Promises in situations with uncertain outcomes

January 13, 2015

Abstract

In this study we investigate how the commitment effect of a promise interacts with uncertainty about the decision situation. In the existing studies on promises subjects always know the options they can choose from before making a promise. Subjects play out the different scenarios when they make the promise and have already decided if they will keep the promise in the later decision stage. By introducing uncertainty about the possible outcomes we are able to disentangle (stated) intentions from the final manoeuvring in the decision situation. We also measure the heart rate variability to test for emotional balancing between guilt aversion (promise keeping) and the temptation of a material gain (promise breaking). Our results show that a free form message reduces significantly behavioural variation caused through changes in the decision environment.

keywords: promise, moral manoeuvring, outcome uncertainty

1 Introduction

People are moral relativists, concluded the psychologists Goldstone and Chin after observing how honest employees at the University Michigan self reported copies made at a communal copy machine. The copies were costly and the authors find that intermediate degrees of dishonesty are widespread, in 40% of the cases people under-reported the actual amount of copies made, but only a minority cheated maximally and reported no copies. This early and prominent experiment in Psychology laid the ground for many following studies examining how external incentives and internal constraints form actions an the consequences in economic and organizational contexts**???**. Individuals sacrifice financial gains in order to comply with social norms, moral values, even in anonymous settings, to obtain a positive self-image (**??**Mazar et al., 2008). Social and moral behaviour is hereby not a constant but rather the result of an internal manoeuvring process(Shalvi et al., 2011), in which the individual tries to serve the self-interest, but also satisfy the desire to act morally appropriate. Socially oriented behaviour is therefore sensitive to material gains and costs of a decision situation.

In our experiment we investigate how an ex ante made promise for social behaviour influences the variability of social behaviour. Does a promise make the tradeability of social behaviour more inherent? In a trust game, the social optimum outcome can be reached when the first mover chooses send his endowment to the counter player.The amount gets multiplied, and the second mover has to decide how to allocate the gains. Two choice options are offered in our design, an equal split and an unequal split, whereby the material gains for the second mover, so to say the material temptation to behave asocial, vary over the course of the game and are randomly realized. Before interactions started, second movers had the possibility to commit themselves in a non-binding statement to a certain action, when trusted. The message

was transmitted to each of the counter players before the trusting decision needed to be made. When the first mover decided to trust, a random draw determined the two available options the second mover could choose from. Promise-makers, who committed themselves to split equally, consequently did not know the material benefits they need to forgoe in order to keep the promise. This characteristic contributes to the existing literature in the following way.

To our knowledge, all existing studies on promises present the entire strategy and outcome space to subjects prior the promise making stage. Hence, participants are able to simulate the choice of the two options, promise keeping and promise breaking, and outweigh the material and emotional consequences, before making the promise. But the character of foreseeability lacks in our opinion external validity. Our study, in contrast, introduces uncertainty about the outcome choices. The promise making is thus disentangled from the manoeuvring process whether to split equally, and keep the potential promise, or follow the material temptation and choose the unequal split.

In our understanding, this scenario is closer to reality; in daily situations people make a promises to facilitate e.g. exchange, without having perfect knowledge of the future state of the world the promise will apply to. Our results show that with a free form promise the variability in behaviour is significantly reduced and makes the behaviour of a second mover more predictable. The promise increases the moral obligation to act fairly, after the public commitment

By measuring the heart rate variability of subjects we gain profound insights in the internal decision-making process. Our results suggest that subjects indeed outweigh their decisions and the moral calculus is mirrored in the heart rate.

The article proceeds as follows; Section 2 gives a short overview of promises

as a commitment device and summarises the relevant studies on promises and uncertainty as well as related studies that also captured physiological reactions during the experiment. The experimental design and procedure are described in Section 3, which also gives an introduction in the measurement of heart rate variability. Section 4 presents the results and Section 5 summarizes and concludes.

In opposite to contracts, informal commitments like promises have not direct consequences. However economic agents have an incentive to live up to their promise to secure future interactions and maintain a good reputation in settings of repeated interactions. If the interaction in contrast is one shot and the identity anonymous, a purely in payoffs interested agent has no incentive to keep its word. A wide branch of laboratory studies (**?**Vanberg, 2008; Hurkens and Kartik, 2009, e.g.) discount this prediction derived from standard economic theory. A substantial percentage of subjects keep their promise, even in the abstract and anonymous setting of a laboratory experiment. The literature explains this compliance with internal costs. Letting the expectations of others down (Charness's belief based guilt aversion) or deviating from the social/moral norm of truth telling (Vanberg, 2008; Binmore, 2006) produces internal tension that the promisor wants to avoid. Both explanations have in common that breaking a promise is linked to emotional costs.

All exiting studies on promises present, to our knowledge, hereby the entire strategy and outcome space to subjects prior the promise making stage. Hence, the participants are able to simulate the choice of the two options, promise keeping and promise breaking, and outweigh the material and emotional consequences, before making the promise. The character of foreseeability lacks in our opinion external validity.

Our study contains uncertainty how costly it will be to keep the promise. After the promise is made a random draw implements two alternative op-

tions the promisor can choose from. One option will let him or her keep the promise, by choosing the other option the promisor breaks the promise. We disentangle the promise making moment from the decision whether to keep of break the promise. Following the literature the promisor faces an inner conflict at this stage and outweighs the costs and benefits from promise breaking. We assume that this process causes emotional stress which we capture by measuring the heart rate variability of subjects. We find evidence that subjects indeed outweigh their decision and the moral calculus is mirrored in the heart rate. The article proceeds as follows; Section 2 gives a short overview of promises as a commitment device and summarises the relevant studies on promises and uncertainty as well as related studies that also captured physiological reactions during the experiment. The experimental design and procedure are described in Section 3, which also gives an introduction in the measurement of heart rate variability. Section 4 presents the results and Section 5 summarizes and concludes.

2 Literature

Standard economic theory predicts that all sorts of cheap talk, communication without direct consequences, should not affect the strategic behaviour of an agent. However, a wide range of empirical studies provide evidence for opposing behaviour. Communication can foster cooperation (see Balliet, 2009, for metastudy). A fast emerging branch in the literature is investigating how statements about intended behaviour can serve as a commitment device and why people keep a promise against the predictions of standard economic theory. Participants in laboratory experiments are willing to sacrifice material gains to avoid lying, even in anonymous one shot interactions (Gneezy, 2005). The manifestation of lying aversion is hereby found to be heterogeneous among but also within an agent (**??**). Participants are more likely to lie when the benefit is increasing or the harm to others is decreasing (Gneezy, 2005; Lundquist et al., 2009). Shalvi et al. (2011) use the term ethical manoeuvring in respect to lying behaviour and suggest the concept of an internal cost- benefit- analysis. A sufficient amount of profit is needed to compensate for internal discomfort caused by lying. On the other side if the lie is too big, the emotional/moral discomfort outweighs the material gain. In consequence people avoid lying a lot but also a little. We assume to see a similar balancing process for promise keeping, breaking. Where the motivation to keep a promise has its origin is controversial. Some scholars see belief dependent guilt aversion as the reason for promise keeping; the promise raises the expectations of the counter player, so that the promisor fears emotional costs, guilt, from letting the other person down (Charness and Dufwenberg, 2006). The second explanations ties in with a social norm perspective; keeping one's word is a moral duty (Bowles, 2008; Vanberg, 2008) . In both argumentation the promisor bears internal costs from promise breaking.
Our study ties in with the existing literature in two ways. First we investigate

how the commitment of a promise affects behaviour in a changing environment. How doe the commitment interplay with uncertainty? . Schweitzer and Hsee (2002) showed that under vagueness about the state of the world it is easier to stretch the truth. People are more likely to lie under uncertainty. In the study participants were in the role of a car seller and learned how much miles they have driven with the car. In one treatment the range of driven miles was between 74 and 76 k, in another treatment the mile level lied between 60 and 90k. The authors measured how the width of the range corresponds with lying about the level reported towards a potential buyer. People were more likely to lie when the range was wider.

? investigated how imperfect information in an ultimatum game influences the willingness to deceive others in a pre-communication phase. If recipients did not know the size of the pies, lies on side of the allocators were more common.

Serra-Garcia et al. (2011) studied which impact vagueness in communication has on the frequency of lies. The authors found that when information about an intended contribution to a public good game needed to be precise, lies can be observed frequently. However in the treatment where vague messages were allowed, inconvenient outcomes are disguised by vague messages.

Also the prominent study of (Charness and Dufwenberg, 2006) on the effect of promises in a trust game included an unforeseeable move by nature. After the first mover did the transfer, a second mover had to decide whether to roll a dice or not. While the not rolling decision provided a higher payoff for himself, the roll choice increased also the payoffs of the first mover with a high chance. The payoffs of the second mover were hereby constant over the possible nature outcomes, no uncertainty existed about his own payoffs.

Our study in contrast incorporates uncertainty about the own possible payoffs and in specific about the costs of keeping a promise. It is unknown to the promisor as well as to the promisee how much material gain the promisor

has to sacrifice to keep his promise later in the game. Previous studies showed that lying/ promise breaking behaviour reacts on changes in potential payoffs. In treatments where lying was more materially beneficial or the caused material harm for the counter part was smaller, higher rates of lying could be observed (Gneezy, 2005). The comparison was carried out in a between treatment design. In our study subjects face varying options from a linear function of benefit and harm over time. The costs of promise keeping increase sequentially over the range of possible outcomes. Hence we are able to observe manouvering behaviour of the promisors depending on the varying costs of promise keeping. A feature that might provide empirical evidence for the theoretical claim about the convexity of lying costs (Kartik, 2009).
During the experiment we measured the heart rate variability of participants. We add to the few studies who observe physiological change in lying situations. Assuming that the information about the actual costs of promise keeping causes a rethinking in the actors we want to test how physical stress/ emotional costs interact with the decision whether to keep of break a promise.
Wang et al. (2010) measured the eye movement in a sender receiver game with a pre-communication stage. The study makes use of the fact that the pupil dilates under stress, but also under cognitive difficulty and arousal or pain. While the pupil dilated more when senders deceived their counterpart, it stays unclear if the movement was because the sender felt guilty or because if the situation was cognitively more demanding for the deceiver.
Wibral et al. (2012) tested how an artificial increase in testosterone level affects self-serving lying. Numerous studies have shown that men are more willing to lie in order to obtain material benefits than women **???**.
Participants in this experiment were either part of the testosterone administered or the placebo group. Both groups played, unconscious about their group allocation, the die roll paradigm. Subjects were asked to roll a six sided die and report the result. The self-reported outcome determined the payoff

from the game. The procedure was double blind, so that the experimenter was not able to detect lying. Surprisingly the results show that subjects in the testosterone group were reporting more often honestly, and sacrificed a material gain in favour of truth. The finding is even more counterintuitive as the player did not have a counterpart, except the experimenter, that was affected by the choices. Hence, according to the authors, pro social status seeking can be ruled out as the motivator. The authors suggest concerns for self-image and pride as an explanation for the higher reporting in the testosterone group.

The third study using physiological measures to examine dishonesty behaviour is a fMRI study of **?**. This work is most related to the research program we pursue in the following. The authors investigate the neural correlates of promise making and promise keeping, respectively promise breaking. In a simplified trust game Player 1 in role of the receiver can state their intended behaviour in an antecedent promise stage. When Player 2 decides to trust the first player and sends money, the receiver decides whether to keep or break the promise (decision stage) by sending half of the money back. The entire strategy space is known by all players before the game starts. Hence subjects in the role of the receivers are able to simulate the choice of keeping or breaking the promise and the respective material and emotional costs, before actually making a promise. In the case of promise breaking increased activations in the brain areas for emotional conflict (dorsolateral prefrontal cortex, anterior cingulated cortex, amygdale, ventral striatum) could be observed during the experiment. Furthermore the authors identify two types of subjects, the honest and dishonest promise makers (compare with Hurkens and Kartik (2009)) . Dishonest promise makers show an increased activation in brain regions for conflict and negative emotional processing. The authors argue that the decision to fool is therefore made prior, in the promise stage. The later behaviour in the decision stage can be predicted with the activa-

tion in the prior stage. We argue that this is not a surprising observation, as the subjects can foresee the later decision situation perfectly and play out all possible outcomes before making the promise, financially as well as emotionally. We have chosen an alternative setup in our study that includes uncertainty; subjects cannot foresee the later decision situation. In consequence promise makers commit to an action while not knowing the costs of keeping the promise.

3 Experimental design and procedure

During the period of March 2013 and March 2014 we ran 12 computerized experimental sessions. Participants were students of the Queensland University of Technology and were invited using the database software ORSEE (**?**). The invitation email informed students that the heart rate will be measured during the experiment. An extra compensation of 5 AUD was offered, so that the average earning was about 18 AUD. The experiment lasted on average 45 minutes and 180 students participated in total. The experiment was programmed in the experimental software CORAL (Schaffner, 2013).

When participants arrived, they were seated in front of visually isolated computer terminals. A description about the heart rate monitors was handed out. Students wired themselves with the heart rate monitors. While the experimenter was checking the monitors, the written instructions were handed out. Participants could read the instructions in their own pace and questions were answered privately.

The used trust game is referring to the design of Baumgartner et al. (2009), see Figure **??**. In our modification players do not know the strategy choices they face later in the game. 20 AUDs needed to be distributed in the second stage if the first mover decided to trust the second mover. Two options are available: Option A, the social option, offered always an equal split between the two players, option B, teh self-serving option, offered an unequal split with an unknown realization of 10 + x/ 10 - x with x [-2,10]. The game lasted for 20 rounds and the players were matched each round with a new different partner (stranger matching). Figure 1 displays the course of the game.

In the first four rounds subjects played the game without an antecedent promise stage. These rounds served as practice rounds and to capture heterogeneity in the heart rate. From the fifth round on the second mover could

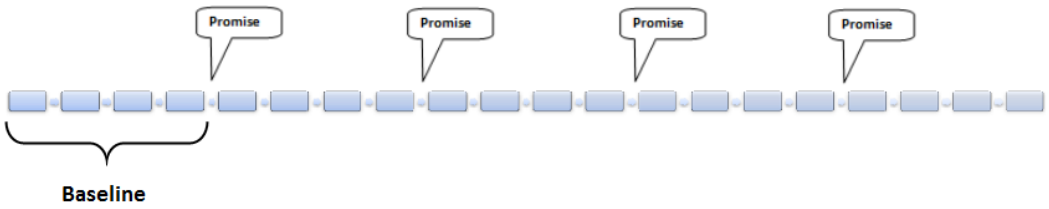

Figure 1: Course of Game

make a promise about his intended behaviour. This promise was valid for the consequent four rounds. Each round the promise was transmitted to the first mover, who then could decide whether to trust the second movers or choose the outside option. If the first mover decided to trust, a nature move determined the realization of the unequal split in this round. Both players learned the realization. Player 2 had to decide which option to choose and respectively whether to keep or break the previously made promise. We asked Player 1 about his or her belief which option Player 2 will choose. Figure 3 describes the decision stages.

Our experiment obtains two treatment variations. In the first group a free form message was offered in which second movers could express their intended future behaviour. In the following we call this treatment free promise. The alternative treatment offered a 7 step scale[1], with which the Player 2 could ddefine the generalizability of the promise[2]. Hence the second treatment incorporated a mutable element in the promise responding to the uncertainty about the outcome choices.We will call this treatment select promise in the following.

In total our data compromises 3600 observations, from 180 participants. Each participant played 20 rounds of the game. 5 sessions were played with the free form promise; 43 subjects in the role of the receiver could make every

[1]I promise to choose always/usually/frequently/sometimes/rarely/ never the equal split. + I do not promise anything.

[2]Instructions and screenshots can be found in the Appendix.

Figure 2: Decision Tree in this study

(3.5,12) –(8.5,12); (5.9,11) node**Player 2**; (3.5,10) –(8.5,10);
(3.5,12)–(3.5,10); (8.5,12) – (8.5,10);
[-¿] (6.0,10) – (6,7.5); (2.5,8.7) nodePROMISE;
(3.5,7.4)–(8.5,7.4); (5.9,6.4) node**Player 1**; (3.5,5.4)–(8.5,5.4);
(3.5,7.4)–(3.5,5.4); (8.5,7.4) – (8.5,5.4) ;
[-¿] (5.5,5.4) – (3,2); (2,4) nodeIN;
[-¿] (6.5,5.4) – (8,2); (8,4) nodeOUT;
(-1,1.9)–(4,1.9); (1.5,0.9) node**Player 2**; (-1,-.1)–(4,-.1);
(-1,1.9)–(-1,-.1); (4,1.9) – (4,-.1);
(6,1.9)–(11,1.9); (8.5,1.1) nodePlayer 1: $ 4; (8.5,0.5) nodePlayer 2: $ 0;
(6,-.1)–(11,-.1);
(6,1.9)–(6,-.1); (11,1.9) –(11,-.1);
[-¿] (0.5,-.1) – (-1.5,-3.4); (-3,-1.5) nodeEQUAL SPLIT;
[-¿] (1.5,-.1) – (3,-3.4); (4.5,-1.5) nodeUNEQUAL SPLIT;
(-0.5,-3.5)–(-5.5,-3.5); (-3,-4.3) nodePlayer 1: $10 ; (-3,-4.9) nodePlayer 2: $10;
(-0.5,-5.5)–(-5.5,-5.5);
(-0.5,-3.5)–(-0.5,-5.5); (-5.5,-3.5) – (-5.5,-5.5);
(1.5,-3.5)–(6.5,-3.5); (4,-4.3) nodePlayer 1: $10 - x; (4,-4.9) nodePlayer 2: $10 +
x; (1.5,-5.5)–(6.5,-5.5);
(1.5,-3.5)–(1.5,-5.5); (6.5,-3.5) – (6.5,-5.5);

four rounds a statement about their intended behaviour for the next following rounds[3] . The select treatment comprises 6 sessions. 47 subjects in the role of the receiver could choose one of seven statements from a given promise scale to describe their intention for the next four rounds. Table 1 provides an overview.

Table 1: Treatment overview

Treatment	Free Promise		Select Promise	
Roles	Player 1	Player 2 *(promisor)*	Player 1	Player 3 *(promisor)*
Observations	43	43	47	47
Total	86		94	

[3]From the 5th round on.

Heart rate measurements

In our study we collected electrocardinal data from the participants to measure the physiological reactions in different decision situations through the experiment. We wired the participants before they started the actual game. HRV measurements have been used in economic experiments before (Falk et al., 2011; **?**; Brandts and Garofalo, 2012; Dulleck et al., 2014), but are traditionally used in medical research and psychological research regarding emotional and mental processes. A detailed discussion of heart rate measurements and methods in economic experiments can be found in (**?**). The measurement of the heart rate variability gives us insights into the Autonomous Nervous System (ANS) which is responsible for the unconscious operations of the human body, which of course includes the heartbeat. The ANS balances activity of its two major subsystems, the sympathetic and parasympathetic systems. The first is dominant in fight and fly responses while the latter is more dominant in rest situations. Both systems have mechanism to regulate the heart rate, but can be distinguished by the frequency of changes introduced by either system. The sympathetic system has a relatively long reaction time of between 2 and 20 seconds (low frequency LF changes), whereas the parasympathetic system regulates the heart beat within much short intervals between 0.5 and 4 seconds (high frequency HF changes). It has been shown that the relative activity of the systems as measured in the ratio of low frequency changes and high frequency changes in the heart rate (LF/HF ratio) is a reliable indicator of mental stress (Appelhans and Luecken, 2006). The LF/HF values, used in the analysis, are normalised by the individual standard deviation.

Hypotheses

Two options are available for Player B in our design, choosing the equal split which leaves both players with $10 or choosing the unequal split, which realization changes from round to round. Each possible split sums up to $20. Choosing the unequal split and earning more than in the equal split means that the other player will consequently receive less than he would have received in the equal split.

According to the assumptions of anonymity and rationality a second mover purely interested in monetary outcomes should always choose the unequal split, except when the upcoming unequal split is 8/12. We included this option to test for lying aversion that is not related to material harm. This option also serves as a control moment for the heart rate measurements. In contrast, considering recent findings of the literature on promises, we expect that offering the possibility to make a promise leads to a lower ratio of unequal split choices and therefore to more cooperation. Various studies have shown that communication is able to promote prosocial behaviour (**?**). Promises in particular have been shown to be perceived as a serious commitment device. People are willing to sacrifice material gain to keep their promise. Receivers anticipate this correctly, and in consequence a higher cooperation level can be reached by a promise. Findings in social psychology suggest that especially voluntarily given and individually formulated statements have a strong commitment effect (Kiesler, 1971; Charness and Dufwenberg, 2010).

Behavioural Hypothesis: *The promise enhances internal constraints. Lower rates of unequal splits are observed when a subject made a promise. The counter player anticipates the commitment effect correctly and trusts more likely when a promise is made. The commitment effect is stronger for own formulated messages than predetermined promises.*

For high unequal splits we expect the ratio of unequal choices to be relatively low. Subjects back off due to the consequences the choice bears for the other player and the emotional costs the choice imposes on themselves.

Physical Hypothesis: *Emotional reactions of the promisor are strongest when the promisor breaks the promise at high unequal splits.*

Regarding the treatment variations we expect that by offering a range of promises (select treatment) the mental manoeuvring takes already place in the promise making stage. The presentation of statements with different degrees of generalizability encourages a foreplay of situations where it would be appropriate to break the promise. In comparison with the free promise treatment we expect less physical reaction in the decision stage. All eventualities have been thought through prior in the promise stage. Also the counter player is better informed what to expect for certain upcoming outcomes and shows less reactions to deviations that can be rationalized.

Physical Hypothesis: *The physiological reactions to similar situations are weaker in the free promise treatment than in the select treatment.*

Lundquist et al. (2009) suggest that the aversion to lie is enhanced with the size of a lie, but also with the strength of a promise. Ergo we expect a low share of unequal splits for subjects that have chosen a strong promise

(always) in the select treatment. And vice versa an increasing ratio of unequal splits for statements that allow for deviation in the intended behaviour.

Physical Hypothesis: *Physical reactions are less strong when a subject breaks a promise and made a vague promise than when the agent made a comprehensive promise.*

Wang et al. (2010) mentioned that strong physical reactions can also be observed when choices are more demanding. The physical reactions can be caused by the temptation of the material gain and the internal reasoning the subjects are going through.

4 Results

In both treatments participants played the first four rounds of the game without making any promise. In 69.17% of the cases the first player chose IN and an interaction between the both players occurred. The second player chose the unequal split in 66.91% of the cases[4][5]. The first rounds served to control for learning effect and to capture physical heterogeneity in the heart rates; this observation wants to be mentioned but is not focus of the examination of this paper.

Free promise

In 49% the subjects chose to use the possibility to send a message and stated the intention to choose the equal split. In 17.80% no message was send to the other person. In the remaining cases subjects used communication, but did not reveal clear intentions for the decision stage. The coding whether a message implied a promise about future behaviour has been done independently by two research assistants and the main author.

The ratio of promise messages was relative stable over the course of the game (5th round: 48.84%, 9th round: 41.86% , 13th round: 51.17%, 17th round: 53.49%). A relatively low percentage (10.63%) of the promises included an element that conditioned the statement in its comprehensiveness, e.g.: *"I probably split equally"* or *"Willing to grant us both $10 or at least you more*

[4]Through the entire game 15 subjects have always chosen the unequal split and 9 participants decided always for the equal split. One could argue that those participants have to be excluded from the analysis. We run the analysis without them and could not find substantial changes in the results. We therefore prefer to keep those subjects in the data set.

[5]Please note that rounds with an unequal split implying a potential negative surplus for the receiver (8/12) have been excluded from the analysis on the split choice. In these rounds the promisor loses money from promise breaking/deviating. We included this possibility in the design as a control moment for the heart rate variability.

than $7". Following our categorization we can analyse which message type generated the most trust, measured in percentage of first movers that chose IN after they read the text. Facing a promise for an equal split motivated 74.70% of the first movers to interact with the promisor. This means that for player pairs in which the receiver made a promise an interaction happened 17% more often compared with player pairs in which no promise was made. The promise facilitated significantly trusting behaviour. (MWT[6]: p¡0.001, z= - 4.78). If the promise included an element that narrowed the statement the trusting ratio lower, about 63.24%. Communication that was just based on friendly messages reached an IN move rate of 56.87% and empty messages a ratio of 67.46%.

Result 1: *Messages that state clear intentions about future cooperative behaviour promote cooperation.*
This result is line with past findings in the field. Promises seem to change the infrequence of the exchange partner in a way, that an interaction can be relied upon.

But is this trust reciprocated? Do promise-makers keep their promise and choose always the equal split? Or at which realization of the unequal split are promisors willing to break the promise for the sake of a material gain? Players, who made a promise choose the unequal split in 48% of the cases, this is 20% less than the frequency in the receiver group who decided not to give a statement about their intended behaviour (68.15%). The difference is statistically significant on a 1% level (MWT: p=0.001, z=3.81) and in line with the literature. These results are based on the average of all upcoming

[6]The Mann-Whitney U is a non-parametric test that is a widely used significance test in Experimental Economics. It allows a comparison between two independent groups or conditions without that the assumption of a normal distribution is needed.

splits. But for our initial question regarding the moral manoeuvring it is more relevant how a promisor behaves throughout the different realizations of the unequal splits. Figure 3shows the frequency of unequal splits, over the potential surplus by choosing the unequal split.

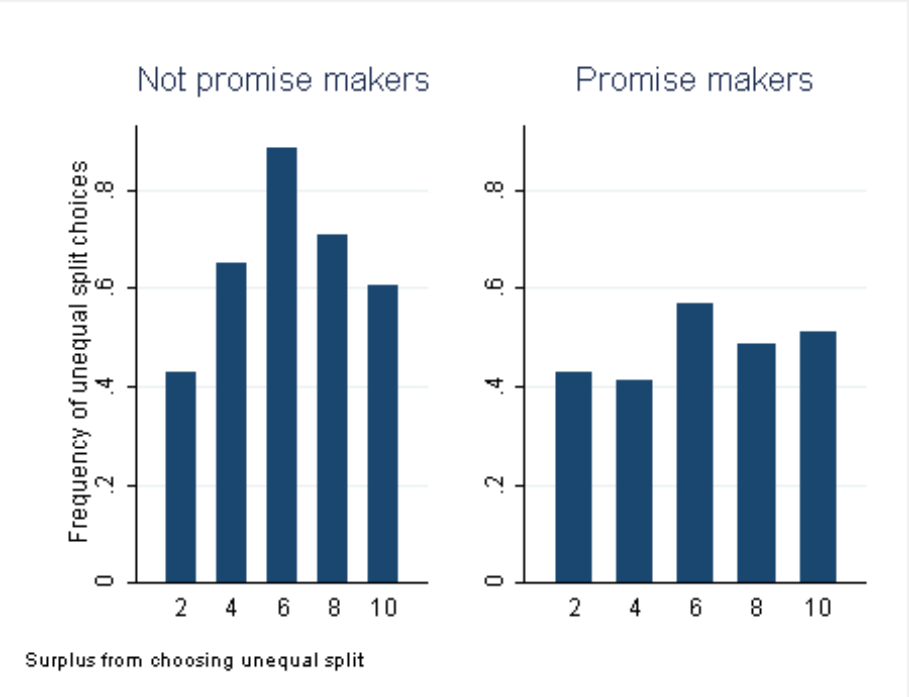

Figure 3: Free Promise

4 shows fitted values and is based on the regression below. For the promise makers (red) the x axis captures the varying potential gains from breaking the promise and choosing the unequal split . While for the non -promise makers (blue) the x axis represents only a trade-off between own gain and cost to the other player.

For the subjects who decided not to make a promise we see an inverse u shape in the frequency of unequal spilt choices over the course of upcoming unequal splits. If the surplus from choosing the unequal split is low, more than half of the subjects are splitting equally. But when the possible gain is increasing subjects, who did not make a promise, choose to take the unequal

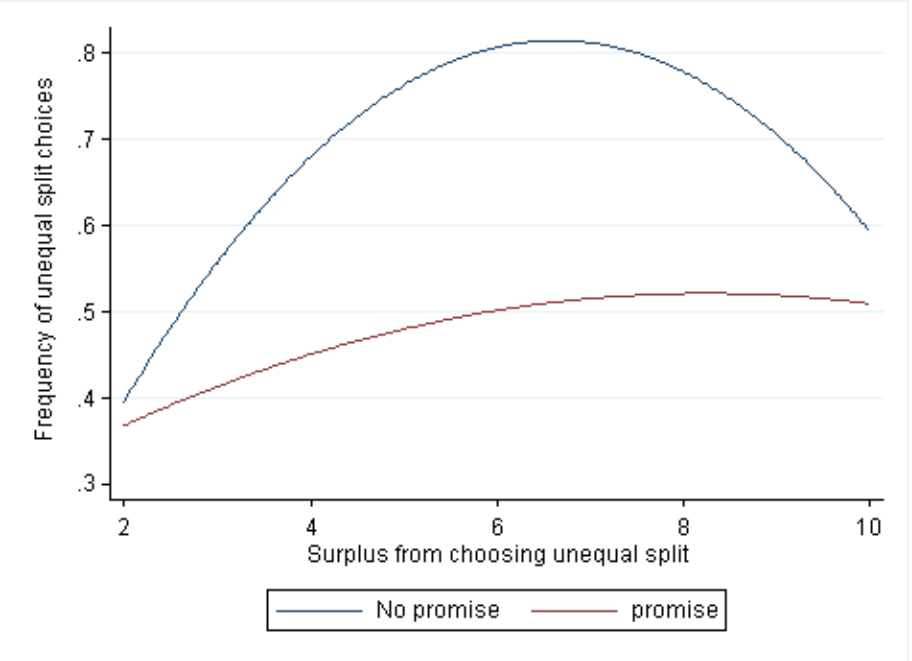

Figure 4: Free Promise

split more frequently. The highest ratio is reached when the interaction partner receives the same value than he or she would have had with the outside option (16/4). 88.24 % of the non-promise makers choose here to take the unequal split. For unequal splits with stronger differences between the resulting payoffs the ratio of unequal splits decreases. At the end of the span only 60.14% of the receivers choose the unequal split which leaves the interaction partner with \$0. For the participants who decided not to make a promise anything a balancing process between own benefits and harm imposed on others can be observed. However our main interest is how the promise makers behave over the course of unequal splits. The estimation below considers the effect of the promise as well as the magnitude of the splits on the behaviour. The second estimation takes the interaction terms between promise and the two later variables into account and served for the graph in Figure 4. The OLS regressions estimate the probability that the

unequal split is chosen depending on the material gain and whether a promise has been made previously. Choosing an unequal split for a promise maker means to break the previously made promise to choose an equal split. All estimations are clustered on an individual level.

Table 2: Regression Free: Promise-makers

	(1)	(2)
	Unequal split choice	Unequal split choice
surplus	0.153*	0.259***
	(0.0592)	(0.0697)
surplus × surplus	-0.0110*	-0.0195***
	(0.00431)	(0.00489)
promise to split equally	-0.207*	0.298
	(0.0855)	(0.352)
promise to split equally × surplus		-0.194+
		(0.109)
promise to split equally × surplus × surplus		0.0156+
		(0.00807)
Constant	0.228	-0.0429
	(0.203)	(0.253)
Observations	357	357
Adjusted R^2	0.053	0.056

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

It can be seen that the frequency of unequal split choices is significantly lower for the promise makers at all upcoming splits (p=0.0072, H0: surplus = surplus*promise). But also the course over the different upcoming unequal split is statistically different between promise makers and non-promise makers (p=0.004, H0: c.surplus*c.surplus = equalpromise*c.surplus*c.surplus).

The promise reduces the variation in the behaviour over the different upcoming splits.

Result 2: *For subjects who made a promise the behavioural manoeuvring in response to changes in the costs of promise keeping is significantly lower.*

In the next step we link the choice behaviour in the free treatment with the individual's heart rate variability.
The analysis of the heart rate variability in the promise stage reveals that also we can predict the later behaviour or promise makers in the decision situation. (Baumgartner et al., 2009) could see a higher activity in brain regions that are associated with emotional conflict for later promise breakers. We measure a higher LF/HF ratio in the promise stage for subjects who keep their promise in the later decision stage and split equally. (MWT: p=0.001, z= 3. 27) and not existent for subjects who decided not to make a statement about their intended behaviour. We therefore conclude that the higher heart rate variability is an effect of making a promise and a sign that the subjects take the commitment serious.
Result 3: *Promise makers who demonstrate a high stress level (LFHF) when making the self-formulated promise are more likely to keep the promise.*

In the decision situation we distinguish between promisors who keep or break their promise by choosing the respective split. Figure 6 shows that for subjects who keep their promise no obvious difference is visible in the heart rate variability over the various upcoming unequal splits. Whereas subjects, who break the promise, have stronger physical reactions with an increasing magnitude of unequal splits. The more the realized payoff deviates from the stated intention, the stronger are the emotional costs a promise breaker experiences with his or her choice. One could argue that the higher LF/HF

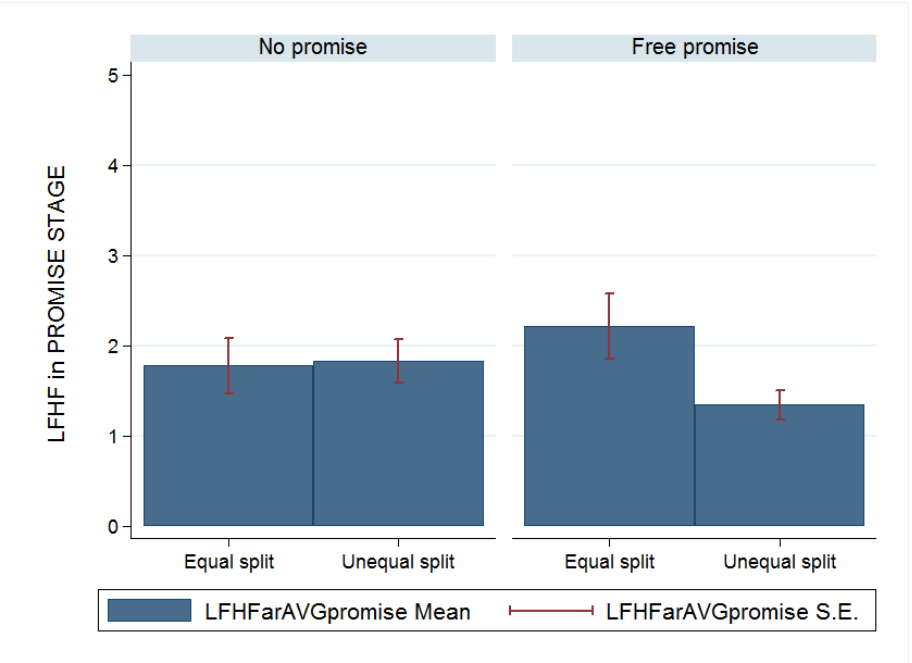

Figure 5: HRV in Promise stage

ratio is due to the harm the choice imposes on the interaction partner and the physiological reaction can be explained. If this would be the case we should see the same trend in the data for the participants who decided not to make a promise. This is clearly not the case. In the group of not promise makers a slightly higher reaction can be measured for subjects who choose the unequal split, but no variation over the course of upcoming unequal splits is apparent. In the choice data we observed for subjects who did not make a promise strong variation in form of inverse u-shape over the unequal split choices. We were arguing that a balancing of material gain and emotional costs is taking place. The consistency of the physiological reactions shows that the participants have balanced the opposing forces sucessfully before making the decisions.

Result 4: *The emotional costs of breaking a promise and the corresponding physiological reaction increase with the more the choice deviates from the*

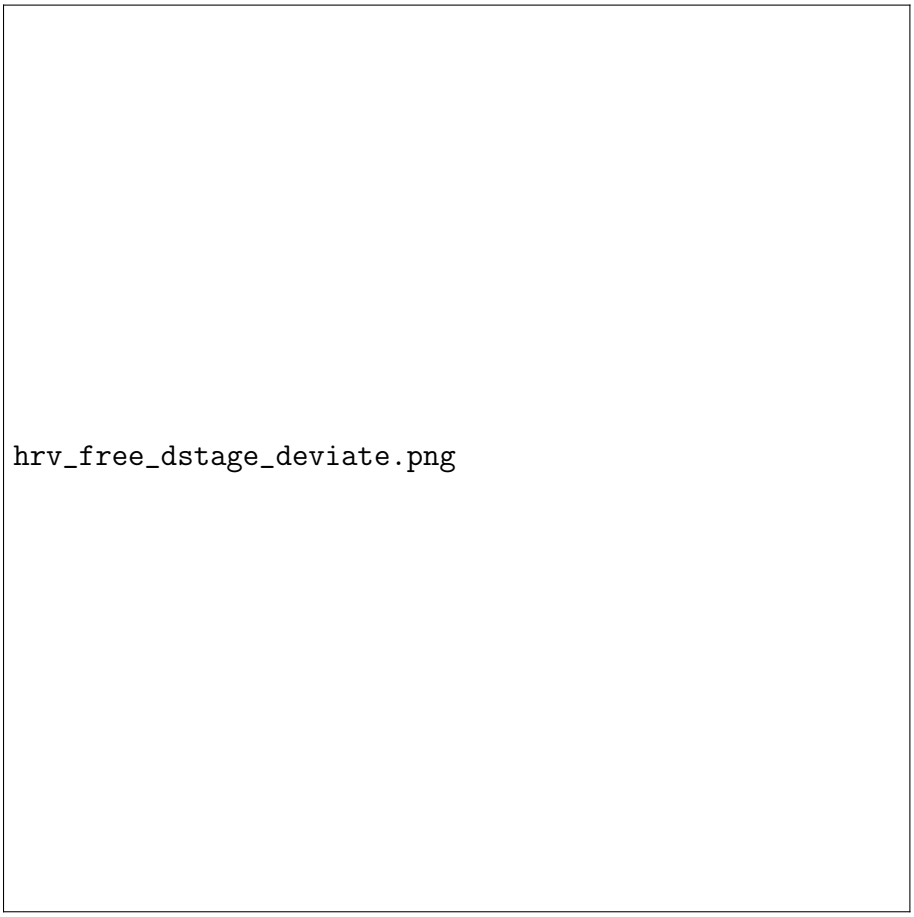

Figure 6: HRV Decision stage

promised intention.

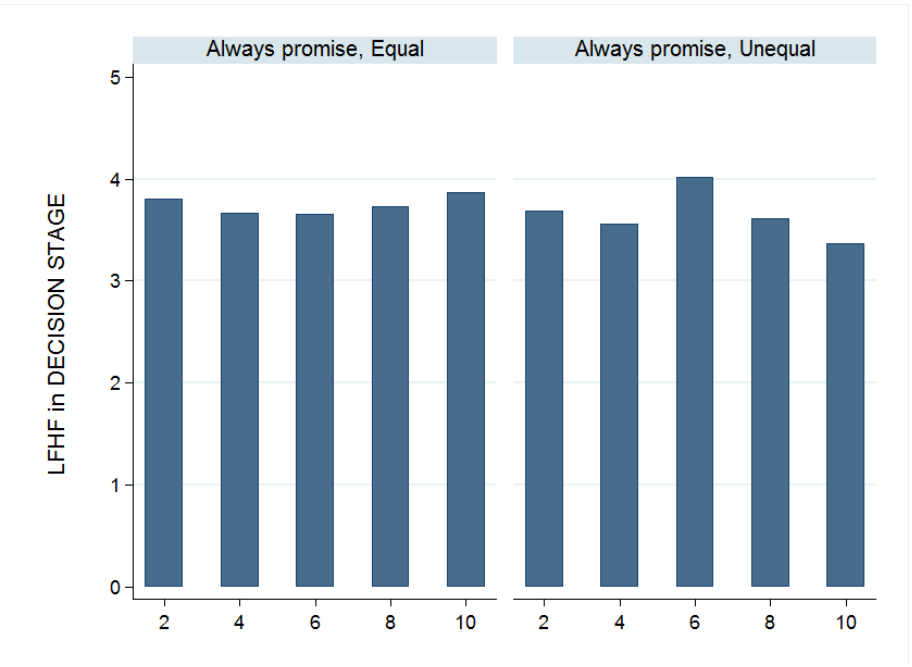

Figure 7: HRV Decision stage for each split

4.1 Results Select Promise

In a second treatment subjects were provided with a set of messages they could choose from to describe their future intended behaviour. The message varied in the strength of the commitment and corresponded with the uncertainty about the upcoming alternatives. In about half of the cases, 48.94%, the receivers chose to send the strongest promise to the interaction partners, *I promise to always choose the equal split.*. Second popular was the promise that indicated that Player 2 will usually choose the equal split, this statement was send in 28.72% of the cases. Frequently and sometimes were chosen less often, in 11.70% and 5.32% of the cases. The promise to never split equally and the option not to promise anything were chosen rarely, 1.06% and 4.26%. In the further analysis we will concentrate on the always promise and promises which narrow the intention to split equally.
In 81.25% when an always promises was send, the counter player decided to trust and chose to interact with the promisor. For promises that narrowed the commitment the interaction rate was substantially lower and decreased with the vagueness of the message: usually 68.52% (sd: 46.96), frequently 53.41% (sd: 50.24), sometimes -42.5% (sd: 50.40).

Result 5: *The more comprehensive a promise the more does an interaction partner trust.*

Did the first movers anticipate correctly the trustworthiness of the promisor, depending on the chosen statement? Promisors who indicated that they will *always* split equally deviated in 42.08% of the cases from this promise and chose the unequal split. Promisors who chose a more vague promise (usually, frequently and sometimes) chose the unequal split in 56.67% of the cases. Interactions with vague promises are pooled, as no difference is observable in the corresponding behaviour of the promisors. In the further analysis we

compare how the choice of a comprehensive promise affects the commitment in comparison with statements that indicate exceptions in the promised behaviour. First, the always promise generated significantly more trust (MWT: p¡0.001, z = -5.05) and the promisors lived more often up to this expectations (MWT: p=0.005, z = 2.81) compared with the vague promises.

Result 6: *A comprehensive promise leads to higher trustworthiness.*

Table 3: Regression Select: Promise- makers

	(1)	(2)
	Unequal split choice	Unequal split choice
surplus	0.109*	0.0995
	(0.0478)	(0.0827)
surplus × surplus	-0.00447	-0.00369
	(0.00380)	(0.00669)
promise to split always equally	-0.170*	-0.208
	(0.0799)	(0.268)
promise to split always equally × surplus		0.0157
		(0.102)
promise to split always equally × surplus × surplus		-0.00132
		(0.00827)
Constant	0.134	0.156
	(0.132)	(0.216)
Observations	411	411
Adjusted R^2	0.083	0.078

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

But how much does the responding behaviour on varying payoffs differ between subjects that made comprehensive or restricted promises? The estimations **??** show that the only difference between always and vague promises is the level of promise keeping. Always promisors keep their promise more frequently. But there is no significant difference in the reaction to the magnitude of the unequal splits. Both functions are linear and increasing in the material benefit of the promisor.
This is an interesting contrast to the behaviour in the free promise treatment. When subjects freely choose the variation in response to a changing environment could be lowered. Participants who did not make a promise about their intention are balancing the material benefit and harm on others in their choice. Whereas in the select promise treatment participants given a predetermined set of promises. The set contained promises that indicated that there may be exceptions to the promised behaviour. Deviations from the initial intention can then be justified in extreme case, like at the upper end of the unequal outcome distribution.

Result 7: *A self- phrased promise is more effective in supporting consistent behaviour in an unforeseeable, changing environment compared to a preformulated promise.*

We broaden herewith the finding of Charness and Dufwenberg (2010), which showed that bare promises have only a little effect. Free form messages were significantly more efficient in promoting cooperation.
The analysis of the physical data in the select treatment reveals that if the LF/HF ratio is higher in the promise stage the *always*-promisors is more likely to break the promise in the decision situation. However the effect is only weakly significant (MWT: $p = 0.09$, $z = -1.69$). Interestingly this effect goes in the opposite direction of what we have observed in the free promise

treatment, where subject who kept the promise showed a higher LF/HF ratio in the promise stage. We argued that freely chosen commitment create a stronger obligation for the promisor. In the select treatment on the other hand we suspect activation due to strategic considerations.

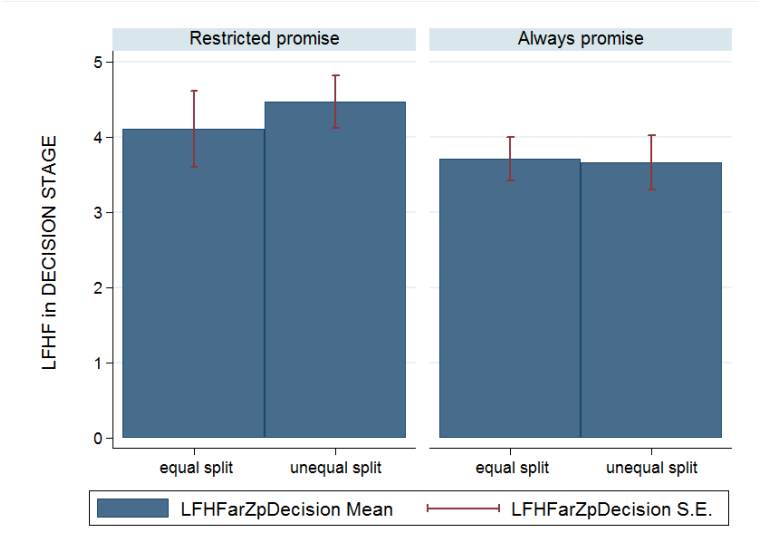

Figure 8: HRV Select Decision stage

Strategic players, who used the promise as a signal to raise the expectations of first mover, but do not intend to follow it, will use the always promise as it is the strongest in this setting. Five players in the role of the receiver took always the unequal split, three of them choose to send the always promise in all rounds.

Figure 8 shows the behaviour in the decision stage. We see that subjects who break the always promise have a lower LF/HF ratio over all possible surplus magnitudes in comparison to the restricted promise makers. One may argue that the decision whether to keep of break a promise is more cognitive demanding for promises limiting the generalizability of the commitment. For each upcoming split the promisor needs to evaluate the surplus by promise breaking and decide whether it is worthwhile to break the promise.

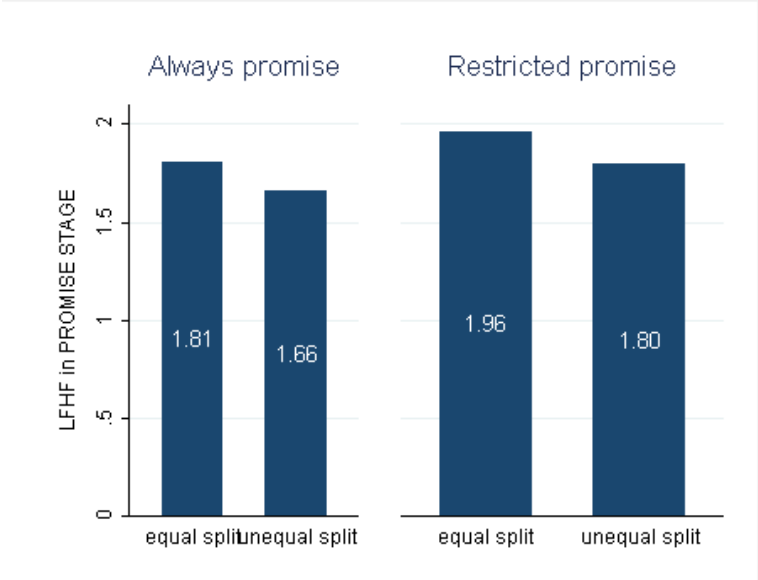

Figure 9: HRV Select Promise stage without

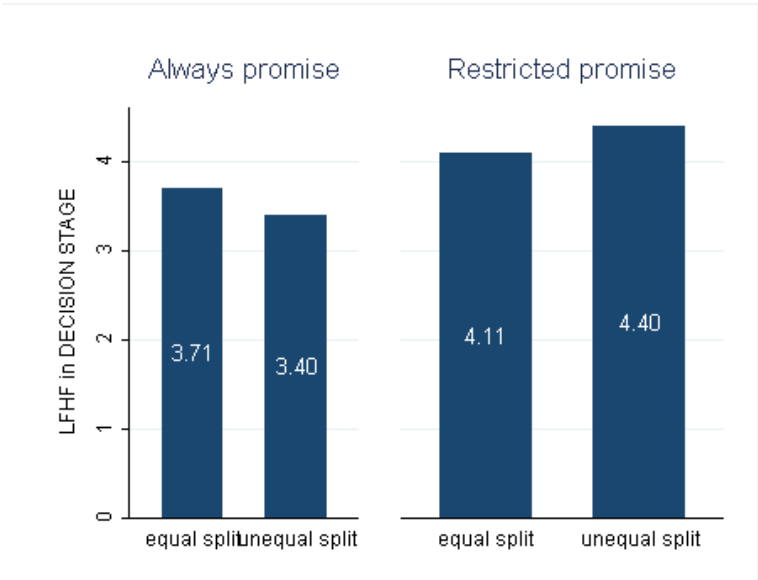

Figure 10: HRV Select Decision stage without

In the always category no difference in the LFHF ratio is apparent for subjects who keep and break the promise (MWT: p =0.55, z=0.60). For the restricted promise, however the difference in the heart rate variability is statistically sig-

nificant between promise breaking and promise keeping subjects (MWT: p= 0.03, z= -2.18). Whereas the significant difference vanishes (MWT: p=0.10, z= -1.63) when the five subjects, who always deviated, are excluded from the data set (compare Figure 10). For the always promise makers the difference between promise keeping and promise breaking subjects is further on non significant (MWT: p=0.20, z=1.29).

Foreseeability

In general it should be mentioned that the heart rate variability in the decision stage is substantially higher than in the promise stage, for all decision makers and regardless for the treatment group. This suggests that subjects are not able to perfectly simulate the decision prior in the promise stage. And the uncertainty indeed shifts the decision process to the later stage.

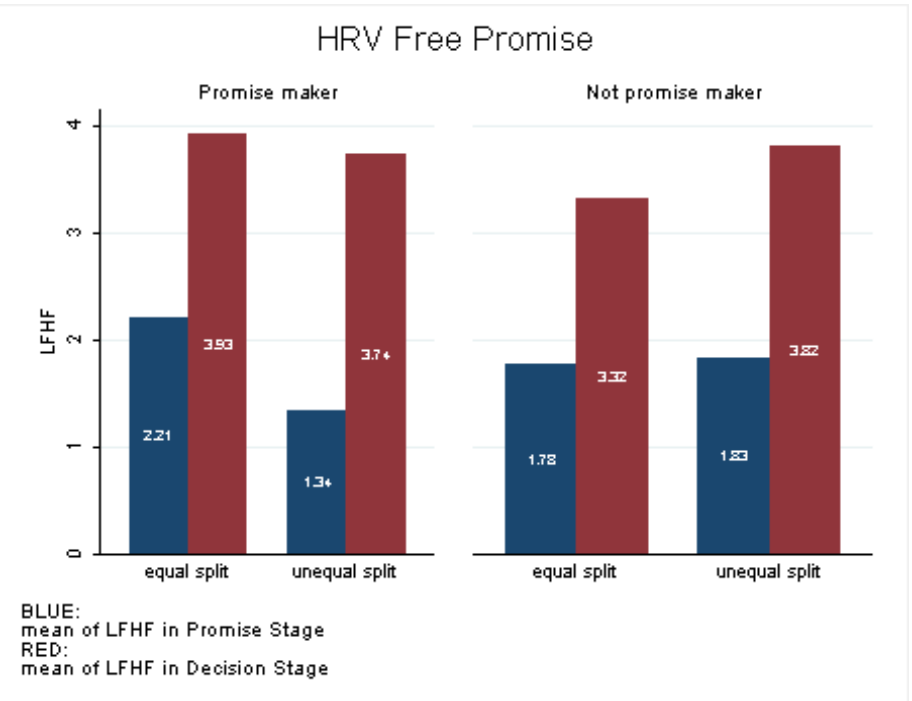

Figure 11: HRV Free: Promise vs. Decision stage

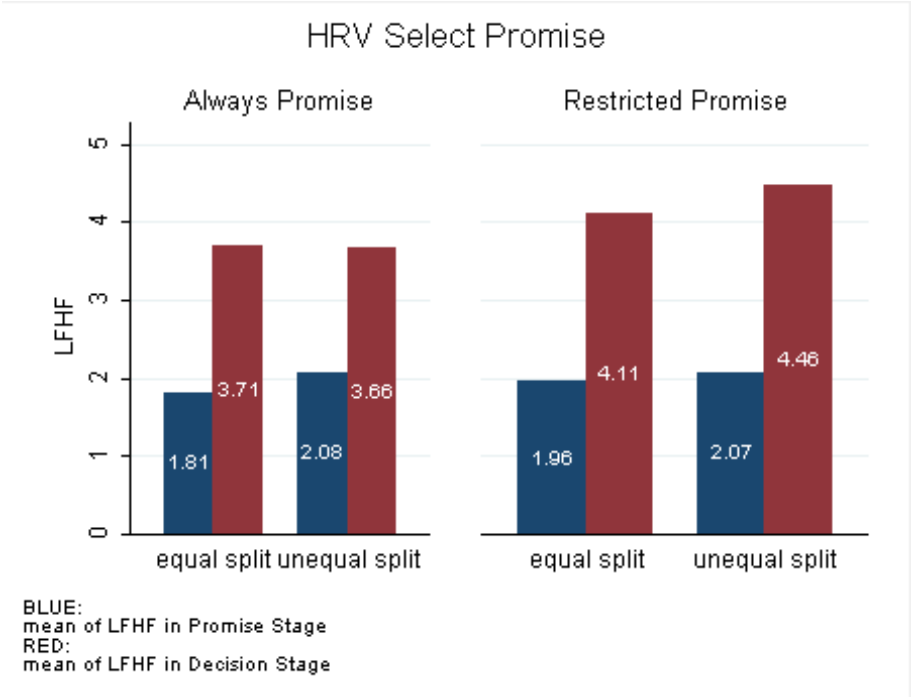

Figure 12: HRV Select : Promise vs. Decision stage

5 Conclusion

We designed a trust experiment with a pre-play communication stage, where messages could be sent about intended behaviour. The game structure contained uncertainty about the choice alternatives the promisor faced in the decision situation. The choice alternatives, respective the costs of promise keeping were unknown at the promise making stage. Emotional and material consequences related with the choice options could not be foreseen. With a random draw over the possible choice options we disentangled the decision to make a promise from the decision to keep a promise. This is, to our knowledge, a novel element not present in any existing promise studies. Choice options and consequences are usually deterministically known before a promise is made. However, in real life when promises are made, it is unusual that the person knows the circumstances they face later on. Thus they cannot perfectly foresee how costly it will be to keep the promise. We argued that moral manoeuvring in such situations takes places. Agents balance the material gains from promise breaking with moral costs related to breaking the promise. In our design the promise was about an equal split resulting from trusting behaviour of the first mover in a trust game. The promisor faced each round an altering unequal split, with different associated surpluses. During the experiment we measured the heart rate variability of the participants. We can obtain a correlate of mental stress, the LF/HF ratio. The physiological data allowed us to control for emotional involvement while subjects were making the promise and during the decision situation whether to keep or break the promise. The LF/HF was generally higher at the decision stage than the promise stage, indicating that participants were indeed not able to foreplay the decision. We saw a balancing of own benefits and harm the choice imposed on other players. The frequency of unequal split choices took the form of an inverse u shape over the magnitude

of surpluses that could be gained by choosing the unequal split. When a self-formulated promise came into play the variation in the behaviour was significantly reduced. The unequal split ratio was still about 40% and the promise is respectively broken, but the ratio stayed practically constant over the possible surpluses. Whereas when the promise was predetermined, the unequal choice ratio was steadily increasing with the possible surplus. The physical data supported the assumption that for the promisors in the free treatment ascribed importance to the promise. So could we see higher amplitudes in the promise stage for participants who decided to make a promise and kept it later on? For promise breaking participants stronger physical reactions could be observed the more the payoffs of the realized unequal split deviated from the promise equal split.

The limitation of this study lies in the payoff dimension of the unequal splits. Material gain means in our design automatically harm for the other player. We therefore cannot identify internal motives for promise keeping. The promisors may want to avoid imposing harm on the other person or act of promise breaking itself implies internal costs for him. Vanberg (2008) tried to answer this question and the results of his study support the later hypothesis. But we leave it to future research to discuss this issue in more detail. Support for our design we find in many real life situations, where a promise is made to overcome coordination problems (Charness and Dufwenberg, 2010). The promise brings the trustor in a vulnerable position by believing the promise from the interaction partner. Promise breaking means consequently exploiting the trustee. Our results provide evidence that participants, who choose to make a promise in an uncertain environment, reduce the variation in behaviour and their choice is less dependent on environmental changes. With the physiological measurements we could open the black box a bit and see that prior made commitments affect the decision maker also physically/emotionally. We found support that emotional costs increase

with the deviation from the promised action. But our study shows also that the choices and the physiological/emotional reactions following a promise depend strongly on the form of the message given. Our results confirm that free form communication is most effective as a coordination device; even in decision environments with uncertainty.

6 Appendix

6.1 Instructions (FREE)

Thank you for participating in this experiment. The purpose of this experiment is to study how people make decisions. In case you should have questions at any time, please raise your hand. Please do not speak to other participants during the experiment and please turn off your mobile phone. We also ask you not to reveal any details about the experiment after you have participated. You will receive $5 for arriving on time and $5 for wearing the heart rate monitor. Depending on your decisions and the decisions of other participants, you will receive an additional amount (see below). At the end of the experiment, you can collect the final amount in cash privately in a sealed envelope from the School of Economics and Finance front office.
This session consists of 20 rounds. In each round you will interact with another randomly chosen participant. No participant will ever know the identity of the person with whom he or she is paired. At the end of the session one round is randomly chosen as the round that will be paid.

TASK DESCRIPTION

At the start of the experiment each participant will be randomly assigned to role Player A or Player B. You will stay in the same role for the duration of the experiment. In each round one participant with the role Player A is paired with another participant with the role Player B. The amount of money you earn depends on the decision made within your pair. Player A is the first mover. They have to decide whether they want to interact with Player B or not. The consequences of these decisions are as follows:

- If **Player A decides not to interact** (move OUT) the game is over. Player A then receives $7 from this round, while Player B receives nothing.

- If **Player A decides to interact** (move IN) then Player B is asked to decide between two options, EQUAL SPLIT or UNEQUAL SPLIT.
  The consequences of the two are then as follows:
  - EQUAL SPLIT: each player receives $10
  - UNEQUAL SPLIT: $20 will be unequally allocated between the two players. The unequal split valid for the respective round is randomly determined by the computer. Each of the following allocations comes up with the same probability:
    (Player A / Player B):
    ($0 / $20), ($2 / $18), ($4 / $16) ,($6 / $14), ($8 / $12), ($12 / $8)

When Player A decides between IN and OUT they do not know which unequal payoff is valid for the respective round. The computer determines the unequal split valid for the respective round after Player A has made their choice.
Player B sees first Player A's choice and then learns about the unequal split effective for the respective round. If Player A has chosen OUT Player B has no decision to make and both players learn their payoffs (Player A $7 and Player B $0) for this round. If Player A has chosen IN Player B now decides between LEFT and RIGHT knowing the consequences of each of these options for both players. After Player B has made his decision, the interaction in this round is over and both players learn their payoffs from this round.
A **new pair of players** will then be matched together for the next round.

MESSAGE

Before the 5th round starts Player B can make a promise.

This message is valid for 4 rounds and will be displayed to each Player A before she makes the decision whether to interact or not. Before the 9th, the 13th and the 17th round starts Player B can make a new statement which is then again valid for the next four rounds and will again be displayed to the Player A of the respective round. Please recall that a new pair of players is matched together for each round.

The promise option will be given from the 5th round on, which means that the first 4 rounds will be played without a promise.

Please note that you have for each of your choices 30 sec. time to make a decision (except for the message stage). If you do not make a decision in this time, the computer automatically distributes a zero payoff for you and a compensation payoff for the other player.

6.2 Instructions (SELECT)

Thank you for participating in this experiment. The purpose of this experiment is to study how people make decisions. In case you should have questions at any time, please raise your hand. Please do not speak to other participants during the experiment and please turn off your mobile phone. We also ask you not to reveal any details about the experiment after you have participated. You will receive $5 for arriving on time and $5 for wearing the heart rate monitor. Depending on your decisions and the decisions of other participants, you will receive an additional amount (see below). At the end of the experiment, you can collect the final amount in cash privately in a sealed envelope from the School of Economics and Finance front office.
This session consists of 20 rounds. In each round you will interact with another randomly chosen participant. No participant will ever know the identity of the person with whom he or she is paired. At the end of the session one round is randomly chosen as the round that will be paid.

TASK DESCRIPTION

At the start of the experiment each participant will be randomly assigned to role *Player A* or *Player B*. You will stay in the same role for the duration of the experiment. In each round one participant with the role Player A is paired with another participant with the role Player B. The amount of money you earn depends on the decision made within your pair. Player A is the first mover. They have to decide whether they want to interact with Player B or not. The consequences of these decisions are as follows:

- If **Player A decides not to interact** (move OUT) the game is over. Player A then receives $7 from this round, while Player B receives nothing.
- If **Player A decides to interact** (move IN) then Player B is asked to decide between two options, EQUAL SPLIT or UNEQUAL SPLIT.

The consequences of the two are then as follows:

- EQUAL SPLIT: each player receives $10
- UNEQUAL SPLIT: $20 will be unequally allocated between the two players. The unequal split valid for the respective round is randomly determined by the computer.

Each of the following allocations comes up with the same probability:
(Player A / Player B):
($0 / $20), ($2 / $18), ($4 / $16) ,($6 / $14), ($8 / $12), ($12 / $8)

When Player A decides between IN and OUT they do not know which unequal payoff is valid for the respective round. The computer determines the unequal split valid for the respective round after Player A has made their choice.
Player B sees first Player A's choice and then learns about the unequal split effective for the respective round. If Player A has chosen OUT Player B has no decision to make and both players learn their payoffs (Player A $7 and Player B $0) for this round. If Player A has chosen IN Player B now decides between LEFT and RIGHT knowing the consequences of each of these options for both players. After Player B has made his decision, the interaction in this round is over and both players learn their payoffs from this round.
A new pair of players will then be matched together for the next round.

MESSAGE

Before the 5th round starts Player B can make a promise.
The following 7 messages are available:

- I promise to choose always the equal split.
- I promise to choose usually the equal split.
- I promise to choose frequently the equal split.
- I promise to choose sometimes the equal split.
- I promise to choose rarely the equal split.
- I promise to choose never the equal split.
- I do not promise anything.

This message is valid for 4 rounds and will be displayed to each Player A before she makes the decision whether to interact or not. Before the 9th, the 13th and the 17th round starts Player B can make a new statement which is then again valid for the next four rounds and will again be displayed to the Player A of the respective round. Please recall that a new pair of players is matched together for each round.
The promise option will be given from the 5th round on, which means that the first 4 rounds will be played without a promise.
Please note that you have for each of your choices 30 sec. time to make a decision (except for the message stage). If you do not make a decision in this time, the computer automatically distributes a zero payoff for you and a compensation payoff for the other player.

References

Appelhans, B. M. and Luecken, L. J. (2006). Heart rate variability as an index of regulated emotional responding. *Review of General Psychology*, 10(3):229.

Balliet, D. (2009). Communication and cooperation in social dilemmas: A meta-analytic review. *Journal of Conflict Resolution*, 54(1):39–57.

Baumgartner, T., Fischbacher, U., Feierabend, A., Lutz, K., and Fehr, E. (2009). The neural circuitry of a broken promise. *Neuron*, 64(5):756–770.

Binmore, K. (2006). Why do people cooperate? *Politics, Philosophy & Economics*, 5(1):81–96.

Bowles, S. (2008). Policies designed for self-interested citizens may undermine" the moral sentiments": Evidence from economic experiments. *science*, 320(5883):1605–1609.

Brandts, J. and Garofalo, O. (2012). Gender pairings and accountability effects. *Journal of Economic Behavior & Organization*, 83(1):31–41.

Charness, G. and Dufwenberg, M. (2006). Promises and partnership. *Econometrica*, 74(6):1579–1601.

Charness, G. and Dufwenberg, M. (2010). Bare promises: An experiment. *Economics Letters*, 107(2):281–283.

Dulleck, U., Schaffner, M., and Torgler, B. (2014). Heartbeat and economic decisions: Observing mental stress among proposers and responders in the ultimatum bargaining game. *PLoS ONE*, 9(9).

Falk, A., Menrath, I., Verde, P., and Siegrist, J. (2011). Cardiovascular consequences of unfair pay. Technical report.

Gneezy, U. (2005). Deception: The role of consequences. *The American Economic Review*, 95(1):384–394.

Hurkens, S. and Kartik, N. (2009). Would i lie to you? on social preferences and lying aversion. *Experimental Economics*, 12(2):180–192.

Kartik, N. (2009). Strategic communication with lying costs. *The Review of Economic Studies*, 76(4):1359–1395.

Kiesler, C. A. (1971). *The psychology of commitment: Experiments linking behavior to belief*. Academic Press New York.

Lundquist, T., Ellingsen, T., Gribbe, E., and Johannesson, M. (2009). The aversion to lying. *Journal of Economic Behavior & Organization*, 70(1):81–92.

Mazar, N., Amir, O., and Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, 45(6):633–644.

Schaffner, M. (2013). Programming for experimental economics: Introducing coral - a lightweight framework for experimental economic experiments. QuBE Working Papers 016, QUT Business School.

Schweitzer, M. E. and Hsee, C. K. (2002). Stretching the truth: Elastic justification and motivated communication of uncertain information. *Journal of Risk and Uncertainty*, 25(2):185–201.

Serra-Garcia, M., van Damme, E., and Potters, J. (2011). Hiding an inconvenient truth: Lies and vagueness. *Games and Economic Behavior*, 73(1):244–261.

Shalvi, S., Dana, J., Handgraaf, M. J., and De Dreu, C. K. (2011). Justified ethicality: Observing desired counterfactuals modifies ethical perceptions

and behavior. *Organizational Behavior and Human Decision Processes*, 115(2):181–190.

Vanberg, C. (2008). Why do people keep their promises? an experimental test of two explanations1. *Econometrica*, 76(6):1467–1480.

Wang, J. T.-y., Spezio, M., and Camerer, C. F. (2010). Pinocchio's pupil: using eyetracking and pupil dilation to understand truth telling and deception in sender-receiver games. *The American Economic Review*, 100(3):984–1007.

Wibral, M., Dohmen, T., Klingmüller, D., Weber, B., and Falk, A. (2012). Testosterone administration reduces lying in men. *PloS one*, 7(10):e46774.