

Does Overconfidence Lead to Bargaining Failures?*

Paola Colzani

University of Lausanne

Luís Santos-Pinto

University of Lausanne

February 15, 2021

Abstract

We use a laboratory experiment to study the causal impact of self-confidence on bargaining with joint production. We exogenously manipulate the self-confidence of subjects regarding their relative performance by employing easy and hard tasks. Subjects are randomly matched into pairs and each pair bargains over a joint surplus which can be either high or low. The size of the joint surplus depends on the pair's relative performance on the task. Our main experimental findings are as follows. First, the percentage of bargaining failures when subjects perform the easy task is more than triple than when they perform the hard task. Second, there is a remarkably high percentage of bargaining failures when subjects perform the easy task and bargain over a low surplus. Third, when subjects perform the easy task and bargain over a high surplus, all pairs reach an agreement and most settle on the equal split. Our findings shed light on the conditions and mechanisms under which overconfidence causes bargaining failures.

JEL Classification: C79; C91; C92; D91

Keywords: Overconfidence; Bargaining; Joint Production; Laboratory Experiment

*Corresponding author: Luís Santos-Pinto, University of Lausanne, Faculty of Business and Economics (HEC Lausanne), Internef 535, CH-1015, Lausanne, Switzerland, Ph: 41-216923658, Fax: 41-216923435, LuisPedro.SantosPinto@unil.ch; Paola Colzani, University of Lausanne, Faculty of Business and Economics (HEC Lausanne), Switzerland, Paola.Colzani@unil.ch

We are thankful to Jordi Brandts, Adrian Bruhin, Luís Cabral, Rustamdjan Hakimov, João Montez, Christian Thoni, and Christian Zehnder. We are also thankful to Tess Bardy and Noah Boren for their outstanding work as research assistants. This research was supported by grant #176446 of the Swiss National Science Foundation (SNSF).

Does Overconfidence Lead to Bargaining Failures?

1 Introduction

One of the enduring puzzles in bargaining is why there is disagreement in cases where both parties would appear to be better off reaching an agreement. For example, the “gas wars” between Russia and Ukraine often lead to breakdowns in transit or supply that are inefficient (Tingle 2015).

One prominent explanation for bargaining failures is asymmetric information (Fudenberg and Tirole 1981, Roth and Murnighan 1982, Sobel and Takahashi 1983, Bebchuk 1984, Kennan and Wilson 1993). Another explanation is judgment biases (Bazerman and Neal 1982, Neal and Bazerman 1983, Neal and Bazerman 1985, Farber and Bazerman 1989, Thompson and Loewenstein 1992, Babcock et al. 1995, Babcock and Loewenstein 1997).

If the parties involved in a legal dispute are mutually optimistic about their chances of winning in court they often fail to agree on a settlement (Priest and Klein 1984, Waldfogel 1995 and 1998, Farmer et al. 2004, Merlo and Tang 2019).¹ Disagreement might also occur due to self-serving biases which lead people to overestimate their own contributions to joint tasks. For example, married couples overestimate their individual contributions to various household tasks they are responsible for such as making breakfast, cleaning house, shopping for groceries, and caring for children (Ross and Sicoly 1979).²

Yet, another judgment bias that might explain bargaining failures is overconfidence. Experimental evidence from Psychology and Economics shows that most people are overconfident, that is, they tend to overestimate their absolute skills, overplace themselves relative to others, and overestimate the precision of their private information, estimates, and forecasts.³ Overconfidence affects behavior in goods markets (DellaVigna and Malmendier 2006, Grubb 2009) and in labor markets (Gervais and Goldstein 2007,

¹The parties’ excessive optimism about their bargaining power in the future is a prominent explanation for costly delays before reaching an agreement (Hicks 1932, Farber and Katz 1979, Shavell 1982, Yildiz 2011). Yildiz (2003) shows that under bilateral bargaining over a deterministic surplus, optimism cannot cause bargaining delays. However, Ali (2006) finds that optimism can lead to costly delays when there are more than two bargainers, the surplus is deterministic, and bargainers are extremely optimistic. More recently, Ortner (2013) shows that optimism can lead to costly delays under bilateral bargaining when the size of the surplus follows a stochastic process. Yildiz (2011) summarizes these results as follows “(...) optimism plays a subtle role in bargaining, and optimism alone cannot explain the bargaining delays.”

²Van den Steen (2004) shows that optimistic agents in a cooperative venture tend to attribute success to their own choice of action and failure to the action choice of their associate. Hence, optimistic and self-serving biases can be closely related.

³Moore and Healy (2008) distinguish between three types of overconfidence: (i) overestimation of one’s absolute skills or performance, (ii) overestimation of one’s relative skills or performance (overplacement or the “better-than-average” effect), and (iii) excessive confidence in the precision of one’s private information, estimates, and forecasts (overprecision or miscalibration). We use the term overconfidence in the sense of overplacement.

DellaVigna 2009, Santos-Pinto 2010, Spinnewijn 2013, Kőszegi 2014, Spinnewijn 2015, Santos-Pinto and de la Rosa 2020). Overconfidence plays a role in strategic decisions such as market entry (Camerer and Lovo 1999) and mergers and acquisitions by CEOs (Malmendier and Tate 2005).

This paper uses a laboratory experiment to investigate how overconfidence affects bargaining with joint production. Does overconfidence cause bargaining failures? Does it cause bargaining delays? If so, what are the mechanisms through which this happens? For example, does overconfidence cause disagreement in subjective entitlements and/or in opening proposals? Are overconfident bargainers less willing to make concessions?

In our experiment subjects start by performing a general knowledge quiz. To study the causal impact of self-confidence on bargaining with joint production we manipulate subjects' self-confidence. We do that using a between-subjects design that assigns randomly half of the subjects to an easy quiz and the other half to a hard quiz. The self-confidence manipulation is successful: subjects who perform the easy (hard) quiz are, on average, overconfident (underconfident).⁴ Next, we match subjects randomly into pairs. A pair's average rank in the quiz determines that pair's joint surplus which can be either low or high. A pair bargains over a low surplus when the pair's average rank in the quiz is higher than the average rank of their group, otherwise the pair bargains over a high surplus. Combining the two quiz treatments – easy and hard– with the two surplus size conditions – low and high– our experiment has a total of four conditions.

Subjects are informed about the size of their surplus and whether they are the better or worse performer in their pair. After this, each pair bargains anonymously over a computer network by sending proposals that consist of an amount for themselves and an amount for the partner. If an agreement is reached in 10 minutes the joint surplus is divided as agreed. If agreement is not reached by the end of the 10 minutes, subjects receive a zero payoff. This unstructured (or free-form) bargaining protocol follows Roth and Malouf (1979), Roth and Murnighan (1982), Gächter and Riedl (2005), and Karagözoğlu and Riedl (2014). Free-form bargaining allows us to analyze a rich set of layers of the bargaining process: agreements, opening proposals, concessions, and bargaining duration.

Our main findings are as follows. First, the percentage of bargaining failures when subjects take the easy quiz – 15% – is more than triple than when they take the hard quiz – 4%. Second, there is a remarkably high percentage of bargaining failures – 28% – when subjects take the easy quiz and bargain over a low surplus. Third, when subjects take the easy quiz and bargain over a high surplus, all pairs reach an agreement and most settle on an equal split.

The experiment also shows that overconfidence has a causal impact on bargaining dynamics. First,

⁴This manipulation uses the empirical finding that subjects tend to display overplacement in easy tasks and underplacement in hard tasks (Kruger 1999, Kruger and Dunning 1999, Krueger and Mueller 2002, Moore and Kim 2003, Moore and Healy 2008).

there is more disagreement in subjective entitlements in the easy quiz than in the hard quiz. Second, there is more disagreement in opening proposals when subjects take the easy quiz and bargain over a high surplus than when subjects take the hard quiz. Third, concessions are the lowest (highest) when subjects take the easy quiz and bargain over a low (high) surplus.

Our results shed light on the conditions and mechanisms under which overconfidence causes bargaining failures. Overconfident bargainers feel entitled to a larger share of the surplus regardless of its size. When the surplus being negotiated is low, overconfident bargainers display a relatively low level of disagreement in opening proposals but are reluctant to make concessions and often fail to reach an agreement. In contrast, when the surplus being negotiated is high, overconfident bargainers display a high level of disagreement in opening proposals but are able to make concessions until an agreement is reached. Interestingly, our main results are in line with Bénabou and Tirole's (2009) predictions about the effects of overconfidence on bargaining with joint production.⁵

Our study contributes to the experimental literature on how judgement biases affect bargaining behavior. As we have seen, optimistic and self-serving biases are a prominent explanation for bargaining failures. This literature is composed of experimental studies (Bazerman and Neal 1982, Neal and Bazerman 1983, Thompson and Loewenstein 1992, Neal and Bazerman 1985, Farber and Bazerman 1989, Babcock et al. 1995, Babcock and Loewenstein 1997) and empirical studies (Waldfogel 1995 and 1998, Farmer et al. 2004, Merlo and Tang 2019). For example, Neal and Bazerman (1985) investigate the effects of optimism on labor-management negotiations and find that optimistic negotiators are less likely to display concessionary behaviors and reach an agreement than realistic ones. Thompson and Loewenstein (1992) find that self-serving, biased attention to available information in a conflict affected the parties' perceptions of fairness and the length of a strike in a simulated labor dispute. In Babcock and Loewenstein (1997), subjects are assigned the role of plaintiff or defendant in a tort case. Subjects need to reach a settlement and are asked to make a prediction regarding the judicial awards to the parties involved in the tort case. Babcock and Loewenstein (1997) find that self-serving judgements of fairness lead to a discrepancy in predictions regarding the judicial award which, in turn, lead to a settlement impediment. Waldfogel (1998) attempts to distinguish one-sided asymmetric information from optimistic bias explanations using civil litigation data and finds support for the optimistic explanation. Farmer et al. (2004) analyze the causes of settlement failure using final-offer arbitration data from Major League Baseball and find that their results more consistent with an optimism model explanation than with one-sided asymmetric information models of arbitration.

Our study also contributes to the experimental literature on bargaining with joint production. In the studies cited in the previous paragraph, subjects bargain over an exogenous surplus and get to read

⁵We discuss Bénabou and Tirole's (2009) model and its predictions in Section 5.

a description of the case before or after they have been told which role they will play in it. Hence, these studies do not endogenize the actions that lead to bargaining and litigation. Yet, most real-life bargaining situations involve the division of a surplus that is produced by the bargainers themselves (Karagözoğlu 2004). For example, business partners deciding how to share the profits of a partnership, labor negotiating with management over wages and states bargaining with each other over the ownership and sale of strategic resources (e.g., gas, oil, water). As Karagözoğlu (2004) points out “From a purely standard theoretical point of view, whether the surplus is produced by the bargaining parties or not should not make a difference since the costs incurred by the bargaining parties due to the production of the pie (e.g., cost of effort, investment, contribution) are sunk at the time when they sit at the bargaining table and thus should not affect bargaining behavior.” However, experimental evidence shows that individuals’ valuations and bidding behavior are heavily influenced by sunk costs they previously incur (Phillips et al. 1991, Hackett 1993). Hence, knowing if overconfidence affects bargaining with joint production is important for understanding real-life negotiations.

Within the experimental literature on bargaining with joint production, the study that is closest to ours is Karagözoğlu and Riedl (2014). Karagözoğlu and Riedl (2014) analyze the impact of performance information and production uncertainties on bargaining over a jointly produced surplus. They manipulate performance information by telling subjects whether they are the better or the worse performer in their pair. The study finds that without performance information pairs tend to settle on the equal split while with performance information pairs reach asymmetric agreements in favor of the best in the pair and there are bargaining delays. Karagözoğlu and Riedl (2014) do not find bargaining failures. In stark contrast to Karagözoğlu and Riedl (2014), our study manipulates self-confidence and finds a remarkably high percentage of bargaining failures when subjects take the easy quiz and bargain over a low surplus.

Finally, our study contributes to the growing literature on the impact of overconfidence on strategic interactions. Overconfidence has been shown to have both damaging (Camerer and Lovo 1999; Malmendier and Tate 2005) and beneficial effects (Bénabou and Tirole 2002; Compte and Postlewaite 2004; Gervais and Goldstein 2007; Santos-Pinto 2010). Our study considers the impact of overconfidence on bargaining over a jointly produced surplus and identifies the conditions and mechanisms under which overconfidence leads to bargaining failures.

The rest of the paper is organized as follows. Section 2 describes the experimental design. Section 3 presents the research hypotheses. Section 4 reports the results. Section 5 discusses our findings. Section 6 concludes the paper.

2 Experimental Design and Procedures

The experiment consists of eight parts which are explained in detail below. First, subjects perform a general knowledge quiz. Second, the size of the joint surplus is determined and subjects are informed about it. Third, subjects' beliefs about own rank and partner's rank are elicited. Fourth, subjects are informed whether they are the better or worse performer in their pair. Fifth, subjects' subjective entitlements are elicited. Sixth, subjects bargain over the joint surplus. Seventh, subjects' risk preferences are elicited. Finally, subjects fill in a demographic questionnaire and are paid.

2.1 Quiz

The experiment is composed of two treatments which exogenously manipulate self-confidence using a general knowledge quiz. Half the subjects perform an easy quiz and the other half a hard quiz. Easy tasks induce overconfidence (overplacement) and hard tasks induce underconfidence (underplacement) due to the "hard-easy" effect (Kruger and Dunning 1999, Moore and Kim 2003, Moore and Small 2007, Moore and Healy 2008, Dargnies et al. 2019). Subjects believe they have lower (higher) ranks than their peers in easy (hard) tasks failing to realize that other subjects are facing the same level of difficulty. We label the two treatments EASY and HARD.

The quiz is composed of 46 questions which are divided into 6 different general knowledge topics: Science, Geography, Movies, Music, History, and Switzerland. We use questions from Moore and Healy (2008), update some of their questions, and add questions about Switzerland as we run the experiment at the University of Lausanne. Subjects have 20 minutes to complete the quiz. Groups are composed of 24 subjects. The number of correct answers in the quiz determines subjects' ranks within each group.⁶

2.2 Joint Surplus

After completing the quiz, subjects are randomly and anonymously matched into pairs. The surplus of each pair is either high or low. The high surplus is worth CHF 39 and the low surplus is worth CHF 20.⁷

The joint surplus of pair (i, j) , denoted S_{ij} , is obtained as follows:

$$S_{ij} = \begin{cases} 39 & \text{if } \frac{r_i+r_j}{2} \leq 12 \\ 20 & \text{if } \frac{r_i+r_j}{2} > 12 \end{cases} ,$$

⁶In case of a tie, subjects who gave the same number of correct answers were assigned the same rank. Overall our data set, only two ties among members of the same pair appeared hence the impact of ties is negligible.

⁷The experimental points for the high and low surpluses, 2710 and 1370, respectively, are identical to those in Karagözoğlu and Riedl (2014). However, payments in Karagözoğlu and Riedl (2014) are in Euros (€), while in our case, payments are in Swiss Francs (CHF). In Karagözoğlu and Riedl (2014), 2710 and 1370 points correspond to 17.6 € and 9.0 €, respectively. This amounts are lower than the ones in Swiss Francs we pay. Hence, points being constant, we pay higher amounts to comply with laboratory payments norms at University of Lausanne.

where r_i and r_j are the ranks of subjects i and j , respectively. Note that the average rank of a group of 24 subjects is equal to 12.5. Hence, if a pair's average rank in the quiz, $(r_i + r_j)/2$, is smaller (greater) than the average rank of the group, then the pair bargains over a high (low) surplus. This rule implies that the surplus is high when either both partners' ranks are above average or one partner's rank is above average and the other is not but the average of the partners' ranks is above average. The opposite is true for a low surplus. Importantly, the joint surplus is observable to the pair but exact individual contributions are not.

Hence, the experiment has two surplus size conditions which we label LOW and HIGH. Combining the two treatments with the two surplus size conditions we have a total of four conditions: EASY-LOW, EASY-HIGH, HARD-LOW and HARD-HIGH.

2.3 Elicitation of Beliefs about Own Rank and Partner's Rank

After subjects are informed about the surplus size, we ask them to estimate their own rank and their partner's rank on the quiz. This allows us to measure self-confidence with respect to the group and with respect to the partner. Self-placement with respect to the group, b_{ii} , is the difference between a subject's rank and her estimate of own rank:

$$b_{ii} = r_i - E_i(R_i|S_{ij} = s),$$

where r_i is subject i 's rank, $E_i(R_i|S_{ij} = s)$ is i 's estimate of own rank conditional on the pair's surplus. Self-placement with respect to the group is zero when a subject correctly estimates her own rank, is positive (there is overplacement with respect to group) if the estimate is lower than her real rank, and negative (there is underplacement with respect to the group) if the estimate is higher than her real rank.

Since partners' ranks matter for the surplus assignment and for the bargaining stage we measure how subjects compare themselves with respect to their partners as well. Indeed, beliefs regarding relative performance in the pair and hence regarding relative contributions to the jointly produced surplus are likely to influence bargaining behavior. Following Moore and Healy (2008), self-placement with respect to the partner, Δb_{ij} , is measured by

$$\Delta b_{ij} = b_{ii} - b_{ij} = [r_i - E_i(R_i|S_{ij} = s)] - [r_j - E_i(R_j|S_{ij} = s)],$$

where r_j is subject j 's rank (i 's partner), and $E_i(R_j|S_{ij} = s)$ is i 's estimate of j 's rank conditional on the pair's surplus.

The estimates $E_i(R_i|S_{ij} = s)$ and $E_i(R_j|S_{ij} = s)$ are elicited with binarized scoring rules (Hosain and Okui 2013). The binarized scoring rule induces truth telling irrespective of subject's EU risk preference and even if subjects are non-EU maximizers. This is the main advantage of the binarized

scoring rule over other scoring rules (e.g., the quadratic scoring rule) given the substantial evidence on heterogeneity in risk preferences (Hey and Orme 1994, Harless and Camerer 1994, Starmer 2000), which suggests that there is a majority of non-EU maximizers and a minority of EU maximizers (Bruhin et al. 2010, Conte et al. 2011, Bruhin et al. 2019).

Subject i 's payoff for her estimate of her own rank is

$$\Pi_{ii} = \begin{cases} 2 & \text{if } [E_i(R_i|S_{ij} = s) - r_i]^2 \leq k \\ 0 & \text{if } [E_i(R_i|S_{ij} = s) - r_i]^2 > k \end{cases},$$

where $E_i(R_i|S_{ij} = s) - r_i$ is i 's prediction error about her own rank, and k is a random number drawn from the uniform distribution with support on $[0, (n - 1)^2]$, where n is the number of subjects in the group. Hence, subject i earns CHF 2 for her estimate of own rank if her prediction error squared is lower than k and otherwise earns CHF 0.

Subject i 's payoff for her estimate of her partner's rank is:

$$\Pi_{ij} = \begin{cases} 2 & \text{if } [E_i(R_j|S_{ij} = s) - r_j]^2 \leq k \\ 0 & \text{if } [E_i(R_j|S_{ij} = s) - r_j]^2 > k \end{cases},$$

where $E_i(R_j|S_{ij} = s) - r_j$ is i 's prediction error about her partner's rank. Hence, subject i earns CHF 2 for her estimate of her partner's rank if her prediction error squared is lower than k and otherwise earns CHF 0.

2.4 Information about Relative Performance

After we ask subjects to estimate their own rank and their partner's rank, we inform them whether they are the better or worse performer in their pair. This information is a noisy signal about the individual contributions to the joint surplus. We provide this information to subjects because Karagözoğlu and Riedl (2014) show that without it most pairs agree on an equal split.

2.5 Elicitation of Subjective Entitlements

We define "subjective entitlement" as the fraction of joint surplus that a subject believes it is fair to keep for herself. For instance, if a subject believes that it is fair to equally split the joint surplus, her subjective entitlement is equal to 50%. To elicit subjects' subjective entitlements we use the formulation from Gächter and Riedl (2005) and ask subjects: "*According to your opinion, what would be a 'fair' distribution of the jointly produced surplus from the vantage point of a non-involved neutral arbitrator? (Please use exact amounts; no intervals! The amounts have to sum up to the jointly produced surplus!)*." Subjects are asked to enter on the computer screen which amount they think it is fair to keep for themselves and which amount they think it is fair to give to their partner. The two amounts need to sum up to

the joint surplus. We divide the amount they think it is fair to keep for themselves by the surplus to have percentage measure and thus be able to compare subjects who receive high and low surpluses.

2.6 Bargaining

We implement an unstructured (or free-form) bargaining protocol following Roth and Malouf (1979), Roth and Murnighan (1982), Gächter and Riedl (2005), and Karagözoğlu and Riedl (2014). An unstructured bargaining protocol has three main advantages compared to a structured one (Nash 1953, Rubinstein 1982). First, it allows us to analyze a rich set of layers of the bargaining process: agreements, opening proposals, concessions, and bargaining duration. Second, most bargaining in the world is unstructured, without the cut-and-dried rules of structured bargaining protocols. Third, it avoids exogenous first-mover effects.

However, with free-form bargaining, there is no precise theoretical prediction for what the bargaining dynamics and outcomes will be. Still, classical cooperative bargaining solutions (e.g. Nash 1950) are often employed to predict agreements in free-form bargaining. In our free-form bargaining protocol the disagreement point is symmetric (both players get zero). If bargainers have the same utility function over money, $u(0) = 0$, then the bargaining problem is symmetric and the Nash (1950) bargaining solution predicts an equal split.

2.7 Elicitation of Risk Preferences

Risk preferences can play a large role in bargaining (Murnighan et al. 1988). If the randomization of subjects into treatments is successful there are no systematic differences in risk preferences across treatments. Still, given the importance of risk preferences, we measure them using Crosetto and Filippin (2013)'s Bomb Risk Elicitation Task (BRET).⁸ Subjects earn up to CHF 2 in the task depending on their choice. The BRET and the payoff allocation are described in detail in Appendix H.

2.8 Subjects and Payments

The experiment was computerized and programmed with the software z-Tree (Fischbacher 2007) and conducted at the LABEX (Laboratory for Behavioral Experiments) in the University of Lausanne. Subjects were recruited via ORSEE (Greiner 2015).

⁸This task has five advantages compared to others. First, it allows to estimate both risk averse and risk seeking preferences very precisely. Second, it has a good trade-off between precision and comprehensibility. Third, it is defined on the gain domain and hence it does not suffer from loss aversion as a potential confound. Fourth, it does not provide endogenous reference points against which some outcomes could be perceived as losses. Fifth, it imposes a unique choice which prevents its results from being biased by violations of the Reduction Axiom.

In total 190 subjects participated in 8 randomized experimental sessions. We had 7 sessions with 24 participants each and one with 22 participants.⁹ 94 subjects participated in the EASY treatment and 96 subjects in the HARD treatment. In total we have 95 bargaining pairs. Sessions lasted on average 70 minutes. Most of the subjects were undergraduate students from different faculties at University of Lausanne and EPFL. At the end of the experiment subjects filled in a demographic questionnaire and were paid their earnings in cash individually and confidentially. The average earnings were CHF 28 including a show-up fee of CHF 10. The experimental instructions can be found in Appendix J.

3 Research Hypotheses

This section presents our hypotheses on how self-confidence affects bargaining outcomes and dynamics. We expect that overconfidence leads to bargaining failures when subjects negotiate over a jointly produced surplus. Hence, we expect more bargaining failures when subjects take the easy quiz than when they take the hard quiz.

Hypothesis 1: *There are more bargaining failures when subjects take the easy quiz than when they take the hard quiz.*

At the beginning of the bargaining stage, when subjects exchange their first proposals, we can observe initial disagreement. We use two measures of initial disagreement: subjective entitlements and opening proposals. We expect that overconfident subjects are more likely to disagree at the start of the bargaining stage. Hence, we expect more disagreement in subjective entitlements and opening proposals when subjects take the easy quiz than when they take the hard quiz.

Hypothesis 2: *Disagreement in subjective entitlements is higher when subjects take the easy quiz than when they take the hard quiz.*

Hypothesis 3: *Disagreement in opening proposals is higher when subjects take the easy quiz than when they take the hard quiz.*

During the bargaining stage, subjects may overcome initial disagreement by making concessions. We expect overconfident subjects to concede less than underconfident ones. Hence, we expect to observe less concessionary behavior when subjects take the easy quiz than when they take the hard quiz.

Hypothesis 4: *Concessions are lower when subjects take the easy quiz than when they take the hard quiz.*

Considering the existence of initial disagreement and the reluctance to concede part of the surplus in order to reach an agreement, we expect bargaining duration to be longer when subjects take the easy

⁹In the session with 22 participants we informed subjects that only 22 participants were in the room and explained that ranks would go from 1 to 22 and that neither the surplus assignment nor the payoff for belief elicitation would be modified.

quiz.

Hypothesis 5: *Bargaining duration is longer when subjects take the easy quiz than when they take the hard quiz.*

4 Results

This section presents our results and is organized as follows. Section 4.1 reports results on the self-confidence manipulation. Section 4.2 presents results on bargaining outcomes. Section 4.3 presents the results on bargaining dynamics.

4.1 Self-Confidence Manipulation

As expected, the easy quiz resulted in higher scores ($M = 30.7$ out of 46, $SD = 7.15$) than did the hard quiz ($M = 8.4$ out of 46, $SD = 4.95$). The mean percentage of correct answers in the easy quiz, 64%, is highly significant greater than in the hard quiz, 17.5% (p -value < 0.01 , 1-sided, t -test).¹⁰ As mentioned above, we measure overconfidence as self-placement with respect to the group and self-placement with respect to the partner.

4.1.1 Self-Placement with Respect to the Group

The top panel of Figure 1 depicts mean self-placement with respect to the group in the two treatments. The mean self-placement with respect to the group in the EASY treatment is equal to 2.04 and to -1.61 in the HARD treatment.¹¹ Hence, we have overplacement with respect to the group in the EASY treatment and underplacement with respect to the group in the HARD treatment.

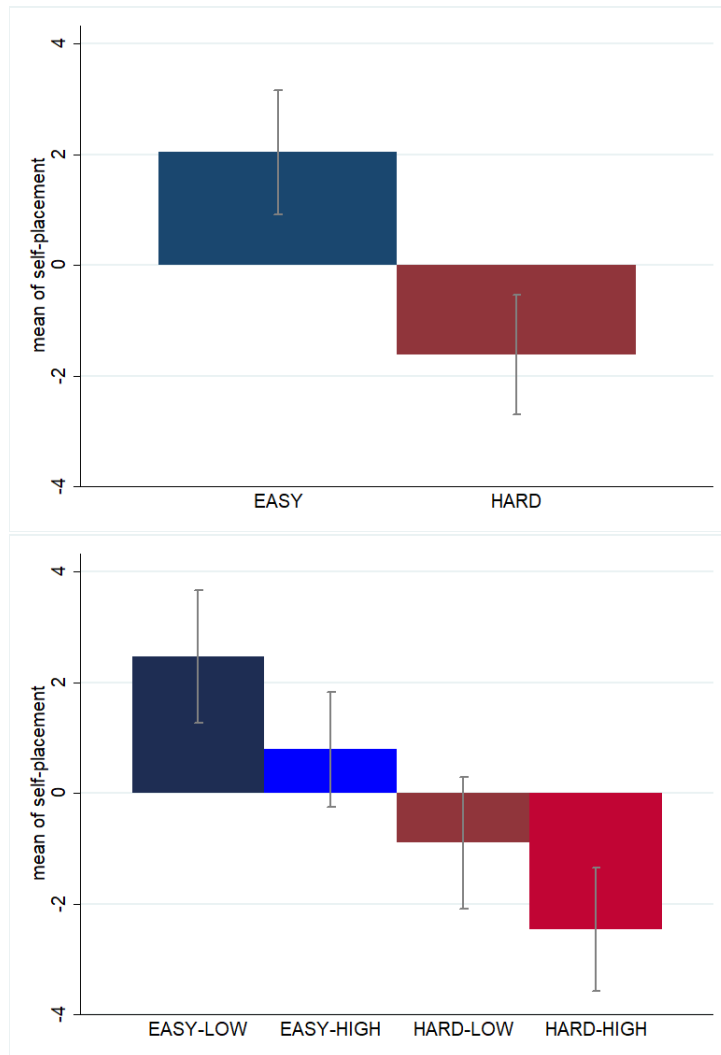
The distributions of self-placement with respect to the group in the two treatments are significantly different (p -value < 0.01 , 1-sided, Kruskal-Wallis test).¹²

¹⁰From now on “highly significant” means with a p -value less than 1%, “significant” means with a p -value less than 5%, and “weakly significant” means with a p -value less than 10%.

¹¹The mean estimated rank in the EASY treatment is equal to 10 and the mean estimated rank in the HARD treatment is equal to 14. The mean estimated rank in the EASY treatment is highly significant lower than the mean estimated rank in the HARD treatment (p -value < 0.01 , 1-sided, t -test). We find a highly significant differences in the distributions of estimated ranks in the two treatments (p -value < 0.01 , 1-sided Kruskal-Wallis test). The mean rank in the EASY treatment is equal to 12.04 and the mean rank in the HARD treatment is equal to 12.39. A 1-sided t -test and a 1-sided Kruskal-Wallis test highlight no differences in means nor in distributions in ranks.

¹²Note that self-placement with respect to the group can be compatible with Bayesian updating (Benoît and Dubra 2011). Appendix A shows that overplacement with respect to the group in the EASY treatment is incompatible with Bayesian updating whereas underplacement with respect to the group in the HARD treatment is compatible with Bayesian updating.

Figure 1: Mean self-placement with respect to the group



Note: The top panel shows the means of self-placement with respect to the group in the EASY and HARD treatments; the bottom panel shows the means of self-placement with respect to the group in the EASY and HARD treatments and across LOW and HIGH surpluses. Spikes depict 95% confidence intervals.

Result i: *There is overplacement with respect to the group when subjects take the easy quiz and underplacement when subjects take the hard quiz.*

Being informed whether they will bargain over a high or a low surplus is a signal that subjects may use to update beliefs about their ranks. Knowing the surplus is high, a subject might revise her estimate of own rank towards higher relative performance on the quiz (towards the top ranks). Knowing the surplus is low, a subject might revise her estimate of own rank towards lower relative performance on the quiz (towards the bottom ranks). This leads us to expect higher overplacement with respect to the group in the EASY-HIGH condition than in the EASY-LOW condition and lower underplacement with respect to the group in the HARD-HIGH condition than in the HARD-LOW condition. However, low

skill (bottom rank) subjects tend to overplace themselves whereas high skill (top rank) subjects tend to underplace themselves (Kruger and Dunning 1999). The way the joint surplus is determined implies that pairs that did worse on the quiz are more likely to bargain over a low surplus than pairs that did better on the quiz. Hence, a pair that bargains over a low surplus is more likely to be composed of partners who overplace themselves whereas a pair that bargains over a high surplus is more likely to be composed of partners who underplace themselves. Since this effect works in the opposite direction to the first it might well be that surplus size does not affect the self-confidence manipulation.

The bottom panel of Figure 1 depicts mean self-placement with respect to the group in the four conditions. Self-placement with respect to the group is, on average, equal to 2.80 in the EASY-LOW condition, to 1.18 in the EASY-HIGH condition, to -0.67 in the HARD-LOW condition, and to -2.28 in the HARD-HIGH condition.^{13,14} Hence, we observe overplacement with respect to the group in the EASY treatment and underplacement in the HARD treatment independently of surplus size.

4.1.2 Self-Placement with Respect to the Partner

The top panel of Figure 2 depicts the means of self-placement with respect to the partner in the two treatments. It shows that there is overplacement with respect to the partner in the EASY treatment and underplacement in the HARD treatment. The mean self-placement with respect to the partner is equal to 2.01 in the EASY treatment and to -2.29 in the HARD treatment. The distributions of self-placement with respect to the partner are highly significant different between the two treatments (p -value < 0.01 , 1-sided, Kruskal-Wallis test).

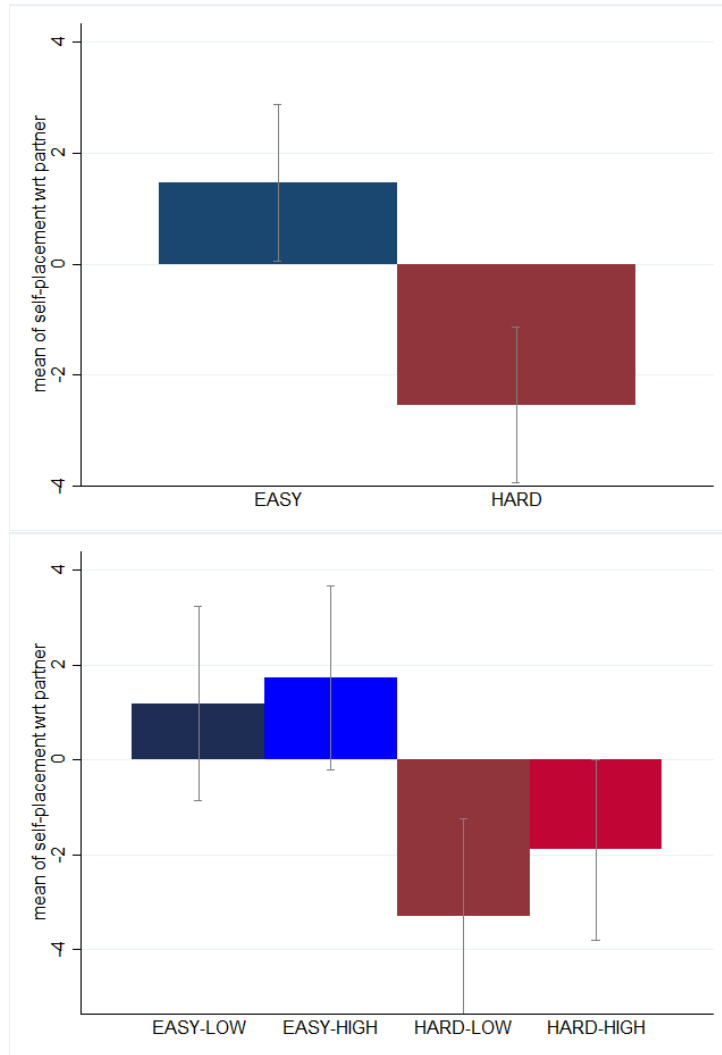
Result ii: *There is overplacement with respect to the partner when subjects take the easy quiz and underplacement with respect to the partner when subjects take the hard quiz.*

The bottom panel of Figure 2 depicts mean of self-placement with respect to the partner in the four conditions. Self-placement with respect to the partner is equal to 2.08 in the EASY-LOW condition, to 1.93 in the EASY-HIGH condition, to -3.7 in the HARD-LOW condition, and to -1.28 in the HARD-

¹³The mean estimated rank is equal to 12.48 in the EASY-LOW condition, to 7.18 in the EASY-HIGH condition, to 18.23 in the HARD-LOW condition, and to 10.98 in the HARD-HIGH condition. There are highly significant differences in means and distributions among all conditions (p -value < 0.01 , 1-sided, t -test; p -value < 0.01 , 1-sided, Kruskal-Wallis test in all cases besides the comparison among EASY-LOW and HARD-HIGH for which we find significant differences: p -value = 0.029, 1-sided, t -test; p -value = 0.031, 1-sided, Kruskal-Wallis test). The mean rank is equal to 15.28 in the EASY-LOW condition, to 8.36 in the EASY-HIGH condition, to 17.55 in the HARD-LOW condition, and to 8.69 in the HARD-HIGH condition. Further information on individual rank and pair composition can be found in Appendix F.

¹⁴There are highly significant differences among EASY-LOW and HARD-LOW, EASY-HIGH and HARD-HIGH, EASY-LOW and HARD-HIGH (p -value < 0.01 , 1-sided, Kruskal-Wallis test). The difference among HARD-LOW and HARD-HIGH is weakly significant (p -value = 0.09, 1-sided, Kruskal-Wallis test).

Figure 2: Mean of self-placement with respect to the partner



Note: The top panel shows the means of self-placement with respect to the partner in the EASY and HARD treatments; the bottom panel shows the means of self-placement with respect to the partner in the EASY and HARD treatments and across LOW and HIGH surpluses. Spikes depict 95% confidence intervals.

HIGH condition.¹⁵

Taken together, Results i and ii indicate that our manipulation worked: we observe overconfidence in the EASY treatment and underconfidence in the HARD treatment.

4.2 Bargaining Outcomes

This section reports results on how self-confidence affects bargaining outcomes.

To measure the percentage of pairs that failed to reach an agreement we use the variable “bargaining

¹⁵There are highly significant differences in distributions among EASY-LOW and HARD-LOW (p -value < 0.01 , 1-sided, Kruskal-Wallis test) and among EASY-HIGH and HARD-LOW (p -value < 0.01 , 1-sided, Kruskal-Wallis test) while the difference among EASY-HIGH and HARD-LOW is weakly significant (p -value = 0.08, 1-sided, Kruskal-Wallis test).

failure.” The variable is a dummy equal to 1 if a pair failed to reach an agreement within 10 minutes and equal to 0 otherwise. We have one observation per bargaining pair and hence a total of 95 observations.

The top panel of Figure 3 displays the percentage of bargaining failures in the EASY and HARD treatments. In the EASY treatment 15% of pairs fails to reach an agreement. This is highly significant greater than zero (p -value < 0.01 , 1-sided, t -test). In the HARD treatment only 4% of pairs fail to reach an agreement; this is weakly significant greater than zero (p -value = 0.08, 1-sided, t -test). Hence, the percentage of bargaining failures in the EASY treatment is more than the triple that in the HARD treatment. The percentage of bargaining failures in EASY treatment is weakly significant greater than in the HARD treatment (p -value = 0.076, 1-sided, Kruskal-Wallis test).

Result 1a: *The percentage of bargaining failures when subjects take the easy quiz—15%—is more triple than when they take the hard quiz—4%.*

Result 1a confirms our main hypothesis that self-confidence has a causal effect on bargaining outcomes. Subjects who take the easy quiz become, on average, overconfident and those who take the hard quiz become, on average, underconfident. This in turn, causes a higher percentage of bargaining failures among bargainers who take the easy quiz than among those who take the hard quiz. In other words, we find that overconfident bargainers often fail to reach an agreement whereas underconfident bargainers do not.

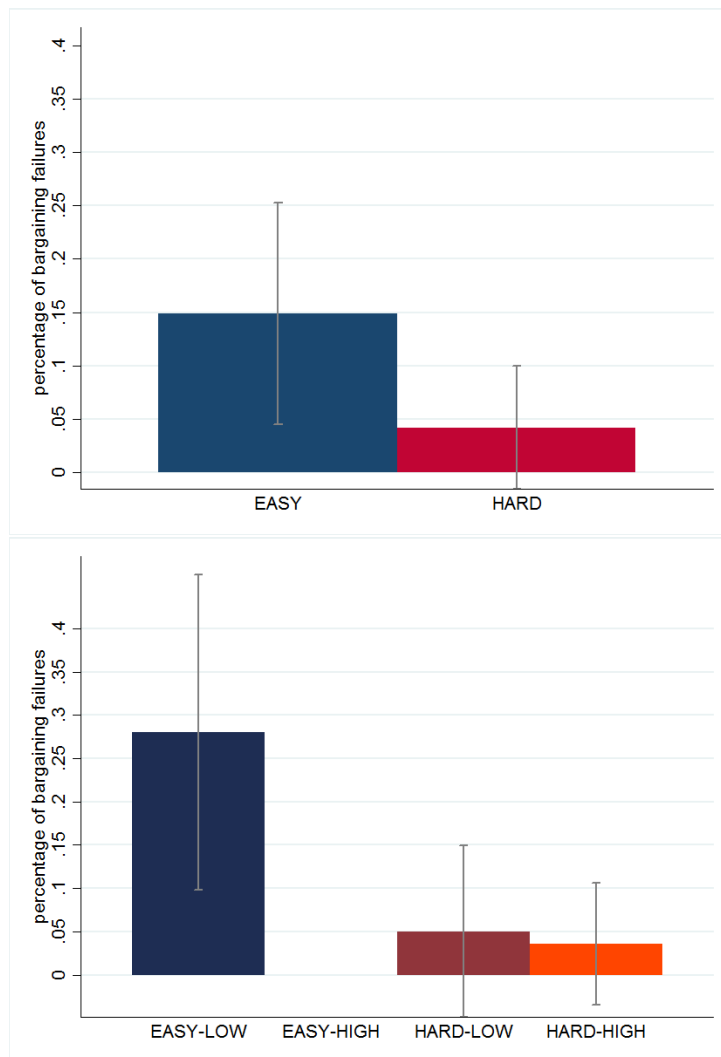
Next, we analyze whether surplus size affects bargaining failures. The bottom panel of Figure 3 displays the percentage of bargaining failures in the four conditions. The percentage of bargaining failures in the EASY-LOW condition (28%) is remarkably higher than the percentage of bargaining failures in the other three conditions. There are 0% bargaining failures in the EASY-HIGH condition, 5% in the HARD-LOW condition, and 3.6% in the HARD-HIGH condition. The percentage of bargaining failures in the EASY-LOW condition is significantly greater than zero (p -value < 0.01 , 1-sided, t -test) whereas this is not the case in the other conditions (p -value > 0.3 , 1-sided, t -test). In addition, the distribution of the percentage of bargaining failures in the EASY-LOW condition is significantly different from those of the other conditions (p -value < 0.05 , 1-sided, Kruskal-Wallis test).

Result 1b: *There is a remarkably high percentage of bargaining failures (28%) when subjects take the easy quiz and bargain over a low surplus.*

Result 1b tells us that overconfident bargainers who negotiate over a low surplus often fail to reach an agreement.

Finally, we analyze how pairs split the surplus across treatments. To avoid surplus size confounds we refer to percentages of the jointly produced surplus. Hence, the sum of the surplus splits of subject i and subject j in a pair is equal to 1. We have one observation per subject and a total of 190 observations.

Figure 3: Percentage of bargaining failures

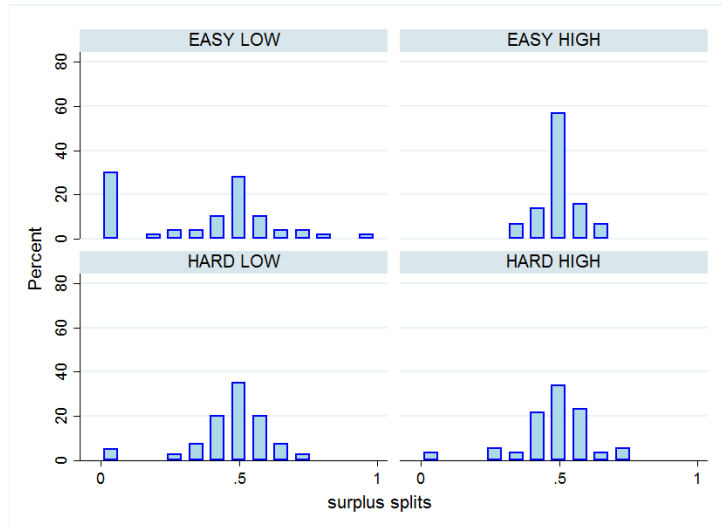


Note: The top panel shows the percentage of bargaining failures in the EASY and HARD treatments; the bottom panel shows the percentage of bargaining failures in the EASY and HARD treatments and across LOW and HIGH surpluses. Spikes depict 95% confidence intervals with standard errors clustered at pair level.

Figure 4 displays the distributions of surplus splits in the four conditions. We observe more unequal splits in the EASY-LOW condition than in the other three conditions. In the EASY-LOW condition the distribution of surplus splits is bimodal with the 28% of pairs not being able to reach an agreement and 28% settling on the equal split. The other three conditions show a unimodal distribution of surplus splits around the equal split.¹⁶ The percentage of pairs who settled for the equal split is equal to 54.5% in the EASY-HIGH condition, to 35% in the HARD-LOW condition and 28.5% in the HARD-HIGH

¹⁶There are significant differences in distributions among the EASY-LOW and the EASY-HIGH condition (p -value < 0.05, 1-sided, Kruskal-Wallis test) and among the EASY-HIGH and the HARD-HIGH condition (p -value < 0.05, 1-sided, Kruskal-Wallis test); the difference in distributions of surplus splits in the EASY-HIGH and the HARD-LOW condition is weakly significant (p -value = 0.074, 1-sided, Kruskal-Wallis test).

Figure 4: Surplus splits



Note: The graph shows the distribution of surplus splits in the EASY and HARD treatments across LOW and HIGH surpluses. Each band has 0.05 width.

condition.¹⁷ Hence, most of the agreements in the EASY-HIGH condition are on the equal split.

Result 1c: *When subjects take the easy quiz and bargain over a high surplus, all pairs reach an agreement and most settle on an equal split.*

Result 1c tells us that overconfident bargainers who negotiate over a high surplus are always able to reach an agreement and most settle on the equal split. Appendix B shows using an OLS regression with condition dummies that Results 1a, 1b and 1c are robust to gender, risk aversion, nationality, and other demographics we collected.¹⁸

4.3 Bargaining Dynamics

This section reports results on how self-confidence affects bargaining dynamics.

¹⁷Since subjects were bargaining over values and not percentages, it is possible that many pairs did not precisely settle on the equal split but in a neighborhood of it. Hence we intend for equal split the 5% neighborhood around the 50-50 split. For instance, in the EASY-HIGH condition the 7% of pairs settles exactly on the equal split but almost the 30% of pairs settles on a 5% neighborhood of the equal split. To take this into consideration, in figure 4 each band has 0.05 width.

¹⁸Conditional on agreement, subjects who are the better performer in their pair get, on average, 56% of the surplus while subjects who are the worse performer in their pair get, on average, only 44% of the surplus. There is a highly significant difference in the distribution of surplus splits for subjects who are the better or the worse performer in their pair (p -value < 0.01 , 1-sided, Kolmogorov-Smirnov test). This result is in line with Karagözoğlu and Riedl (2014). Appendix C reports an OLS regression with surplus splits as the dependent variable and condition dummies as well as demographic controls as explanatory variables. The regression confirms that subjects who are the better performer in their pair get, on average, a higher share of the surplus than subjects who are the worse in their pair.

Table 1: Disagreement in subjective entitlements

		LOW	HIGH	
EASY	m	1.15	1.16	1.16
	sd	(0.15)	(0.15)	(0.15)
	n	25	22	47
HARD	m	1.06	1.09	1.08
	sd	(0.13)	(0.12)	(0.12)
	n	20	28	48

4.3.1 Disagreement in Subjective Entitlements

To measure disagreement in subjective entitlements we sum the fraction of the joint surplus each partner believes it is fair to keep for herself. The sum goes from 0, when each partner believes it is fair to give the full joint surplus to the partner, to 2, when each partner believes it is fair to keep the full joint surplus for herself. When the sum is greater than 1 there is disagreement in subjective entitlements. When the sum is equal to 1 there is agreement.¹⁹ We have one observation per pair and a total of 95 observations.

Table 1 reports the means and standard deviations of disagreement in subjective entitlements in the EASY and HARD treatments and across HIGH and LOW surplus as well as the number of observations in each condition. The mean disagreement in subjective entitlements is highly significant greater than 1 in the EASY and HARD treatments (p -value < 0.01 , 1-sided, t -test). The mean disagreement in subjective entitlements is significantly greater in the EASY than in the HARD treatment (p -value < 0.01 , 1-sided, t -test) and the distributions of disagreement in subjective entitlements in the EASY and HARD treatments are significantly different (p -value = 0.012, 1-sided, Kruskal-Wallis test).

Result 2: *Disagreement in subjective entitlements is higher when subjects take the easy quiz than when they take the hard quiz.*

Result 2 suggests that overconfidence leads to disagreement in subjective entitlements. Overconfident bargainers feel entitled to a larger share of the surplus than underconfident bargainers. Result 2 is in line with our second hypothesis.

Next, we analyze whether surplus size affects disagreement in subjective entitlements. The mean disagreement in subjective entitlements is significantly greater than 1 in each of the four conditions (p -value < 0.05 , 1-sided, t -test). The distributions of disagreement in subjective entitlements in the EASY-HIGH and EASY-LOW conditions are not significantly different (p -value = 0.88, 1-sided, Kruskal-

¹⁹We sum fractions and not to amounts in order to avoid a surplus size confound.

Table 2: Disagreement in opening proposals

		LOW	HIGH	
EASY	m	1.24	1.31	1.27
	sd	(.29)	(.25)	(.27)
	n	25	22	47
HARD	m	1.18	1.18	1.18
	sd	(.21)	(.15)	(.18)
	n	20	28	48

Wallis test). This is also the case in the HARD-HIGH and HARD-LOW conditions (p -value = 0.19, 1-sided, Kruskal-Wallis test). Hence, we conclude that surplus size has no effect on disagreements in subjective entitlements.

Appendix D performs additional analysis on subjective entitlements. It discusses how being the best or the worst in a pair affects subjective entitlements. It reports a Tobit regression with surplus splits as a dependent variable and condition dummies as explanatory variables (and demographic controls). The regression shows that subjects who are the best in a pair are more likely to have higher subjective entitlements than subjects who are the worst in a pair.

4.3.2 Disagreement in Opening Proposals

At the beginning of the bargaining stage, each subject sends an opening proposal indicating which part of the surplus she is willing to keep for herself and which part of the surplus she is willing to give to her partner. To measure disagreement in opening proposals we sum the fraction of the joint surplus each partner wants to keep for herself in her opening proposal. The sum goes from 0 to 2. When the sum is higher than 1, there is disagreement in opening proposals.²⁰ We have one observation per pair and a total of 95 observations.

Table 2 reports the means and standard deviations of disagreement in opening proposals in the EASY and HARD treatments and across HIGH and LOW surpluses as well as the number of observations in each condition. The mean disagreement in opening proposals is statistically greater than 1 in the EASY and HARD treatments (p -value < 0.01, 1-sided, t -test). However, contrary to our fifth hypothesis, the distributions of disagreement in opening proposals in the EASY and HARD treatments are not significantly different (p -value = 0.13, 1-sided, Kruskal-Wallis test).

Result 3a: *There are no significant differences in disagreement in opening proposals across easy and*

²⁰Again, we refer to fractions and not to amounts in order to avoid surplus size confounds.

hard quizzes.

Next, we analyze whether surplus size affects disagreement in opening proposals. The mean disagreement in opening proposals is highly significant greater than 1 in each of the four conditions (p -value < 0.01 , 1-sided, t -test). The distributions of disagreement in opening proposals in the EASY-HIGH and EASY-LOW conditions are not significantly different (p -value = 0.14, 1-sided, Kruskal-Wallis test). Similarly, the distributions of disagreement in opening proposals in the HARD-HIGH and HARD-LOW conditions are not significantly different (p -value = 0.66, 1-sided, Kruskal-Wallis test). However, the distributions of disagreement in opening proposals in the EASY-HIGH and HARD-LOW conditions are significantly different (p -value < 0.05 , 1-sided, Kruskal-Wallis test) and weakly significant different in the EASY-HIGH and HARD-HIGH conditions (p -value = 0.06, 1-sided, Kruskal-Wallis test).

Result 3b: *Disagreement in opening proposals is higher when subjects take the easy quiz and bargain over a high surplus than when they take the hard quiz.*

The differences among disagreement in subjective entitlements and disagreement in opening proposals may stem from the fact that subjective entitlements are non-binding verbal statements while opening proposals can be binding if the partners accept them.

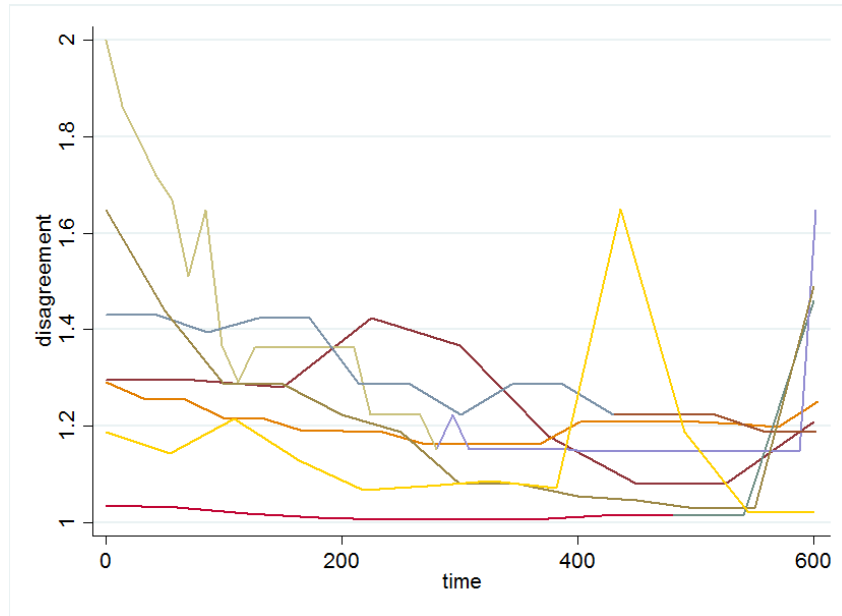
Figure 5 displays how disagreement in proposals evolves over time for the pairs who failed to reach an agreement in the EASY-LOW condition.²¹ As one can see, disagreement in proposals decreases during the first 5 minutes (300 seconds) and stays more or less flat between minutes 5 and 9. Interestingly, during the 10th minute, the disagreement in proposals of some pairs increases.

4.3.3 Concessions

We define as concession the difference between the fraction of the surplus a subject demands for herself in her opening proposal (independently of who was the first mover in a pair) and the fraction of the surplus the subject obtains. A concession is positive if a subject has to give up on part of her opening proposal in order to reach an agreement. A concession is negative if a subject obtains a higher fraction of the surplus than her opening proposal in order to reach an agreement. A concession is equal to zero if a subject belongs to a pair that did not reach an agreement. We have one observation per subject and a

²¹Note that a few assumptions were needed to realize this graph. First, since we implement free-form bargaining, one partner can send to the other multiple proposals without waiting for a counter offer. Whenever this happened, we computed disagreement as the sum of the current standing proposal of one partner and the former proposal of the other one. In practice, we interpret the absence of a new proposal as standing by the very same proposal (while the other partner is sending new offers). Second, we failed to record the exact timing of each single proposal. However, we could precisely record bargaining duration and the number of proposals sent. Hence we approximate the timing of each proposal simply dividing the bargaining duration (600 seconds in case of disagreement) by the number of proposals sent by each partner. The second assumption also implies a third one: that proposals were sent at regular time intervals.

Figure 5: Disagreement over time



Note: The graph shows the dynamics of disagreement for those pairs who failed to reach an agreement in the EASY-LOW condition. Each series refers to one pair.

total of 178 observations since 12 subjects accepted their partner's opening proposal (for their partners the concession is equal to zero since they obtained exactly what they asked for).

Table 3 reports the means and standard deviations of concessions in the EASY and HARD treatments and across HIGH and LOW surpluses as well as the number of observations in each condition. Mean concessions are highly significant greater than zero in the EASY and HARD treatments (p -value < 0.01 , 1-sided, t -test). However, contrary to our sixth hypothesis, the distributions of concessions in the EASY and HARD treatments are not significantly different (p -value = 0.71, 1-sided, Kruskal-Wallis test).

Result 4a: *There are no significant differences in concessions across easy and hard quizzes.*

Next, we analyze whether surplus size affects concessions. Mean concessions are highly significant greater than zero in each of the four conditions (p -value < 0.01 , 1-sided, t -test). Table 3 shows, in line with our seventh hypothesis, that concessions are the lowest in the EASY-LOW condition. There are significant differences in distributions among the EASY-LOW condition and the other three conditions (p -value < 0.05 , 1-sided, Kruskal-Wallis test). There are significant differences in distributions among the EASY-HIGH condition and the other three conditions (p -value < 0.05 , 1-sided, Kruskal-Wallis test).²²

²²We do not find significant differences among the HARD-LOW and HARD-HIGH condition (p -value = 0.89, 1-sided, Kruskal-Wallis test).

Table 3: Concessions

		LOW	HIGH	
EASY	m	6%	16%	11%
	sd	(16%)	(16%)	(17%)
	n	47	42	89
HARD	m	8%	9%	9%
	sd	(11%)	(11%)	(11%)
	n	38	51	89

Result 4b: *Concessions are the lowest (6%) when subjects take the easy quiz and bargain over a low surplus. Concessions are the highest (16%) when subjects take the easy quiz and bargain over a high surplus.*

Results 1, 2, 3, and 4 shed light on the conditions and mechanisms under which overconfidence leads to bargaining failures. Overconfident bargainers feel entitled to a larger share of the surplus regardless of its size. When the surplus being negotiated is low, overconfident bargainers display a relatively low level of disagreement in opening proposals but are reluctant to make concessions and often fail to reach an agreement. In contrast, when the surplus being negotiated is high, overconfident bargainers display a high level of disagreement in opening proposals but are able to make concessions until an agreement is reached.

4.3.4 Bargaining Duration

Bargaining duration the time that each pair takes to reach an agreement. Table 4 reports the means and standard deviations of bargaining duration in the EASY and HARD treatments and across LOW and HIGH surpluses. Contrary to our fifth hypotheses, the mean bargaining duration is not significantly different across easy and hard quizzes (and across low and high surpluses) (in all comparisons we observe p -value > 0.3 , 1-sided, Kruskal-Wallis test).

Result 5: *There are no significant differences in bargaining duration across easy and hard quizzes (and across high and low surpluses).*

Result 5 is surprising given the high percentage of bargaining failures in the EASY-LOW condition. To make sense of this result we analyzed bargaining duration across three bargaining outcomes: bargaining failure, agreement on the equal split, and agreement on other splits. The analysis shows that in the EASY-LOW condition, the longer bargaining time of pairs who fail to reach an agreement (600 seconds) is compensated by a shorter bargaining time of pairs who reach an agreement on the equal split

Table 4: Mean bargaining duration (in seconds)

		LOW	HIGH	
EASY	m	304	347	325
	sd	(249)	(216)	(232)
	n	25	22	47
HARD	m	302	301	301
	sd	(230)	(227)	(226)
	n	20	28	48

(10 seconds), and of pairs reach an agreement on other splits (188 seconds).

5 Discussion

This section discusses our findings and is organized as follows. Section 5.1 considers the implications of our main results for real-life bargaining. Section 5.2 shows that our main findings are in line with Bénabou and Tirole (2009). Section 5.3 argues that mood effects are not a plausible explanation for our findings. Section 5.4 rules out differences in pair-rank composition across conditions as an explanation for our results. Finally, Section 5.5 excludes explanations like gender and risk attitudes.

5.1 Implications for Real-Life Bargaining

Our main findings (Results 1a, 1b, and 1c) show that overconfidence can cause bargaining failures when the surplus that is produced by the bargainers themselves. This result has important implications for real-life bargaining as most of it concerns negotiations over jointly produced surpluses. In addition, we find that overconfidence only leads to bargaining failures when the surplus under negotiation is low. This suggests that overconfidence is more likely to lead to bargaining failures in real-life negotiations that do not involve high stakes.

5.2 Bénabou and Tirole (2009)

Bénabou and Tirole (2009) study the effects of overconfidence on bargaining with joint production. In their model, two matched agents decide whether to keep or destroy their match. If the match is kept, they bargain over their joint surplus. Bargaining is a standard Nash demand game (Nash 1953). Each agent has either low or high skill. The joint surplus is high if both agents have high skill and low otherwise. Importantly, agents are uncertain about their skills and individual contributions to the joint surplus are

imperfectly recalled. Finally, agents derive utility not only from the share of the joint surplus but also from having high beliefs about their skills (ego utility).

Bénabou and Tirole (2009) show that in a symmetric, pure-strategy Perfect Bayesian equilibrium, agents who bargain over a high surplus always reach an agreement and share the surplus equally. In contrast, agents who bargain over a low surplus only reach an agreement when the weight of ego utility is low. If the weight of ego utility is high, agents who bargain over a low surplus fail to reach an agreement. The intuition behind this result is as follows. Agreeing to inferior terms in a low surplus pair entails a loss in ego utility since it implies that at least one of the agents has low skill. When the surplus is low and ego utility concerns are important, the ego utility benefit from refusing to settle – maintaining a high self-confidence – is greater than the monetary cost of disagreement – the lost share of the low surplus.

Our experimental design is similar to Bénabou and Tirole (2009)'s model in three key aspects. First, subjects are uncertain about their relative performance on the quiz and therefore about their individual contributions to the joint surplus. Second, pairs are assigned a high or a low surplus. Pairs whose average rank is smaller (greater) than the average rank of their group bargain over a high (low) surplus. Hence, if a pair is assigned a high (low) surplus, it is more likely that both partners are high (low) skill. Third, subjects receive a noisy signal about their contributions to the joint surplus which mimics Bénabou and Tirole's (2009) assumption that agents imperfectly recall their contributions to the joint surplus.

If individuals have ego utility from believing they have a low rank, then our main findings are in line with Bénabou and Tirole's (2009) predictions. Our finding that bargaining failures are highest when subjects take the easy quiz and bargain over a low surplus is consistent with Bénabou and Tirole's (2009) intuition that overconfident bargainers might fail to reach an agreement when the surplus is low because the ego utility from refusing to settle – maintaining high self-confidence – is greater than the monetary cost of disagreement – the lost share of the low surplus. In addition, our finding that most pairs agree on the equal split when subjects take the easy quiz and bargain over a high surplus is consistent with Bénabou and Tirole's (2009) prediction that overconfident agents who bargain over a high surplus always reach an agreement and share the surplus equally.

5.3 Mood Effects

It has been shown that a positive mood increases negotiators' tendency to reach a cooperative bargaining outcome (Kramer et al. 1993, Forgas 1998) and their frequency of arriving at agreements that enhance joint gains (Barsade 2002). We did not control for mood so we cannot rule out that mood effects could influence some of our findings. However, we believe that it is hard to explain all of our findings with a mood effect.

Our experiment could potentially induce mood effects through either the type of quiz or the size of the surplus. The easy quiz could induce a positive mood and the hard quiz could induce a negative mood (or a neutral one). The high surplus could induce a positive mood and the low surplus could induce a negative mood (or a neutral one). If that were the case, then we would have observed the highest percentage of bargaining failures in the HARD-LOW condition which is the one that induces the lowest mood. However, this was not the case.

5.4 Differences in Pair-Ranks Composition Across Conditions

Our main results could be due to differences in pair-ranks composition across conditions. Let us assume that disagreement is more likely to happen the further apart are the ranks of the two subjects in a pair. In addition, let us assume that equal-splits are more likely to happen the closer are the ranks of the two subjects in a pair. If the ranks of the pairs in the EASY-LOW condition are far apart, then this could explain the higher percentage of bargaining disagreement in that condition. If the ranks of the pairs in the other conditions are close together, then this could explain the higher percentages of equal-splits in those conditions. Appendix F shows that this is not the case. We find no statistically significant differences in mean ranks, standard deviations of ranks, and rank distributions across the two treatments. Furthermore, we find that the distribution of the square of the difference of ranks is not significantly different across the four conditions

5.5 Gender and Risk Attitudes

Appendix G shows that our main results are not driven by gender differences. We find no gender differences in either bargaining outcomes or dynamics. Still, we find two gender differences. First, males perform slightly better in the quiz than females. Second, the percentage of risk averse females (77%) is significantly higher than that of risk averse males (55%).

Appendix H shows that our main results cannot be explained by differences in risk attitudes across conditions. We find no significant differences in the proportions of risk adverse, risk seeker, and risk neutral subjects across the four conditions. Appendix I shows that the randomization of subjects across treatments was successful.

6 Conclusion

This experiment shows that self-confidence has a causal effect on bargaining with joint production. Subjects who take the easy quiz become, on average, overconfident and those who take the hard quiz become, on average, underconfident. This in turn, causes a higher percentage of bargaining failures

among bargainers who take the easy quiz than among those who take the hard quiz. Hence, we show that overconfident bargainers often fail to reach an agreement whereas underconfident bargainers do not.

Furthermore, we find that in all cases where overconfident bargainers fail to reach an agreement, the surplus under negotiation was low. The bargaining dynamics clarify why this is the case. Overconfident bargainers who negotiate over a low surplus are reluctant to make concessions and, as a consequence, often fail to reach an agreement. In contrast, overconfident bargainers who bargain over a high surplus are willing to make concessions and reach an agreement.

This study contributes to the literature on bargaining by being the first to look at the causal effect of self-confidence on bargaining with joint production. The remarkably high level of bargaining failures that we find when subjects take the easy quiz and bargain over a low surplus stands in stark contrast with previous experiments on bargaining with joint production which hardly find bargaining failures.

References

- Ali, S. Nageeb M. (2006). "Waiting to Settle: Multilateral Bargaining with Subjective Biases," *Journal of Economic Theory*, Vol. 130, 109-137.
- Babcock, Linda, George Loewenstein, Samuel Issacharoff, and Colin Camerer (1995). Biased Judgments of Fairness in Bargaining, *American Economic Review*, Vol. 85, No. 5, 1337-1343.
- Babcock, Linda, and George Loewenstein (1997). "Explaining Bargaining Impasse: The Role of Self-serving Biases," *Journal of Economic Perspectives*, Vol. 11, 109-126.
- Barsade, S. (2002). "The Ripple Effect: Emotional Contagion and Its Influence on Group Behavior," *Administrative Science Quarterly*, Vol. 47, 644-675.
- Bazerman, Max H., and M. A. Neal (1982). "Improving Negotiator Effectiveness under Final Offer Arbitration: The Role of Selection and Training," *Journal of Applied Psychology*, Vol. 67, 543-548.
- Bebchuk, Lucian A. (1984). "Litigation and Settlement Under Imperfect Information," *The RAND Journal of Economics*, Vol. 15, No. 3, 404-415.
- Bénabou, Roland, and Jean Tirole (2002). "Self confidence and personal motivation," *The Quarterly Journal of Economics*, Vol. 99, No. 2, 459-65.
- Bénabou, Roland, and Jean Tirole (2009). "Over my dead body: Bargaining and the price of dignity," *American Economic Review*, Vol. 117, No. 3, 871-915.
- Benoît, J. P., and J. Dubra (2011). "Apparent overconfidence," *Econometrica*, Vol. 79, No. 5, 1591-1625.
- Bruhin, A., Fehr-Duda, H., and T. Epper (2010). "Risk and Rationality: Uncovering Heterogeneity in Probability Distortion," *Econometrica*, Vol. 78, No. 4, 1375-1412.

- Bruhin, A., Manai, M., and L. Santos-Pinto (2019). "Risk and Rationality: The Relative Importance of Probability Weighting and Choice Set Dependence," Working Paper, University of Lausanne.
- Burks, S. V., Carpenter J. P, Goette L., and A. Rustichini (2013) "Overconfidence and social signalling," *The Review of Economic Studies* , Vol. 80, No. 3, 949â€“983.
- Camerer, C., and D. Lovo (1999). "Overconfidence and excess entry: An experimental approach," *American Economic Review*, Vol. 89, No. 1, 306-318.
- Compte, O., and A. Postlewaite (2004). "Confidence-enhanced performance," *American Economic Review*, Vol. 94, No. 5, 1536-1557.
- Conte, A., Hey, J. D., and P. G. Moffatt (2011). "Mixture Models of Choice under Risk," *Journal of Econometrics*, Vol. 162, No. 1, 79-88.
- Crosetto, P., and A. Filippin (2013): "The 'bomb' elicitation task," *Journal of Risk and Uncertainty*, Vol. 47, No. 1, 31-65.
- Dargnies, M., Hakimov, R., and D. Kübler (2019): "Self-confidence and unraveling in matching markets," *Management Science*, Vol. 65, No. 12, 5603-5618.
- DellaVigna, S., and U. Malmendier (2006): "Paying Not to Go to the Gym," *American Economic Review*, Vol. 96, No. 3, 694-719.
- DellaVigna, S. (2009). "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature*, Vol. 47, No. 2, 315-372.
- Eckel, C. C., and Grossman, P. J. (2008). "Men, women and risk aversion: Experimental evidence, " *Handbook of Experimental Economic Results*, Vol. 1, 1061-1073.
- Farber, Henry S., and Harry C. Katz (1979). "Interest Arbitration, Outcomes, and the Incentive to Bargain," *Industrial and Labor Relations Review*, Vol. 33, No. 1, 55-63.
- Farber, Henry S., and Max H. Bazerman (1989). "Divergent Expectations as a Cause of Disagreement in Bargaining: Evidence From a Comparison of Arbitration Schemes," *The Quarterly Journal of Economics*, Vol. 104, No. 1, 99-120.
- Farmer, Amy, Paul Pecorino, and Victor Stango (2004). "The Causes of Bargaining Failure: Evidence from Major League Baseball," *The Journal of Law & Economics*, Vol. 47, No. 2, 543-568.
- Filippin, A., and P. Crosetto (2016). "A reconsideration of gender differences in risk attitudes," *Management Science*, Vol. 62, No. 11, 3138-3160.
- Fischbacher, U. (2007). "z-tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, Vol. 10, No. 2, 171-178.
- Forgas, J. P. (1998). "On Feeling Good and Getting Your Way: Mood Effects on Negotiator Cognition and Bargaining Strategies," *Journal of Personality and Social Psychology*, Vol. 74, No. 3, 565-577.

- Fudenberg, Drew, and Jean Tirole (1981). "Sequential Bargaining with Incomplete Information," *The Review of Economic Studies*, Vol. 50, 221-248.
- Gächter, Simon, and Arno Riedl (2005). "Moral Property Rights in Bargaining with Infeasible Claims," *Management Science*, Vol. 51, No. 2, 249-263.
- Gervais, S., and I. Goldstein (2007). "The positive effect of biased self-perceptions in firms," *Review of Finance*, Vol. 11, No. 3, 453-496.
- Greiner, B. (2015). "Subjects pool procedures: Organizing experiments with ORSEE," *Journal of the Economic Science Association*, Vol. 1, No. 1, 114-125.
- Grubb, M. D. (2009). "Selling to Overconfident Consumers," *American Economic Review*, Vol. 99, No. 5, 1770-1807.
- Hackett, S. C. (1993). "Incomplete Contracting: A Laboratory Experimental Analysis," *Economic Inquiry*, Vol. 31, No. 2, 274-297.
- Harless, D. W., and Colin F. Camerer (1994). "The Predictive Utility of Generalized Expected Utility Theories," *Econometrica*, Vol. 62, No. 6, 1251-1289.
- Hey, John D., and C. Orme (1994). "Investigating Generalizations of Expected Utility Theory Using Experimental Data," *Econometrica*, Vol. 62, No. 6, 1291-1326.
- Hicks, John (1932). *The Theory of Wages*. New York: Macmillan.
- Hossain, T., and R. Okui (2013). "The binarized scoring rule," *Review of Economic studies*, Vol. 80, No. 3, 984-1001.
- Johansson Stenman, O. and Nordblom, K. (2010). "Are men really more overconfident than women? A natural field experiment on exam behavior," Working Paper, University of Gothenburg.
- Karagözoğlu, Emin (2004). "Bargaining Games with Joint Production," Oxford Handbook of Economic Conflict Resolution, (eds. R. Croson and G. Bolton), September 2012, Ch. 24, 359-372, Oxford University Press, USA.
- Karagözoğlu, E., and A. Riedl (2014). "Performance information, production uncertainty, and subjective entitlements in bargaining," *Management Science*, Vol. 61, No. 11, 2611-2626.
- Kennan, John, and Robert Wilson (1993). "Bargaining with Private Information," *Journal of Economic Literature*, Vol. 31, No. 1, 45-104.
- Kőszegi, Botond (2014). "Behavioral Contract Theory," *Journal of Economic Literature*, Vol. 52, No. 4, 1075-1118.
- Kramer R. M., E. Newton, and P. L. Pommerenke (1993). "Self-enhancement biases and negotiation judgment: effects of self-esteem and mood," *Organizational Behavior and Human Decision Processes*, Vol. 56, 110-133.

- Kruger, Justin, and David Dunning (1999). "Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-assessments," *Journal of Personality and Social Psychology*, Vol. 77, No. 6, 1121-1134.
- Krueger, Joachim, and Ross A. Mueller (2002). "Unskilled, Unaware, or Both? The Better-than-average Heuristic and Statistical Regression Predict Errors in Estimates of Own Performance," *Journal of Personality and Social Psychology*, Vol. 82, No. 2, 180-188.
- Kruger, Justin (1999). "Lake Wobegon Be Gone! The 'Below-average effect' and the Egocentric Nature of Comparative Ability Judgments," *Journal of Personality and Social Psychology*, Vol. 77, No. 2, 221-232.
- Malmendier, U., and G. Tate (2005). "Ceo confidence and corporate investment," *The Journal of Finance*, Vol. 60, No. 6, 2661-2700.
- Merlo, Antonio, and Xun Tang (2019). "Bargaining with Optimism: Identification and Estimation of a Model of Medical Malpractice Litigation," *International Economic Review*, Vol. 60, No. 3, 1029-1061.
- Moore, Don A., and D. A. Small (2007). "Error and bias in comparative judgment: on being both better and worse than we think we are," *Journal of Personality and Social Psychology*, Vol. 85, No. 6, 1121-1135.
- Moore, Don A., and P. J. Healy (2008). "The trouble with overconfidence," *Psychological Review*, Vol. 115, No. 2, 502.
- Moore, Don A., and T. G. Kim (2003). "Myopic Social Prediction and the Solo Comparison Effect," *Journal of Personality and Social Psychology*, Vol. 85, No. 6, 1121-1135.
- Murnigan, J. K., Roth, A. E., and F. Schoumaker (1988). "Risk aversion in bargaining: An experimental study," *Journal of Risk and Uncertainty*, Vol. 1, No. 1, 101-124.
- Nash, John F. (1950). "The Bargaining Problem," *Econometrica*, Vol. 18, 155-162.
- Nash, John F. (1953). "Two-person Cooperative Games," *Econometrica*, Vol. 21, 128-140.
- Neal, M. A., and Max H. Bazerman (1983). "The Effect of Perspective Taking Ability under Alternate Forms of Arbitration on the Negotiation Process," *Industrial and Labor Relations Review*, Vol. 36, 378-388.
- Neal, M. A., and Max H. Bazerman (1985). "The Effects of Framing and Negotiator Overconfidence on Bargaining Behaviors and Outcomes," *The Academy of Management Journal*, Vol. 28, No. 1, 34-49.
- Niederla, M. A., and L. Vesterlund (2007). "Do women shy away from competition?," *The Quarterly Journal of Economics*, Vol. 122, No. 3, 1067-1101.
- Ortner, Juan (2013). "Optimism, Delay and (In)efficiency in a Stochastic Model of Bargaining," *Games and Economic Behavior*, Vol. 77, 352-366.

- Park, Y. J., and L. Santos-Pinto (2010). "Overconfidence in tournaments: Evidence from the field," *Theory and Decision*, Vol. 69, No. 1, 144-166.
- Pulford, B. D., and A. M. Colman (1997). "Overconfidence: Feedback and item difficulty effects," *Personality and individual differences*, Vol. 23, No. 1, 125-133.
- Phillips, O. R., R. C. Battalio, and C. A. Kogut (1991). "Sunk and Opportunity Costs in Valuation and Bidding," *Southern Economic Journal*, Vol. 58, No. 1, 835-856.
- Priest, George L., and Benjamin Klein (1984). "The Selection of Disputes for Litigation," *Journal of Legal Studies*, Vol. 13, 1-56.
- Ross, Michael, and Fiore Sicoly (1979). "Egocentric Biases in Availability and Attribution," *Journal of Personality and Social Psychology*, Vol. 37, No. 3, 322-336.
- Roth, Alvin E., and Michael W. K. Malouf (1979). "Game Theoretical Models and the Role of Information on Bargaining," *Psychological Review*, Vol 86, No. 6, 574-594.
- Roth, Alvin E., and J. Keith Murnighan (1982). "The Role of Information in Bargaining: An Experimental Study," *Econometrica*, Vol. 50, No. 5, 1123-1142.
- Rubinstein, A. (1982). "Perfect equilibrium in a bargaining model," *Econometrica*, Vol. 50, No. 1, 97-109.
- Santos-Pinto, Luís (2010). "Positive Self-Image in Tournaments," *International Economic Review*, Vol. 51, No. 2, 475-496.
- Santos-Pinto, Luís, and Enrique L. de la Rosa (2020). "Overconfidence in Labor Markets," K. Zimmermann (Ed), *Handbook of Labor, Human Resources and Population Economics*. Springer-Verlag.
- Shavell, S. (1982). "Suit, Settlement, and Trial: A Theoretical Analysis under Alternative Methods for the Allocation of Legal Costs," *Journal of Legal Studies*, Vol. 11, 55-81.
- Sobel, Joel, and Ichiro Takahashi (1983). "A Multi-Stage Model of Bargaining," *The Review of Economic Studies*, Vol. 50, 411-426.
- Soll, J., and Klayman, J. (2004). "Overconfidence in interval estimates," *Journal of Experimental Psychology: learning, memory and cognition*, Vol. 30, No. 2, 299.
- Spinnewijn, J. (2013). "Insurance and Perceptions: How to Screen Optimists and Pessimists," *The Economic Journal*, Vol. 123, 606-633.
- Spinnewijn, J. (2015). "Unemployed but Optimistic: Optimal Insurance Design with Biased Beliefs," *Journal of the European Economic Association*, Vol. 13, 130-167.
- Starmer, Chirs (2000). "Developments in Non-expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk," *Journal of Economic Literature*, Vol. 38, No. 2, 332-382.

- Svenson, O. (1981). "Are we all less risky and more skillful than our fellow drivers?," *Acta psychologica*, Vol. 47, No. 2, 143-148.
- Thompson, L., and G. Loewenstein (1992). "Egocentric Interpretations of Fairness and Interpersonal Conflict," *Organizational Behavior and Human Decision Processes*, Vol. 51, 17-197.
- Tingle, David (2015). "Bargaining Practice and Negotiation Failure in Russia-Ukraine Gas Relations," Cambridge Working Papers in Economics 1504, Faculty of Economics, University of Cambridge.
- Van den Steen, Eric (2004). "Rational Overoptimism (And Other Biases)," *American Economic Review*, Vol. 94, No. 4, 1141-1151.
- Yildiz, Muhamet (2003). "Bargaining without a Common Prior-An Immediate Agreement Theorem," *Econometrica*, Vol. 71, 793-811.
- Yildiz, Muhamet (2011). "Bargaining with Optimism," *Annual Review of Economics*, Vol. 3, 451-478.
- Waldfogel, Joel (1995). "The Selection Hypothesis and the Relationship between Trial and Plaintiff Victory," *Journal of Political Economy*, Vol. 103, No. 2, 229-260.
- Waldfogel, Joel (1998). "Reconciling Asymmetric Information and Divergent Expectations Theories of Litigation," *The Journal of Law & Economics*, Vol. 41, No. 2, 451-476.

Appendix

A. Self-Placement and Bayesian Updating

Self-placement with respect to the group can be compatible with Bayesian updating (Benoît and Dubra 2011). Burks et al.(2013) propose a simple rule to test whether overplacement (and underplacement) are compatible with Bayesian updating. This rule is called the "allocation function." According to the allocation function, when subjects are asked to estimate the skill level they most likely belong to, the largest (modal) group of subjects believing they belong to a certain skill level should be included in that skill level. When this allocation function is violated, Bayesian updating is rejected. In Burks et al. (2013) subjects are asked to indicate the most likely skill level (or the mode of their distribution of beliefs about skill). However, in our experiment, subjects are asked to estimate their mean rank not their modal rank. The allocation function is valid under the assumption that subjects have belief distributions where the mean and mode are not very far apart.

Tables 5 and 6 show how estimates of own ranks are related to the actual ranks in the EASY and HARD treatments, respectively (see also Figure 6). Subjects who display overplacement are below the diagonal since they estimate to have a higher rank than their actual rank. Subjects who display underplacement are above the diagonal since they estimate to have a lower rank than their actual rank.

Table 5: Relations among estimates and actual ranks, EASY treatment

EASY		estimates				n
		BEST			WORST	
actual		{1; 6}	{7; 12}	{13; 18}	{19; 24}	
BEST	{1; 6}	15	8	1	0	24
	{7; 12}	4	16	4	0	24
	{13; 18}	2	13	8	1	24
WORST	{19; 24}	1	7	14	0	22
n		22	44	27	1	94

Table 5 shows that there is overplacement in the EASY treatment: 39 observation are on the diagonal, 41 observations are below the diagonal, and 14 observations are above the diagonal. Table 6 shows that there is underplacement in the HARD treatment: 44 observation are on the diagonal, 20 observations are below the diagonal, and 32 observations are above the diagonal.

Tables 5 and 6 allow us to investigate whether self-placement with respect to the group is compatible with Bayesian updating. If this is the case, then the largest number of subjects who estimate their rank to be in the first quartile should belong to the first quartile of actual ranks, the largest number of subjects who estimate their rank to be in the second quartile should belong to the second quartile of actual ranks, and so on. Table 5 shows that overplacement in the EASY treatment is incompatible with Bayesian updating. The allocation function in the EASY treatment is violated by subjects who believe their ranks belong to the third quartile. Among the 27 subjects who believe to be in the third quartile, only 8 belong to that quartile. This is less than the 14 subjects who belong to fourth quartile. In contrast, Table 6 shows that underplacement in the HARD treatment is compatible with Bayesian updating. The allocation function in the HARD treatment is not violated by subjects who believe their ranks belong to the first, second, third, and fourth quartiles.

B. Bargaining Failures

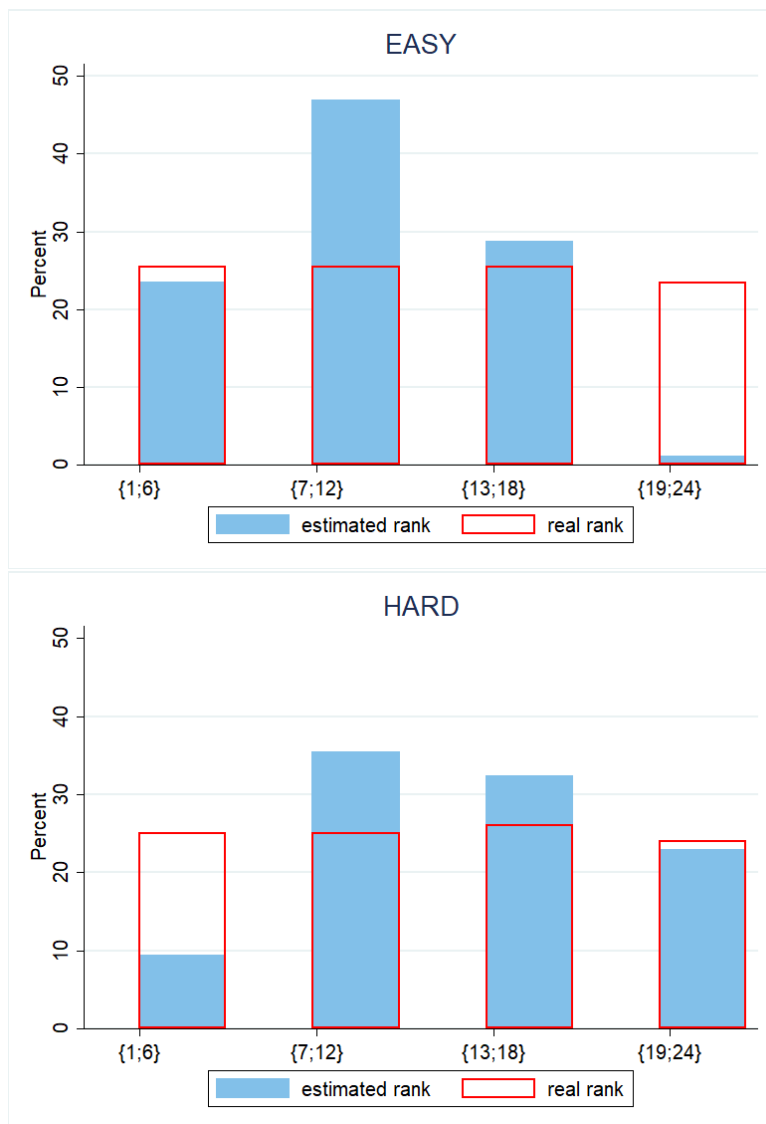
Table 7 presents an OLS model regression with bargaining failures as dependent variable. Bargaining failures is a dummy variable which equals 1 if the pair did not reach an agreement, and 0 otherwise. The explanatory variables are three dummies, one for condition EASY-LOW, another one for condition HARD-LOW, and the third one for condition. The variable BEST equals 1 if a subject is the best in the pair, and 0 otherwise. We add demographic controls which are described in Appendix I. Robust standard errors are clustered at pair level. We have 95 observations, one observation per pair.

Table 7 shows that the estimated coefficient for the EASY-LOW dummy is positive which provides

Table 6: Relations among estimates and actual ranks, HARD treatment

HARD		estimates				n
		BEST {1; 6}	{7; 12}	{13; 18}	WORST {19; 24}	
BEST	{1; 6}	6	10	8	0	24
	{7; 12}	3	13	6	2	24
	{13; 18}	0	8	11	6	25
WORST	{19; 24}	0	3	6	14	23
n		9	34	31	22	96

Figure 6: Real Rank vs Estimates - EASY and HARD



support for Result 2: The percentage of bargaining failures is higher in the EASY-LOW condition than in the other conditions.²³

Table 7: OLS Regression: Bargaining failures on condition dummies and controls

	Barg. Failures	
	Coeff.	Std. Error
EASY-LOW	0.265***	(0.0895)
HARD-LOW	0.0349	(0.0752)
HARD-HIGH	0.0403	(0.0520)
BEST	-0.0331	(0.0654)
Risk Averse	0.0382	(0.0557)
Gender	-0.0278	(0.0563)
Swiss Nat.	-0.0463	(0.0686)
Unil	-0.0854	(0.0843)
Age	0.00141	(0.0233)
Grades	0.00452	(0.0611)
Bachelor	-0.0209	(0.0936)
Grad. Parents	0.0129	(0.0771)
onlychild	0.00117	(0.0940)
Big Town	0.0685	(0.0932)
People Known	-0.0189	(0.0235)
Constant	0.0393	(0.530)
R^2	0.216	
Adjusted R^2	0.067	
Observations	95	

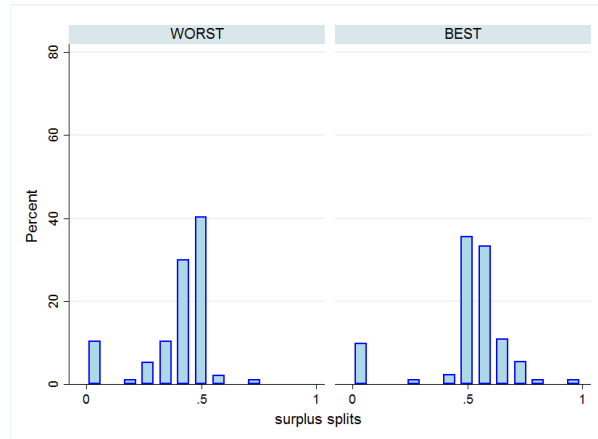
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

C. Surplus Splits

Table 8 presents a Tobit regression with surplus splits as shares of the jointly produced surplus as dependent variable. The explanatory variables are three dummies, one for condition EASY-LOW, another one for condition HARD-LOW, and the third one for condition. The variable BEST equals 1 if a subject is

²³A LOGIT regression would have been a suitable model for bargaining failures, however we are unable to implement it since the dummy variable for the EASY-HIGH condition only assumes the value zero given that all pairs reached an agreement.

Figure 7: Surplus splits



Each band has 0.05 width.

the best in the pair, and 0 otherwise. We add demographic controls which are described in Appendix I. Robust standard errors are clustered at pair level. We exclude bargaining failures from the analysis and thus have 172 observations.

Table 8 shows that the dummy BEST plays a significant role: subjects who are the best in a pair obtain on average higher surplus splits than subjects who are the worst in a pair.²⁴ as mentioned in 4.2. Figure 7 reports the distributions of surplus splits for best and worst partner in the pair: the distribution of surplus splits is asymmetric in favor of the best in the pair.

This result is in line with Karagözoğlu and Riedl (2014) since they find that subjects who are the best in a pair display higher subjective entitlements than subjects who are the worst in a pair. This in turn implies that subjects who are the best in a pair get higher surplus splits than subjects who are the worst in the pair. We discuss subjective entitlements in Appendix D.

D. Subjective Entitlements

As mentioned, the paper closest to ours is Karagözoğlu and Riedl (2014). Karagözoğlu and Riedl (2014) study the impact of performance information and production uncertainties on bargaining with joint production. Karagözoğlu and Riedl (2014) show that performance information affects bargaining outcomes through subjective entitlements. When subjects are informed about their relative contributions to the joint surplus, they feel entitled to split the surplus accordingly. Hence, subjects who are the best in a pair have higher subjective entitlements than subjects who are the worst in a pair. As a consequence, surplus splits are in favor of the best in a pair.

In line with Karagözoğlu and Riedl (2014), we find that the average subjective entitlement for subjects who are the best in a pair is higher and equal to 63% and the average subjective entitlement for

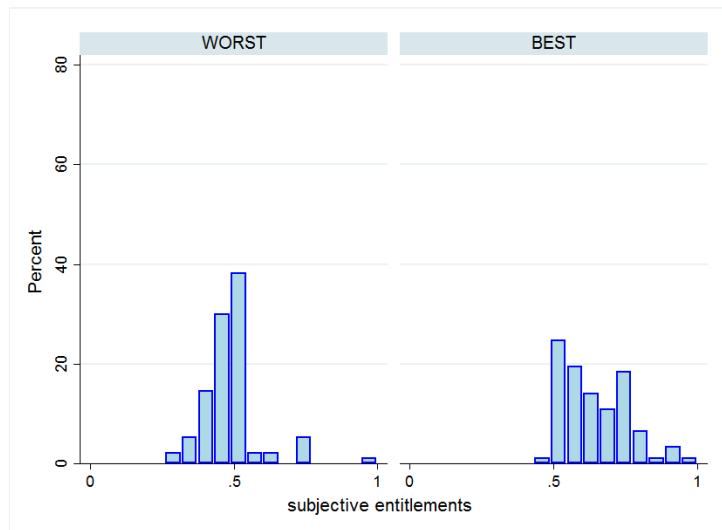
²⁴Equivalent results are found running OLS regressions instead of Tobit regressions.

Table 8: Tobit Regression: Surplus splits on condition dummies and controls, conditional on agreement

Surplus Splits		
	Coeff.	Std. Error
EASY-LOW	0.00275	(0.00848)
HARD-LOW	-0.00525	(0.00728)
HARD-HIGH	-0.00655	(0.00754)
Risk Averse	-0.0270	(0.0189)
BEST	0.125***	(0.0189)
Gender	-0.0163	(0.0126)
Swiss Nat.	-0.00511	(0.0137)
Unil	-0.0120	(0.0141)
Age	-0.00322	(0.00277)
Grades	0.00858	(0.00660)
Bachelor	0.00372	(0.0156)
Grad. Parents	-0.000654	(0.0188)
onlychild	-0.0182	(0.0259)
Big Town	0.0258	(0.0213)
People Known	-0.0180**	(0.00877)
Constant	0.506***	(0.0730)
sigma	0.0860***	(0.0134)
Pseudo R^2	-0.295	
Observations	172	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 8: Subjective entitlements



Note: the graph shows the distribution of subjective entitlements distinguishing among subjects who are the worst (left panel) and subjects who are the best (right panel) in a pair. Each band has 0.05 width.

subjects who are the worst in a pair is lower and equal to 49%. Figure 8 shows that the distribution of subjective entitlements is highly significant different across subjects who are the best and the worst in a pair (p -value < 0.01 , 1-sided, Kruskal-Wallis test).

Table 9 presents a Tobit regression with subjective entitlements as shares of the jointly produced surplus as dependent variable. The explanatory variables are three dummies, one for condition EASY-LOW, another one for condition HARD-LOW, and the third one for condition. The variable BEST equals 1 if a subject is the best in the pair, and 0 otherwise. We add demographic controls which are described in Appendix I. Robust standard errors are clustered at pair level. We have 190 observations.

Table 9 shows that that the dummy BEST plays a significant role: subjects who are the best in a pair have on average higher subjective entitlements than subjects who are the worst in a pair, confirming that information about who is the best in a pair has an impact on subjective entitlements. Neither the self-confidence manipulation nor the surplus size have an impact on subjective entitlements.

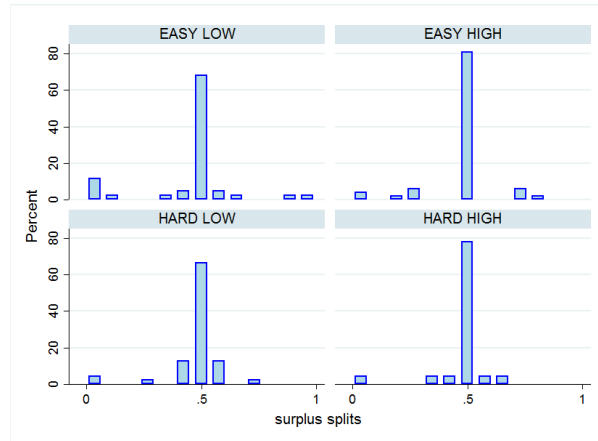
Karagözoğlu and Riedl (2014) run a treatment where subjects do not receive information about who is the best and worst in a pair. They find that in this case most subjective entitlements are equal to the 50%. As a consequence, bargaining often leads to equal splits. Appendix E reports the results of introducing this performance information manipulation in our design.

Table 9: Tobit Regression: Subjective entitlements on condition dummies and controls

	Subjective Entitlements	
	Coeff.	Std. Error
EASY-LOW	-0.000354	(0.0226)
HARD-LOW	-0.0450	(0.0230)
HARD-HIGH	-0.0295	(0.0213)
Risk Averse	0.00922	(0.0171)
BEST	0.150***	(0.0173)
Gender	0.00597	(0.0167)
Swiss Nat.	0.00131	(0.0155)
Unil	-0.00786	(0.0165)
Age	0.000988	(0.00410)
Grades	-0.00637	(0.0124)
Bachelor	0.00516	(0.0240)
Grad. Parents	-0.0196	(0.0195)
onlychild	0.00557	(0.0235)
Big Town	0.0255	(0.0190)
People Known	0.000824	(0.00890)
Constant	0.504***	(0.129)
sigma	0.103***	(0.00809)
Pseudo R^2	-0.370	
Observations	190	

* $p < 0.10$, ** $p < 0.01$, *** $p < 0.001$

Figure 9: Surplus splits



Note: The graph shows the distribution of surplus splits in the EASY and HARD treatments when there is no information about who is the best in a pair across LOW and HIGH surpluses. Each band has 0.05 width.

E. Performance Information

To investigate whether the absence of performance information may have an impact in our design, we run an additional treatment in which we do not provide information about who is the best or the worst in a pair (without any further design modifications). We now discuss how performance information affects bargaining outcomes and bargaining dynamics.

We observe a lower percentage of bargaining failures without performance information. Similarly to Result 2, the highest number of bargaining failures is in the EASY-LOW condition, 9% , while in the other three conditions it is below 5%. The percentage of bargaining failures in the EASY-LOW condition is weakly significant greater than zero (p -value = 0.08, 1-sided, t -test).

Figure 9 shows the distribution of surplus splits without performance information across the four conditions.²⁵ Figure 9 shows clearly that without performance information 69.5% of pairs settle on the equal split whereas with performance information only 37% of pairs settles on the equal split. Hence, when there is no performance information, fairness concerns are predominant with respect to self-confidence.

Without performance information 60% of subjective entitlements are the equal split whereas with performance information only 30% of subjective entitlements coincide with the equal split. In line with Karagözoğlu and Riedl (2014), disagreement in subjective entitlements is highly significant greater with

²⁵As mentioned in 4.2, since subjects were bargaining over values and not percentages, it is possible that many pairs did not precisely settle on the equal split but in a neighborhood of it. In figure 9 each band has 0.05 width so that the equal split bin includes the 5% neighborhood around the equal split.

Table 10: Subjects ranks in the four conditions

		LOW	HIGH	
EASY	m	15.28	8.36	12.04
	sd	(5.92)	(6.10)	(6.91)
	n	50	44	94
HARD	m	17.55	8.70	12.38
	sd	(5.12)	(5.66)	(6.97)
	n	40	56	96
	m	16.29	8.55	
	sd	(5.67)	(5.83)	
	n	90	100	190

performance information than without (1.11 and 1.06 on average, respectively; p -value < 0.01 , 1-sided, Kruskal-Wallis test). Without performance information, there are neither significant differences in disagreement in opening proposals among EASY and HARD treatment nor among the four conditions. In all treatments and conditions disagreement in subjective entitlements is significantly greater than 1 (p -value < 0.05 , 1-sided, t -test). We do not find significant differences in disagreement in opening proposals, concessions, and bargaining duration across the two performance information treatments. Moreover, results for disagreement in opening proposals, concessions, and bargaining duration are similar but hold without significance.

F. Relative Performance: Individual Ranks and Pair Composition

Differences in rank distributions may generate biases both at the individual and at the pair level. At the individual level, for instance, the presence of more ties in one condition than in another may generate confounds in the observed levels of overplacement. At the pair level, differences in rank variations across pairs may lead to different bargaining outcomes; for instance, we may observe more disagreement when there is high variation in partners' ranks and more equal split when there is low variation in partners' ranks.

In what follows, we investigate whether there are significant differences in ranks among the EASY and the HARD treatment and among conditions to exclude biases at the individual level; next we introduce a variation measure to exclude biases in ranks composition at the pair level. Table 10 reports means and standard deviations of ranks as well as the number of observations in the two treatments and in the four conditions. There are no significant differences in mean ranks (2-sided, t -test), standard

deviations of ranks (2-sided *sd*-test), and ranks distributions (Kruskal-Wallis test) across the EASY and HARD treatments. There are no significant differences in ranks standard deviations across LOW and HIGH (2-sided *sd*-test). Since the surplus size depends on partners' mean relative performance, there must be differences in ranks means and distributions by construction. Indeed, we find highly significant differences in mean ranks and ranks distributions across surplus size (p -value < 0.01 , 2-sided, t -test; p -value < 0.01 , 1-sided, Kruskal-Wallis test). Finally, there are no significant differences in ranks standard deviations across the four conditions (2-sided, *sd*-test). There are highly significant differences in means and distribution across all the four conditions (p -value < 0.01 , 2-sided, t -test; p -value < 0.01 , 1-sided, Kruskal-Wallis test) with the exception of the difference among EASY-LOW and HARD-LOW that is weakly significant different (p -value = 0.06, 2-sided, t -test; p -value = 0.07, 1-sided, Kruskal-Wallis test) and of EASY-HIGH and HARD-HIGH condition that do not differ significantly (p -value = 0.77, 2-sided, t -test; p -value = 0.63, 1-sided, Kruskal-Wallis test). However, since there are differences among LOW and HIGH by construction, what matters to exclude a ranking bias across conditions is the absence of differences in variation (i.e. in standard deviations) that has been verified.

To make sure that there is no pair composition bias across treatments and conditions, it is important to exclude differences in ranks variation across pairs. To measure ranks variation we define the variable "pair variance", V , as follows:

$$V = \frac{\sum (r_i - r_j)^2}{N}, \quad (1)$$

where r_i is the rank of partner i , r_j is the rank of partner j , N is the number of pairs ij , $(r_i - r_j)^2$ is the difference of partners' ranks squared. There are in total 95 pairs out of which 48 took part to the EASY and 47 in the HARD treatment. In the EASY treatment V is equal to 112.5 and in the HARD treatment V is equal to 88. Given that V takes only one value per treatment, we test whether there are significant differences in the differences of partners' ranks squared, $(r_i - r_j)^2$, among the EASY and the HARD treatment. There are no significant differences in means nor in distributions among EASY and HARD (p -value = 0.2, 2-sided t -test; $p = 0.089$ 1-sided Kruskal-Wallis test).

Out of 95 pairs, 45 pairs were assigned a LOW jointly produced surplus and 50 pairs were assigned a HIGH jointly produced surplus. V is equal to 98 for pairs that were assigned a LOW surplus and to 104 for pairs that were assigned a HIGH surplus. Again, given that V takes only one value per condition, we test whether there are significant differences in the difference of ranks squared $((r_i - r_j)^2)$ among LOW and HIGH. There are no significant differences in means nor in distributions among LOW and HIGH (2-sided t -test, 1-sided Kruskal-Wallis test).

To complete, Figure 10 shows the distribution of the difference of partners' ranks squared in the EASY and in the HARD treatment (top panel) and the distribution of difference of partners' ranks squared in the four conditions (bottom panel).

The distribution of the difference of ranks squared ($(r_i - r_j)^2$) does not differ significantly across the four conditions (however note that the difference in distributions among EASY-HIGH and HARD-LOW is weakly significant (p -value = 0.09, 1-sided, Kruskal-Wallis test)).

Regarding ties in pairs, we only had a tie in two pairs overall the experiment. Regarding ties among all participants in EASY and HARD, note that the distributions of ranks do not differ significantly hence ties should not play a role.

Finally, we want to exclude that bargaining failures are due to differences in sophistication among subjects. We investigate whether beliefs about own and partner ranks are consistent with the information about the surplus. A pair is assigned a high surplus if the average of partners' ranks is lower or equal to 12 (as shown in section 2.2). Beliefs about own rank and partner rank are consistent with the information about the surplus if the average of the estimations is lower or equal to than 12 and the surplus is high or if the average of the estimations is greater than 12 and the surplus is low. Out of 190 observations, 172 subjects hold beliefs that are consistent with the information about the surplus. In other words, the 90.5% of subjects make consistent estimations. Among 18 subjects who belong to a pair that failed to reach an agreement, only 3 subjects show inconsistency. Note that the 3 subjects belong to a pair who was assigned a low surplus. For 2 of the subjects the average of the estimations is equal to 12 (hence, the threshold) and for one subjects the average of estimations is equal to 11 (close to the threshold). Thus we can conclude that overall subjects make consistent estimations and bargaining failures do not stem from differences in sophistication among subjects.

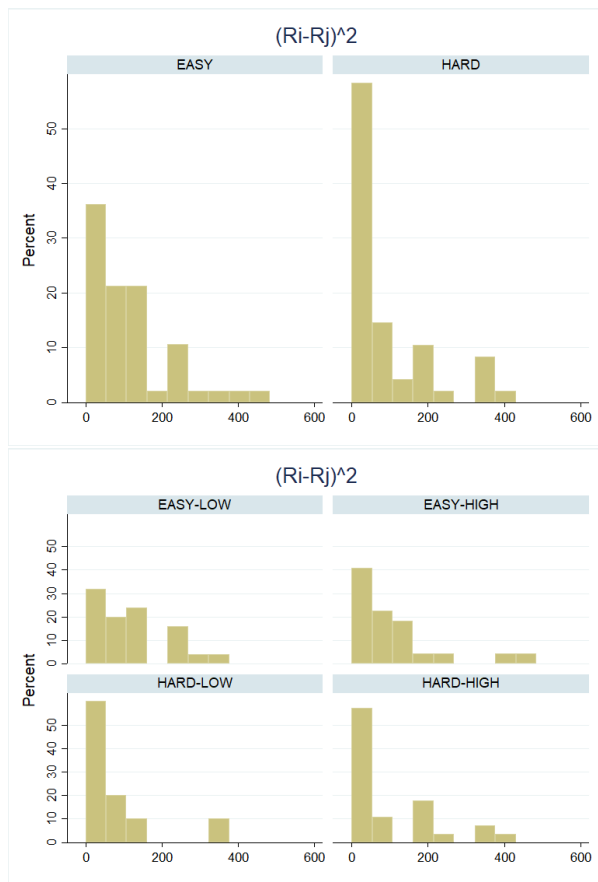
G. Gender Differences

In this section we analyze whether there are gender differences in bargaining outcomes, bargaining dynamics, overconfidence and in risk attitudes. Out of 190 subjects, 100 males and 90 females participated to the experiment.

The regressions reported in Tables 7 and 8 show that there are no significant gender differences in bargaining failures and in surplus splits, respectively. The regression reported in Table 9 shows that there are no significant gender differences in disagreement and subjective entitlements. We also find no significant gender differences in disagreement in opening proposals, concessions, and bargaining duration. Thus, we conclude that there are no gender differences in either bargaining outcomes or dynamics.

The literature finds conflicting results on overconfidence and gender. Some studies find that men are more overconfident than women (Niederle and Vesterlund 2007, Pulford and Colman 1997, Soll and Klayman 2004) while others find that there are no gender differences in overconfidence (Moore and Healy 2008, Johansson Stenman and Nordblom 2010). We do not find significant gender differences in

Figure 10: Difference of ranks squared



Note: the top panel shows the distributions of the difference of partners' ranks squared in the EASY and HARD treatments; the bottom panel shows the distribution of the difference of partners' ranks squared in the EASY and HARD treatments and distinguishes among LOW and HIGH surplus.

overconfidence when we refer to self-placement with respect to the partner (p -value = 0.21, 1-sided, Kruskal-Wallis test; overplacement with respect to a partner is equal to -0.97 for males and to 0.73 for females). However, when we refer to self-placement with respect to the group, we find that females are, on average, more overconfidence than males. Overplacement is, on average, -0.47 for males and 0.93 for females and the difference in distributions is weakly significant (p -value = 0.062, 1-sided, Kruskal-Wallis test).

A difference in performance among male and females may explain this result. Indeed, we find that males perform better than females both in the EASY and in the HARD quiz. In the EASY quiz males replied, on average, correctly to 32 questions while females replied, on average, correctly to 29 questions (the difference is weakly significant, p -value = 0.054, 1-sided, t -test). In the EASY quiz, males and females are equally overconfident (overplacement is 1.8 for males, 2.3 for females, and the difference is not significant, p -value = 0.68, 1-sided, t -test). In the HARD quiz males replied, on average, correctly to 10 questions while females replied, on average, correctly to 6.5 (the difference is highly significant, p -value < 0.01, 1-sided, t -test). In the HARD quiz, males and females are underconfident but males are marginally more underconfident than females (overplacement is equal to -2.6 for males, -0.46 for females, and the difference is significant, p -value = 0.024, 1-sided, t -test). Hence, if we distinguish across EASY and HARD, we do not find gender differences in overconfidence in the EASY treatment but we find that males are marginally more underconfident than females in the HARD treatment.

Even if Filippin and Crosetto (2016) show that the gender differences in risk aversion are task related, in many experimental studies women show more risk aversion than men (Eckel and Grossman 2008) and this is the case in our data. We find that the 77% of females and the 55% of males is risk averse and the difference is highly significant (p -value < 0.01, 1-sided, Kruskal-Wallis test). For the definition of risk aversion we follow Crosetto and Filippin (2013). Moreover, we introduce a continuous measure of risk preferences, “risk attitude” (described in details in Appendix G). The average risk attitude is equal to -13.2 for females and to -7.9 for males; the difference in risk attitude is highly significant (p -value < 0.01, 1-sided, Kruskal-Wallis test).

H. Risk Attitudes and the BRET

As mentioned in Section 2.7 we measure risk attitudes with the Bomb Risk Elicitation Task (BRET, Crosetto and Filippin 2013). The task is as follows. We display a 10x10 matrix (i.e. a square composed of 100 boxes) on the computer screen. Subjects can earn points for each box they decide to collect. The box collection process is automatic: for each second elapsed, a box is collected (during the experiment, boxes pass from dark gray to light gray). Subjects have the possibility to interrupt the collection process at any time pressing a STOP button. Behind one of the 100 boxes hides a “bomb” that destroys every-

thing that has been collected. The bomb can be hidden in any box with the same probability (equal to 1/100). However, subjects do not know which box hides the bomb.

If a subject collects the box hiding the bomb her payoff for the task is equal to zero. If not, her payoff for the task is positive and proportional to the number of boxes collected. Participants learn where the bomb is (and hence if they collected it) only at the end of the task. To make sure that everyone has understood this task, we run a practice round that does not pay the points that participants may have accumulated.

The BRET is so that with 100 boxes subjects choosing to collect 50 boxes are risk neutral, subjects collecting less than 50 boxes are risk averse, and those collecting more than 50 boxes are risk seekers. The minimum possible number of boxes that one can collect is 0, the maximum is 100, and the probability of collecting the bomb increases with the number of boxes collected.

We find that 64% of subjects are risk averse in the EASY-LOW, 59% in the EASY-HIGH, 77% in the HARD-LOW, and 62.5% in the HARD-HIGH condition. We also find that 16% of subjects are risk seekers in the EASY-LOW, 20% in the EASY-HIGH, 12.5% in the HARD-LOW, and 16% in the HARD-HIGH condition. The remaining subjects in each condition are risk neutral. There are no significant differences in the proportions of risk adverse, risk seeker, and risk neutral subjects across the four conditions.

In addition to referring to discrete risk attitude measures suggested in Crosetto and Filippin (2013), we introduce a continuous measure, “risk attitude”, of risk attitudes. We define this variable as the difference among the number of boxes collected minus 50 (the variable is included among -49 , absolute risk aversion, and $+49$, absolute risk seeking behavior). The average risk attitude is equal to -12 in the EASY-LOW condition, to -7 in the EASY-HIGH condition, -15 in the HARD-LOW condition and to -9 in the HARD-HIGH condition. There are weakly significant differences in means and distributions between EASY-HIGH and HARD-LOW (p -value = 0.0516, 2-sided, t -test; p = 0.058, 1-sided, Kruskal-Wallis test) and significant differences between HARD-LOW and HARD-HIGH (p -value < 0.05, 2-sided, t -test; p -value < 0.05, 1-sided, Kruskal-Wallis test). Further investigation regarding risk preferences is carried on with the randomization check in Appendix I.

I. Randomization Check

To make sure that the randomization over the EASY and HARD treatment was successful, we run a multinomial logit regression (table 11). The multinomial regression includes the variable “Risk averse” defined in Appendix G and the demographic controls. Demographic controls include: Gender (dummy equal to 1 if male, 0 if female), Swiss Nationality (dummy equal to 1 if Swiss, 0 otherwise), Age, Unil (dummy equal to 1 if the student is affiliated to UNIL, 0 otherwise), Grades, Bachelor (dummy equal to

1 if bachelor student, 0 otherwise), Grad. Parents (dummy equal to 1 if both parents have a degree, 0 otherwise), Only Child (dummy equal to 1 if only child, 0 otherwise), Big Town (dummy equal to 1 if resident in a big town, 0 otherwise), People Known (number of people known during the lab section). The EASY treatment is the base outcome in the model.

As Table 11 shows, there are no significant differences EASY and HARD treatment and thus we can conclude that our randomization worked.

Table 11: Multinomial logit

	HARD
Risk Averse	0.245 (0.320)
Gender	0.133 (0.319)
Swiss Nat.	0.0162 (0.343)
Unil	0.104 (0.339)
Age	-0.116 (0.0754)
Grades	0.264 (0.267)
Bachelor	0.646 (0.442)
Grad. Parents	0.254 (0.335)
Only Child	0.144 (0.494)
Big Town	0.109 (0.387)
People Known	-0.285 (0.192)
Constant	0.413 (2.212)
r2	
chi2	14.11
p	0.227

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

J. Experimental Instructions

This appendix provides the English translation of the French experimental instructions.

General Explanations of the Experiment

You are about to participate in an economic experiment. The experiment is conducted by the Département d'Econometrie et Economie Politique (DEEP) of the University of Lausanne and funded by the Swiss National Science Foundation (SNSF). It aims at better understanding bargaining behavior.

For your participation in the experiment you will earn a lump sum payment of 10 CHF for sure. You can earn more during the experiment. The experiment consists of **six** parts. In some parts of the experiment you can earn points that depend on your decisions. During the experiment, we will consider points instead of CHF.

Thus it is to your own benefit to read these explanations carefully.

The total number of points you have earned during the experiment will be exchanged into CHF at the end of the experiment. The exchange rate is

$$\mathbf{70 \text{ points} = 1 \text{ CHF}}$$

In other words, each point corresponds to approximately 1.43 cents.

You can take your decisions at your own speed.

It is prohibited to communicate with the other participants during the whole course of the experiment. It is also prohibited to use your mobile phone during the whole course of the experiment. If you do not abide by these rules you will be excluded from the experiment and all payments. However, if you have questions you can always ask one of the experimenters by raising your hand.

Your anonymity is guaranteed

At the end of the experiment, one of the experimenters will give you a payment sheet reporting the amount you will receive. You have to take it with you and bring it to the experimenter outside of the LABEX.

The experimenter outside of the LABEX is not informed about the decisions you have taken during the experiment. The experimenter will pay you in accordance to your payment sheet. After that, you will sign a payment receipt. Given that the receipt does not include your participant number, no experimenter will be able to determine your identity.

The backside of these explanations gives you an overview of the experiment. If you have any questions right now, please raise your hand. Otherwise, you can now proceed with the explanations on the first part of the experiment.

The backside of these explanations gives you an overview of the experiment. If you have any questions right now, please raise your hand. Otherwise, you can now proceed with the explanations on the first part of the experiment.

Overview of the Experiment

- **Part 1:**
Answering General Knowledge Quiz
- **Part 2:**
Receiving Information about the Joint Surplus
- **Part 3:**
Making Two Estimations
- **Part 4:**
Bargaining over the Joint Surplus
- **Part 5:**
Making a Choice under Risk
- **Part 6:**
Questionnaire and Payment

If you have questions please raise your hand.

Part 1: Answering a General Knowledge Quiz

In this part of the experiment you will have 20 minutes to answer a general knowledge quiz. There is a timer to the top right of the screen that indicates the remaining time (in seconds). Note that 20 minutes are equal to 1200 seconds.

The quiz is composed of 46 questions. **The greater the number of questions you answer correctly, the more likely is that you will earn more in the bargaining part of the experiment.** Hence, it is in your best interest to provide the correct answers to as many questions as you can.

Note that the correct answers to some questions involve providing the first and last names of famous people. You can only get full credit for your answers to these questions if you provide us both with the

first and the last name of that famous person. If you only provide us with the first name or the last name we will only give you half credit for your answer.

Unanswered questions count as wrong answers.

Part 2: Receiving Information about the Joint Surplus

During the rest of the experiment you will be randomly paired with another participant in this session but you will not know who that participant is. Neither during nor after the experiment will anybody be informed about who has been paired with whom.

In the experiment, you and the person you are paired with (your partner) will bargain over a joint surplus. Depending on your performance and the performance of your partner on the general knowledge quiz, the size of the joint surplus over which you will bargain in the fourth part of the experiment will be either small (**1390 points**) or large (**2710 points**).

We will now explain how the size of the joint surplus is determined. All 24 participants, including you and your partner, have completed the general knowledge quiz. According to the performance of all participants, each of them is attributed a rank. Rank 1 corresponds to the participant whose performance was the best (or, the participant who answered correctly the highest number of questions), rank 2 to the participant whose performance was the second best, and so on.

The size of the joint surplus is determined by your performance and your partner's performance on the general knowledge quiz. More precisely, the size of the joint surplus depends on the sum of your rank and your partner's rank in the general knowledge quiz, as follows:

1. If the sum of your rank and your partner's rank is from **2 to 24**, then the joint surplus will be **2710 points**.
2. If the sum of your rank and your partner's rank is from **25 to 47**, then the joint surplus will be **1390 points**.

The four examples that follow illustrate how the size of the joint surplus is determined.

Example 1: Suppose that your rank is 1 and that your partner's rank is 1 as well (you both performed the best). In this case, the sum of the ranks is 2. Since 2 is greater than 2 and less than 24, then the joint surplus would be 2710 points.

Example 2: Suppose that your rank is 5 and that your partner's rank is 10. In this case, the sum of the ranks is 15. Since 15 is greater than 2 and less than 24, then the joint surplus would be 2710 points.

Example 3: Suppose that your rank is 12 and that your partner's rank is 23. In this case, the sum of the ranks is 35. Since 35 is greater than 25 and less than 47, then the joint surplus would be 1390 points.

Example 4: Suppose that your rank is 23 and that your partner's rank is 23 as well (you both performed the worst). In this case, the sum of the ranks is 46. Since 46 is greater than 25 and less than 47, then the joint surplus would be 1390 points.

Part 3: Making Two Estimations

In this part of the experiment you will be asked to provide us with two estimates. The first one is the estimate of your rank in the quiz. The second one is the estimate of your partner's rank in the quiz. We will now explain in detail how you should indicate your **estimate of your rank in the quiz** and how this estimate influences your earnings.

- a) How to indicate your estimate of your rank?

All 24 participants, including you and your partner, have completed the quiz. According to the performance of all participants, each of them is attributed a rank. Rank 1 corresponds to the participant whose performance was the best (or, the participant who answered correctly the highest number of questions), rank 2 to the participant whose performance was the second best, and so on.

We want you to tell us your estimate of your rank as an integer between 1 and 24.

- b) How does your estimate of your rank influence your earnings?

The more precise your estimate of your rank is, the higher is the probability that you will earn 140 points. In other words, the likelihood of earning the 140 points is higher, the closer your estimate of your rank is to your true rank in the quiz.

Your earnings are obtained as follows:

- First, the computer randomly draws a number between 0 and 529. Every number between 0 and 529 is equally likely.
- Second, the difference between your estimate of your rank and your true rank is the prediction error. If the prediction error, multiplied by itself, is not larger than the random number drawn by the computer, then you will earn **140 points**. Otherwise, you will earn **0 points**.

Important: You may wonder why we have chosen this payment rule. The reason is that this payment rule makes it optimal - for you - to state precisely your estimate of your rank.

Example: Your estimate of your rank is 13, however given your performance in the quiz, your true rank is 10. Thus in this case the prediction error is $(13-10) = 3$. The prediction error multiplied by itself is 9. If the random number drawn by the computer is greater than or equal to 9, for example 26, then you will earn 140 points. If the random number drawn by the computer is smaller than 9, for example 8, then you will earn 0 points.

We will now explain in detail how you should indicate your **estimate of your partner's rank in the quiz** and how this estimate influences your earnings.

a) How to indicate your estimate of your partner's rank?

All 24 participants, including you and your partner, have completed the quiz. According to the performance of all participants, each of them is attributed a rank. Rank 1 corresponds to the participant whose performance was the best (or, the participant who answered correctly the highest number of questions), rank 2 to the participant whose performance was the second best, and so on.

We want you to tell us your estimate of your partner's rank as an integer between 1 and 24.

b) How does your estimate of your partner's rank influence your earnings?

The more precise your estimate of your partner's rank is, the higher is the probability that you will earn 140 points. In other words, the likelihood of earning the 140 points is higher, the closer your estimate of your partner's rank is to your partner's true rank in the quiz.

Your earnings are obtained as follows:

- First, the computer randomly draws a number between 0 and 529. Every number between 0 and 529 is equally likely.
- Second, the difference between your estimate of your partner's rank and your partner's true rank is the prediction error. If the prediction error, multiplied by itself, is not larger than the random number drawn by the computer, then you will earn **140 points**. Otherwise, you will earn **0 points**.

Important: You may wonder why we have chosen this payment rule. The reason is that this payment rule makes it optimal - for you - to state precisely your estimate of your partner's rank.

Example: Your estimate of your partner's rank is 3, however, given his/her performance in the quiz your partner's true rank is 14. Thus, in this case, the prediction error is $(3-14) = -11$. The prediction error multiplied by itself is 121. If the random number drawn by the computer is greater than or equal to 121, for example 200, then you will earn 140 points. If the random number drawn by the computer is

smaller than 121, for example 75, then you will earn 0 points.

Before providing us with your estimations, we will ask you to answer a few comprehension question. Your answers to these questions will have no consequences on the experiment nor on your final payment. The experiment will continue as soon as all the participants will have answered the questions correctly.

Part 4: Bargaining over the Joint Surplus

Information about relative performance in the quiz.

Remark: the following part is used only in the INFO treatments.

Next, you will receive information on your screen about who was the best performer in the general knowledge quiz among you and your partner.

If you have answered correctly more questions in the general knowledge quiz than your partner, then you are the best in the pair and your partner is the worst in the pair.

If you have answered correctly less questions in the general knowledge quiz than your partner, then you are the worst in the pair and your partner is the best in the pair.

If you and your partner have the same number of correct answers in the general knowledge quiz, then you and your partner are equally performing.

Bargaining

You will have a maximum of 10 minutes to reach an agreement on the distribution of the joint surplus. You do not have to use up all the bargaining time but must not exceed it. **If you do not agree on a distribution of the joint surplus within 10 minutes, then you will earn nothing from this bargaining stage! If you do agree on a distribution of the joint surplus then you will earn the points you and your partner agreed on.**

The bargaining takes place via computer. During bargaining you will work with a screen that consists of four parts, which we will explain in what follows. Hereby a screen-shot.

remaining time [sec]: 483				
who	kept by partner	given by partner	you keep	you give
partner	2000	710		
you			2600	110
partner	2000	710		
you			2500	210
partner	1900	810		
you			2400	310

Fourth part: bargaining over the joint surplus

the joint surplus you are bargaining over is equal to 2710

new offer

points you want to keep

points you want to give

your partner's current proposal

your partner keeps 1900

your partner gives you 810

In the **upper-right part** of the screen the joint surplus you are bargaining over is displayed. The timer right on the top shows how much bargaining time (in seconds) is still remaining. Note that 10 minutes are equal to 600 seconds.

In the **lower-left part** of the screen you can enter a new proposal. You will need to enter the points you want to keep for yourself and the points you want to give to your partner. There is a SEND button to confirm and send proposals.

In the **upper-left part** of the screen a table shows all previous proposals and the identity of proposers (you or your partner). Each proposal is listed in the table in chronological order. Every time you make a proposal, the table will show that you have made the proposal and display the proposal you have made. Every time your partner makes a proposal, the table will show that he/she has made the proposal and display the proposal he/she has made.

In the **lower-right part** of the screen you see your partner's currently valid proposal. The ACCEPT button allows you to accept your partner's currently valid proposal. If your partner has not made any proposal to you yet, this part of the screen is empty. Similarly, if you have not made any proposal, this part of the screen in your partner's computer is empty. To make a first (or new) proposal you have to fill in two boxes in the lower-left part of the screen with corresponding points of the joint surplus for yourself and for your partner. The points you fill in have to add up to the joint surplus. Thereafter, you need to press the SEND button to send your proposal. The following rules apply:

1. The sum of points for yourself and for your partner cannot be exceed the joint surplus nor be lower than the joint surplus.
2. Only offers with integer points are allowed.

3. A sent offer is binding, that is, if your partner accepts your proposal, bargaining is finished and both of you earn the points on which you have agreed upon. The same holds if you accept a proposal of your partner. You can only accept the current proposal; earlier proposals are not valid any more.

Hence, as long as you have not pressed the SEND button you can still change the offer. A sent proposal is binding. You can always make a new proposal, provided that neither you nor your partner have accepted a proposal and provided that there is still time left.

If you want to accept a currently valid proposal, you have to press the ACCEPT button. Once you or your partner accept a proposal, bargaining is over and each of you will receive the agreed share of the joint surplus.

Part 5: Making a Choice under Risk

On your computer screen you will see a square composed of 100 boxes.

You earn 1.4 points for every box that you decide to collect. The box collecting process is automatic: for every second that elapses, a box changes color. The boxes start disappearing from the top-left corner of the screen and number of boxes collected is updated accordingly.

Behind one of these boxes hides a “bomb” that destroys everything that has been collected.

The bomb can be in any box with equal probability (the probability the bomb is in one particular box is equal to $1/100$). However, **you do not know behind which box the bomb is.**

Your task for this stage is to choose when to stop the box collecting process. You can do it by hitting the STOP button at any time.

If you collect the box that contains the bomb, the bomb will explode and you will earn zero points. If you stop the box collecting process before collecting the box that contains the bomb, the bomb will not explode and you will earn the points accumulated that far.

Note that you will only know if one of the boxes you collected contains the bomb at the end of the task; indeed if you collect the box that contains the bomb, the bomb only explodes at the end of the task: this means that you can collect the box that contains the bomb without knowing it.

We will start this stage with a practice round. The goal of the practice round is to show you how this task works. After the practice round is over, the task starts. The practice round is just an example: you will not earn the points accumulated in this part.

Part 6: Questionnaire and Payment

At the end of the experience we will ask you to answer a questionnaire. Next, we will proceed with your payment.

Your final payment includes your gains for each part of the experiment in CHF and the lump sum payment²⁶. The amount you will be paid will be shown to you at the end of the experiment.

Note that 70 points in the experiment correspond to 1 CHF.

If you have questions please raise your hand.

²⁶Show up fee of 10 CHF