	**34**	**33**
missing beliefs	2	1
beliefs adding up to more than 100%	-	4
useable beliefs	**32**	**28**

Source: own calculations based on experimental evidence from 12/05.

Table 8: Useable actions and beliefs in treatment S-P

Student-Pilot	A	B
participants	58	116
missing actions	-	-
antireciprocical actions	-	4
useable actions	**58**	**112**
missing beliefs	1	3
beliefs adding up to more than 100%	2	13
useable beliefs	**55**	**96**

Source: own calculations based on experimental evidence from 12/05.

Table 9: Useable actions and beliefs in treatment H-L

H-L	A	B
participants	49	67
missing actions	-	-
antireciprocical actions	-	1
useable actions	**49**	**66**
missing beliefs	-	-
beliefs adding up to more than 100%	2	7
useable beliefs	**47**	**59**

Source: own calculations based on experimental evidence from 12/05.

Table 10: Useable actions and beliefs in treatment P-P

Pilot-Pilot	A(H+L)	B(H+L)
participants	25+34	24+33
missing actions	1+2	1+0
antireciprocical actions	-	0+1
useable actions	**24+32**	**23+32**
missing beliefs	-	-
beliefs adding up to more than 100%	1+1	3+3
useable beliefs	**23+31**	**20+29**

Source: own calculations based on experimental evidence from 12/05.