

ONLINE APPENDIX

This Online Appendix complements the paper “Building an Equilibrium: Rules versus Principles in Relational Contracts” by Gibbons, Grieder, Herz, & Zehnder. The Appendix contains additional results and analyses (sections A and B), as well as the instructions provided to experimental participants (sections C and D) and the coders (section E).

A Summary Statistics and Correlations for Key Variables

Table 1: Summary Statistics for key variables

Variable	Mean	Std. Dev.	N
Principle	0.58	0.5	121
Nudge Treatment	0.5	0.5	121
Post-Shock Performance	1472.15	475.06	121
Efficient Equilibration Score	0.74	0.95	121
Normative-Behavioral Consistency (by both parties)	1.41	1.22	121
Efficient Normative Consensus	1.26	1.07	121

Table 2: Correlations between key variables

Variables	(1)	(2)	(3)	(4)	(5)	(6)
(1) Principle	1.00					
(2) Nudge Treatment	0.56 (0.00)	1.00				
(3) Post-Shock Performance	0.39 (0.00)	0.03 (0.77)	1.00			
(4) Efficient Equilibration Score	0.30 (0.00)	0.06 (0.49)	0.49 (0.00)	1.00		
(5) Normative-Behavioral Consistency (by both parties)	0.37 (0.00)	0.20 (0.03)	0.38 (0.00)	0.61 (0.00)	1.00	
(6) Efficient Normative Consensus	0.21 (0.02)	0.04 (0.63)	0.32 (0.00)	0.43 (0.00)	0.05 (0.59)	1.00

Notes: The table displays correlation coefficients between variables. Significance levels in parentheses. $N = 121$

B Additional Analyses

B.1 Observable Characteristics of Pairs with Principles and Pairs with Rules

In this section, we analyze whether observable characteristics (e.g., demographics) allow predicting whether a pair in the experiment ended up with a rule or a principle and whether the observable characteristics of pairs with principle-based agreements differed between the Baseline and the Nudge Treatment. The available data on observable characteristics were elicited in the exit questionnaire after the experiment. Specifically, we collected data on participants’ age, sex, whether they were German native speakers (the language in which the experiment was conducted), the number of participants in the experimental session they knew before the session, the number of experiments they had participated in during the last year, and, finally, their field of study.

Table 3 below reports the means of these variables split by whether a pair was coded as having a principle or a rule only. As the p -values reported in column (3) of Table 3 indicate, in the Baseline there were no significant differences between the pairs with rules only and the pairs with principles in terms of observable characteristics. For the Nudge Treatment, the p -values reported in column (6) of Table 3 show a significant difference for non-native speakers. However this seems mainly due to the overall low number of non-native speakers in the Nudge treatment and none of them being in one of the few pairs with an exclusively rule-based agreement. In addition, there is a marginally significant effect for STEM students, which were underrepresented in rule-based pairs in the Nudge. This could be an indication of the negative selection into rules in the Nudge treatment.

Table 4 complements the analyses reported in Section 4.3.2 in the paper, where we compare pairs with principle-based agreements in the Baseline with pairs with principle-based agreements in the Nudge treatment. Table 4 considers the same comparison by showing how these pairs differed in terms of observable characteristics. The only significant difference we find is for the proportion of non-native speakers, which is higher in the pairs with principle-based agreements in the Baseline than in the Nudge (22% versus 6%). However, this difference is driven by a random difference in the share of non-native speakers in the Baseline versus the Nudge treatment overall: in the Baseline, 17% of participants were non-native speakers, compared to 6% in the Nudge. As treatments were randomly assigned to participants within an experimental session, this difference can only be random. Note from the results in Table 3 that actually all non-native speakers in the

Table 3: Observable Characteristics of Individuals in Pairs with Rules Only vs. Pairs with Principles

	Baseline			Nudge		
	(1) Rule only	(2) Principle	(3) p	(4) Rule only	(5) Principle	(6) p
Age	21.45	22.33	0.15	21.71	21.88	0.78
Female participant	0.54	0.58	0.64	0.43	0.53	0.50
Non-native speaker	0.13	0.22	0.27	0.00	0.06	0.01**
STEM student	0.57	0.54	0.81	0.38	0.65	0.10*
Econ student	0.09	0.06	0.55	0.14	0.06	0.41
# of participants known in session	0.46	0.31	0.33	0.29	0.31	0.88
# of experiments participated last year	2.78	3.31	0.39	3.43	2.96	0.66
Observations	82	36		14	104	

Notes: The table reports means of observable characteristics of participants. Columns (1) to (3) only consider data from the Baseline, columns (4) to (6) only consider data from the Nudge Treatment. Observations are at the participant level.

Nudge Treatment were in principle-based pairs. Thus, the difference in the proportion of non-native speakers in pairs with principle-based agreements between the Baseline and the Nudge seems mainly driven by the random difference in the base-rate of non-native speakers in the two treatments, and we therefore believe it does not constitute a relevant effect.

Table 4: Comparison of Observable Characteristics of Individuals in Pairs with Principles in Base vs. Nudge

	Base	Nudge	p
Age	22.33	21.88	0.47
Female participant	0.58	0.53	0.58
Non-native speaker	0.22	0.06	0.03**
STEM student	0.54	0.65	0.29
Econ student	0.06	0.06	0.96
# of participants known in session	0.31	0.31	0.99
# of experiments participated last year	3.31	2.96	0.56
Observations	36	104	

Notes: The table reports means of characteristics of participants in pairs with principle-based agreements. Observations are at the participant level.

In column (1) of Table 5, we provide additional results from regression analyses entering the variables we collected in the post-experimental questionnaire on demographic and other observable characteristics as potential predictor variables of whether a pair ended up with a principle or not (considering data from both the Baseline and the Nudge Treatment and thus entering a Nudge dummy in the regression). None of the coefficients of the observable characteristics variables are

Table 5: Predicting the Emergence of Principles with Pairs' Observable Characteristics

	(1) Principle Anywhere	(2) Principle in Statement
1 female participant in pair	−0.04 (0.11)	0.00 (0.16)
2 female participants in pair	0.13 (0.14)	0.13 (0.18)
1 STEM student in pair	0.16 (0.11)	−0.21 (0.15)
2 STEM students in pair	0.11 (0.11)	−0.02 (0.16)
1 econ student in pair	−0.08 (0.11)	0.35** (0.17)
1 non-native speaker in pair	0.04 (0.10)	0.08 (0.16)
2 non-native speakers in pair	0.10 (0.22)	0.29*** (0.09)
Age (mean in pair)	0.03 (0.02)	−0.07** (0.03)
# of participants known in session (mean in pair)	−0.04 (0.06)	−0.17 (0.11)
# of experiments participated in last year (mean in pair)	0.02 (0.03)	−0.01 (0.03)
Nudge Treatment	0.56*** (0.08)	0.27* (0.14)
Constant	0.14 (0.14)	0.55** (0.23)
R^2	0.38	0.25
Observations	121	70

* $p < .10$, ** $p < .05$, *** $p < .01$

Notes: Column (1) is based on data from all pairs and the dependent variable is a dummy variable indicating whether a pair had a principle. Column (2) restricts the sample to pairs that had a principle and the dependent variable is a dummy variable indicating whether the principle was formulated in the statement. OLS estimates (linear probability models). Robust standard errors are in parentheses. The variable "Age" was centered at the sample mean. There were no pairs with 2 econ students, which is why this regressor is omitted.

statistically significant, and there are also no meaningful effects that are close to conventional thresholds of statistical significance. (We return to the analysis reported in column (2) of Table 5 in Section B.2 below.)

B.2 Principles in Chats vs. Principles in Statements

In this section we differentiate between principles that were coded as part of the written statements that the participants in a pair explicitly agreed upon at the end of the communication phase and principles that were only coded as being present in the chats. Specifically, we test whether there are performance differences between pairs with principles in the statements vs. pairs with principles in the chats only.

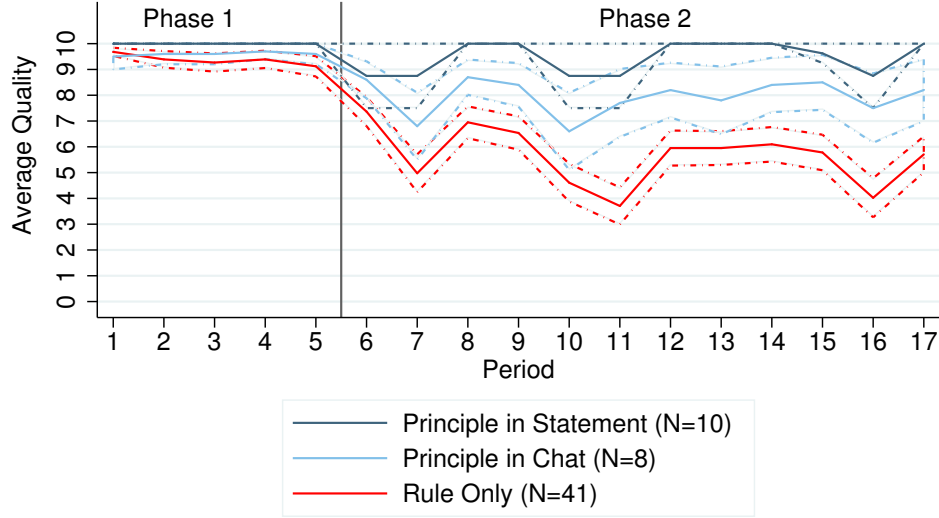
In the Baseline there were 8 pairs with a principle manifested in the statement and 10 pairs with a principle that was present only in the chat codings. In the Nudge Treatment, 38 pairs formulated a principle in their statements, whereas 14 pairs only did so in the chat.

There is a difference in Phase 2 performance between pairs with principles that were written down in the statements compared to principles that were only agreed upon in the chats. In Figure 1 below we reproduce Figure 3 from the paper, distinguishing by whether a principle was coded as being present in the statement or only in the chats. The data show that pairs with a principle in the statement performed better than pairs with a principle in the chat only, which were, however, still better than pairs with a rule only.

The results from t -tests indicate that in the Baseline the difference in long-run performance in Phase 2 between pairs with principles in statements and pairs with rules only was highly significant ($t = 7.02$, $p < 0.01$). Also the difference between pairs with principles in the chat only compared to pairs with rules only was (marginally) significant ($t = 2.12$, $p = 0.05$). The difference between principles in chats vs. principles in statements turned out to be just non-significant ($t = 1.72$, $p = 0.12$), however the statistical power is low because of the small number of observations.

Table 6 displays alternative specifications in the form of regressions analyzing whether there were differences in Phase 2 profits between pairs with principles in the statement and pairs with principles in the chats only. Column (1) considers Baseline data only, column (2) considers the data from the Nudge Treatment. The results indicate that both in the Baseline and in the Nudge Treatment the pairs with principles in the statement tended to perform better in Phase 2 than the

Figure 1: Quality in Pairs with Principles in Chats or Statements vs. Rules (Baseline Only)



Notes: Only data from the Baseline included. One of the 60 pairs in the Baseline had neither a principle nor a rule and is thus not included in the graph ($n = 59$). The solid lines display average quality in a given period. The dashed lines represent plus/minus one standard error of the mean.

pairs with a principle only in the chats. The difference is captured by the coefficient for “Principle in statement,” which turns out to be marginally significant in the regression specification for the Baseline ($p = .08$) but not in the Nudge Treatment ($p = .70$).

Table 6: Effects of Principles in Statements and Chats on Phase 2 Performance

	(1) Base	(2) Nudge
Principle anywhere	319.70** (141.64)	643.48*** (149.74)
Principle in statement	221.02* (124.80)	57.15 (149.60)
Constant	1333.90*** (71.32)	900.67*** (73.93)
R^2	0.19	0.25
Observations	60	61

* $p < .10$, ** $p < .05$, *** $p < .01$

Notes: Phase 2 profit is the dependent variable. OLS estimates. Robust standard errors are in parentheses.

In the Baseline, pairs who managed to write down a principle in the statement summarizing

their agreement were thus more successful than pairs who only agreed on the principle in the chats but did not include it in the statement. A potential reason could be that they had thought more thoroughly about these principles or that these principles were clearer than the principles that pairs only agreed upon in the chats. As an effect of the Nudge Treatment, more pairs managed to write down a principle in the statement. However, in the Nudge Treatment these statement principles were not significantly more successful anymore than the principles that were only formulated and agreed upon in the chats. The statement principles in the Nudge Treatment were also significantly less successful than the statement principles in the Baseline (average Phase 2 profits: 1,601.29 vs. 1,874.63; $t = 3.54$, $p = .01$). Thus, while the Nudge did cause pairs to try to come up with principles in the statements, they did not manage to create as successful principles as the pairs with principles in the statements in the Baseline.

In a next step, we again use observable characteristics to try to predict whether a pair wrote down a principle in the statement or only agreed on a principle more informally in the chats. The corresponding analysis is presented in column (2) of regression Table 5 that we introduced in Section B.1 above. The dependent variable in column (2) of Table 5 is a dummy for whether a principle was formulated in the statement and the sample is restricted to pairs with a principle coded as present anywhere (i.e., in the statement or in the chat). There is a significant positive effect of the dummy variable capturing the case of a pair with one economics student. In addition, there is a significant effect for pairs composed of 2 non-native speakers. However, there was only one such pair in the entire sample, and the result is thus driven by that single pair. Finally, age had a significantly negative impact on whether a principle was coded in the statement or only in the chat. However, given that the experiment was conducted with a student sample, the age range in the participant sample was very small, and we are therefore cautious in interpreting this effect. Overall it seems that there are no clear and meaningful associations between observable characteristics and whether the principle was written down in the statement or was only present in the chat codings.

B.3 Additional Codings of Chats

In this section we report further analyses of the contents of the chats and how they differ between the Baseline and the Nudge Treatment, as well as between pairs with principles and pairs with

rules only. The aim of these additional codings was to analyze in more depth how principles came about. In a second step, we can use these additional codings as potential explanatory variables for performance, agreement type and equilibration.

We categorize our additional text codings into four categories: (1) Phase 1 strategy; (2) references to promises and trust; (3) talk about Phase 2, and (4) the mood in the interaction. Phase 1 strategy variables provide an indication of the general understanding of the structure of the game and the ability of groups to find agreement on the efficient and fair solution when the strategic environment is known. References to promises and trust give an indication of the use of soft factors and attempts to activate social preference related contract enforcement mechanisms. Talk about Phase 2 indicates the ability of group members to anticipate challenges for their trading relationship. Finally, the mood variables capture soft social skills that indicate the overall atmosphere in the group.

Summary statistics for all coded variables, split by treatment, are displayed in Table 7 below. Column (1) considers all pairs, column (2) only pairs with principles, and column (3) only pairs with rules. Reported p -values refer to t -tests.

Table 7 reveals a few interesting patterns concerning the structure of conversations, the treatment and the endogenous selection of a principle or a rule:

- First, regarding the *Phase 1 strategy*, the large majority of pairs reached an agreement on a quality of 10 and a price of 70, and on average such an agreement was reached within 3.2 minutes. In about a third of the pairs, a suboptimal proposal was made, and in about 20% of the groups, one side had to offer an explanation as to why choosing a quality of 10 and a price of 70 in Phase 1 of the experiment is a sensible choice. While there is no statistically significant difference in any of the Phase 1 strategy variables across treatments, it is notable that the need for explanation was marginally significantly more likely for principles in the Nudge Treatment than for principles in the Baseline. This is a first indication that principles in the Nudge might be of lower caliber than in the Baseline, because the degree of independent mutual understanding appears to be lower.
- Regarding *promises and trust*, we observe that promises are marginally significantly more frequent in the Baseline than in the Nudge and also mentions of trust appear more frequently in

Table 7: Differences in Chat Codings by Treatment and Agreement Type

	(1) All Pairs			(2) Principles			(3) Rules		
	Base	Nudge	<i>p</i>	Base	Nudge	<i>p</i>	Base	Nudge	<i>p</i>
Phase 1 strategy:									
Suboptimal proposal	0.32	0.34	0.75	0.33	0.33	0.96	0.32	0.29	0.88
Explanation necessary for 10-70	0.17	0.20	0.67	0.06	0.21	0.06*	0.20	0.14	0.75
Reached agreement on 10-70	0.88	0.92	0.53	0.94	0.96	0.78	0.88	0.86	0.89
Time to agreement on 10-70	3.11	3.30	0.70	2.73	3.37	0.27	3.29	2.76	0.69
Promises and Trust:									
Trust mentioned	0.27	0.18	0.26	0.22	0.21	0.93	0.29	0.00	0.00***
Promise	0.18	0.07	0.05*	0.17	0.08	0.37	0.20	0.00	0.00***
Threats	0.07	0.10	0.53	0.11	0.12	0.96	0.05	0.00	0.16
Appeals to honesty and fairness	0.27	0.26	0.96	0.22	0.29	0.58	0.29	0.14	0.37
Phase 2:									
Foresee misunderstandings	0.02	0.11	0.03**	0.06	0.13	0.29	0.00	0.00	.
Mention Phase 2 uncertainty	0.18	0.57	0.00***	0.33	0.65	0.02**	0.12	0.14	0.89
Deviations discussed	0.18	0.21	0.68	0.11	0.23	0.22	0.20	0.14	0.75
Mood:									
# of Greetings	1.58	1.39	0.16	1.56	1.35	0.30	1.59	1.86	0.15
Smalltalk	0.12	0.13	0.81	0.17	0.15	0.90	0.10	0.00	0.04**
Positive emotions expressed	0.60	0.69	0.31	0.67	0.69	0.85	0.59	0.71	0.54
# of positive emoticons	2.93	3.02	0.89	3.61	3.33	0.74	2.71	1.00	0.04**
# of negative emoticons	0.05	0.00	0.18	0.11	0.00	0.33	0.02	0.00	0.32
Observations	60	61		18	52		41	7	

Notes: The table shows means of the additionally coded chat variables and *p*-values of *t*-tests testing whether there were significant differences between the Baseline and the Nudge Treatment.

the Baseline (although the latter difference does not surpass conventional significance levels). Column (3) shows that attempts to leverage social preferences to manage the relationship are almost non-existent in pairs with rules only in the Nudge Treatment, and mentions of trust and promises are significantly lower than in the Baseline. But even for principles, this pattern between Nudge Treatment and Baseline holds (albeit clearly not statistically significant). It therefore appears that the use of social capital as a contract enforcer is less frequent in the Nudge Treatment than in the Baseline treatment.

- Rather unsurprisingly, explicit chat about *Phase 2* is more common in the Nudge Treatment, in particular explicit mentions of uncertainty regarding Phase 2 are more frequent. Despite this, groups rarely talk about reactions to deviations and potential misunderstandings regarding Phase 2.
- Finally, the *mood* of the conversations seems relatively similar in the Baseline and the Nudge.

The only noticeable difference is between groups with rules, which again shows a potential lack of soft skills in pairs with rules in the Nudge Treatment.

Turning to the impact of these variables on performance and equilibration, Table 8 presents pairwise correlation coefficients between all coded variables and Phase 2 performance as well as efficient equilibration.

Table 8: Correlations between Chat Codings and Phase 2 Performance and Efficient Equilibration

	(1) Phase 2 Profit	(2) Eff. Equilibration Score
Phase 1 strategy:		
Suboptimal proposal	-0.24***	-0.25***
Explanation necessary for 10-70	-0.09	-0.10
Reached agreement on 10-70	0.36***	0.26***
Time to agreement on 10-70	-0.20**	-0.20**
Promises and Trust:		
Trust mentioned	0.06	-0.06
Promise	0.19**	0.13
Threats	0.10	0.05
Appeals to honesty and fairness	0.14	0.12
Phase 2:		
Foresee misunderstandings	0.07	0.18*
Mention Phase 2 uncertainty	0.15	0.07
Deviations discussed	-0.02	0.05
Mood:		
# of Greetings	0.11	0.08
Smalltalk	0.24***	0.15*
Positive emotions expressed	0.28***	0.04
# of positive emoticons	0.22**	0.15*
# of negative emoticons	-0.00	-0.10
Observations	121	121

Notes: This table displays pairwise correlation coefficients between Phase 2 profits respectively efficient equilibration and additional codings of the chat conversations. “Time to agreement on 10-70” is conditional on having agreed on 10-70. The sample is reduced to 109 observations for this correlation.

- *Phase 1 strategy:* The mere presence of a suboptimal proposal is predictive of lower Phase 2 profits (column 1) and of a lower efficient equilibration score (column 2). Reaching the quality 10 / price 70 agreement for Phase 1 is predictive of significantly higher Phase 2 earnings, as well as of a significantly lower efficient equilibration scores. Moreover, The time it takes groups to reach the 10-70 agreement also matters. The longer it takes, the worse is the efficient equilibration score and performance. These correlations support the view that

the ease with which groups agree on the fair and efficient outcome in Phase 1 is a strong predictor of performance in Phase 2.

- *Promises and trust*: The making of promises significantly predicts performance in Phase 2 and is also positively (albeit non-significantly) associated with efficient equilibration. Appeals to honesty and fairness are also positively correlated with profits and efficient equilibration, but standard significance levels are not reached ($p = 0.12$ and $p = 0.18$, respectively). It therefore appears that attempts to mobilize social preference related contract enforcement mechanisms are also, at least weakly, predictive of performance in Phase 2 of our experiment.
- *Phase 2*: Foreshadowing misunderstandings significantly predicts efficient equilibration, but the positive correlation with performance remains insignificant. Mentioning uncertainty also has a sizeable correlation coefficient with performance and efficient equilibration, but it remains statistically insignificant ($p = .10$ and $p = 0.46$, respectively). Overall, our indicators for the anticipation of challenges are not significantly predictive of performance.
- *Mood*: Finally, a positive mood in the chat is predictive of performance and efficient equilibration. Both smalltalk and positive emoticons positively correlate with performance and equilibration. It therefore appears that soft skills and a positive atmosphere within a group are predictive of performance.

How does this analysis inform the failure of the Nudge to significantly increase performance in Phase 2? Let us consider again what kind of variables from the extended chat codings are correlated with performance in Phase 2, and what they tell us about the groups: First, whether suboptimal proposals were made, explanations were necessary and the time it took groups to reach the 10-70 agreement seems mostly like a proxy of the general understanding and intellectual capability in a pair. We could not reasonably expect that the Nudge Treatment affects these variables (and there is also no evidence of that, see Table 7). The mood variables that are positively correlated with performance capture to some degree the soft skills present in the group. Again, this is not something that the Nudge Treatment targeted directly. Finally, promises were significantly correlated with Phase 2 profits and also appeals to fairness and honesty or mentions of trust were positively, albeit not significantly, correlated with Phase 2 performance. Looking back at Table 7, the incidence

of promises was actually reduced in the Nudge Treatment. The reduced incidence of expressions of trust and the making of promises may in fact indicate that the Nudge may have fostered the appearance of principles in the statements and chats, but not really triggered a discussion and commitment to adhere to these principles (as indicated by a promise), and reduced the incidence of social preference related mechanisms to enforce the principles. This may be one explanation for the lack of treatment effect and decline in performance among groups with principles that should be explored in further research.

B.4 Exploring Reasons behind the Lack of Increased Price Equilibration in the Nudge Treatment

In order to explore the reasons behind the lack of price equilibration in the Nudge Treatment, Table 9 shows whether any of the additionally coded variables capturing chat content has predictive power for whether or not pairs that achieved quality equilibration also achieved price equilibration. The sample only includes pairs who were equilibrated on quality in scenario 1 (the one actually implemented in Phase 2). As can be seen, all these groups had reached the 10-70 agreement. While columns (2) and (3) split the sample into Baseline and Nudge, we focus our interpretation on column (1), because otherwise sample sizes become very small. The following results emerge:

- It can be seen that the need for explanation of the 10-70 strategy for quality and price negatively correlates with price equilibration, as does the time it took to reach the 10-70 agreement. Threats and appeals are marginally significantly correlated with price equilibration. Pairs that achieved price equilibration reached the 10-70 agreement significantly faster. Both these differences point to selection: pairs that had less trouble finding the optimal and fair solution in Phase 1, and were able to do so faster, were more likely to achieve price equilibration.
- Moreover, we find differences between pairs that achieved price equilibration versus those who did not for mentions of threats as well as for appeals to fairness and honesty. These soft factors, seem to enable price equilibration.
- As we have previously seen, the factors predictive of price equilibration were not significantly affected by our treatment, again suggesting that the Nudge did not enable groups to focus

on the right aspects to foster equilibration and performance.

Table 9: Differences in Price Equilibration in Scenario 1 Conditional on Quality Equilibration

Price equilibration reached?	(1) All Pairs (N=61)			(2) Nudge (N=39)			(3) Baseline (N=22)		
	no	yes	<i>p</i>	no	yes	<i>p</i>	no	yes	<i>p</i>
Phase 1 strategy:									
Suboptimal proposal	0.32	0.15	0.13	0.28	0.19	0.54	0.40	0.08	0.11
Explanation necessary for 10-70	0.25	0.06	0.05**	0.28	0.10	0.16	0.20	0.00	0.17
Reached agreement on 10-70	1.00	1.00	.	1.00	1.00	.	1.00	1.00	.
Time to agreement on 10-70	3.89	2.40	0.04**	3.82	2.56	0.16	4.03	2.12	0.16
Promises and Trust:									
Trust mentioned	0.18	0.18	0.97	0.17	0.10	0.53	0.20	0.33	0.50
Promise	0.11	0.18	0.41	0.06	0.10	0.65	0.20	0.33	0.50
Threats	0.00	0.12	0.04**	0.00	0.19	0.04**	0.00	0.00	.
Appeals to honesty and fairness	0.14	0.36	0.05**	0.11	0.33	0.10*	0.20	0.42	0.29
Phase 2:									
Foreshadowing misunderstandings	0.11	0.12	0.87	0.17	0.14	0.84	0.00	0.08	0.34
Mention Phase 2 uncertainty	0.50	0.42	0.56	0.67	0.57	0.55	0.20	0.17	0.85
Deviations discussed	0.07	0.21	0.11	0.11	0.24	0.30	0.00	0.17	0.17
Mood:									
# of Greetings	1.61	1.58	0.86	1.44	1.52	0.74	1.90	1.67	0.29
Smalltalk	0.11	0.24	0.17	0.11	0.24	0.30	0.10	0.25	0.37
Positive emotions expressed	0.75	0.67	0.48	0.78	0.71	0.66	0.70	0.58	0.59
# of positive emoticons	3.61	3.97	0.72	3.72	4.10	0.77	3.40	3.75	0.84
# of negative emoticons	0.00	0.00	.	0.00	0.00	.	0.00	0.00	.
Observations	28	33		18	21		10	12	

Notes: The table shows means of the additionally coded chat variables split by whether pairs that had reached quality equilibration also reached price equilibration (in the scenario implemented in Phase 2). The *p*-values stem from *t*-tests testing for significant differences between the respective splits.

In addition to the chat codings displayed in Table 9, we also coded again all statements for whether they mentioned prices and / or quality. For both price and quality we differentiated whether the statement mentioned it either explicitly (i.e., in numerical terms) or more in a general fashion (e.g., “the buyer pays a fair price”). Table 10 shows the results both in terms of treatment differences (Base vs. Nudge), as well as in terms of price equilibration.

We find that the pairs in the Baseline were significantly more likely to mention both price and quality in an explicit (numerical) manner in their statements, whereas pairs in the Nudge Treatment were directionally (but not significantly) more likely to make general references to price and quality, likely driven by the push towards principles by the Nudge. However, splitting the data by whether a pair achieved price equilibration shows that pairs who were successful in terms of price equilibration were actually significantly less likely to have mentioned an explicit price (and

an explicit quality) in their statement. Thus, it does not seem to be the case that this difference in the way prices were mentioned in the statements (either explicitly or in a more general fashion) drives the difference in price equilibration between Baseline and Nudge pairs.

Table 10: Differences in Mentions of Price and Quality in Statements by Treatment and Price Equilibration

	(1) Treatment Comparison			(2) Price Equilibration		
	Base	Nudge	p	No	Yes	p
Price explicit	0.95	0.78	0.02**	0.96	0.70	0.01**
Price general	0.11	0.22	0.14	0.19	0.22	0.74
Quality explicit	0.91	0.78	0.08*	0.96	0.63	0.00***
Quality general	0.13	0.22	0.23	0.15	0.26	0.32
Observations	56	46		27	27	

Notes: The table shows the means of additional codings of the statements for whether price and / or quality were mentioned explicitly or in a general fashion. The p -values stem from t -tests testing for significant differences between the respective splits.

Taken together, our interpretation of these results is that a successful treatment should foster the making of (credible) promises, appeals to honesty and fairness, and foster understanding of the basic structure of the game. Such a treatment may in turn increase equilibration and performance. Our Nudge Treatment did not sufficiently move these levers to generate a significant effect.

B.5 Ratings of Agreements by Participants

In the post-experimental questionnaire, participants responded to a number of questions asking them to judge the caliber of their agreements. To the extent that (at least in retrospect) the participants themselves are probably the best judges of what was good and what was bad about their agreements, this data can provide interesting insights about whether there was indeed a caliber difference between principles in the Baseline and principles in the Nudge Treatment that could explain the differences in performance. We analyze it in this section.

Table 11 shows the means of participants' ratings of statements about the caliber of their agreements or the consequences the agreements had. All statements always referred directly to the agreement a pair had made at the beginning of the study. Column (1) again provides a treatment comparison. We can see that the Nudge did lead to agreements that were judged as significantly more flexible and able to adapt to the new situation in Phase 2 (or potentially also to other situations). This is unsurprising as the Nudge basically told participants to come up with such

agreements. We can also see, however, that this flexibility may have come at the cost of clarity, as the agreements in the Baseline were judged as significantly more clearly formulated. This could be, of course, because there were more rules than principles in the Baseline, and it is easier to formulate a rule clearly than it is to do the same for a principle. Columns (2) and (3) are therefore insightful, as they separate the pairs into principles and rules groups.

Table 11: Participants’ Ratings of Agreements in Exit Questionnaire

	(1) All Pairs			(2) Principles			(3) Rules		
	Base	Nudge	<i>p</i>	Base	Nudge	<i>p</i>	Base	Nudge	<i>p</i>
Helpful in Phase 2	3.36	3.54	0.46	4.08	3.87	0.47	3.07	1.50	0.00***
Led to misunderstandings in Phase 2	2.60	2.47	0.58	2.00	2.24	0.42	2.88	4.00	0.00***
Differing interpretations	2.18	2.18	0.99	1.56	1.99	0.09*	2.46	3.14	0.14
Clearly formulated	3.88	3.41	0.01**	4.25	3.56	0.01***	3.77	2.64	0.01**
Flexible enough to adapt to Phase 2	3.10	3.66	0.01**	3.89	3.91	0.93	2.78	2.07	0.07*
Would also have been flexible enough to adapt to other situations	3.02	3.54	0.01***	3.69	3.68	0.96	2.74	2.43	0.45
Partner respected agreement	3.63	3.65	0.93	4.42	3.90	0.13	3.32	2.36	0.00***
Subject respected agreement	3.58	3.87	0.21	4.36	4.13	0.43	3.27	2.36	0.04**
Led to conflicts	2.80	2.73	0.81	1.83	2.46	0.11	3.23	4.14	0.02**
Observations	60	61		18	52		41	7	

Notes: The table displays means of participants’ ratings of their agreements in the post-experimental questionnaire. Participants indicated the extent to which a given statement applied to the agreement they had made with their partner on 1-5 Likert scales (with 1 meaning “does not apply” and 5 meaning “applies”). The *p*-values stem from *t*-tests testing for significant differences between the Baseline and the Nudge Treatment.

Interestingly, the results in column (2) show that also when comparing principles in the Baseline vs. principles in the Nudge Treatment, participants judged the principles as having been more clearly formulated in the Baseline than in the Nudge Treatment. This is a further indication that indeed, the Baseline principles were of higher caliber than the Nudge principles. Some further results in column (2) indicate the consequences of this difference in clarity of the principles in the Baseline compared to the Nudge. We see that the Nudge principles were marginally significantly more prone to differing interpretations within a pair. They also tended to lead to more conflicts and participants were less likely to judge that the partner had respected the agreement (likely a consequence of differing interpretations), albeit these two latter effects are not statistically significant.

The results in column (3) show very clearly that the pairs that did not manage to come up with a principle in the Nudge Treatment (despite having been told to do so), and only managed to

come up with a rule, had significantly lower ratings on almost all variables capturing the calibre of the agreement than the rule only pairs in the Baseline. This pattern could be caused by at least two things. First, there was a negative selection into the rules only group in the Nudge Treatment (i.e., pairs who did not manage to come up with a principle even though they were told to do so). Second, in the Baseline there were some pairs that we coded as having a rule only (because what they wrote down looked like a rule), but their answers indicate that the pair actually had rather a principle in mind.

In sum, the analysis of participants' ratings of their agreements provides clear indications that there was indeed a difference in caliber between the principles in the Baseline compared to the Nudge Treatment. Whereas the Nudge did lead to more principles, it failed to trigger principles that were clear enough to be helpful. Our data are limited, however, in allowing us to identify what exactly is needed to create such clarity in a principle. The results discussed in Section B.3 above (see Tables 7 and 8) provide some indications that pairs in the Nudge Treatment spent less time on the building of trust and the making of promises than pairs in the Baseline. However, our current evidence in this regard is rather weak and ultimately lacks statistical power. We believe that investigating more deeply how exactly principles need to be built in order to ensure clarity and achieve stable cooperation in unstable environments is an exciting avenue for future research.

C Instructions First Trading Phase

Starting on the next page, we reproduce the original experimental instructions for the first trading phase handed out to participants. We provide the instructions given to sellers as an example. Instructions for buyers contained the exact same information and were presented in the same way. The experimental manipulation in the instructions in the Nudge treatment is contained on the final page of these instructions.

General Instructions for Sellers

You are now taking part in an economic study funded by various research funding agencies. Depending on your decisions, you can earn money. It is therefore very important that you read these explanations carefully.

These instructions are for your private information only. **There is an absolute ban on communication during the study.** If you have any questions, please contact us. Non-compliance with this rule leads to exclusion from the study and from all payments.

At the beginning of the study you will receive a show-up fee of 10 Dollars. During the study, we are not talking about Dollars, but about points. In the course of the study, you can earn a further amount of money by earning points. Your total income will therefore be calculated in points first. The total number of points you earn during the study is then converted into Dollars at the end, with

1 Point = 0.04 Dollars.

At the end, you get the points earned during the study plus the 10 Dollars show-up fee in cash.

At the beginning of the study, the participants were randomly divided into two groups: Buyer and seller. **You are a seller throughout the entire study.** Each seller was randomly assigned to a buyer. **The study consists of 2 parts.** The first part consists of 5 rounds. The second part consists of an unknown, randomly determined number of rounds. You will be assigned to the same buyer in all rounds, both in part 1 and in part 2. None of the other study participants will know your exact assignment or your decisions. Anonymity will therefore be preserved.

On the following pages you will first be informed about the exact procedure of the first part of the study. Information on the second part of the study will be provided towards the end of the instructions.

Course of the first part of the study

At the beginning of the study, each seller is randomly assigned to a buyer. Part 1 of the study comprises 5 rounds. You will be assigned to the same buyer in all 5 rounds of the first part, and also in the following second part of the study.

In each round, each seller carries out a transaction with his assigned buyer: The seller determines the quality of a product for which the buyer pays a price. The buyer makes a profit from the transaction if she pays less for the product than the product is worth to her. The value of the product to the buyer depends on the quality of the product. The seller makes a profit on the transaction if he receives a price that exceeds his production costs. Production costs depend on product quality. Higher quality is always associated with higher production costs.

Before the first part of the study starts, you also have the opportunity to **formulate an agreement in the form of a written statement together with the buyer**. How you can draft this statement will be explained later.

Course of a round

You and your assigned buyer complete one transaction in each of the 5 rounds. The process is organized as follows:

- Stage 1: You choose the product quality that you want to deliver to your buyer. At the same time, your buyer chooses the price she wants to pay you, without knowing what product quality you deliver.
- Stage 2: The buyer receives the product in the quality chosen by you and you receive the price determined by your buyer. Both will be informed about the income in the current round. Then the next round begins.

Stage 1: Determination of quality and price

At the beginning of each round you determine the product quality. The quality must be an integer value between 0 and 10:

$$0 \leq \text{Quality} \leq 10.$$

Quality and production cost

The choice of product quality is associated with costs. The higher the quality, the higher the production cost. The production cost for each product quality is given by the table below.

Product Quality	0	1	2	3	4	5	6	7	8	9	10
Cost of Production	0	1	3	6	9	13	18	23	28	33	40

Hence, a product quality of 0 costs nothing. A product quality of 1 costs one point. A product quality of 10 costs 40 points. The values for all other qualities are given in the table above. You choose the product quality and thus the production cost without knowing the price that the buyer will pay.

You can select the quality using the following screen display:

Stage 1: Choice of Quality

Please indicate the quality you wish to provide.

Quality

Please note the different costs for the different quality levels (see table in the written instructions).

At the same time, the buyer chooses the price she wants to pay.

Price

While you choose the product quality, the buyer determines what price she wants to pay, without knowing the quality you will deliver. The price cannot be lower than 0 and not higher than 100:

$$0 \leq \text{Price} \leq 100.$$

The buyer enters the price via the following screen display:

Stage 1: Choice of Price

Please enter the price you would like to pay.

Price

At the same time, the seller chooses the quality he wants to provide.

Stage 2: Calculation of incomes

Your income depends on the price the buyer pays and the quality of the product you choose to provide.

Your income in a round is calculated as follows:

$$\text{Your Income} = \text{Price} - \text{Production Cost of the Provided Quality}$$

The higher the price chosen by the buyer, the higher your income will be. In addition, the lower the production cost, the higher your income will be. As described above, more quality always costs more. The exact values can be found in the table above.

The buyer's income depends on the price she has paid and the value of the product for her. The value of the product is determined by the product quality you choose to provide as follows:

$$\text{Value of the Product} = 10 * \text{Provided Quality}.$$

With a quality of 0, the value of the product for the buyer is 0. Each additional quality step increases the value of the product by 10 points. For a quality of 1, the value is 10 points, for a quality of 2, the value is 20 points, and so on. For a quality of 10, the value is 100 points.

The income of the buyer is thus as follows:

$$\begin{aligned} \text{Buyer's Income} &= \text{Value of the product} - \text{Price} \\ &= 10 * \text{Provided Quality} - \text{Price} \end{aligned}$$

The higher the quality of the product you deliver, the higher the buyer's income, as a higher quality results in a higher value of the product. At the same time, the lower the price the buyer pays for the product, the higher her income will be.

Note that buyers and sellers can also make losses in any period. These must be paid out of the respective show-up fee or income earned in other periods.

Examples

The following (arbitrarily chosen) examples illustrate how incomes are determined.

Example 1:

The seller chooses a quality of 8. The buyer pays a price of 50.

Buyer's income: $8 \cdot 10 - 50 = 30$.

Seller's income: $50 - 28 = 22$.

Example 2:

The seller chooses a quality of 4. The buyer pays a price of 33.

Buyer's income: $4 \cdot 10 - 33 = 7$.

Seller's income: $33 - 9 = 24$.

Example 3:

The seller chooses a quality of 7. The buyer pays a price of 30.

Buyer's income: $7 \cdot 10 - 30 = 40$.

Seller's income: $30 - 23 = 7$.

At the end of each round, you and the buyer can see a summary of the current round's outcomes on the screen. This includes

- The quality you chose to provide.
- The price paid by the buyer.
- Your income.
- The buyer's income.

After that, the current round is completed and the next round begins. In total, the first part of the study consists of 5 rounds. After the fifth round, the second part of the study begins.

The Second Part of the Study

In the second part of the study you are still assigned to the same buyer. You therefore continue to interact with the same person as in the first part. The second part of the study consists of an unknown number of rounds. In the second part of the study, some elements concerning the course of a round and the calculation of income will change. You will only be informed about the concrete changes at the beginning of the second part of the study.

The second part of the study ends randomly. At the end of each round there is a 90% probability that another round will be carried out. With 10% probability, the second part of the study ends and so does the study.

The Agreement

Before the first round of the first part of the study, you have the opportunity to make a non-binding agreement with the buyer assigned to you. It is a written statement, in which you can describe how you and the buyer make their decisions in this study.

You can draft the written statement on the following screen:

Remaining Time [Sec]: 895	
Chat:	Joint Statement:
	The joint statement appears here as soon as the buyer sends a draft to you.

The left window is a chat window. Here you can chat with the buyer and discuss your agreement. You have a total of 15 minutes to talk to the buyer and find an agreement. **If you do not come up with a written statement within these 15 minutes, the chat window will automatically close and the study will continue without a statement.**

Limit yourself to discussions concerning this study. Chat messages that have nothing to do with the study or identify a study participant may lead to exclusion from the study.

As soon as you have reached an agreement via chat, **the buyer writes a draft of the statement describing the agreement.** The buyer should write down your agreement as precisely as possible. Once the buyer has written the statement, it will be displayed in the right window. You can then accept or return it if you do not agree with the wording of the statement. In this case, you should again discuss adjustments to the statement via chat. The buyer can then adjust the statement accordingly. This process can be repeated within the 15 minute time limit until a statement of the agreement is drafted that is accepted by you and the buyer.

As soon as both you and the buyer have accepted the statement, or if the 15 minutes have expired, the first round of the first part of the study begins. You can display the statement on the decision screens as needed in each round of the first and second parts of the study.

[In Nudge Treatment only:

Please note:

When finding an agreement, you should bear in mind that you do not yet know the exact situation you will encounter in the second part of the study. It may therefore be helpful to consider not only the first part of the study that you already know about, but also the second part that you do not yet have information about. For example, you could think about the principles on which you and the buyer would generally like to act during the study.]

You have arrived at the end of these instructions. Thank you for your attention. **We now ask you to solve the exercises displayed on your screen.** The study can only begin when all participants have correctly completed the exercises. If you have any questions, please raise your hand.

D Instructions Measure of Shared Understanding and Second Trading Phase

Starting on the next page, we reproduce the original experimental instructions for the measure of shared understanding and the second trading phase handed out to participants. We provide the instructions given to sellers as an example. Instructions for buyers contained the exact same information and were presented in the same way. The instructions were identical for the baseline and the Nudge treatment.

Instructions for the Second Part of the Study for Sellers

In the second part of the study you are still assigned to the same buyer. In principle, the process of a round remains the same as in the first part of the study (the seller chooses the quality and the buyer decides on the price at the same time). However, some elements will change compared to the first part.

Before we begin with the second part of the study, we present three possible scenarios of how the situation might change compared to the first part. One of these scenarios is then actually implemented in the second part, i.e. in the second part you will interact with the buyer within one of the described scenarios.

Before the second part begins, and before you know which of the scenarios will actually be implemented, we ask you to answer a few questions for each of the scenarios. The first questions are about **what you and the buyer *should* do in a given scenario as agreed at the beginning of the study:**

1. a) **What quality should you deliver according to the agreement?**
 b) **What price should the buyer pay according to the agreement?**

In your answers, we ask you to **interpret your agreement in a broader sense**. It is possible that you and the buyer did not discuss a particular scenario or decision. In such a case, please consider what the overall meaning and content of the agreement was and base your answers on it.

If your answers to questions 1.a) and b) match those of the buyer, you will receive a bonus: For each question that you answer in the same way as the buyer, you will receive a bonus of 20 points. So think carefully about what decisions you and the buyer should make according to the agreement.

So while questions 1.a) and b) relate to what you and the buyer *should* do according to the agreement, the next questions, 2.a) and b) relate to **what you *will actually do* in the scenario and what you *believe the buyer will actually do*.**

2. a) **What quality do you want to deliver in this scenario in the first round of the second part of the study?**
 b) **What price do you think the buyer in this scenario will pay in the first round of the second part of the study?**

Your decision regarding quality in the first round of the second part of the study is binding and will be implemented exactly as you indicated if the corresponding scenario occurs in the second part. Think carefully about your answers.

Your answers to questions 1) and 2) may be the same, or they may differ. For each of the two questions we will also ask you to indicate how confident you are about your answer. Your statement about the certainty of your answers does not affect the bonus or your income in the study in general.

On the following pages we present the three scenarios in detail.

After you have answered the above questions, the second part of the study begins, in which one of the three scenarios is actually implemented. **The second part of the study ends randomly. At the end of each round, there is a 10% probability that the second part of the study will end and the study will be over. With a probability of 90%, the study continues and another round follows.**

As soon as random chance determines that no more rounds will be carried out, the study is over. You will then receive your income from the five rounds of the first part of the study, your income from the bonus questions (1.a) and b)) via the three scenarios, your income from all rounds of the second part of the study, and your show-up fee of 10 Dollars in cash.

In order for the screens for entering your answers to the various scenarios to appear, please enter the following code on your computer: 31-72-93

Scenario 1: Higher Value of the Product and Alternative Source of Supply for the Buyer

Please read the scenario in its entirety (including the examples at the end) before entering your answers on the computer screen!

In this scenario, the value of the product for the buyer increases and doubles compared to the first part:

$$\text{Value of the Product} = 20 * \text{Quality.}$$

In this scenario, the price must not be less than 0 and not more than 200:
 $0 \leq \text{Price} \leq 200$.

In addition, in this scenario the buyer **now has the option of purchasing the product from a different source than you**. The income that the buyer earns when choosing the alternative source of supply fluctuates randomly. Two situations can arise:

- Situation 1:
The alternative source provides a **quality of 0 at a price of 0**, so the **buyer's income is 0 if he chooses the alternative source** ($20*0 - 0 = 0$). This situation occurs with a **probability of 2/3**.
- Situation 2:
The alternative source provides a **quality of 10 at a price of 40**, so the **buyer's income is 160 if he chooses the alternative source** ($20*10 - 40 = 160$). This situation occurs with a **probability of 1/3**.

Important: Which of the two situations occurs is **decided randomly by the computer at the beginning of each period and you and the buyer are both informed before you make your decisions**.

So at the beginning of each period you and the buyer first learn how high the buyer's income is in that period if he chooses the alternative source of supply. Then, as in the first part, you determine the product quality without knowing whether the buyer will purchase the product from you or choose the alternative source. **The production costs for you are the same as in the first part.**

The buyer decides at the same time whether he wants to purchase the product from you or from the alternative source. If he purchases the product from you, he sets a price (which must be between 0 and 200) as in the first part. **If he obtains the product from the alternative source, you will receive a price of 0**, i.e. you will not receive any compensation in this case. **In this case, the buyer receives the income randomly determined by the computer.** All other parameters remain the same as in the first part of the study.

Your income will be calculated exactly the same as in part 1 of the study:

Your income = Price – Production cost of the provided quality

You always incur the cost of the chosen quality, regardless of whether the buyer chooses the alternative source of supply or purchases the product from you. If the buyer chooses the alternative source, the price you receive is 0.

The buyer's income depends on whether she purchases the product from you or chooses the alternative source:

Purchase of the product from you:

Buyer's Income = value of the product – price
 = 10 * provided quality – price

Alternative source of supply:

Situation 1 (quality = 0; price = 0): buyer's income = $20 \cdot 0 - 0 = 0$

Situation 2 (quality = 10; price = 40): buyer's income = $20 \cdot 10 - 40 = 160$

Examples

The following (arbitrarily chosen) examples illustrate how incomes are determined in this scenario.

Example 1:

The computer randomly determines that the buyer's income is 0 (quality of 0, price of 0) if the buyer chooses the alternative source of supply. Buyer and seller are both informed of this. The seller chooses a quality of 5. Without knowing the quality provided by the seller, the buyer decides to purchase the product from the seller and pays a price of 70.

Income of the buyer: $5 \cdot 20 - 70 = 30$.

Seller's income: $70 - 13 = 57$.

Example 2:

The computer randomly determines that the buyer's income is 160 (quality of 10, price of 40) if the buyer chooses the alternative source of supply. Buyer and seller are both informed of this. The seller chooses a quality of 3. Without knowing the quality provided by the seller, the buyer chooses the alternative source.

Buyer's income: $20 \cdot 10 - 40 = 160$.

Seller's income: $0 - 6 = -6$.

PLEASE ENTER NOW YOUR ANSWERS TO THE QUESTIONS CONCERNING THE ABOVE SCENARIO ON THE COMPUTER SCREEN!

Scenario 2: Fluctuating value of the product

Please read the scenario in its entirety (including the examples at the end) before entering your answers on the computer screen!

In this scenario, the value of the product varies. Two different situations can occur:

- **Situation 1:**

The value of the product to the buyer triples compared to the first part:

$$\text{Value of the product} = 30 * \text{Quality}.$$

This situation occurs with a **50% probability**.

- **Situation 2:**

The product is worthless for the buyer for each quality provided:

$$\text{Value of the product} = 0 * \text{Quality}.$$

This situation occurs with a **50% probability**.

Both cases are equally likely in each round, and the computer randomly determines which situation occurs in each round.

As in the first part, you continue to choose the quality you want to provide.

Important: You make this choice without knowing the value of the product for the buyer in this round.

As in the first part, the buyer again determines the price without knowing what quality you are providing.

Important: In contrast to you, the buyer knows whether situation 1 or 2 has occurred in the current round and how high the value of the product is (depending on the quality).

In this scenario, the price must not be lower than 0 and not higher than 300:

$$0 \leq \text{Price} \leq 300.$$

All other parameters, especially the production costs, remain the same as in the first part of the study.

Your income is calculated exactly the same as in part 1 of the study:

$$\text{Your income} = \text{Price} - \text{Production cost of the provided quality}$$

The buyer's income now depends on the realized value situation:

$$\begin{aligned}\text{Situation 1: Income of the buyer} &= \text{value of the product} - \text{price} \\ &= 30 * \text{provided quality} - \text{price}\end{aligned}$$

$$\begin{aligned}\text{Situation 2: Income of the buyer} &= \text{value of the product} - \text{price} \\ &= 0 - \text{price}\end{aligned}$$

Examples

The following (arbitrarily chosen) examples illustrate the determination of income in this scenario.

Example 1:

The seller chooses a quality of 3. Situation 1 occurs and the value of the product is three times as high as in the first part. The buyer knows the value situation (but not the quality chosen by the seller) and pays a price of 35.

The buyer's income: $3*30-35 = 55$.

Seller's income: $35-6 = 29$.

Example 2:

The seller chooses a quality of 7. Situation 2 occurs and the value of the product is 0. The buyer knows the value situation that has occurred (but not the quality chosen by the seller) and pays a price of 40.

Income of the buyer: $0-40 = -40$.

Seller's income: $40-23 = 17$.

PLEASE ENTER NOW YOUR ANSWERS TO THE QUESTIONS CONCERNING THE ABOVE SCENARIO ON THE COMPUTER SCREEN!

Scenario 3: Higher value of the product and possible loss of price

Please read the scenario in its entirety (including the examples at the end) before entering your answers on the computer screen!

In this scenario, the value of the product for the buyer increases and doubles compared to the first part:

$$\text{Value of the product} = 20 * \text{Quality.}$$

The price in this scenario must not be less than 0 and not higher than 200:

$$0 \leq \text{price} \leq 200.$$

In this scenario, it is also possible that the price paid by the buyer to the seller will not reach the seller.

All other parameters, especially production costs, remain the same as in the first part of the study.

So there are two possible situations in this scenario. One in which the price paid by the buyer does not arrive at the seller and another in which the price paid by the buyer flows into the seller's income (as in the first part).

- **Situation 1:**

The price paid by the buyer **does not reach the buyer**. I.e. the buyer pays the price, but you do not receive it. This situation occurs with **50% probability**.

In this situation, your income is calculated as follows:

$\text{Your income} = 0 - \text{production cost of the provided quality}$

The buyer's income is calculated similarly to the first part:

$\begin{aligned} \text{Buyer's income} &= \text{value of product} - \text{price} \\ &= 20 * \text{provided quality} - \text{price} \end{aligned}$

- **Situation 2:**

The price paid by the buyer **arrives as in the first part and increases your income**. This situation is **50% likely** to occur.

In this situation, your income is calculated as in the first part:

$$\text{Your income} = \text{price} - \text{production cost of the provided quality}$$

In this situation, the buyer's income is calculated exactly as before:

$$\begin{aligned}\text{Buyer's income} &= \text{value of the product} - \text{price} \\ &= 20 * \text{provided quality} - \text{price}\end{aligned}$$

Both cases are equally likely in each round, and the computer randomly determines which situation occurs in each round.

As in the first part, you continue to choose the quality you want to provide and the buyer determines the price (between 0 and 200) he pays (without knowing the quality you provide).

Important: Before you make your decisions, neither you nor the buyer know whether situation 1 or 2 will occur and whether the price paid by the buyer will reach you.

Examples

The following (arbitrarily chosen) examples illustrate the determination of income in this scenario.

Example 1:

The seller chooses a quality of 5. The buyer pays a price of 60 (without knowing the quality provided by the seller). Situation 1 occurs and the price paid does not reach the seller.

Income of the buyer: $5 * 20 - 60 = 40$.

Seller's income: $0 - 13 = -13$.

Example 2:

The seller chooses a quality of 8. The buyer pays a price of 110 (without knowing the quality provided by the seller). Situation 2 occurs and the price paid arrives at the seller.

Income of the buyer: $8 * 20 - 110 = 50$.

Seller's income: $110 - 28 = 82$.

PLEASE ENTER NOW YOUR ANSWERS TO THE QUESTIONS REGARDING THE ABOVE SCENARIO ON THE COMPUTER SCREEN!

E Instructions to Coders

Starting on the next page, we reproduce the instructions provided to coders containing definitions of the different coding categories.

Coding of Statements and Chats

Background: Structure of the Experiment

In the experiment, each participant was randomly assigned a role, buyer or seller, and buyer-seller pairs were randomly formed. These pairs remained together throughout all periods of the experiment. In each period of the experiment, the buyer-seller pairs could complete a transaction, i.e. the buyer could purchase a product from the seller. The seller determined what quality he wanted to deliver (the higher the quality, the higher the production costs for the seller and the higher the value for the buyer) and the buyer determined what price she wanted to pay for the product.

The experiment was divided into two parts. For the first part, the participants received detailed instructions describing in detail all the rules and parameters of the buyer-seller interaction. Please read these instructions carefully, so that you know exactly how the experiment worked and what level of information the participants had. The participants did not receive any detailed information about the second part at the beginning of the experiment. They only knew that they would interact with the same buyer or seller in the second part.

Before the first part started, the participants could chat with each other for 15 minutes to find an agreement on how they would like to act during the experiment (see the relevant sections at the end of the instructions) and to write this down in a joint statement. Your task is to encode these chats and the statements.

Definitions of Coding Categories

Below are the definitions of the different categories according to which you should encode chats and joint statements. Before you start coding, the definition of each of these categories must be clear and familiar. The definitions refer to the possible content of the final agreements and previous chats. So you have to decide if there is any content in an agreement or chat that falls into one of these categories. The categories are not mutually exclusive, i.e. it is possible that an agreement or chat content contains multiple categories.

Fixed rule:

Unique rule that clearly defines numerically a quality and the price to be paid for it.

Examples:

- *The seller delivers a quality of 10 and the buyer pays a price of 70.*
- *Maximum quality and price of 80.*
- *Quality = 7, Price = 44*

Flexible rule:

Rule that numerically defines quality and price, but leaves room for manoeuvre by leaving an acceptable range for quality and / or price.

Examples:

- Quality from 6-8, price = 44
- The quality must never fall below 2 and the price must always be higher than 50.

Suboptimal rule:

Agreed rule is not welfare maximizing, i.e. the rule provides for a quality < 10 (does not apply to flexible rules where 10 is a possible, but not the only possible, quality).

Example:

- All cases with quality < 10.

Unequal rule:

Agreed rule leads to unequal distribution of profit between buyer and seller. Attention: with certain qualities, due to the cost structure (and because the price could only be chosen as an integer), it is not possible to divide the profit exactly equally. In such cases, the code should only be used if the difference in profit is >1. In the case of flexible rules where the same distribution is a possible solution, please do not code here.

Examples:

- Price ≠ 70 with a quality of 10
- Price ≠ 61 or 62 with a quality of 9 (NB: this is a case in which an exactly equal distribution is not possible). At a price of 61 or 60, however, the difference in prize is only one point. 62 or 61, with a quality of 9, should therefore not be coded as unequal, but all other prices with a quality of 9 should be).
- Price ≠ 54 with a quality of 8

Principle:

The participants formulate a principle that defines how to act in general. It is not based solely on a numerical definition of the quality to be delivered or the price to be paid (but can be supplemented by such numerical definitions), but provides overarching, general guidelines for action that can be applied in various situations.

Examples:

- We always provide each other with the greatest possible benefit.
- We minimize the risk for the seller.
- We maximize the existing value and divide the profit equally.
- We look at the situation in Part 2 and then adjust quality and price so that it is acceptable to both.
- Mention of Fairness ("FairErw")
- Participants mention fairness considerations in the agreement / chat. This includes avoiding inequality or striving for equality or discussing what is unfair or fair behaviour.

Mention of Fairness:

Participants mention fairness considerations in the agreement / chat. This includes avoiding inequality or striving for equality or discussing what is unfair or fair behaviour.

Examples:

- We try to be fair to each other.
- We want to balance the profits.
- If someone acts unfairly, he will be punished.
- Not paying the agreed price would be unfair.

Mention of efficiency

The participants mention efficiency considerations in the agreement / chat. This is particularly about welfare maximization discussions, i.e. when participants discuss implicitly or explicitly that they want to maximize the sum of payoffs.

Examples:

- We optimize the sum of the profits.
- Optimal quality where marginal costs = marginal benefit

Mention of the second phase

The still unknown second phase of the experiment is mentioned or discussed.

Examples:

- In the second phase we will do xy.
- What happens later in the experiment?

No agreement

"No agreement concluded" (to be coded only for agreements).

Examples:

- You have not made an agreement
- Empty field

Coding Units and Data Structure

As mentioned before, you have to encode both statements and chats. These are located in different folders in the Excel sheet you will receive. Each statement and each chat is uniquely assigned to a group by a "Group_id" number. Please read and encode first the final agreement of a group and then the chat.

For the joint statements there is one line per group and the codes must be entered into the columns in the same line to the right of the statement. Please enter a "1" in the corresponding column if an agreement contains elements that fall into one of the given categories. You do not have to enter a "0" if the corresponding category does not appear in the agreement.

The chats have multiple lines per group. One line corresponds to one chat message. It is about encoding the chat of a group as a whole and not the individual chat messages. I.e. it is not absolutely necessary for us to know in which line (at which position in the chat) a certain coded element occurs, but only whether it occurs somewhere in the chat. Since we consider the chats as a whole, it is also

important that the encodings correspond to the "consensus" that has developed in the chat and that individual statements are not encoded in isolation. For example, if a participant suggests a suboptimal rule, but the two participants realize during the chat that there are better solutions, then "suboptimal rule" should not be coded for this chat. To ensure this, it is best if you read the chat completely before you add the codes and that you enter the codes in the last line of the chat. This of course means that you may have to read a chat (or certain parts of it) several times. Please do this if necessary. It is very important that the encoding is accurate and consistent.