Promise Competition

Andreas Born[*]

September 19, 2018

Click here for the latest version with or without tables in the text.

Abstract

This paper studies competition when sellers cannot perfectly commit to their offers. Two sellers compete by promising service-quality to a one-time customer. Both have private information about the own cost of supplying quality and of breaking promises. In any refined equilibrium, sellers pool at a single promise. Pooling prevents positive selection of better sellers yet competition induces sellers to promise higher quality than they would otherwise provide. Therefore, competition raises average service-quality despite non-binding contracts and private information. A laboratory experiment confirms these predictions. While participants distinguish themselves by their promises initially, they learn to pool their promises such that it becomes impossible to select better seller-types. I find evidence that promise competition increases buyer welfare in the short run, moreover an increase in promises is correlated with an increase of quality-provision by sellers over repetitions of the experiment.

Keywords: promising, competition, signaling, selection.

1 Introduction

Standard models of imperfect competition presume legally binding promises. For example, sellers perfectly commit to their offered prices or qualities. Real markets

[*]Stockholm School of Economics, andreas.born@phdstudent.hhs.se.

often fall short of this ideal. Quality may be difficult to verify, or the legal authorities may be unwilling or incapable of enforcing the promises. In absence of repeated purchases or other informal enforcement mechanisms, the buyer must ultimately rely on the seller's goodwill and honesty. For example, the markets for car mechanics, complex procurement contracts, and single-term politicians all share this characteristic.

Here, I propose a simple model of non-binding promise competition and test the predictions in a laboratory experiment. Two sellers compete for the custom of a single buyer. The sellers have private information about their good will and honesty. A seller may be either selfish and dishonest, selfish and honest, or (intrinsically) motivated and honest.[1] The buyer has to choose one of the two sellers on the basis of non-binding promises only. Promises may serve as signals about the quality of the service delivered by the sellers. Two central features differ from conventional signaling models. Firstly, seller-heterogeneity spans two dimensions - intrinsic motivation and cost of promise breaking. Secondly, the model involves two simultaneous signals, one promise from each competing seller. To analyze the model, I use the refinement Criterion D1 of Cho and Kreps (1987). This study is the first to investigate seller competition - competition when promises are not perfectly binding. Beyond that it contributes to literature about non-binding electoral promises and generally the behavioral literature about promise keeping.

The model predicts that sellers pool their promises which makes it impossible to infer the type of a seller from the promise.[2] The pooling occurs on a level of quality that is equal to or larger than the quality a motivated seller provides in absence of promises and competition. This induces honest sellers to increase the quality they provide to the buyer. Hence, on average promise competition increases the welfare of buyers compared to a case without communication.

I test these predictions in a laboratory experiment. Based on non-binding promises about their intentions, participants choose a sender to play dictator game with.[3]

[1] For example Bénabou and Tirole (2006) discuss a similar type of motivation.

[2] It is remarkable that Criterion D1 selects pooling equilibria in this model, as it usually selects separating equilibria - the Riley outcome in particular.

[3] The dictator game includes a multiplicator of 2 applied on all points send to the receiver to avoid a single focal point of promising to send 50% of the endowment to the receiver.

To allow equilibrium play to emerge, participants repeat the game ten times with stranger matching. They receive feedback about the actions of their group, specifically about both promises, which promise got selected, and the actual decision of the selected promisor. Finally, participants play a regular one-shot dictator game, randomly either in the beginning or end of the experiment.

In the experiment participants are able to select better senders based on their promises when senders are inexperienced. After gathering experience over a few repetitions however, participants pool their promises around a level of 50 points after four repetitions. Then the distribution of promises is single peaked and around 50% of participants promise five or less points away from the modal promise of their session. Moreover, selected and not-selected promisors do not differ significantly in the mean or distribution of the amount they send. In line with the second prediction of the model, participants give more to the receiver in the first round of the promise game than the dictator game. Later rounds of the promise game are hard to compare to a one-shot game as I find a significant drop in the amount given over the repetitions of the promise game.[4] In addition I observe that giving rebounces to its original level when the dictator game is played after the promise game, this is similar to the restart effect in repeated public good games first described by Andreoni (1988). However - supporting the predicted mechanism - I find that a change of promise is correlated with an according change in giving in all periods.

This study contributes in several ways to previous literature. Firstly, the model is relevant to the literature on electoral competition[5] by showing that a second dimension of private information, here intrinsic motivation (a lower cost of providing quality), is relevant to understand the effects of promise competition. I show that the second dimension alters the set of promises all sellers make. The motivation dimension may also be of interest for self-selection into cheap talk situations as studied by Fehrler et al. (2018). In the equilibria of this model, not only sellers who are dishonest but also those who are intrinsically motivated gain higher utility on

[4] I argue this suggests that an underlying factor (such as a norm or moral obligation) changes with repetitions and feedback similar to decline of cooperation in a public good game due to imperfect conditional cooperation as described by Fischbacher and Gächter (2010).

[5] See Corazzini, Kube, Maréchal, and Nicolò (2014) and Fehrler, Fischbacher, and Schneider (2018).

expectation, hence might be more likely to self-select.

Secondly, previous literature has studied how multiple dimensions of private information can lower the information transmitted by signals.[6] In particular Frankel and Kartik (2017) study how unobservable heterogeneity in two dimensions can 'muddle' the informativity of market signals if market participants try to infer the type of an agent based on a single signal. This study demonstrates that signals can become entirely un-informative if there exists only one type who can break her promise without a cost, i.e. a type who can 'game' the signal for free.[7] The finding contrasts the model of Callander and Wilkie (2007) who analyze electoral competition if politicians cannot perfectly commit to left - right campaign platforms and show that in such environment dishonest politicians are selected more often as they can make more central promises.

Thirdly this study contributes to the experimental literature that investigates promising in a competitive environment.[8] This study is the first to investigate whether competing promises allow a selection of better promisors. It also is the first and to evaluate the combined effect of promise competition on quality provision in comparison to a setting absent competition and promising.[9] Both experimental contributions base on clear predictions of my theoretical framework.

Finally, this study is the first to investigate quality competition of sellers when offers are not perfectly binding but instead are promises, costly to break for some but not all sellers.[10]

The experimental part of this study focuses on one particular motivation for

[6] For example Bernheim and Kartik (2014) investigate self-selection into corrupt political systems in a model with two-dimensional private information, yet they don't regard promises in their analysis.

[7] Frankele and Kartik rule this out by assumption. Note also that there is a second difference in the scenario I study, which is that the promise itself may alter the value of a seller to the buyer. This is so as honest sellers may increase their quality provision after promising high quality.

[8] See Geng, Weiss, and Wolff (2011); Corazzini et al. (2014); Feltovich and Giovannoni (2015); Born, van Eck, and Johannesson (2018); Casella, Kartik, Sanchez, and Turban (2018); Fehrler et al. (2018).

[9] Previous studies have established that an election increases the promise-keeping ceteris paribus Geng et al. (2011); Born et al. (2018) and that people make lower promises absent electoral competition Corazzini et al. (2014). There is also evidence that the effect of competition vanishes when providing the receiver with an outsite option (Casella et al., 2018).

[10] There is literature capturing the idea that sellers misrepresent the true quality of the good they are selling, this literature focuses on the effect of reputation instead of promises however, see for example Jin and Kato (2006).

the cost of promise breaking - a psychological dis-utility from breaking a promise. Importantly the model does not require the cost to be of psychological nature and applies to situations in which such cost stem from different sources as well. Alternative sources for such cost are legal constraints, reputation costs, or fabrication costs. Whereas the economic literature traditionally saw promises as cheap talk, a large body of recent empirical literature finds that some people incur a psychological cost when misrepresenting the truth[11] which leads Abeler, Nosenzo, and Raymond (2018) to conclude that people lie surprisingly little when reviewing the experimental literature on misrepresentation. The literature also finds a substantial heterogeneity regarding the preference for honesty as well as promise-keeping[12] which supports the model's assumption that seller types differ in their honesty. In the model allows a functional form of promise-breaking that includes both a fixed cost of promise breaking and a cost that increases the further an agent deviates her promise. While research hasn't yet investigated the relation between a cost of lying and promise breaking, it seems conceivable that the functional form of these costs should be similar. For this reason, the model introduced in this paper allows for a functional form as suggested by Abeler et al. (2018) and Gneezy et al. (2018).[13]

2 A Model of Promise Competition

A principal can select one out of two agents. Call the principal *promisee* and the agents *promisors* from here on. The two promisors $i \in \{1, 2\}$ simultaneously choose a promise $p_i \in [0, \infty)$ about the service quality $x_i \in [0, \infty)$ they will provide to the promisee if they are selected. The promisee observes the promises and decides which promisor to select. Let a denote the promisee's (mixed) strategy regarding any two promises and $a(p_1, p_2)$ the promisee' probability to select promise p_1 over promise

[11] Some influential studies are Ellingsen and Johannesson (2004); Gneezy (2005); Charness and Dufwenberg (2006); Mazar, Amir, and Ariely (2008); Vanberg (2008); Sutter (2009); Lundquist, Ellingsen, Gribbe, and Johannesson (2009); Fischbacher and Föllmi-Heusi (2013); Abeler, Becker, and Falk (2014).

[12] For example Gibson, Tanner, and Wagner (2013); Gneezy, Rockenbach, and Serra-Garcia (2013); Gneezy, Kajackaite, and Sobel (2018) find this in the domain of lying and Corazzini et al. (2014); Born et al. (2018) in the domain of promise keeping.

[13] These studies assume that people experience both, a direct cost of lying and a cost of being seen as liar. This motivation is plausible in my context, too.

p_2. A promisor who gets selected may freely choose the quality x_i she provides. Promisors cannot condition their decisions on whether they are promisor 1 or 2 and neither can the promisee condition her selection decision on the promisors' index,[14] so henceforth we can omit i. The ex-post utility of a promisor who gets selected is,

$$u(x, p, \alpha, \rho) = 1 - x + \alpha \cdot f(x) - \rho \cdot g(p, x), \tag{1}$$

where $\alpha \in \{0, \overline{\alpha}\}$ expresses the promisor's intrinsic motivation, and $\rho \in \{0, \overline{\rho}\}$ expresses the promisor's cost of promise breaking. A promisor who does not get selected receives zero utility. The promisee only cares about the service-quality x and her utility, $v(x)$, is strictly increasing in x.

Assumption 1. *The function $f(x)$ is twice continuously differentiable and satisfies $\overline{\alpha} f'(0) > 1$ and $f''(x) \leq 0$ for all x.*

The cost of promise-breaking takes the form,

$$g(x, p) = \begin{cases} G(|x-p|/p) + \nu & \text{if } p \neq x; \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Assumption 2. *The function $G(|x-p|/p)$ satisfies $G(0) = 0$, $G(|x-p|/p) \geq 0$, and $G''(|x-p|/p) > 0$.*

A promisor type is described by the pair $\tau = (\alpha, \rho)$. The main analysis focuses on the case in which the type $(\overline{\alpha}, 0)$ does not exist, so the type space is $\mathrm{T} = \{(0,0), (0, \overline{\rho}), (\overline{\alpha}, \overline{\rho})\}$. Call the types "bad", "honest", and "good" and define $\tau_b = (0,0)$, $\tau_h = (0, \overline{\rho})$, and $\tau_g = (\overline{\alpha}, \overline{\rho})$ accordingly. The type τ is private information of each promisor and unobservable to the promisee. Let ϕ_τ denote the the prior probability of type τ. In the baseline setting we assume $\phi_\tau = 1/3$ for all $\tau \in T$.

Define $x^0(\tau) := \arg\max_x \; 1 - x + \alpha f(x)$ as the natural provision of type τ. For simplicity write $x^0(\tau_g)$ as $\overline{x}^0$ and $x^0(0, \cdot)$ as $\underline{x}^0 = 0$. Denote a promisor's optimal provision of quality,

$$x^*(p, \tau) := \arg\max_x u(x, p, \alpha, \rho).$$

[14] A promisee observes two times the same promise picks either with equal likelihood, formally $a(p_1, p_2) = .5$ if $p_1 = p_2$.

Denote the degree of promise-fulfillment $\kappa_\tau(p)$. Thus $x^*(p,\tau) = \kappa_\tau(p)\cdot p$. Finally, we make the following assumption.

Assumption 3. *The parameters are such that,*

$$\frac{\phi_{\tau_g}}{\phi_{\tau_g}+\phi_{\tau_l}}\overline{x}^0 \geq x^*(\overline{x}^0,\tau_h).$$

The assumption implies that type τ_h breaks any promise of $\overline{x}^0$ or more.

To define an equilibrium we introduce more notation. Let s_τ denote the mixed strategy with regard to promises of promisor type τ and s the strategy profile of all types. That is, $s_\tau(p)$ is the probability that type τ promises p. Let μ denote the system of beliefs of the promisee. When the promisee observes a promise p, let $\mu(\tau|p)$ denote the promisee's belief about the probability that promise p is coming from type τ. Given beliefs and a promise p we can write the promisee's expected quality selecting that promise as $E\big[x|p,\mu\big] = \sum_{\tau\in T}\mu(\tau|p)\cdot x^*(p,\tau)$.

We denote the expected utility of a promisor τ as $E[u(p,\tau,a,\rho)|s,a]$, where s represents the strategy of the competing promisor. In a perfect Bayesian Equilibrium (PBE) a promisor's choice of quality-provision is optimal given the own promise and type, and the promise strategy maximizes expected utility given the promisee's strategy. The promisee's strategy maximizes expected utility given the promisee's associated beliefs, and these conform with Bayes' rule whenever it applies. More formally, (s^*,x^*,μ^*,a^*) form a PBE if and only if the following conditions hold.

i For all $\tau\in \mathrm{T}$ and $p\in[0,\infty)$, $x^*(p,\tau) = \arg\max_x u(p,x,\tau)$.

ii For all $\tau\in\mathrm{T}$, if $s^*_\tau(p)>0$, then $p\in\arg\max_p E[u(x^*,p,\alpha,\rho)|s^*,a^*]$.

iii For all p_1,p_2, $a^*(p_1,p_2) = \arg\max_{a(p_1,p_2)} E\big[x|p,\mu\big]\cdot a(p_1,p_2) + E\big[x|p,\mu\big]\cdot(1-a(p_1,p_2))$.

iv For all $\tau\in\mathrm{T}$ and $p\in[0,\infty)$,

$$\mu^*(\tau|p) = \frac{Pr(\tau)\cdot s^*(p,\tau)}{\sum_{\mathrm{T}} Pr(\tau)\cdot s^*(p,\tau)},$$

$$\textit{if}\ \sum_{\mathrm{T}} Pr(\tau)\cdot s^*(p,\tau) > 0,$$

and

$\mu^*(\tau|p)$ is any probability distribution on T otherwise.

3 Analysis

This section describes the Perfect Bayesian Equilibria of the model and refines them. Let $\overline{p}^0$ describe the promise for which type τ_g receives zero ex-post utility, the highest promise this type would consider. The following proposition establishes that any promise lower or equal to $\overline{p}^0$ can be the sole equilibrium promise in a perfect Bayesian Equilibrium.

Proposition 1. *For every promise p in $\left[0, \overline{p}^0\right]$, there exist beliefs such that the following strategies form a perfect Bayesian equilibrium,*

$$s^*_\tau(p) = 1 \textit{ for all } \tau,$$

$$a^*(p^1, p^2) = \begin{cases} 1 & \textit{if } p^1 = p \textit{ and } p^2 \neq p; \\ 0 & \textit{if } p^1 \neq p \textit{ and } p^2 = p; \\ 0.5 & \textit{otherwise,} \end{cases}$$

$$x^*(p, \tau) \textit{ as defined in Equation (2) for all } \tau.$$

Proof. Consider any promise p^* with $s^*_\tau(p^*) = 1$ for all τ. The following beliefs,

$$\mu(\tau_l|p) = \begin{cases} \phi_{\tau_l} & \text{if } p = p^*; \\ 1 & \text{if } p \neq p^*, \end{cases}$$

$$\mu(\tau_h|p) = \begin{cases} \phi_{\tau_h} & \text{if } p = p^*; \\ 0 & \text{if } p \neq p^*, \end{cases}$$

$$\mu(\tau_g|p) = 1 - \mu(\tau_h|p) - \mu(0, 0|p),$$

obey Bayes' rule. Given these beliefs, the expected quality implied by any other promise $p \neq p^*$ equals zero. Hence, the promisee's decision $a^*(p^1, p^2)$ maximizes the promisee's expected utility. Given the promisee's decision, the following strategy maximizes a promisor's expected utility:

$$s_\tau^*(p) = \begin{cases} 1 & \text{if } p = p^*; \\ 0 & \text{otherwise,} \end{cases}$$

and $x^*(p,\tau) = \arg\max u(p,x,\tau)$ as defined in Equation (2). □

The next proposition establishes that no separating equilibria exist in which all types separate from each other by making distinct promises. The intuition is that a promisee tries to avoid the bad type τ_l. Yet this type doesn't face any cost to mimic other promises thus always mimics the most successful promise in terms of selection-probability.

Proposition 2. *There exists no fully separating equilibrium.*

Proof. Suppose there is an equilibrium in which all types separate by making a different promise. W.l.o.g. let type τ_l promise p'. Then, by the fact that all types separate, there exists a promise $p'' > 0$ made by one of the other types with $p'' \neq p'$. For any promise p', type τ' provides quality $\underline{x}^0$ after selection. For any promise p'', either type τ_g or τ_h provides $x > \underline{x}^0$. Since beliefs respect Bayes' Rule, $\mu(\tau_l|p') = 1$ and $\mu(\tau_l|p'') = 0$. Therefore, the promisee strictly prefers p'' to p'. But then $a^*(p', p'') = 0$ and type τ' profits from deviating to promise p''. Hence, this cannot be an equilibrium. □

The model does admit semi-separating equilibria, however. In these, types τ_g and τ_l promise quality p_H and τ_h separates to a lower promise, p_L. The promisee prefers p_H over p_L and selects the latter only if p_H is not available. Proposition 3 establishes this kind of equilibrium.

Recall that ϕ_τ denotes the probability that nature draws a candidate of type τ. Since the promisee prefers p_H, the likelihood to get selected with either promise is,

$$S(p_H) = 1 - 0.5\left(\phi_{\tau_g} + \phi_{\tau_l}\right),$$

and,

$$S(p_L) = 0.5 - 0.5\left(\phi_{\tau_g} + \phi_{\tau_l}\right).$$

Define,

$$S(L/H) = S(p_L)/S(p_H) = 0.25.$$

The following proposition establishes the existence of semi-separating equilibria.

Proposition 3. *For all p_L and p_H such that,*

$$\begin{aligned}
&i.\ u\big(x^*, p_H, \tau_g\big) \geq u\big(x^*, p_L, \tau_g\big) \cdot S(L/H),\\
&ii.\ u\big(x^*, p_H, \tau_h\big) \leq u\big(x^*, p_L, \tau_h\big) \cdot S(L/H),\\
&iii.\ x^*(p_H, \tau_g)\frac{\phi_{\tau_g}}{\phi_{\tau_l} + \phi_{\tau_g}} \geq x^*(p_L, \tau_h),
\end{aligned}$$

there exist beliefs such that the following strategies form a perfect Bayesian equilibrium,

$$a(p_1, p_2) = \begin{cases} 1 & \textit{if } p_1 = p_H \textit{ and } p_2 \neq p_H;\\ 1 & \textit{if } p_1 = p_L \textit{ and } p_2 \notin \{p_H, p_L\};\\ 0 & \textit{if } p_2 = p_H \textit{ and } p_1 \neq p_H;\\ 0 & \textit{if } p_2 = p_L^* \textit{ and } p_1 \notin \{p_L, p_H\};\\ 0.5 & \textit{otherwise,} \end{cases}$$

$$s^*_{\tau_h}(p) = \begin{cases} 1 & \textit{if } p = p_L;\\ 0 & \textit{otherwise,} \end{cases}$$

$$s^*_{\tau_l}(p) = \begin{cases} 1 & \textit{if } p = p_H;\\ 0 & \textit{otherwise,} \end{cases}$$

$$s^*_{\tau_g}(p) = \begin{cases} 1 & \textit{if } p = p_H;\\ 0 & \textit{otherwise.} \end{cases}$$

Proof. Consider the following system of beliefs, which assigns correct probabilities to each type when observing p_H, p_L and full probability to type τ_l when observing any off-equilibrium promise,

$$\mu(0,0|p) = \begin{cases} 0 & \text{if } p = p_L; \\ \frac{\phi_{\tau_l}}{\phi_{\tau_l}+\phi_{\tau_g}} & \text{if } p = p_H; \\ 1 & \text{otherwise,} \end{cases}$$

$$\mu(\tau_h|p) = \begin{cases} 1 & \text{if } p = p_L; \\ 0 & \text{otherwise,} \end{cases}$$

$$\mu(\tau_g|p) = 1 - \mu(\tau_h|p) - \mu(0,0|p).$$

These beliefs follow Bayes' Rule where applicable. By Condition iii, the promisee prefers a promisor with promise p_H over a promisor with p_L given these beliefs. Accordingly the promisee's strategy is optimal. Given the promisee' strategy, no promisor type has incentive to deviate to another promise: Type τ_g maximizes her utility by Condition i, type τ_h maximizes her utility by Condition ii, and type τ_l makes the promise that yields the highest probability of selection, which is optimal as the type faces no cost of promise-breaking. This concludes the proof. □

Lastly, there exists semi-separating equilibria in which τ_h promises quality p_L, τ_l promises a higher quality p_H, and τ_g mixes between both such that the promisee is indifferent. In turn the promisee mixes such that τ_g is indifferent between both promises. Define $\hat{a}$ as the probability distribution over promises p_L and p_H such that τ_g is indifferent between either promise,

$$u\big(x^*, p_H, \tau_g\big) \cdot \hat{a}(p_H) = u\big(x^*, p_L, \tau_g\big) \cdot \hat{a}(p_L).$$

And define $\hat{s}$ as the probability distribution over promises p_L, p_H such that the promisee is indifferent,

$$\frac{\phi_{\tau_g} \cdot \hat{s}(p_H) \cdot x^*(p_H, \tau_g)}{\phi_{\tau_l} + \phi_{\tau_g} \cdot \hat{s}(p_H)} = \frac{\phi_{\tau_g} \cdot \hat{s}(p_L) \cdot x^*(p_L, \tau_g) + \phi_{\tau_h} \cdot x^*(p_L, \tau_h)}{\phi_{\tau_h} + \phi_{\tau_g} \cdot \hat{s}(p_L)}.$$

The following proposition establishes the existence of these semi-separating equilibria.

Proposition 4. *For all p_L and p_H such that,*

$$u\big(x^*, p_H, \tau_g\big) \geq u\big(x^*, p_L, \tau_g\big),$$

there exist beliefs such that the following strategies form a perfect Bayesian equilibrium:

$$a(p_1, p_2) = \begin{cases} \hat{a}(p_H) & \text{if } p_1 = p_H \text{ and } p_2 = p_L; \\ \hat{a}(p_L) & \text{if } p_1 = p_L \text{ and } p_2 = p_H; \\ 1 & \text{if } p_1 \in \{p_L, p_H\} \text{ and } p_2 \notin \{p_L, p_H\}; \\ 0 & \text{if } p_2 \in \{p_L, p_H\} \text{ and } p_1 \notin \{p_L, p_H\}; \\ 0.5 & \text{otherwise,} \end{cases}$$

$$s^*_{\tau_h}(p) = \begin{cases} 1 & \text{if } p = p_L; \\ 0 & \text{otherwise,} \end{cases}$$

$$s^*_{\tau_l}(p) = \begin{cases} 1 & \text{if } p = p_H; \\ 0 & \text{otherwise,} \end{cases}$$

$$s^*_{\tau_g}(p) = \begin{cases} \hat{s}(p_H) & \text{if } p = p_H; \\ \hat{s}(p_L) & \text{if } p = p_L; \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Consider the following system of beliefs, which assigns correct probabilities to each type when observing p_H, p_L and full probability to type τ_l when observing any off-equilibrium promise,

$$\mu(\tau_l|p) = \begin{cases} 0 & \text{if } p = p_L; \\ \frac{\phi_{\tau_l}}{\phi_{\tau_l}+\phi_{\tau_g}} & \text{if } p = p_H; \\ 1 & \text{otherwise,} \end{cases}$$

$$\mu(\tau_h|p) = \begin{cases} 1 & \text{if } p = p_L; \\ 0 & \text{otherwise,} \end{cases}$$

$$\mu(\tau_g|p) = 1 - \mu(\tau_h|p) - \mu(\tau_l|p).$$

These beliefs follow Bayes' Rule where applicable and the promisee is indifferent between a promisor with promise p_H or p_L. Hence, the promisee's strategy is optimal. Given the promisee' strategy, no promisor type gains utility from deviating to another promise: Type τ_g is indifferent between both equilibrium promises, hence maximizes her utility, type τ_h prefers the lower promise, hence maximizes her utility, and type τ_l promises the quality that yields the highest probability of selection, which is optimal as the type faces no cost of promise-breaking. This concludes the proof. No promisor can gain from deviation to an off-equilibrium promise as these are never selected. $\square$

In a nutshell, Propositions 1, 3, and 4 establish the existence of a wide range of Bayesian Equilibria. In particular, any promise may be part of a pooling equilibrium. This is driven by the fact that the Bayesian Equilibrium concept allows promisees to have arbitrary beliefs about any promise never made in a particular equilibrium. As a result, the Bayesian Equilibrium concept cannot make strong predictions in this framework and hence is unsatisfactory. The following section of this paper, applies Criterion D1 to refine the set of equilibria in order to address this issue.

3.1 Refinement

This section refines the perfect Bayesian Equilibria by applying Criterion D1 of Cho and Kreps (1987). As it turns out, the refinement eliminates all equilibria except a range of pooling equilibria at the high natural action and above.

Let $\hat{U}(\tau)$ define the equilibrium expected utility of a promisor of type τ who promises $\hat{p}$ while the other types promise $\hat{p}_{-\tau}$, and $U(\tau, a, p, \hat{p}_{-\tau})$ the expected utility of a promisor of type τ who promises p, given the promisee strategy a and the other types' promise-strategies $\hat{p}_{-\tau}$. Let $MBR(\mu, p)$ define the set of the promisee's mixed strategy best responses to p given beliefs μ. The set of all mixed promisee best responses strategies that make promisor type τ strictly prefer p to $\hat{p}$ is

$$D(\tau, p) := \cup_\mu \{a \text{ in } MBR(\mu, p) | U(\tau, a, p, \hat{p}_{-\tau}) > \hat{U}_P(\tau)\}. \tag{3}$$

Let $D^0(\tau, p)$ be the set of best response strategies such that promisor type τ is indifferent between p and $\hat{p}$.

Definition 1. *(Criterion D1) A type τ is deleted for strategy p under Criterion D1 if there is a type τ' such that*

$$\{D(\tau, p) \cup D_0(\tau, p)\} \subset D(\tau', p)\}.$$

Intuitively, a promisee belief regarding the promisor type a non-equilibrium promise comes from satisfies D1 if the type deviates to this promise for the largest set of promisee decisions, compared to all other types. An equilibrium satisfies D1, if all assigned beliefs about non-equilibrium promises satisfy D1.

D1 permits a range of equilibria that all share the same properties and all make very similar predictions. Hence, the set of refined equilibria is completely sufficient to make the main predictions of the model. Still it is possible to refine that set further. To that end I apply a second refinement, the Undefeated Equilibrium due to Mailath, Okuno-Fujiwara, and Postlewaite (1993), to refine the set of equilibria which satisfy D1 further. That means I apply the undefeated Equilibrium refinement only to the set of equilibria satisfying D1.[15] which reduces the set of equilibria to a single pooling equilibrium, as I show below.

Let $\sigma(s, x, \mu, a)$ denote a Bayesian Equilibrium and, in a slight abuse of notation, let $u(\sigma, \tau)$ denote the expected utility of seller type τ in equilibrium σ. Then an equilibrium is undefeated if it is not defeated by any other equilibrium, as described in the following definition.

Definition 2. *An equilibrium $\sigma(s, x, \mu, a)$ defeats another equilibrium $\sigma'(s', x', \mu', a')$ if there exists a $p \in [0, \infty)$ such that:*

i. $\forall \tau \in \mathrm{T}: \; s'_\tau(p) = 0,$ *and* $K \equiv \{\tau \in \mathrm{T} | s_\tau(p) > 0\} \neq \emptyset;$

ii. $\forall \tau \in K: u(\sigma, \tau) \geq u(\sigma', \tau),$ *and* $\exists t \in K: \; u(\sigma, \tau) > u(\sigma', \tau);$ *and*

iii. $\exists t \in K: \mu'(t|m) \neq \phi_\tau \; \pi(t) / \sum_{t' \in \mathrm{T}} \phi_{\tau'} \; \pi(\tau')$ *for any* $\pi: T \to [0, 1]$

satisfying

$\tau' \in K$ *and* $u(\sigma, \tau') > u(\sigma', \tau') \Rightarrow \pi(\tau') = 1,$ *and*

[15] This deviates from the original definition which requires the refinement to compare all sequential equilibria.

$\tau' \notin K \Rightarrow \pi(\tau') = 0.$

An equilibrium σ is undefeated, if there does not exist a σ' that defeats σ.

In the following I discuss Criterion D1 first before I finally apply the Undefeated equilibrium concept of Definition 2 to the remaining set of equilibria.

Proposition 5. *No semi-separating equilibrium satisfies Criterion D1.*

Proof. I prove this by contradiction. First iI show that D1 eliminates all equilibria in which one type entirely separates (as described in Proposition 3) which permitted by D1. Second I show that D1 also eliminates the remaining semi-equilibrium (described in Proposition 4).

Suppose that there exists an equilibrium in which one type separates from the others with any two promises p', p''. Without loss of generality let p' be the quality which the separating type promises.

(i) Suppose the equilibrium is such that type τ_l separates and promises quality p'. A promisee selecting this promise receives $\underline{x}^0 = 0$. For any p'' type τ_h provides greater or equal quality and τ_g provides a strictly greater quality. Hence a promisee strictly prefers any other promise $p'' \neq p'$ for which $\mu(\tau_g|p'') \geq 0$ to p'. Hence this is not an equilibrium.

(ii) Suppose the equilibrium is such that type τ_g separates and promises promise p'. The promise p' has to be selected with equal or lower probability than p'', otherwise τ_l would deviate. Thus the promisee has to weakly prefer p'' to p'. For any two promises $p' < p''$, types τ_l and τ_h provide a strictly lower quality than τ_g, hence $p'' > p'$ has hold. But then for any promise p'' there exists a $\varepsilon > 0$ such that the promise $p = p'' - \varepsilon$ increases the ex-post utility of τ_h whereas it doesn't alter the utility of τ_l. This means the set of election probabilities at which τ_h deviates to p is larger than the same of τ_l and Criterion D1 eliminates type τ_l for p. Under these beliefs a promisee prefers p to p'' hence this is not an equilibrium.

(iii) Suppose the equilibrium is such that type τ_h separates and promises promise p'. Promise p' cannot be selected over p'' otherwise τ_l would make this promise, too. At the same time τ_h has to prefer p' over p'' given the equilibrium promisee strategy

a. That means $p' < p''$. However then there exists a promise $p = p' + \varepsilon$ for an $\varepsilon > 0$ such that type τ_h is willing to diverge to that promise for a selection probability less than twice than the probability with which a promisor promising p_L gets selected. For that probability type τ_l is not willing to diverge from p_H which means Criterion D1 deletes τ_l for p. Regardless of the promisee's beliefs about the other types, the promisee strictly prefers p to p'. This means τ_h prefers p to p', hence this is not an equilibrium. This concludes the proof that no semi-separating equilibria in which one type separates completely survive criterion D1.

Second, I prove that there exists no semi-separating equilibrium as described in Proposition 4. Suppose the opposite. In such equilibrium τ_g is indifferent between p_L and p_H, τ_h weakly prefers p_L, and τ_l weakly prefers p_H.

Then there exists a promise $p \in (p_L, p_H)$ such that τ_g is willing to deviate to p for the largest set of promisee strategies a. To see this, note that τ_g gains ex-post utility from promising closer to her natural quality provision $\overline{x}^0$ whereas τ_l receives equal ex-post utility from all promises and τ_l gains ex-post utility from promising less. Due to the difference in $\alpha \cdot f(x)$, the marginal gains of τ_g and τ_h are different with an exception of single promises in which the marginal utilities crosses.

Given that, Criterion D1 requires the belief $\mu(\tau_g|p) = 1$. Under this belief the promisee prefers p to p_L and p_L, p_H are not an equilibrium. This contradiction concludes the proof. □

The proposition establishes that Criterion D1 eliminates equilibria with selection based on promises. Note however that this result partly depends on the continuity of the promise space. If the promise space is discrete, D1 permits two adjacent promises to form a semi-separating equilibrium of the kind described in Proposition 4. In these equilibria τ_h promises the lower quality and τ_l the higher quality. τ_g mixes between both promises such that the promisee is indifferent and selects each promise with equal likelihood. These equilibria can explain the findings of Casella et al. (2018), namely mixing between the promises $'5'$ and $'6'$ in the presence of completion.

Complementing the finding that separation is not possible in equilibrium, Propositions 6 and 7 establish that a reduced set of pooling equilibria at a single promise survive Criterion D1. The range of permissible promises starts at the natural service quality of a motivated promisor, $\overline{x}^0$, and ranges to some higher quality I denote as p^{max}.

We define $p^{max} = \min\{\tilde{p}, \hat{p}\}$, where $\tilde{p}$ denotes the promise at which the marginal ex-post utility from the promise relative to total ex-post utility is equal for τ_g and τ_h, and $\hat{p}$ is the promise at which the promisee receives the same expected utility from a mixture of the types τ_g and τ_l at prior probabilities or τ_h with certainty. Formally,

$$u'_p(x^*(\tilde{p}), \tilde{p}, \tau_g)/u(x^*(\tilde{p}), \tilde{p}, \tau_g) = u'_p(x^*(\tilde{p}), \tilde{p}, \tau_h)/u(x^*(\tilde{p}), \tilde{p}, \tau_h),$$

and,

$$\frac{\phi_{\tau_g}}{\phi_{\tau_g} + \phi_{\tau_l}} \, x^*(\hat{p}, \tau_g) = x^*(\hat{p}, \tau_h).$$

Observe that, $p^{max} < \overline{p}^0$. Which means that the highest possible equilibrium is lower than the highest promise τ_g would want to make.

Proposition 6. *There exist beliefs satisfying Criterion D1 such that for p^* in $[\overline{x}^0, p^{max}]$ the following strategies form a perfect Bayesian (pooling) equilibrium:*

$$s_\tau(p^*) = 1 \textit{ for all } t,$$

$$a^*(p^1, p^2) = \begin{cases} 1 & \textit{if } p^1 = p^* \textit{ and } p^2 \neq p^*; \\ 0 & \textit{if } p^1 \neq p^* \textit{ and } p^2 = p^*; \\ 0.5 & \textit{otherwise}, \end{cases}$$

$$x^*(p, \tau) \textit{ as defined in Equation (2) for all types } \tau,$$

where $p^{max} = \min\{\tilde{p}, \hat{p}\}$.

Proof. We begin by showing that there exist beliefs satisfying D1 that support this equilibrium. Let p^* be any equilibrium promise in $[\overline{x}^0, p^{max}]$. Consider a downward deviation to any promise $p' < p^*$, by choice of p^{max}, type τ_h is the type that gains highest expected utility from deviating to p'. Accordingly, D1 deletes all other types, and promisee beliefs are $\mu(\tau_h|p') = 1$ for all $p' < p^*$. Under these beliefs no one will deviate to ρ'.

Secondly, consider any promise larger than the equilibrium promise $p' > p^*$. Both τ_g and τ_h loose ex-post utility from increasing their promise, while τ_l does not. Again D1 deletes these types for promisee beliefs and only the belief $\mu(\tau_l|p') = 1$ is permissible. Hence beliefs satisfying D1 support the equilibria. Finally, given these beliefs a^* is optimal and so is $s_\tau(p^*) = 1$, which concludes the proof. □

The next proposition establishes that Criterion D1 eliminates all other pooling equilibria.

Proposition 7. *The equilibria described in Proposition 6 are the only equilibria that satisfy Criterion D1.*

Proof. First we show that there exist no pooling equilibria in pure strategies other than described in Proposition 6. Suppose there exists an equilibrium with promise p not in the set of equilibria $E = [\overline{x}^0, p^{max}]$. For any p smaller than $\overline{x}^0$, the type that gains highest ex-post utility from deviating to $\overline{x}^0$ is τ_g, hence the only belief not eliminated by D1 is $\mu(\tau_g|\overline{x}^0) = 1$ which does not support p as equilibrium. For any p larger than p^{max}, by construction of p^{max}, the type gaining most from deviating to a marginally lower promise p' is τ_g. Again D1 eliminates all beliefs other than $\mu(\tau_g|p') = 1$ thus p cannot be an equilibrium.

Second, focus on equilibria that involve more than one promise and in which all types mix with equal probabilities between the equilibrium-promises. To see that no such pooling equilibrium exist, consider the utility the different types get from winning with any two promises $p^1 < p^2$. Type τ_h receives higher ex-post utility from the former, whereas τ_l receives equal ex-post utility from either. That means, both types cannot be indifferent between the promises for the same promisee-strategy a. Accordingly there cannot exist a pooling equilibrium in which both types mix between the two promises.

Third, regard equilibria that involve more than one promise in which Promisor types mix differently bewtween these promises. Without loss of generality consider an equilibrium including any two promises $p^1 < p^2$. As established before, τ_h and τ_l do not bewteen two promises for the same selection-likelihood $a(p^1, p^2)$. By Assumption 3 and 1 type τ_g has to mix to all involved promises if any promise is above or equal to $\overline{x}^0$, otherwise there is a promise which is strictly preferred by the voters. Assume this case first. Then there might be an equilibrium in which the voters choose a to set τ_g indifferent, whereas τ_h strictly prefers the lower promise and τ_l is indifferent or prefers the higher promise. As τ_h strictly prefers the p^1 there exists a promise $p = p^2 - \varepsilon$ for some $\varepsilon > 0$ such that τ_g prefers deviation to this promsie for a larger set of a than the other types. Criterion D1 requires promisee beliefs $\mu(\tau_g|p) = 1$, hence they prefer p to p^2 and this cannot be an equilibrium. Now assume $p^1 < p^2 < \overline{x}^0$, type τ_g gains the most ex-post utility from a marginal deviation to $p = p^2 + \varepsilon$ for some $\varepsilon > 0$. D1 requires promisee beliefs to be accordingly, hence this cannot be an equilibrium either. Which concludes the proof. □

Criterion D1 reduces the set of equilibria to a class of pooling equilibria which all generate the same predictions. Even though this class is completely satisfactory for generating predictions, in the following I investigate the possibility to refine that set even further by using the Undefeated Equilibrium refinement. Proposition 8 shows that a larger set of equilibria satisfy the Undefeated Refinement compared to the set that satisfies Criterion D1.[16] Hence in itself the refinment cannot reduce the set of equilibria.

This changes however, when applying the Undefeated refinement only to the set of equilibria that satisfy D1. Proposition 9 shows that only a single equilibrium satisfies the iterated application of Criterion D1 and the Undefeated Refinement.

Proposition 8. *For every promise in* $[\overline{x}^0, \overline{p}^0]$ *exists a Bayesian pooling equilibrium with that promise which is undefeated by any other Bayesian pooling equilibrium.*

Proof. First note that a pooling equilibrium not necessarily satisfying D1, $\sigma(s, x, \mu, a)$, with an equilibrium promise p at or above $\overline{x}^0$, cannot be defeated by another such equilibrium $\sigma'(s', x', \mu', a')$ if $p' > p$. This is so as no promisor type strictly prefers to make a higher promise, hence Condition i of Definition 2 is never fulfilled.

Now suppose an equilibrium $\sigma'(s', x', \mu', a')$ with equilibrium promise p'. Note that for any $p > p'$ there exists an equilibrium $\sigma(s, x, \mu, a)$ with $\mu(\tau|p') = \sigma_\tau / \sum_{\mathrm{T}} \sigma_\tau$

[16] I do not apply the Undefeated Equilibrium to the Semi-separating equilibria, as it is not possible to compare the utility levels of promisors in pooling and semi-separing equilibria without further assumptions about the functional form of the utility function.

for all τ and $p' < p$. Then there exists no $\sigma'(s', x', \mu', a')$ that can defeat $\sigma(s, x, \mu, a)$, as the promisee in σ has correct beliefs about the distribution of types that promise p' in equilibrium σ' but prefers to remain in σ. This violates Condition *iii* of the Definition 2. This means no pooling equilibrium can defeat σ accordingly it is undefeated. □

Proposition 9. *There exist exactly one equilibrium satisfying Criterion D1 that is undefeated by any other equilibrium satisfying D1:*

$$p^* = \overline{x}^0,$$

$$s_\tau(p^*) = 1 \textit{ for all } t,$$

$$a^*(p^1, p^2) = \begin{cases} 1 & \textit{if } p^1 = p^* \textit{ and } p^2 \neq p^*; \\ 0 & \textit{if } p^1 \neq p^* \textit{ and } p^2 = p^*; \\ 0.5 & \textit{otherwise}, \end{cases}$$

$$x^*(p, \tau) \textit{ as defined in Equation (2) for all types } \tau,$$

$$\mu(\tau_h, p) = \begin{cases} 1 & \textit{if } p < p^*; \\ \phi_{\tau_h} & \textit{if } p = p^*; \\ 0 & \textit{if } p > p^*, \end{cases}$$

$$\mu(\tau_l, p) = \begin{cases} 1 & \textit{if } p > p^*; \\ \phi_{\tau_l} & \textit{if } p = p^*; \\ 0 & \textit{if } p < p^*, \end{cases}$$

$$\mu(\tau_g, p) = 1 - \mu(\tau_l, p) - \mu(\tau_h, p).$$

Proof. Consider any other equilibrium σ that satisfies Criterion D1 as described in Proposition 6. I I show that the equilibrium described above, σ^*, defeats any such equilibrium. In any σ all types pool at $\rho \neq \rho^*$, hence condition (i) is fulfilled. Furthermore, $u(\sigma, \tau) \geq u(\sigma', \tau)$ which holds with strict inequality for τ_h and τ_g, hence condition (ii) is fulfilled. Finally, $\mu(\tau_h | p^*) = 1$, however in σ^* $\pi(\tau_h) = 1$ and $\pi(\tau_g) = 1$, that means the inequality in (iii.) holds for all σ and hence any σ is defeated by σ^*.

It remains to show that σ^* is undefeated. Consider any pooling equilibrium σ'

that could defeat σ^*. In the class of pooling equilibria, τ_l and τ_g both make the highest utility possible in σ^*. Hence, by condition (i), in σ' τ_h needs to make receive higher expected utility than in σ^*. This can only be achieved by promising $p < p^*$. Yet $\mu^*(\tau_h|p) = 1$ which implies that any $p' < p^*$ violates Condition (iii). This contradiction concludes the proof. □

3.2 Benchmarks

This section discusses benchmarks to assess the equilibria with competition. I refer to the equilibria from the model above as 'the predicted equilibria'. I sketch the equilibria under the benchmark scenario and compare them to the predicted equilibrium in terms of selection (is any promisor type more likely to be selected?) and quality provision.

No Competition

We start with a simple comparison to an equilibrium in which there is no competition. In this case a promisee faces one promisor and has no decision to make - she always selects the promisor.[17] Accordingly, all promisors that face a cost of promise breaking, promise their natural quality. The bad promisor type promises anything and doesn't keep the promise. The promisee meets a random promisor and selects them which means each promisor type gets selected with its prior probability. Thus the selection likelihoods of each type is equal to the competition equilibrium. In contrast the service quality is lower in this benchmark scenario. This is driven by the honest type, who, in this scenario, provides zero quality, as compared to $x^*(p^*, \tau_h)$ in the equilibrium with competition described above. Both other types provide equal quality.

Binding contracts

Here we consider the game if promises are binding. In this case there is no difference between the types τ_l and τ_h. Under binding contracts the promisee always wants to

[17] This also covers the case that the promisee has a 0 outside option, but no additional preferences e.g. regarding fairness.

select the promisor who offers the highest quality, since the offer is binding. Let x_τ^{max} denote the quality for which types $\tau \in \{\tau_h, \tau_g\}$ receive zero ex-post utility, hence this denotes the highest quality the type offers. For the benchmarks I assume that $x_g^{max} > x_h^{max}$. This is the more plausible parameter configuration.[18] In equilibrium τ_l and τ_h promise at p_h^{max}. As τ_g has a $\phi_{\tau_l} + \phi_{\tau_h}$ probability of running against one of the other types, the type does not offer p_g^{max}. Instead the type mixes over an atomic distribution over $(p_h^{max}, p^{highest}]$, where $p^{highest}$ is the promise at which the expected utility under certain selection by the promisee equals the expected utility with promise $p^{highest}$ and selection probability $\phi_{\tau_l} + \phi_{\tau_h}$.

In the sketched equilibrium the quality provided to the promisee is larger than in the predicted equilibria - any promisor provides a quality at or above x_h^{max} which is larger than the highest quality provided in the predicted equilibria. In addition positive selection of τ_g occurs, whereas τ_l and τ_h both are selected with an equal smaller likelihood (See Table 1).

Perfect information

Regard an equilibrium in which all sides have perfect information. As promisees have perfect information, they always select the promisor who, given type and promise, provides the highest quality. As promisors have perfect information as well, they know which type they compete against and hence condition their choice of promise on their competitor's type. Let p_τ^{max} denote the promise for which types $\tau \in \{\tau_h, \tau_g\}$ receive zero ex-post utility, hence this denotes the highest promise this type considers. For the benchmark I assume that $p_g^{max} > p_h^{max}$ as this is the more plausible parameter configuration.[19]

In equilibrium, the two honest types of promisors promise their natural quality if they compete against a bad promisor and the promisee selects them with certainty.

[18] If $f(x)$ is negative for high values of x - a case that is generally compatible with this model - x_h^{max} can be larger than x_g^{max}. As an example think of a setting in which τ_g is motivated to fulfill a quality norm but unmotivated to provide (much) higher quality than the norm prescribes. This has some resemblance to the assumption about the 50-50 norm in Andreoni and Bernheim (2009).

[19] As described in Footnote 18, there are parametrizations of the model for which p_h^{max} can be larger than p_g^{max}.

If an honest type runs against a promisee of the same type, both promise p_h^{max} and the promisee selects either one. If the honest type runs against a the good type, τ_g 'outpromises' τ_h. The latter makes any promise equal or lower than p_l^{max} and τ_g promises p_l^{max}. If two good type promisors run against each other, both promise p_h^{max}. Finally the bad promisor type makes any promise, and the promisee only selects the type only if both promisors are of that type.

The following table compares the selection likelihood of each type in the second benchmark to the predictions of the equilibrium described above. In the perfect information scenario, the promisee selects the good type more often and the bad type less often than in the predicted equilibrium. Thus a positive selection of better promisors takes place. Which might not be surprising given perfect information.

Table 1: Comparison selection likelihood in benchmarks and predicted equilibria

Event	**Perfect Info**	**Binding Prom**	**Pred. Equilibria**
Selection τ_l	$\phi_{\tau_l}^2$	$0.5(\phi_{\tau_l}+\phi_{\tau_h})^2$	ϕ_{τ_l}
Selection τ_h	$\phi_{\tau_h}^2+2\phi_{\tau_h}\phi_{\tau_l}$	$0.5(\phi_{\tau_l}+\phi_{\tau_h})^2$	ϕ_{τ_h}
Selection τ_g	$0.5(\phi_{\tau_l}+\phi_{\tau_h})^2$	$(2\phi_{\tau_l}+2\phi_{\tau_h}+\phi_{\tau_g})\phi_{\tau_g}$	ϕ_{τ_g}

In contrast the movement of the quality provision is more ambiguous. Type τ_l does not change the provided quality of 0. τ_g provides higher (when competing with τ_g, or τ_h) or equal (when competing with τ_l) quality compared to the predicted equilibrium. Finally, τ_h provides higher quality (when competing with τ_h) or worse quality (when competing with τ_l). This does not permit a clear comparison regarded the average quality as it depends on the parameters whether average quality increases or decreases. Albeit it appears that under most parameter-choices average quality increases - the selection likelihood of the bad type decreases and the same of the good type increases in addition to the increased average quality provision by the latter - the conclusion that perfect information increases average provided quality might not hold in all situations (for all parameters of the model). This is due to the fact that not only the promisee but also the promisor has better information in this situation and does not have to compete in case

4 Experimental Design

The experiment I ran consists of two parts and a concluding questionnaire. Participants in the experiment play both parts in random order, that means in each session half of the participants starts with Part A and the other half starts with Part B. For simplicity we call the group who starts with Part A 'treatment 1' and the other group 'treatment 2'. Note that this is primarily to control for order effects and not for treatment comparison.

Part A - Dictator Game

In this part participants participants are randomly matched with a partner. Each group of two plays a one-shot dictator game with a multiplier. In the game a sender receives an endowment of 100 points and may decide how to split this endowment between herself and a receiver. The receiver gets no points apart from those the sender decides to assign to her and cannot pursue any action. Each point the sender sends to the receiver is doubled.

In the experiment, I use the strategy-method such that all subjects make a sender decision and the computer randomly determines who becomes a sender after all decisions are made. Participants learn about the outcome from this part only in the end of the experiment after both parts are concluded.

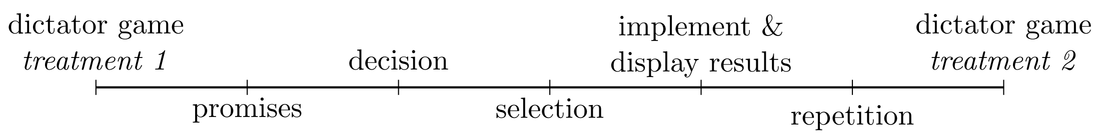
Part B - Promise Game

This part of the experiment is repeated is repeated 10 times. I call each repetition a round. Participants interact in a group of three and are rematched with two random participants in each round. In each group one person is assigned the role of a receiver and two the role of senders.

The receiver is asked to choose one of the two senders to play a dictator game with. The dictator game then is equivalent to the one described in Part A. The sender who is not selected does not participate in the game and receives no points. Before the receiver makes her selection decision. Both senders are asked to make a non-binding promise regarding the amount of points they will send to the receiver if they get selected. Participants enter their promise as a number into the computer interface. The promises are displayed to the receiver before the receiver makes the selection decision. After senders indicate their promise, they decide about the amount they actually want to send on a separate screen. This decision gets automatically implemented if they are selected as sender.

We apply the strategy method in this part of the experiment. That means we ask all three participants to make both sender and receiver decisions. First all participants make promises, after that we ask all participants to decide on the amount to give, finally all participants see the promises of their two group-members and are asked to make a selection-decision. After all participants make these decisions their roles are randomly assigned by the computer, and the according decisions automatically implemented. Each participant then gets feedback on what role they had in that turn, what the promises were in their group, who got selected and what the selected sender actually sent. Figure 1 displays the set-up of the experiment.

Figure 1: Timeline of the experiment.



The two described parts consists of eleven rounds taken together. For the payment, one of the eleven rounds is selected at random and all participants receive 10 cents (USD) for each point they earned during that round. In addition, each participant receives $10 for their participation.

Questionnaire

In the end of the experiment, participants answer a six-item questionnaire with the following questions.

- What is your age?
- What is your gender?
- What is your major?

- Did you incur any problems or questions during the experiemnt that could notbe resolved? If yes please describe!
- What was your rationale for the height of the promises you made?
- What was your rational for selecting the promises you selected?

The questions are not intended for a part of the analysis but as a control that recruiting and the experiment work as intended (i.e. subjects consistently reason in an unexpected way.).

IRB and Pre-registration

The project has been reviewed and approved by the IRB of the University of California San Diego. The registered project number is 180290. The experiment has been pre-registered before conduction at the AEA RCT Registry with the ID AEARCTR-0002952.

Conduction of the Experiment

We conducted the experiment in the beginning of May 2018 at the Incentives Lab at Rady School of Management, UCSD. In total 155 participants took part in the study over 14 sessions.

5 Hypotheses

This section spells out the hypotheses to test in the experiment. The hypothesis are derived from the equilibria in the model.

The equilibria that pass Criterion D1 predict that promisors, senders in the experiment, pool their promise. The equilibria predict this pooling to happen at a single service quality which is higher or equal than the natural quality provided by a motivated type. I recognize that it is unlikely in an experiment to achieve perfect pooling at a single promise, but expect to observe pooling behavior such that a significant share of the participants promises around the same amount larger than

33 points. Furthermore I expect the distribution of promsises to be single peaked around this amount and the distribution of selected and not-selected promises to look similar.

Hypothesis 1 (H1)**.** *Most participants in the sender role pool at the one promise, i.e. that senders in each session promise a similar amount of points.*

The predicted pooling of promises implies that participants are not able to select better senders in equilibrium. The following hypothesis specifies this.

Hypothesis 2 (H2)**.** *Participants who get selected do on average not give more to the receiver than participants who don?t get selected.*

Finally, the model predicts that senders give more to the receivers than they would have done in absence of promise competition. That means I expect senders to give more in the promise game than in the dictator game.

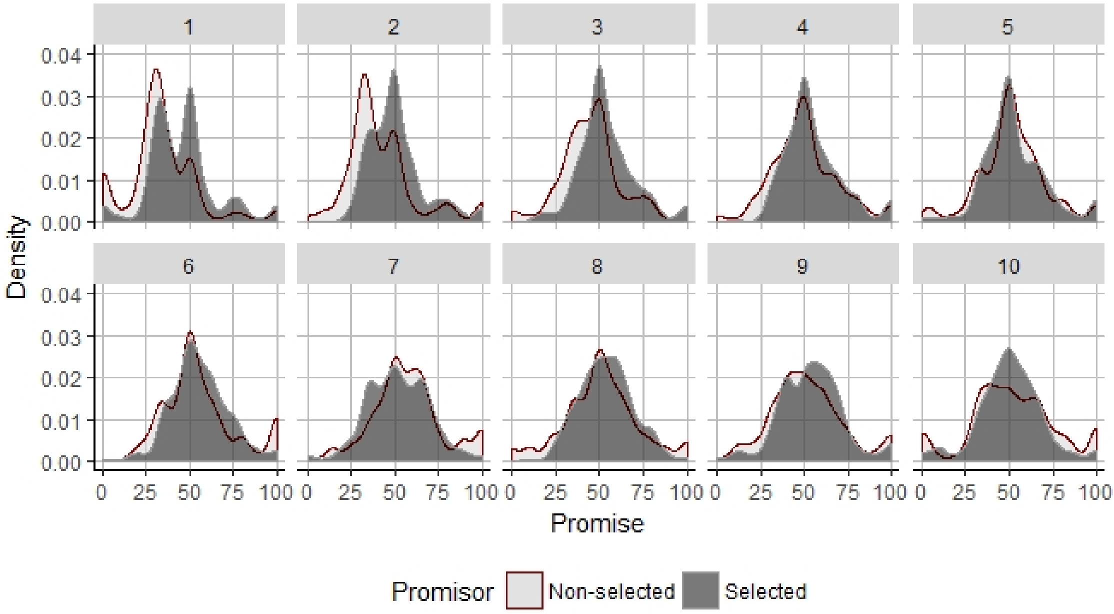
Hypothesis 3 (H3)**.** *Participants give on average more points in the promise game than in the dictator game.*

6 Experimental Results

This section is structured by the hypothesis outlined above. I report p-values based on a t-test if not specified otherwise. Start with the promises participants give over the ten rounds of Part B of the experiment. Figure 2 displays the density of the selected (darker shade) and non-selected (lighter shade) promises by round.

The figure shows that the distribution of promises in the first two rounds of the experiment is double peaked with one peak at 33 and the other at 50 points. Furthermore the distribution of selected and non-selected promises differ, the former have a higher density around 50 whereas the latter have a more density around 33. As a result, the promises that didn't get selected are lower on average. This pattern vanishes after the first three rounds of the experiment, mostly because both selected and non-selected senders focus their promises around 50 points. As a result both distribution approach each other and become single peaked around 50 points and

Figure 2: Density of promises selected and not-selected by round.



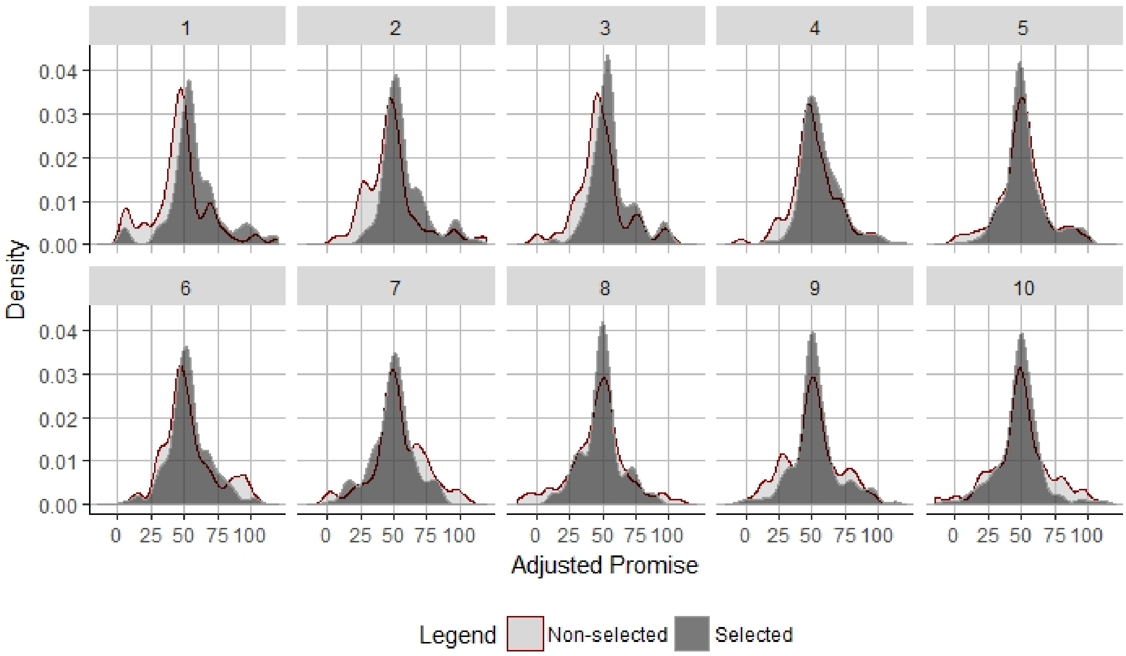
Notes: *Density plot of the promises for each round of the promise game. Separated by selected and non-selected senders. Gaussian kernel with a bandwidth of 4.*

very similar after the first 3 rounds. This suggests that participants - after the first few repetitions - do indeed pool their promises and selected and non-selected senders are indistinguishable by their promises.

While promises are centered around 50 points, there is substantial variation. A potential cause might be that different sessions coordinate on the different promises and do so the more the longer the experiment progresses. This would be in line with the model that allows a range of equilibria. To investigate how much of the variance is due to different promises across sessions, I adjust promises by adding or subtracting a fixed amount to the promises in each round of each session such that the modal promise of each session is precisely at 50 points without changing the variance within a session. The so adjusted promises have to be read as deviations from the modal promise and display intra-session deviations without any variance across sessions.

I define the modal promise as the promise that includes the most other promises in a 5 point radius. To adjust the promises, I add or subtract a fixed amount to

Figure 3: Density of adjusted promises selected and not-selected by round.



Notes: *Density plot of the adjusted promises for each round of the promise game. Separated by selected and non-selected senders. Gaussian kernel with a bandwidth of 4. Promises are adjusted such that the modal promise is 50 for each round of an experimental session. I define the modal promise as the promise that includes the most other promises in a 5 point radius. To adjust the promises, I add or subtract a fixed amount to the promises of each round of a session such that the modal promise is 50. If the modal promise is not unique I adjust such the average mode to 50. The idea is pre-specified in the pre-analysis plan.*

the promises of each round of a session such that the modal promise is 50. If the modal promise is not unique I adjust such the average mode to 50. The idea was pre-specified in the pre-analysis plan.[20] The resulting adjusted promises thus express the deviation from the average modal promise.

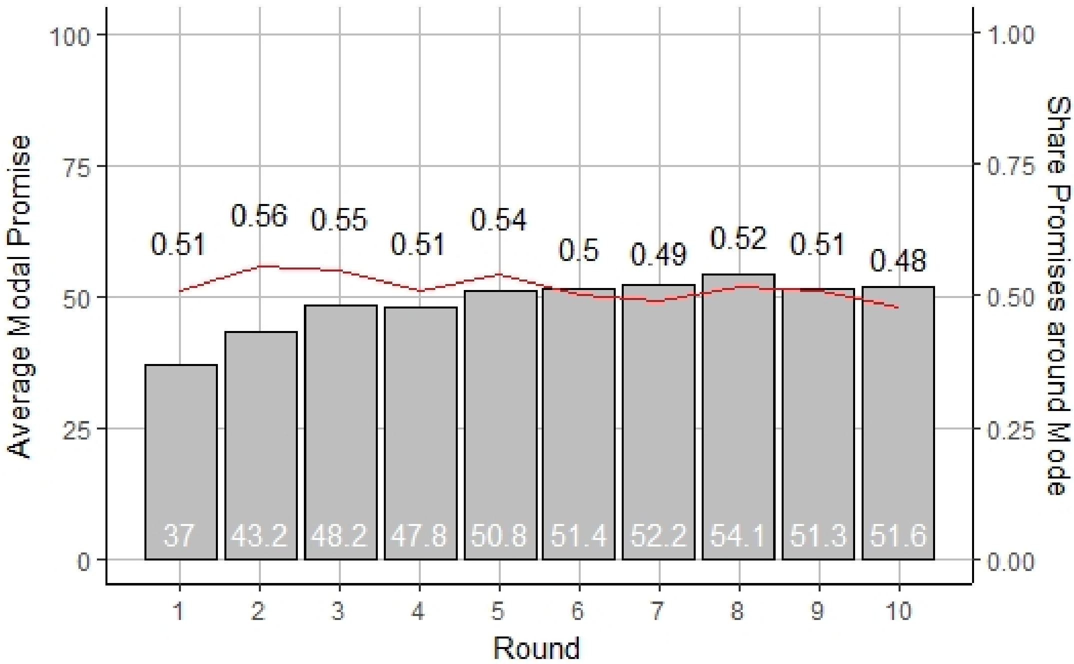
Figure 3 displays the density of selected and non-selected promises with the adjusted promises. Notably the densities are more focused around 50 and fairly stable across all periods. This suggests that a part of the variation in promises observable in Figure 2 is indeed due to different being played in different sessions.

To investigate the share of promises that pool, Figure 4 displays the share of

[20] The plan did not specify how to deal with situations in which more than one modal promise exists. However highest and lowest of these promises are always within a range of 7.5 points and usually within 5 points of each other. The result don't change much if one takes the maximum, average, or minimum of these as the mode to adjust to.

promises around the mode as a red line and the height of the average mode as bars for each repetition of the promise game. The figure demonstrates that the percentage of participants promising at the mode is relatively stable around 50%. In contrast the height of the average modal promise increases over the first rounds from 37 to around 50 points and becomes stable only in the second half of the experiment.

Figure 4: Height of average modal promise and share promising at mode.



Notes: *The line indicates the share of promises that are within 5 points from the modal promise(s) of their session in a particular round. The bars indicate the height of the average modal promise.*

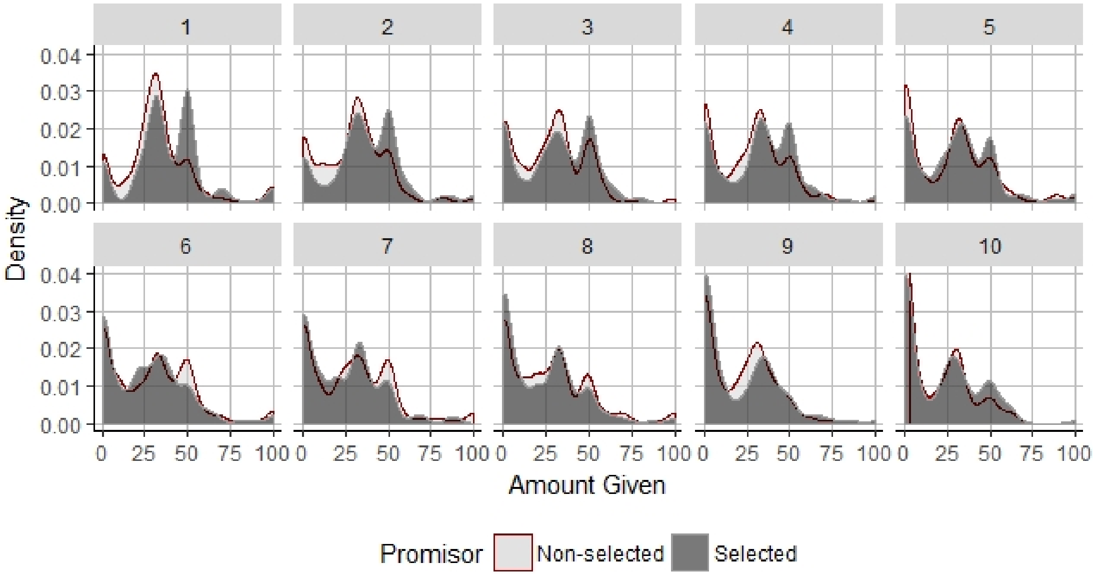
In the pre-analysis plan, I outlined three criteria that would support the idea participants pooling their promises, firstly that about at least 60% make a promise within 5 points of the average mode, that the distribution of promises is single peaked, and that both holds also if only looking at those promises that got selected. As Figure 4 shows the share of participants promising at the mode is slightly smaller, around 50%, however the distribution of promises is single peaked and focuses on that promise. In conclusion I regard this as evidence in favor of the pooling hypothesis of promises yet with some caution regarding the idea that participants only make a single promise. The figure suggests that some participants rather mix around the modal promise however without being more or less likely to be elected. There is a large share of participants that promises close but not exactly at the modal promise.

Finding 1. *The distribution of promises is single peaked around 50 points after the first two rounds of the experiment, with a stable share of participants who promise close to the modal promise of their session. Over the entire duration of the experiment there is a substantial variation promising behavior, which suggests that participants pool around but not precisely at a particular promise. Participants do not pool their promises in the first two rounds, where a share of senders promises around 33 points. These promises are less often selected than promises around 50 points.*

In order to test H2, this part gives provides an overview over the giving-behavior of participants and then proceeds to test whether senders selected differ significantly from those senders not selected in terms of the number of points they send to the receiver. Due to the strategy method I am able to observe the giving decisions of all participants regardless whether they got selected. Notably, the amount participants sent is sharply decreasing over the rounds. In the final round the average amount sent is 52% of the average amount sent in the first round. Figure 5 displays the density the sending decisions for all 10 rounds by selected and not-selection senders. In the first rounds, the distribution of selected promisors has three spikes, one at 50 and one around 33 and a much smaller one at 0 points. This contrasts from the distribution of not selected promisors which has mainly two spikes, a large one at 33 and a smaller one at 0. The difference in distribution reflects the fact that in the first two periods selected promisors often promised 50 points whereas not selected promisors promised 33 more often.

Over the ten periods the number of participants sending 0 points increases steadily while the number at of participants giving 50 decreases. At the same time the difference between selected and not selected senders nearly disappears in line with the hypothesis that participants pool their promises. In period 9 the distributions only show two spikes, one at 0 and one around 33 points and the distributions are very similar for selected and not selected promisors.

Figure 5: Density of amount sent by round and selection.



Notes: *Density plot of the amount sent by participants for each round of the promise game. Separated by selected and non-selected senders. Gaussian kernel with a bandwidth of 4.*

The development of sending decisions, in combination with the distribution of promises displayed above, suggest that over the course of the experiment the norm of living up to ones promise erodes. Some participants however continue to share the amount that leads to an equal ex-post distribution of points.

Table 2 displays the average amount promisors send each by round, and breaks the amounts down into giving by selected and not-selected promisors. The table shows that the overall giving of senders decreases every round of the experiment and particularly so in the first two turns.[21] Turning towards differences between selected and not-selected promisors, selected senders give significantly more (at the 5% level) than the the not-selected senders in rounds 1, 2, and 4. In all other rounds the amounts do not differ significantly. This is in line with the finding that both promises of senders differ by selection status in the first rounds but equalize after four rounds. I conclude that a selection of more generous senders based on the

[21] The average amount sent in round 10 is only 52% of the amount sent in round 1. In the pre-analysis plan I said that the analysis should focus on the first period if the amount decreases by more than 50%. Even-though the decrease is almost as large I think an analysis of all turns is interesting and meaningful.

Table 2: Amount given by round

Round	**senders** all	selected	not-selected	**difference** t-statistic	p-value
1	35.883	38.353	33.007	2.116	0.035
2	32.349	36.490	29.163	3.054	0.002
3	28.970	30.654	27.222	1.443	0.150
4	28.361	31.667	26.549	2.019	0.044
5	27.762	29.536	26.052	1.284	0.200
6	27.135	25.588	28.660	-1.136	0.257
7	25.623	23.575	27.843	-1.659	0.098
8	23.968	22.131	25.869	-1.434	0.153
9	21.565	20.686	21.549	-0.358	0.720
10	18.663	21.046	17.170	1.664	0.097

Notes: The table displays the amount senders give in the promise game by round. The different columns represent all senders or only those who got selected or did not. The final two columns display the test statistic and p-value of a two sided t-test.

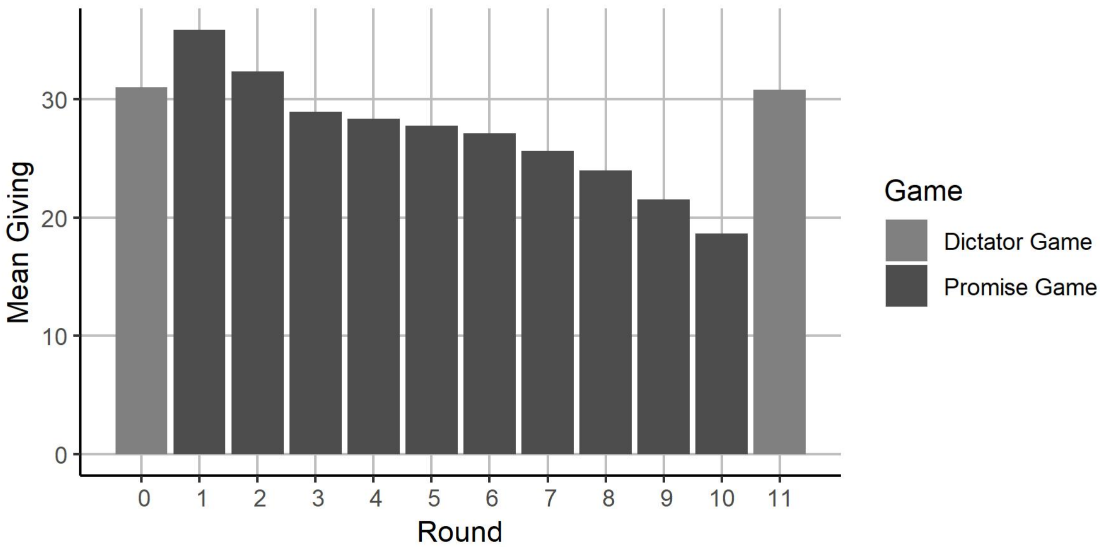
promise is not possible if the senders are experienced. A positive selection of more generous senders occurs with the unexperienced senders in round one and two in contrast.

Finding 2. *After the first four rounds of the experiment, senders selected by the participants do not send significantly more than their not-selected counterpart. This supports the hypothesis no selection of better senders is possible when senders are experienced. However, this is different for unexperienced senders in the first rounds of the experiment.*

To investigate the third prediction (H3), namely that participants send more points to the receiver in the promise game than in the dictator game, we need to compare the giving in the promise game to giving in the dictator game. Figure 6 displays the average points given for each round of the promise game as well as for the dictator game. The first bar represents the average giving of participants who played the dictator game first and the last bar the average amount for participants that played the dictator game last.

The figure shows that participants give around 30 points on average in the dictator game regardless whether they play the game in the beginning or the end of

Figure 6: Density of adjusted promises selected and not-selected by round.



Notes: *Bar graph displaying the mean giving of the participants in each round of the promise game and the dictator game. The dictator game giving is separated by participants who played the dictator game before (Round 0) and after (Round 11) the promise game.*

the experiment. In the first two rounds of the promise game participants give more than in the dictator game, regardless of treatment. However, the amount participants give in the promise game decreases over the repetitions of the promise game and from the fourth round onwards participants give less in the promise game than they do in the dictator game. Perhaps surprisingly the amount participants give in the dictator game in the end of the experiment is much higher than the amount participants give in the last round of the promise game. Moreover, strikingly average giving in the dictator game does not differ at all by the timing of the game (p-value 0.9275). This pattern resembles the restart effect that previous literature finds in a repeated public good game.[22]

The difference between the development of the promise game and dictator game behavior, makes a comparison of behavior in later rounds of the promise game to the one-shot dictator game difficult. Factors that might not be present in a one-shot game appear to change over the repetitions of the promise game - as for example observing participants break their promise might induce negative reciprocity or de-

[22] See Andreoni (1988).

teriorating norms of generosity.[23] For that reason comparing the average giving over all rounds of the promise game to the one-shot dictator game seems likely to be uninformative and or at least biased. Hence, in the preferred specification I compare the giving in the first round of the promise game to the dictator game in Table 3. The table displays both a comparisons for completeness and due to pre-registration.[24] Note that this comparison includes the giving decision of all participants regardless whether they got selected.

Table 3: Comparison promise and dictator game givings

	Mean sending		**Difference**	
Round	Promise Game	Dictator Game	t-statistic	p-value
all	27.141	30.967	-2.086	0.039
1	35.745	30.967	2.417	0.017

Notes: Comparison of average giving in promise and dictator game. Includes both selected and non-selected senders. The two last columns display the result of a two-sided and paired t-test.

Averaging over all rounds the participants give significantly less in the promise game compared to the dictator game (3.8 points). In contrast the first round of the promise game participants give significantly more (4.8 points) in the promise game instead. As argued I regard the latter comparison as evidence that the promise game, at least initially, does increase giving as predicted.

Table 4 provides an additional test of the mechanism predicted by the model. The model predicts that quality provision increases due to the higher promises under promise competition. In other words, higher promises make honest participants give more than they had in the dictator game.[25] The table displays a regression of the difference in giving of round t and $t-1$ on the according difference in promising. In other words, it investigates whether, given an overall decrease of giving, a change in promise, on average, is correlated with an according change in giving. As the table shows, a change of an individuals promise by 10 points corresponds to a 2.65 point

[23] Figure 9 in the appendix suggests that in particular those participants who give 50 points in early rounds are those who decrease their giving subsequently.

[24] Both comparisons have been pre-specified in the pre-analysis plan.

[25] Note that this test was not specified hence should be regarded with as exploratory analysis.

Table 4: Regression of change in giving on change in promising

	$\Delta_{t/t-1}$ Giving		
	(1)	(2)	(3)
$\Delta_{t/t-1}$ Promise	0.265***	0.273***	0.282***
	(0.065)	(0.065)	(0.069)
($\Delta_{t/t-1}$ Promise) sqrt			−0.001
			(0.001)
Constant	−2.179***		
	(0.207)		
Individual FE			X
Round FE		X	X
N	1,377	1,377	1,377
R^2	0.075	0.091	0.115

Notes: Regression of difference of giving in round t to $t-1$ on difference of promise. Round and individual fixed effects. Clustered standard errors (individual level) in parenthesis. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

change in the amount of points the participant sends to the receiver. At the same time the negative intercept reflects the overall decrease of giving with on average 2.2 points per round. Significance and size of the coefficient hold up when including round and individual fixed effects and an additional squared functional form. In a further investigation, Figure 1 and Table 6 both in the Appendix demonstrate that the effect observable in this table largely stems from those participants who keep their initial promise in the first round of the experiment. This group circumvents 2/3 of all participants.

Finding 3. *I find that participants send more points to the receivers in first round of the promise game than in the dictator game. In addition a change in the promise is correlated with an according change in giving in all rounds of the experiment. I argue that later rounds of the promise game are not comparable to the dictator game, due to a decrease in giving and a restart effect between the games similar to repeated public good games.*

6.1 Who decreases giving in the promise game?

To provide a better understanding of the data, this section sheds light on the decrease in points participants send to the receivers over repetitions of the promise game.[26]

Table 5: Regression of giving in t on giving and role in $t-1$.

	give	
	(1)	(2)
Giving $t-1$	0.678***	0.677***
	(0.031)	(0.031)
Giving $t-1$ Select Sender	0.111***	0.087**
	(0.019)	(0.033)
Previous Role: Receiver		−4.648*
		(1.912)
Previous Role: Selected Sender		1.294
		(1.504)
Giving Selected Sender * Receiver		0.111*
		(0.049)
Giving Selected Sender * Selected Sender		−0.032
		(0.041)
Round FE	X	X
N	1,377	1,377
R^2	0.786	0.788

Notes: Regression of giving on giving in previous round. Individual fixed effects. Clustered standard errors on individual level in parenthesis. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

To that purpose, Table 5 displays a regression of giving in period t on giving and the role of the individual participant in period $t-1$. The regression includes dummies for round fixed effects and clusters standard errors on the individual level. The high and significant coefficient of a participant's own giving in $t-1$ underlines that participants' decisions are relatively steady over different periods. In addition also the amount given by the selected sender in round $t-1$ has a significantly positive coefficient. This indicates that participants - on average - react to the giving of others and adjust their giving in the same direction. Specification (2) includes an indicator for the role a participant had in round $t-1$ as well as an interaction with the giving

[26] This section has not been pre-registered and should therefore be seen as exploratory.

of the selected sender in that round. Interestingly, the interaction between being the receiver and the giving of the selected sender is significantly positive.

7 Discussion and Conclusion

This paper studies non-binding promise competition. It predicts that such competition leads seller to pool their promises at a quality equal to or above the quality a good seller-type would have naturally provided. Having made such promise leads sellers with a cost of promise-breaking, who promised above their natural quality provision to increase the same to avoid incurring this cost. By that mechanism, promise competition increases average service quality even though it does not allow buyers to learn the type of seller they are facing.

Here I would like to highlight one key difference to previous models that investigate political promise competition namely Corazzini et al. (2014) and Fehrler et al. (2018): This model includes two dimensions heterogeneity of promisors. The analysis of the model in this paper shows that the second dimension of unobserved heterogeneity, here intrinsic motivation, is relevant to understand the effects of promise competition. In particular in the model presented here, the natural quality provision of motivated sellers determines the height of promises made in equilibrium and the difference in the natural action between motivated and unmotivated types can be relevant to understand how much participants increase their quality provision due to promise competition.

The paper tests the model's predictions in a laboratory experiment. In the experiment, participants do pool their promises around an amount 50 points[27] such that selected senders are indistinguishable from the average sender after a few repetitions of the game. That means a selection of better promises is not possible, in line with the predictions of the model. However, participants need a few rounds of learning to establish this behavior. Particularly, in the first 2 rounds of the experiment the distribution of promises is double peaked (peaks at 33 and 50) and so is

[27] The number of points a sender gives to the receiver is multiplied by two. Accordingly giving 33/100 points leads to an ex-post equal split, while giving 50 points gives the receiver twice as many points as the sender.

the distribution of points given. This suggests that the models results hold mostly in environments in which promisors have a certain degree of experience wheras in a setting with unexperienced promisors, the promise might very well be informative about a promisor's type.

Supporting the second prediction of the model, senders in the first round of the promise game give larger sums of points to the receivers than in a one-shot dictator game. However, I document a strong decline of the giving in the promise game which makes a comparison of later rounds of the promise game to the dictator game difficult. Over the 10 repetitions of the game, the amount of points participants give to the receivers decreases by 48% in the promise game. Previous literature documents a similar decline in public good games.[28] To my knowledge there is no literature documenting decreasing giving behavior in repeated dictator games, which would make it interesting to compare the decrease in the repeated promise game to a repeated dictator game.

Table 5 investigates the decreasing giving in the promise game and suggests that a mechanism similar to conditional cooperation of participants documented in a public good game[29] is at work in this experiment. A hypothesis to explain the results of this table is that a mechanism similar to conditional cooperation is at work in the repeated promise game. In this game participants do not cooperate but might follow a norm or moral rule that prescribes participants to give points to the receiver and/or keep their promise. Some participants never or always adhere to that rule, but many participants conditionally adhere to it. Hence, if participants observe others not to adhere to the rule they in turn will also decrease their giving. Furthermore participants appear to reciprocate the action of a selected sender to stronger degree when the action affects them directly (as the receiver) as compared to a situation where they only observe the action (as a non-selected sender). Testing this hypothesis might be an interesting avenue for future research.

In addition to the decrease of giving in the promise game, I also find a restart effect in a subsequent dictator game. In spite of the fact that participants play with

[28] For example Isaac, Walker, and Thomas (1984) and Plott, Isaac, and McCue (1985).

[29] For example by Keser and Van Winden (2000); Brandts and Schram (2001); Fischbacher, Gächter, and Fehr (2001); Croson (2007); Fischbacher and Gächter (2010).

the same pool of participants, they give much in a dictator game than in the last round of the promise game, and in fact giving in the dictator game does not differ by the timing of the game (before / after the promise game). This appears similar to the increase of contributions after restarting a public good game, first documented by Andreoni (1988). In the light of this findings, it appears to be an interesting question whether giving in a dictator game decreases at the same rate as in the promise game. The observability of broken promises and the fact that one sender does not get selected might trigger negative reciprocity to a larger extend than in a repeated dictator game.

Furthermore an change in promises is positively correlated with an according change giving in over all turns of the experiment.

With regard to the empirical findings of this study, Corazzini et al. (2014) find that introducing elections and promises in combination increases giving in comparison to a setting in which either one is present. In contrast the study of Casella et al. (2018) concludes that competition leads to higher offers without affecting transfers. Whereas this study can confirm the former (higher offers) it also finds increased giving. Yet the setting they study is different in some important respects, i.e. their experiment has an outside option, and no multiplier. In particular the outside option of the sender might discourage following through on.

References

Abeler, J., A. Becker, and A. Falk (2014): "Representative Evidence on Lying Costs," *Journal of Public Economics*, 113, 96–104.

Abeler, J., D. Nosenzo, and C. Raymond (2018): "Preferences for Truth-Telling," *Econometrica*, forthcoming.

Andreoni, J. (1988): "Why Free Ride?: Strategies and Learning in Public Goods Experiments," *Journal of Public Economics*, 37, 291–304.

Andreoni, J. and B. D. Bernheim (2009): "Social Image and the 50–50 Norm: A Theoretical and Experimental Analysis of Audience Effects," *Econometrica*, 77, 1607–1636.

Bénabou, R. and J. Tirole (2006): "Incentives and Prosocial Behavior," *American Economic Review*, 96, 1652–1678.

Bernheim, B. D. and N. Kartik (2014): "Candidates, Character, and Corruption," *American Economic Journal: Microeconomics*, 6, 205–46.

Born, A., P. van Eck, and M. Johannesson (2018): "An Experimental Investigation of Election Promises," *Political Psychology*, 39, 685–705.

Brandts, J. and A. Schram (2001): "Cooperation and Noise in Public Goods Experiments: Applying the Contribution Function Approach," *Journal of Public Economics*, 79, 399–427.

Callander, S. and S. Wilkie (2007): "Lies, Damned Lies, and Political Campaigns," *Games and Economic Behavior*, 60, 262–286.

Casella, A., N. Kartik, L. Sanchez, and S. Turban (2018): "Communication in Context: Interpreting Promises in an Experiment on Competition and Trust," *Proceedings of the National Academy of Sciences*.

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

Cho, I.-K. and D. M. Kreps (1987): "Signaling Games and Stable Equilibria," *The Quarterly Journal of Economics*, 102, 179–221.

Corazzini, L., S. Kube, M. A. Maréchal, and A. Nicolò (2014): "Elections and Deceptions: An Experimental Study on the Behavioral Effects of Democracy," *American Journal of Political Science*, 58, 579–592.

Croson, R. T. (2007): "Theories of Commitment, Altruism and Reciprocity: Evidence from Linear Public Good Games," *Economic Inquiry*, 45, 199–216.

Ellingsen, T. and M. Johannesson (2004): "Promises, Threats and Fairness," *The Economic Journal*, 114, 397–420.

Fehrler, S., U. Fischbacher, and M. T. Schneider (2018): *Honesty and Self-Selection into Cheap Talk*.

Feltovich, N. and F. Giovannoni (2015): "Selection vs. Accountability: An Experimental Investigation of Campaign Promises in a Moral-Hazard Environment," *Journal of Public Economics*, 126, 39–51.

Fischbacher, U. and F. Föllmi-Heusi (2013): "Lies in Disguise—an Experimental Study on Cheating," *Journal of the European Economic Association*, 11, 525–547.

Fischbacher, U. and S. Gächter (2010): "Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments," *The American Economic Review*, 100, 541–556.

Fischbacher, U., S. Gächter, and E. Fehr (2001): "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment," *Economics Letters*, 71, 397–404.

Frankel, A. and N. Kartik (2017): "Muddled Information," *Journal of Political Economy*, forthcoming.

Geng, H., A. R. Weiss, and I. Wolff (2011): "The Limited Power of Voting to Limit Power," *Journal of Public Economic Theory*, 13, 695–719.

Gibson, R., C. Tanner, and A. F. Wagner (2013): "Preferences for Truthfulness: Heterogeneity among and within Individuals," *American Economic Review*, 103, 532–48.

Gneezy, U. (2005): "Deception: The Role of Consequences," *The American Economic Review*, 95, 384–394.

Gneezy, U., A. Kajackaite, and J. Sobel (2018): "Lying Aversion and the Size of the Lie," *American Economic Review*, forthcoming, available at SSRN: https://ssrn.com/abstract=2852055.

Gneezy, U., B. Rockenbach, and M. Serra-Garcia (2013): "Measuring Lying Aversion," *Journal of Economic Behavior & Organization*, 93, 293–300.

Isaac, R. M., J. M. Walker, and S. H. Thomas (1984): "Divergent Evidence on Free Riding: An Experimental Examination of Possible Explanations," *Public choice*, 43, 113–149.

Jin, G. Z. and A. Kato (2006): "Price, Quality, and Reputation: Evidence from an Online Field Experiment," *The RAND Journal of Economics*, 37, 983–1005.

Keser, C. and F. Van Winden (2000): "Conditional Cooperation and Voluntary Contributions to Public Goods," *Scandinavian Journal of Economics*, 102, 23–39.

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

Mailath, G. J., M. Okuno-Fujiwara, and A. Postlewaite (1993): "Belief-Based Refinements in Signalling Games," *Journal of Economic Theory*, 60, 241–276.

Mazar, N., O. Amir, and D. Ariely (2008): "The Dishonesty of Honest People: A Theory of Self-Concept Maintenance," *Journal of Marketing Research*, 45, 633–644.

Plott, C. R., R. M. Isaac, and K. McCue (1985): "Public Goods Provision in an Experimental Environment," *Journal of Public Economics*, 26, 51–74.

Sutter, M. (2009): "Deception through Telling the Truth?! Experimental Evidence from Individuals and Teams," *The Economic Journal*, 119, 47–60.

Vanberg, C. (2008): "Why Do People Keep Their Promises? An Experimental Test of Two Explanations 1," *Econometrica*, 76, 1467–1480.

A Extensions to the model.

A.1 Existence of equilibria with four types

The main analysis is based on a the type-space $\mathrm{T}_{alt} = \{(0,0), (\overline{\alpha}, 0), (\overline{\alpha}, \overline{\rho})\}$. In this section I add type $\tau_d = (\overline{\alpha}, 0)$ to the type-space. Call this type *deceptive.* This section shows that the described pooling equilibria exist even with four types. However,the upper bound of the range p^{max} may be different. We maintain the assumption that each type is drawn with equal probability.

Recall that we define $p^{max} = \min\{\tilde{p}, \hat{p}\}$. As τ_d is equivalent to τ_l with regard to marginal utilities, $\tilde{p}$ remains unchanged. However, $\hat{p}$ now is such that,

$$\frac{\phi_{\tau_g}\ x^*(\hat{p}, \tau_g) + \phi_{\tau_d}\ \overline{x}^0}{\phi_{\tau_g} + \phi_{\tau_l} + \phi_{\tau_d}} \ = \ x^*(\hat{p}, \tau_h).$$

Observe that Assumption 3 implies that,

$$\frac{\phi_{\tau_g} + \phi_{\tau_h}}{\phi_{\tau_g} + \phi_{\tau_l} + \phi_{\tau_d}} \overline{x}^0 \ > \ x^*(\overline{x}^0, \tau_h).$$

It follows that $\hat{p} > \overline{x}^0$.

Proposition 10. *The class of pooling Equilibria that passes Criterion D1 with type-space T_{base}, also exists with type space T_{alt}, albeit the upper bound might be different.*

Proof. Consider any pooling equilibrium described in Proposition (6). Both τ_d and τ_l face no cost of promise breaking, hence are indifferent among all promises for an equal chance of selection. That means both types deviate to any off-equilibrium promise for exactly the same set of mixed best-replies by the promisee, $D(\tau_d, p) = D(\tau_l, p)$ for all p. For that reason the type does not eliminate any beliefs that D1 permits in type-space T. Accordingly the set of beliefs supporting any of these pooling equilibria still exist and the following beliefs follow Bayes' rule and pass Criterion

D1,

$$\mu(\tau_l|p) = \begin{cases} \phi_{\tau_l} & \text{if } p = p^*; \\ 1 & \text{if } p > p^*; \\ 0 & \text{if } p < p^*, \end{cases}$$

$$\mu(\tau_h|p) = \begin{cases} \phi_{\tau_h} & \text{if } p = p^*; \\ 1 & \text{if } p < p^*; \\ 0 & \text{if } p > p^*, \end{cases}$$

$$\mu(\tau_d|p) = \begin{cases} \phi_{\tau_d} & \text{if } p = p^*; \\ 0 & \text{if } p < p^*; \\ 0 & \text{if } p > p^*, \end{cases}$$

$$\mu(\tau_g|p) = \; 1 - \mu(\tau_h|p) - \mu(0,0|p) - \mu(\tau_d|p).$$

The promisee-strategy is optimal under the new type space as long as $p \leq \hat{p}$. Given the promisee-strategy, the promisor strategy is optimal,too. □

A.2 Discussion of Assumption 3

In the model, the existence of equilibria rests partly on Assumption 3. The assumption ensures that the expected quality from a promise that both τ_g and τ_l make is higher than the quality from any lower promise exclusively made by type τ_h. If that is not the case, thus the assumption is violated, and the likelihood of type τ_l is sufficiently large, all pooling equilibria are eliminated by Criterion D1 since promisee prefer to elect a lower promise that is made by τ_h exclusively under these circumstances. The following proposition establishes this.

Proposition 11. *Consider a promise $p \geq \overline{x}^0$ such that $\frac{\phi_{\tau_g}}{\phi_{\tau_g}+\phi_{\tau_l}} x^*(p,\tau_g) < x^*(p,0,\overline{\rho})$. There is no equilibrium that survives Criterion D1 in which this quality is promised with positive probability.*

Proof. We prove this by contradiction. Suppose there exists an equilibrium that survives D1 in which a promise $p \geq \overline{x}^0$ such that $\frac{\phi_{\tau_g}}{\phi_{\tau_g}+\phi_{\tau_l}} x^*(p,\tau_g) < x^*(p,0,\overline{\rho})$, is made with positive probability. Proposition 6 and 7 establish that then this must be a pooling equilibrium in which $s_\tau(p) = 1$ for all τ in T. In such equilibrium type τ_h and τ_g have a larger ex-post utility if they get selected with $p' = p - \varepsilon$ for some $\varepsilon > 0$. In contrast type τ_l receives exactly the same ex-post utility with either promise. Hence, Criterion D1 eliminates type τ_l for promise ρ' and the beliefs μ to be accordingly. Regardless which other type survives this means $E[x|p', \mu] > E[x|p, \mu]$ such that the promisor prefers p' to p. Accordingly p cannot be a promise made in equilibrium which contradicts the assumption. $\square$

That means that relaxing Assumption 3 reduces the set of pooling equilibria that survive criterion D1 up to the point that no pooling equilibrium remains. In other words, if too many bad types are in a market that can lead to a break-down of all equilibria.

Similarly equilibria could break down if the fixed cost of promise-breaking is high enough such that τ_h keeps all potential equilibrium promises. As this type gains the most from downward deviation to a lower promise, D1 requires the promisee to belief such promise comes from τ_h but that eliminates any equilibrium as τ_h keeps that promise.

While it is conceivable that a too large amount of bad types destroys any equilibrium, it imight seem undesirable that the existence of equlibria depends on the fact that the honest type has a low enough cost of promise-breaking. In the following I address this concern by demonstrating that introducing an additional type with an intermediate level of motivation to the model allows for equilibria in which the intermediately motivated and honest type keeps the equilibrium promise while the unmotivated and honest type breaks it.

Assume a type $(\dot{\alpha}, \overline{\rho})$ such that $0 < \dot{\alpha} < \overline{\alpha}$ and assume that $\dot{\alpha}$ and ν are large enough such that $x^*(p, \dot{\alpha}, \overline{\rho})$ for some $p \geq \overline{x}^0$ while at the same time the natural action of that type is lower than $x^0_{\dot{\alpha},\overline{\rho}} < x^0_{\tau_g}$.

Proposition 12. *For $\dot{\alpha}$ large enough, there exist beliefs that conform with Criterion D1 such that for p^* in $[\overline{x}^0, p^{max}]$ the following strategies form a perfect Bayesian* *(pooling) equilibrium,*

$$s_\tau(p^*) = 1 \text{ for all } t,$$

$$a^*(p^1,p^2) = \begin{cases} 1 & \text{if } p^1 = p^* \text{ and } p^2 \neq p^*, \\ 0 & \text{if } p^1 \neq p^* \text{ and } p^2 = p^*, \\ 0.5 & \text{otherwise}, \end{cases}$$

$$x^*(p,\tau) \text{ as defined in Equation (2) for all types } \tau,$$

where $p^{max} = \min\{\tilde{p}, \hat{p}\}$.

Proof. Once we show that D1 deletes the belief that type $(\dot{\alpha}, \overline{\rho})$ for any deviating from the equilibrium, the proof is equivalent to the proof of Proposition 6. To see that consider any equilibrium promise p^* and firstly consider any upward deviation. Type $(\dot{\alpha}, \overline{\rho})$ is clearly loosing ex-post utility from making a higher promise, whereas τ_l is not, hence the former is deleted by D1 for any promise $p > p^*$. Consider any deviation to a lower promise. Here $(\dot{\alpha}, \overline{\rho})$ gains utility from getting selected with a lower promise, as does τ_h. Whereas the former is keeping her promise the latter is breaking it. It is straightforward to see τ_l gains more from a downward deviation if $\dot{\alpha}$ is such that the marginal utility from motivation off-sets the higher loss of utility from keeping a promise as compared to breaking it optimally. Formally, for $\dot{\alpha} f'_x(x^*) > u'_p(p^*, p^*, \tau_h) - u'_p(x^*, p^*, \tau_h)$ D1 deletes $(\dot{\alpha}, \overline{\rho})$ and the rest of the proof is identical to the proof of Proposition 6.

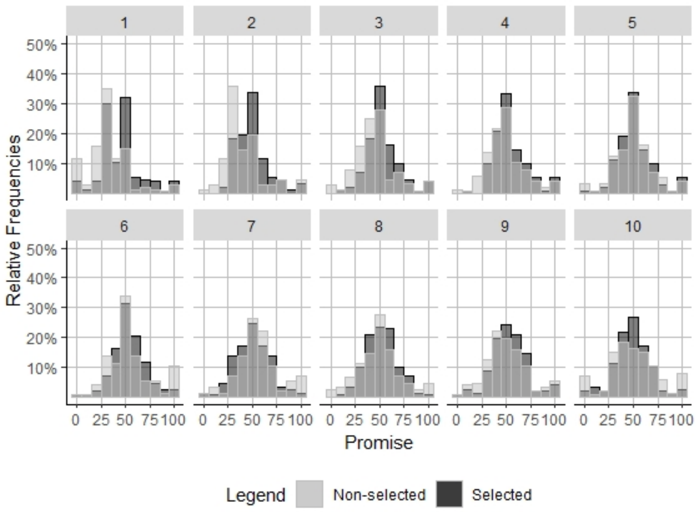
We begin by showing that there exist believes permissible by D1 to support this equilibrium. Let p^* be any equilibrium promise in $[\overline{x}^0, p^{max}]$. Consider a deviation to any promise $p' < p^*$, by choice of p^{max}, type τ_h is the type that gains highest expected utility from deviating to p'. Accordingly, D1 deletes all other types, and promisee beliefs are $\mu(\tau_h|p') = 1$ for all $p' < p^*$ which does not support deviation. Secondly, consider any promise larger than the equilibrium promise $p' > p^*$. Both τ_g and τ_h loose ex-post utility from increasing their promise, while τ_l does not. Again D1 deletes these types for promisee beliefs and only the belief $\mu(\tau_l|p')$ is permissible. Hence beliefs permissible by D1 exist to support the equilibria. Finally, given these beliefs a^* is optimal and so is $s_\tau(p^*) = 1$, which concludes the proof. $\square$

B Supporting Figures and Tables

B.1 Histograms of promises of selected and non-selected senders

To complement the density plots shown in the main part of the paper, the following two figures display show histograms of the promises in each round divided by selection status. Figure 7 displays promises and Figure 8 adjusted promises. The pictures confirm the findings from the density plots above. Promises are centered around 50 and single peaked generally, except for the first two rounds in which there are two peaks, one around 33 and one at 50 points. Notably the adjusted promises are more densely centered around 50.

Figure 7: Density of promises selected and not-selected by round.

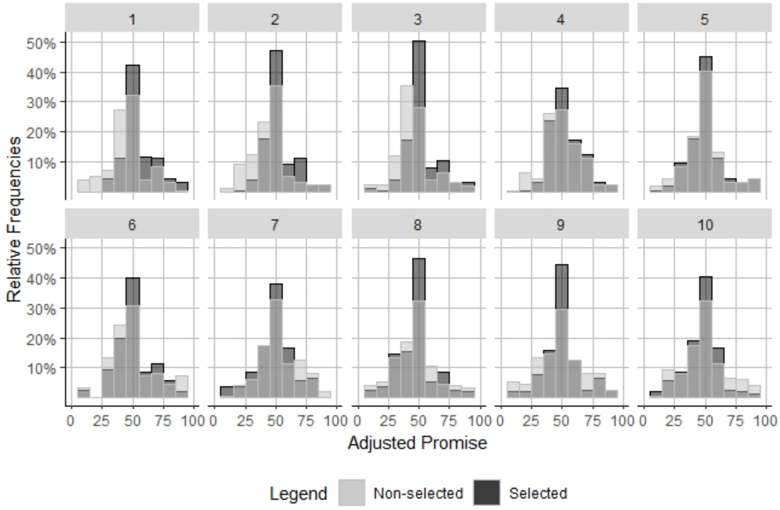


Notes: *Histogram of promises for each round of the promise game. Separated by selected and non-selected senders.*

B.2 Comparison of promises and giving

This Figure provides a comparison of the density of the promises given and the giving in the promise game. Notably densities are quite similar in the first round of the experiment, but over the rounds of the experiment more participants break their promise and give close to nothing to the receiver. Beginning from round 4

Figure 8: Density of adjusted promises selected and not-selected by round.



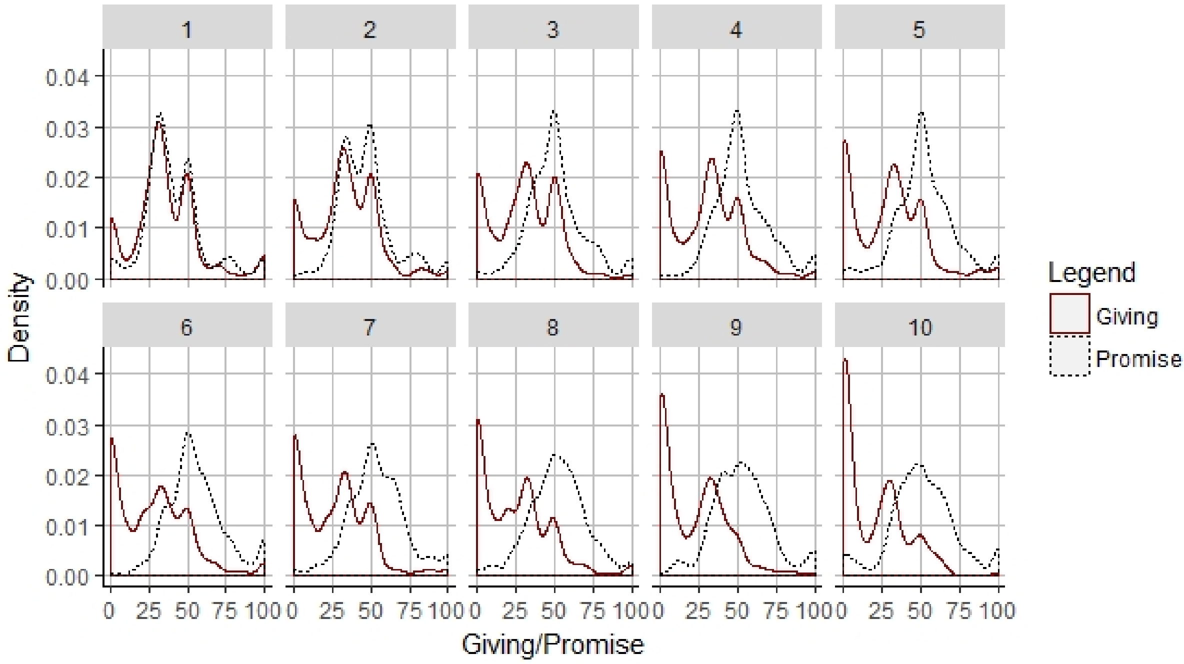
Notes: *Histogram of adjusted promises for each round of the promise game. Separated by selected and non-selected senders.Promises are adjusted such that the modal promise is 50 for each round of an experimental session. I define the modal promise as the promise that includes the most other promises in a 5 point radius. To adjust the promises, I add or subtract a fixed amount to the promises of each round of a session such that the modal promise is 50. If the modal promise is not unique I adjust such the average mode to 50. The idea is pre-specified in the pre-analysis plan.*

participants who give 50 in earlier rounds seem to drive the effect. At the same time the distribution of promises does not change much. This suggests that the effects stems from participants who become less inclined to life up to their promise instead of changing promises. Moreover it is noteworthy that there is a fraction of participants who give around 33 points throughout the entire experiment, regardless of the distribution of promises. This could support the idea that these participants are 'natural givers' who give even if promises get increasingly less followed upon.

B.3 Initial promise-keepers drive effect in Table 4

This section shows that there is heterogeneity in how a change in a participant's promise correlates with a subsequent change in giving. In particular participants can be divided with respect to their decision whether to keep their promise in the first round of the experiment or not. Figure 10 displays a scatter plot of the difference

Figure 9: Density of promises and giving by round.

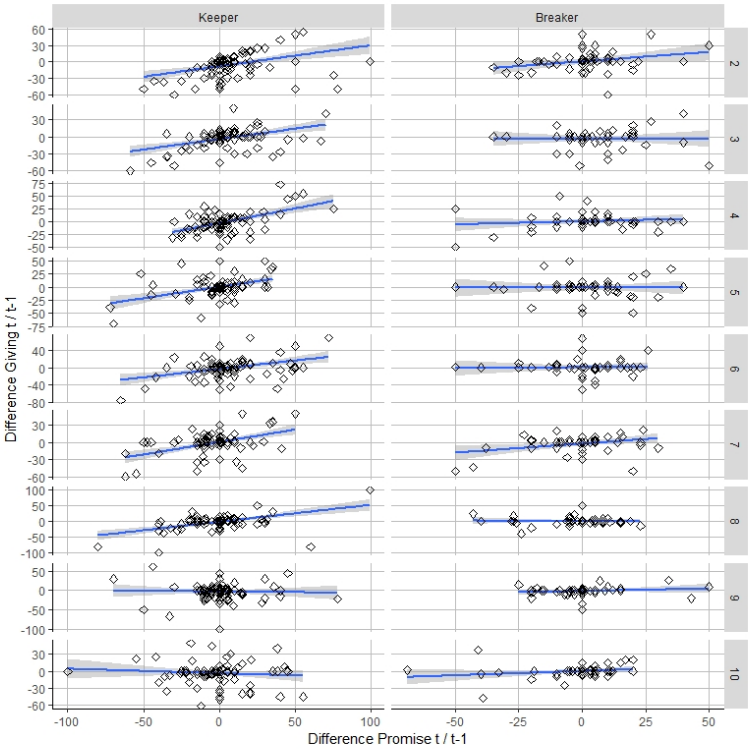


Notes: *Bar graph displaying the mean giving of the participants in each round of the promise game and the dictator game. The dictator game giving is separated by participants who played the dictator game before (Round 0) and after (Round 11) the promise game.*

in promise between period t and $t-1$ on the difference on giving for each period, divided by participants who kept their promise in the first round of the promise game and those who did not. For the group of initial promise keepers a simple linear fit reveals a positive correlation in all but the two last periods, whereas a change in promising does not correspond to a change in giving for the inital promise breakers. To confirm the visual investigation, Table 6 investigates the same in an OLS regression identical to specification (3) in Table 4 but separated for the two groups of participants. The regression controls for individual and round fixed effects and also allows for clustered standard errors on the individual level.

The table shows that the correlation between a change in giving and promising is close to zero for initial promise breakers and statistically insignificant. In contrast the correlation is highly significant for initial promise keepers, on average a 10 point increase in the promise corresponds to a 3.5 point increase in giving.

Figure 10: Change in promise and giving $t/t-1$ by initial promise-keeping.



Notes: *Scatter plot of the difference in the promise between round t and $t-1$ on the difference in amount given between round t and $t-1$. Separated whether participant kept or broke the promise in round 1.*

B.4 Promise Breaking

Table 2 demonstrates that, with an exception for round 1,2, 4 of the experiment, participants do not select senders who give significantly more (or less) to receivers than the senders. This is in line with the finding that in later rounds of the experiment selected and non-selected senders do not differ in their promises. This section investigates whether the same also holds fr promise-breaking. The equilibria of the model predict this, even though this hasn't been the focus of the analysis.[30]

Figure 11 displays the share of broken promises by round. In accordance with the decline in giving over the rounds of the experiment while promises first increase and then remain stable, the share of broken promises increases from 0.346 in the first round to 0.752 in the last round. The increase in the share of broken promise is largest between the first 3 rounds, which are the those rounds in which many

[30] This has been pre-specified as exploratory analysis before the experiment.

Table 6: Regression of change in giving on change in promising, by fulfillment of promise in round 1

	$\Delta_{t/t-1}$ Giving	
	keepers	breakers
$\Delta_{t/t-1}$ Promise	0.345***	0.110
	(0.082)	(0.090)
($\Delta_{t/t-1}$ Promise) sqrt	−0.002	−0.002
	(0.001)	(0.002)
Individual FE	X	X
Round FE	X	X
N	900	477
R^2	0.157	0.050

Notes: Regression of difference of giving in round t to $t-1$ on difference of promise. Regression (1) uses participants that keep their promise in round 1. Regression (2) uses participants that break their promise in round 1. Round and individual fixed effects. Clustered standard errors (individual level) in parenthesis.
* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

participants increase their promise (see Figure 9 in the appendix for a comparison).

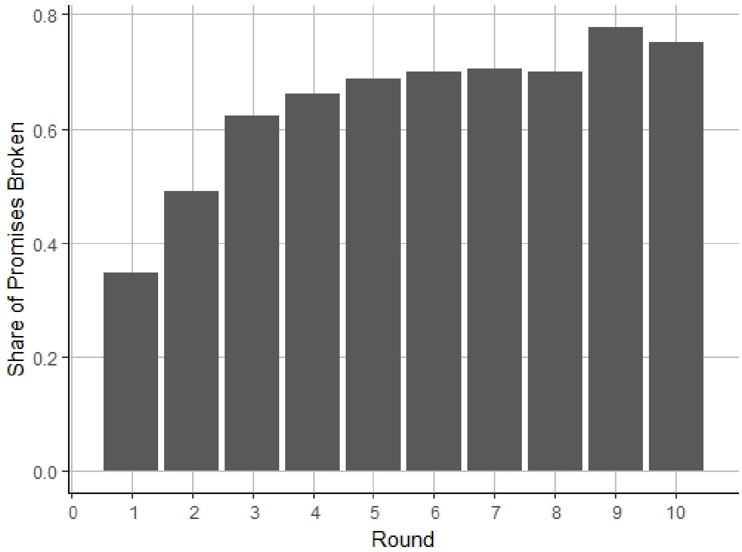
To test whether selected senders differ from the rest, Table 7 displays the shares of broken promises by selection status and round. The table reports the results of a two-sided test of proportions (chi-squared test) testing whether the shares differ by select status. On average selected participants do not break their promises with a different frequency than the other senders. This holds with an exception is period 8, in which selected senders break the promise significantly more often.

B.5 Attribution of Giving Increase

The model predicts that honest promisors who give litte in the dictator game drive the higher giving in the promise game. In the pre-analysis I specify that I run the following regression as exploratory analysis to investigate this,

$$\Delta_i = \alpha + \beta_i d_i + \varepsilon_i,$$

Figure 11: Share of broken promises by round.



Notes: *Bar graph of the share of broken promises in each round.*

Table 7: Share of broken promises by selection status and round

Round	**senders** all	selected	not-selected	Chi-2 test p-value
1	0.346	0.359	0.294	0.273
2	0.490	0.523	0.477	0.493
3	0.621	0.667	0.588	0.193
4	0.660	0.654	0.654	1
5	0.686	0.693	0.686	1
6	0.699	0.739	0.660	0.170
7	0.706	0.719	0.706	0.899
8	0.699	0.778	0.641	0.012
9	0.778	0.797	0.784	0.888
10	0.752	0.725	0.771	0.429

Notes: The table displays the share of senders who break their promise by round of the promise game. The last column displays the p-value of a test of proportions comparing the share of broken promises by selected and not-selected senders.

where $\Delta_i = s_i - d_i$ is the difference between the amount given in the dictator game d_i and the average amount given over all rounds of the promise game s_i of participant i, ε_i is the robust error term.

The following table gives an overview over the results of this regression. Albeit the result confirms the prediction of the model, I caution that, in light of the high R^2 term, the result is likely to be an artifact of regressing the difference of giving on the dictator game giving as the difference contains the dictator game giving. As this part of the analysis is pre-specified, I still present the table below.

Table 8: Regression difference in giving on dictator game giving

	Diff. Giving
Giving Dictator game	-0.715^{***}
	(0.082)
Constant	18.322^{***}
	(2.557)
N	153
R^2	0.499

Notes: Regression of difference in giving between promise and dictator game on giving in the dictator game. Robust standard errors in parenthesis.
* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.