Broken Contracts and Hidden Partnerships: Theory and Experiment

Jingnan Chen [*]   Daniel Houser [†]

February 2, 2015

Abstract

Previous research indicates that unenforceable informal contracts (or commitments) promote trust and reciprocity. Nonetheless, while such contracts may benefit existing exchange, in dynamic environments they may also hinder ones willingness to explore newly emerging Pareto efficient opportunities. This issue arises in both business and social contexts, and includes industry non-compete agreements as well as personal relationship commitment decisions. We report data from an experiment using a novel three-person trust game where, in different treatments, different players are able to communicate with each other. We find that when, between the point of commitment and the point of decision, no new information is received regarding the expected value of commitment, then people overwhelmingly decide in accord with their informal contracts and avoid exploring potentially Pareto improving opportunities. However, when new information arises that reduces the relative value of commitment, and when this occurs following the commitment but before the decision, then people are significantly more likely to deviate from their informal contract and pursue a Pareto improvement. Further, we observe a contingency effect where the likelihood with which people follow an informal commitment declines with the number of contingencies that must occur in order for the contract to be realized. Finally, none of the theories of lying aversion that we explore are able fully to explain our data.

Keywords: informal contract, communication, behavioral game theory, multi-trust game.

[*]Corresponding author. Economics Department, Business School, University of Exeter, Exeter, UK EX4 4PU.Email: j.chen2@exeter.ac.uk

[†]Interdisciplinary Center for Economic Science (ICES) and Department of Economics, George Mason University, Fairfax, VA 22030. Email: dhouser@gmu.edu

1 Introduction

Many forms of partnerships and cooperation rely on informal contracts (*e.g.*, non-binding promises, commitments, or statements of intent), particularly in cases where a formal contract is unavailable or incomplete. Despite the lack of enforcement mechanisms for non-binding contracts/commitment, a growing body of literature in economics and other fields (*e.g.*, psychology and sociology) suggests that informal contracts/commitments tend to discipline behaviors. That is, people honor their contracts/commitments even when faced with losing personal monetary payoffs. The literature further suggests that communication, particularly informal contract/commitment, is crucial for facilitating cooperation and improving efficiency. (see, for example Charness and Dufwenberg, 2006; Charness et al., 2012, 2013; Charness and Dufwenberg, 2011; Miettinen and Suetens, 2008; Vanberg, 2008; Ellingsen and Johannesson, 2004; Kerr and Kaufman-Gilliland, 1994; Loomis, 1959; Sally, 1995). By the same token, the literature indicates that people not only strive to keep their promises, but are also averse to breaking informal contracts and lying to others. These encouraging findings are based on the results of two-player static decision-making environments, which are the focus of the previous literature. However, questions remain as to: (i) whether these efficiency-enhancing and trust-promoting effects of informal contract/commitment persist in more dynamic and multi-player environments such as the world we live in, where new opportunities arise; (ii) whether there are other environments where informal contract/commitments are less effective in disciplining behaviors.

Indeed, in more dynamic contexts, commitments (formal or informal) can and sometimes do constrain people from exploring Pareto improving opportunities, in both business and social contexts. One example is the case of former Microsoft Vice President Kai-Fu Lee, who was hired as the chairman of Googles Chinese branch. Mr. Lees employment with Google was delayed when Microsoft filed a lawsuit claiming that [a]ccepting such a position with a direct Microsoft competitor like Google violates the narrow non-competition promise Lee made when he was hired as an executive. Although Dr. Lee had shown no intention of divulging Microsofts confidential information to Google, Microsoft alleged that Dr. Lees employment with Google threatened the disclosure of Microsofts trade secrets. Both parties eventually settled the case, but ultimately a huge amount of money and time were lost due to the non-competition promise Lee made in his employment contract. Such circumstances are far from rare in the business world. For instance, a startup called CrossGain was forced to lay off

some 20 Microsoft defectors until their non-compete agreements expired, even though CrossGain was not in direct competition with Microsoft. And in the social context, a contract like marriage may similarly act as a constraint towards a better match discovered later in life.

In this paper, we investigate environments inspired by the examples above. Here, we introduce dynamics to the two-person static decision-making situation by adding a third strategic player from whom new beneficial opportunities may arise. Additionally, we allow players to make unbinding informal contracts while varying the path of communication[1].

Our paper makes three key contributions to the literature. First, we provide, to our knowledge, the first empirical evidence on the propensity of people to keep informal contracts in environment where doing so is inefficient. Second, we shed light on the extent to which communication especially in the form of informal contracts can modulate peoples behaviors. We devise four message treatments to systematically explore different paths of communication, by varying the pairing of sender and receiver and the alignment of their monetary interests. Finally, we shed light on the empirical relevance of existing behavioral theories. In particular, we take our data to the three current competing theories: innate preference for honesty models (see, *e.g.*, Ellingsen and Johannesson, 2004; Miettinen and Suetens, 2008; Vanberg, 2008; Gibson et al., 2013), the consequence based preference model (see, *e.g.*, Battigalli and Dufwenberg, 2007; Battigalli et al., 2013; Charness and Dufwenberg, 2006) and simple type models (see, *e.g.*, Hurkens and Kartik, 2009). We derive and summarize testable predictions from these theories and then compare those predictions with our experimental data.

Our main findings are that: (i) when, between the point of commitment and the point of decision, no new information is received regarding the expected value of commitment, then people overwhelmingly decide in accord with their informal contracts and avoid exploring potentially Pareto improving opportunities. However, (ii) when new information arises that reduces the relative value of commitment, and when this occurs following the commitment but before the decision, then people are significantly more likely to deviate from their informal contract and pursue a Pareto improvement. (iii)We observe a contingency effect where the likelihood with which people follow an informal commitment declines with the number of contingencies that must occur in order for the contract to be realized. Finally, none of the theories of lying aversion that we explore are able fully to explain our data.

The remainder of our paper is organized as follows: Section 2 describes the struc-

---

[1]Different path of communication here indicates the different pairings of sender and receiver.

ture of the game and corresponding predictions under various models. Section 3 describes the key hypotheses. Section 4 details the experiment procedure. Section 5 describes our main results. Section 6 explores possible explanations for the observed-but-unaccounted-for behaviors, and the final section summarizes and concludes.

2   The Game and Theory Predictions

This section sets the stage for the subsequent experiment. We first introduce the multi-trust game[2] on which our design is based; clarify various communication treatments; and derive the key predictions from existing theories.

2.1   The Multi-trust Game

We devise a novel three-person game, the extensive form of which is shown in Figure 1. Using backward induction and assuming risk-neutral selfish players, there is a unique subgame perfect Nash equilibrium (In, Out, Right). This equilibrium is inefficient. Therefore, our game G_1 shares the dilemma common to previously studied trust game variants[3].

There are several points to note about our game. First, its structure is closely related to that described by Charness and Dufwenberg (2006). The key difference is that we add a strategic player in place of chance, yet maintain unobservable actions. However, unlike Charness and Dufwenberg (2006), where defection necessarily reduces the trustees payoff, in G_1 defection may have no payoff consequences to the trustor. Second, payoffs in our G_1 differ from sequential trust games with constant multipliers

---

[2]Our environment can be thought of as including two competing (directly or indirectly) firms, Acme (A) and Com (C), and worker Brittany (B). Suppose A considers whether to hire B. A would like to do so, but is concerned about C later hiring B away. The reason is that if C provides B a better opportunity, after B has worked at A, then B can bring information (e.g., big client lists or core technologies) that benefits C but harm A. Moreover, passing information may also result in harm for B (for example, lawsuits or business espionage for worker B, and loss of business for firm A). Now, in order to convince B to leave A, firm C may make an informal commitment not to reveal any information that may be (inadvertently) disclosed by worker B. If B accepts this promise, C could defect, and in doing so, leave both B and C worse off. On the other hand, if C does cooperate, then firm A is not impacted by B's departure, but both B and C are made better off. This provides a possibility to find alternative mechanisms to achieve higher profitability and potentially even better social outcomes.

[3]Such related games are Charness and Dufwenberg (2006) – two-person trust game with a hidden action; Sheremeta and Zhang (2013) & Rietz et al. (2012) – sequential three person trust game; and Cassar and Rigdon (2011) – three person trust game with one trustee two trustor or one trustor two trustee, finally Bigoni et al. (2012) – two person trust game with an add-on dominant solvable game between the trustee and a third player.

3

across trustees (e.g., Sheremeta and Zhang (2013) or Rietz et al. (2012)) in that multipliers in G_1 double with the second trustee. This makes it much more profitable to establish the second partnership. The game below illustrates these ideas.
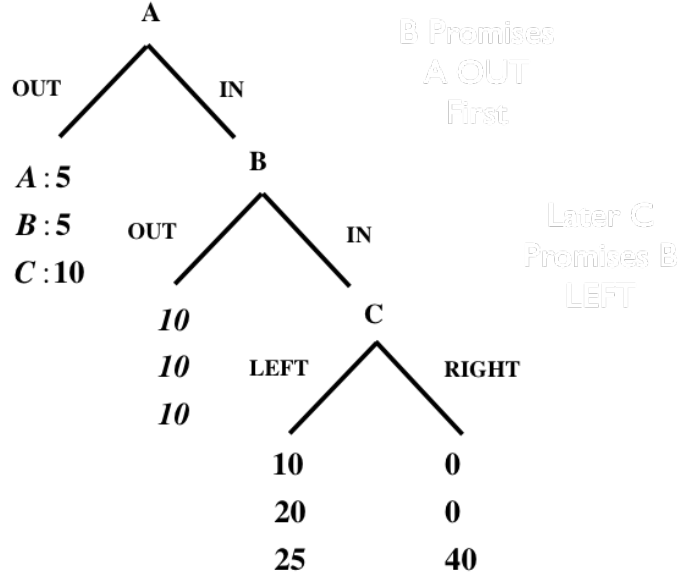

Figure 1: The Multi-trust Game - G_1

A and B consider whether to form a partnership; if no partnership occurs, then both parties receive the outside option payoff of $5. In this case, C receives $10. If a partnership is formed, a trust relationship emerges, and the payoffs to this relationship depend on the B's decision. B is faced with a dilemma – to stay with the current partnership (corresponding to B's *Out* option) or form an additional trust relationship with a third person and enjoy a potentially higher payoff (corresponds to B's *In* option). Note that A is NO better off (perhaps even worse off) by B's choosing *In*; thus, A would always prefer B to choose *Out* and maintain an exclusive partnership. If B chooses to stay with A (corresponding to the strategy profile (*In, Out, Left/Right*), both A and B are better off (with the payoff of $10 for each), and C (who has no move) again earns the outside option of $10. The strategy profile (*In, Out, Left/Right*) corresponds to the situation where an exclusive partnership contract is enforceable. However, such a contract may not be enforceable. Indeed, B's choice may not be observable to A, depending on the Cs decision. Our game captures this as discussed below.

For this case, if B chooses to form a new partnership with C (corresponding to B's *In* option), C can either be cooperative and reciprocal by choosing *Left*, or defect by choosing *Right*. Note that if C chooses *Left*, B's behavior is unknown to A (B's original partner). However, if C chooses *Right*, not only does B receive nothing from the newly

initiated partnership (C takes all), A is also impacted and receives nothing. In this case, A knows B's choice. Note that A may foresee such outcomes and choose not to enter the trust partnership with B. The players choices *Out*, *In* and *Right* describe those possibilities.

2.2 Communication

We next focus on treatments that differ by whether a pre-play communication opportunity is available and how such opportunity is presented. Among all the communication treatments, one player transmits a message to the other player(s) before they play game G_1. If we maintain the assumption that players are selfish, then all the pre-game cheap talk communication should have no effect, and the strategy profile (*In, Out, Right*) remains the unique sub-game perfect solution. However, if there are other concerns that incentivize players, as detailed in the next section, communication will have an impact on behaviors.

To better investigate the effect of communication, inform and verify various existing theories, we consider the following communication treatments denoted as B-A, C-B, C-A, and Double treatments respectively.

In the B-A treatment, prior to the game play, player B can transmit a message to A. In a similar fashion, in the C-B treatments, player C can transmit a message to B prior to the game. Similarly, in the C-A treatments, player C can transmit a message to A.

In the Double treatment, it is common information from the start that role B can send a message to A, and the experimenters collect those messages and pass them on to their matched partners. At this point, for player B, the Double treatment is exactly the same as the B-A treatment. However, after all messages are received, we announce a surprise communication opportunity where role C has the chance to transmit a message to B, and after messages from Cs are transmitted, players play the game. Note that the Double treatment is designed such that we can compare B's behavior in the B-A and Double treatments, and Cs behavior in the C-B and Double treatments holding the players communication opportunity constant.

The next section introduces existing behavioral theories that aim to model the effect of communication on trust and reciprocity.

2.3 The Models and Predictions

Standard economic models of self-interest utility maximization emphasize the role of outcome in dictating agents choices. As a result, these models offer no predictions on how cheap-talk communications will influence behaviors, since cheap-talk is not an enforceable commitment and is therefore impossible to verify. One of the important assumptions in these standard models is that self-interested agents will have no problem lying or defaulting on their words as long as the resulting outcome is preferred. However, honesty and promise-keeping behaviors are frequently reported not only on the news (*e.g.*, the whistleblowers) but also as observations in lab and field experiments. To account for these seemingly puzzling behaviors, researchers have come up with three types of models: intrinsic preference for honesty models, consequence-based models, and type models. The following section discusses those models in detail.

2.3.1 Intrinsic Preference For Honesty Models

Among the models of preference for honesty[4], there are two main varieties. One is the homogeneous aversion to lying model suggested by Ellingsen and Johannesson (2004), Miettinen and Suetens (2008), and Vanberg (2008), where the model assumes that people incur a similar fixed cost in their utilities when caught lying; the other is the heterogeneous cost of lying model, where different people might incur different costs while caught lying (*e.g.*, Gibson et al. (2013)).

Homogeneous aversion to lying model

Ellingsen and Johannesson (2004) proposed a modified model (based on Fehr and Schmidt (1999)) (we now call EJ model) with an added universal cost component l to reflect the universal cost of lying. If there can be no communication, there is no cost of lying. Thus, in the no-communication baseline game, the predictions correspond to the case with selfish preferences described in section 2.1.

We define player i as those who communicate and state a verifiable contract/commitment[5]. Formally, player i has the following simplified utility function:

$$u_i = \begin{cases} m_i - l & \text{if player } i \text{ lies} \\ m_i & \text{otherwise.} \end{cases} \tag{1}$$

---

[4]In some papers, it also called lying aversion, or the cost of lying model (for example, Lundquist et al. (2009)).

[5]For those who choose not to communicate or send non-verifiable communication, their decision problems are modeled with standard self-interest maximization. And (*In, Out, Right*) remains the unique backward-induction solution.

where m_i denotes agent *is* immediate monetary payoff , $-l$ denotes the utility loss from lying. Notice that l is invariant to players therefore implies the assumption that people share a homogeneous cost to lying[6].

G_2, G_3, G_4 in figure 2, 3, 4 model this for B-A, C-B, and Double treatment for players who communicate and indicate a verifiable contract/commitment. For all
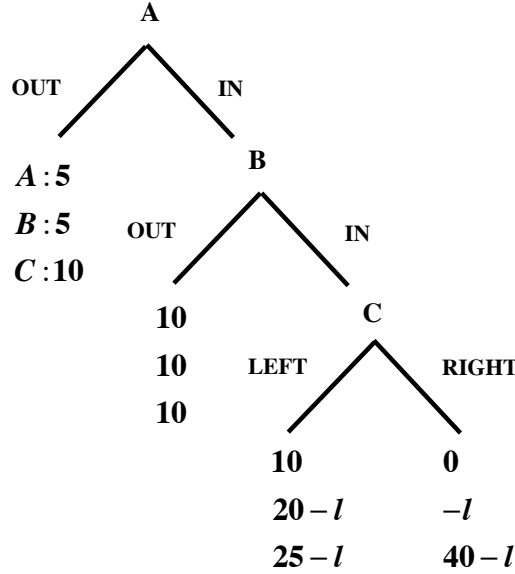

Figure 2: G_2 - B-A treatment

models under consideration, the C-B and C-A treatments are treated exactly the same (although player B has more direct decision relevance to player C than player A). Consequently, in the following sections, we do not explain in depth the predictions for the C-A treatment, they are exactly the same as for the C-B treatment. Notice that game G_2 - G_4 is a nonstandard game in which the utilities are not just numbers (m_i) at the end nodes but rather reflect the adjusted utility (u_i)(this applies to all the games in the following sections).

We denote $_t^p\varphi_s$ the percentage of players p in treatment t choosing strategy s, while $p \in \{B,C\}$, $t \in \{Base, BA, CB, CA, Double\}$, $s \in \{Out, R\}$, where $Base$, BA, CB, CA, $Double$ represent baseline, B-A, C-B, C-A and Double treatment respecitively, and In and R represent strategy In and $Right$ respectively. For example, $_{Base}^B\varphi_{In}$ denotes the percentage of player B choosing In in Baseline treatment.

---

[6]Similarly, Miettinen (2008) introduces similar invariant fixed cost to lying to model the effects of pre-play agreements in contracts. Vanberg (2008) provides supporting laboratory evidence suggesting that lack of lying behaviors in his experiment can be better explained with a simple cost of lying model.
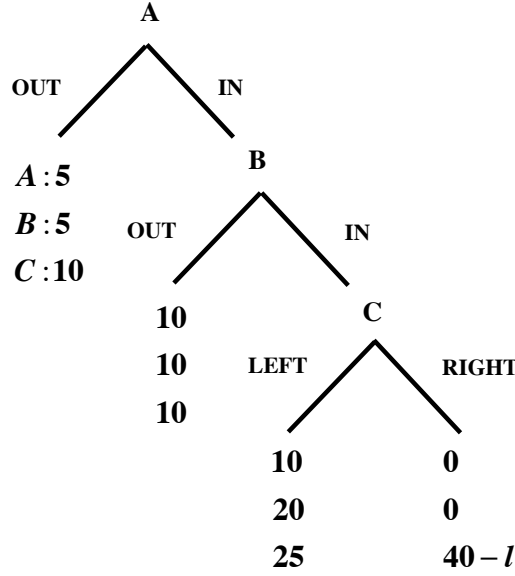
Figure 3: G_3 - C-B(C-A) treatment

Comparing all communication treatments with Baseline, we have the following:

Proposition 1 *(EJ model)*[7] *If* $l < 15$, $_{Base}^{p}\varphi_s = \,_{j}^{p}\varphi_s$, *where* $j \in t$ *and* $j \neq Baseline$; *if* $l \geq 15$, **for player B**, $_{Base}^{B}\varphi_{Out} = \,_{BA}^{B}\varphi_{Out}$, $_{Base}^{B}\varphi_{Out} > \,_{CB}^{B}\varphi_{Out} = \,_{CA}^{B}\varphi_{Out}$, $_{Base}^{B}\varphi_{Out} > \,_{Double}^{B}\varphi_{Out}$; **for player C**, $_{Base}^{C}\varphi_R = \,_{BA}^{C}\varphi_R$, $_{Base}^{C}\varphi_R > \,_{CB}^{C}\varphi_R = \,_{CA}^{C}\varphi_R$, $_{Base}^{C}\varphi_R < \,_{Double}^{C}\varphi_R$.

Proof. 1) In the B-A treatment, Right is still the dominant strategy for C, In and Out remain the best responses for B and A respectively in G_2 as in G_1. Therefore, we expect no treatment difference between B-A and Baseline.

2) In the C-B (C-A) treatment, if $l \geq 15$, *Left* becomes the dominant strategy for C, *(In, In)* is the best response strategy profile for both A and B. As a result, we expect to observe a higher percentage of Bs choosing *In* and a higher percentage of Cs choosing *Left* than the Baseline. If $l < 15$, however, *Right* once again becomes the dominant strategy for C. For A and B, the best responses are *In* and *Out*. In this case, we expect no treatment differences between the C-B and Baseline treatments.

3) For the Double treatment, if $l \geq 15$, *In, Out, Left* describe the best responses for players A, B, C respectively; if $l < 15$, *In, Out, Right* are the best responses instead. Therefore, if $l < 15$, we should not expect to see treatment differences from Baseline,

---

[7]According to Ellingsen and Johanneson (2004), the estimated l is around 21.43 SEK $\approx$ USD 2.64. If we apply this estimates to our game, $l = 2.64 < 15$, we should expect no effects at all across treatments.
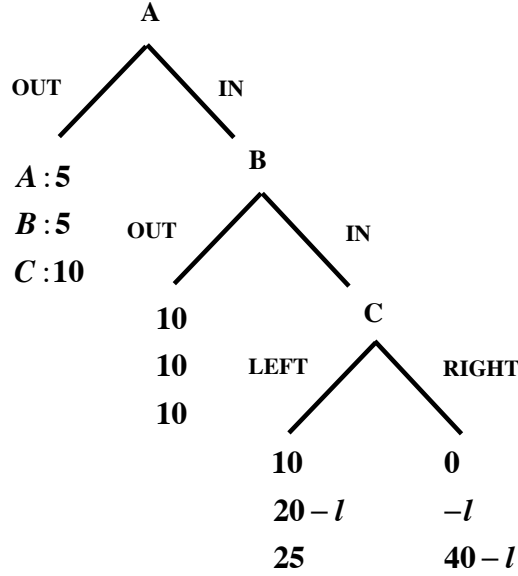
Figure 4: G_4 - Double treatment

and communication in the form of cheap talk has no effect on the behavior of all players. If $l \geq 15$, we should observe an increased rate of player B choosing *In* for Double compared with the Baseline Treatment; for player C, we should observe an decreased frequency of choosing *Right* in the Double treatment compared with Baseline. ∎

In order to make within-treatment comparisons, we have to make one simple assumption:

Assumption 1 *The frequency of informal contract/commitment exchanged is constant from the same role.*

Assumption 1 implies that: 1) the percentage of Bs sending informal contracts in the B-A treatment is comparable to that in the Double treatment; 2) the percentage of Cs exchanging informal commitment in the C-B treatment is comparable to that in the Double treatment; and 3) the percentage of Bs receiving informal contracts in the C-B treatment is comparable to that in the Double treatment. We denote that b and c percent of Bs and Cs send informal contracts respectively, where $b, c \in [0, 1]$.

Proposition 2 *(EJ model)* If $l < 15$, **for player B**, ${}^{B}_{BA}\varphi_{Out} = {}^{B}_{Double}\varphi_{Out} = {}^{B}_{CB}\varphi_{CB}$; **for player C**, ${}^{C}_{BA}\varphi_R \; {}^{C}_{CB}\varphi_R = {}^{C}_{Double}\varphi_R$. If $l \geq 15$, **for player B**, ${}^{B}_{CB}\varphi_{Out} < {}^{B}_{Double}\varphi_{Out} < {}^{B}_{BA}\varphi_{Out}$; **for player C**, ${}^{C}_{BA}\varphi_R > {}^{C}_{CB}\varphi_R = {}^{C}_{Double}\varphi_R$.

Proof. 1) In both the B-A and Double treatments, if B is rational and sends a contract, he will only choose *In* if $p(20 - l) - (1 - p)l > 10$, i.e., $20p - l > 10$, where p is Cs

9

probability of choosing *Left*. Specifically, in the B-A treatment, $p = 0$, regardless of the value of l, it is B's best response to choose *Out* (whether or not they send informal contracts).

2) In the Double treatment, if $l < 15$, Cs always choose *Right* whether they send informal commitment or not, i.e., $p = 0$. In response to that, all Bs choose *Out* whether they receive informal contracts or not. If $l \geq 15$, the situation is more complicated. $1 - c$ percent of Bs do not receive informal contracts, i.e., $p = 0$, the best response them is *Out*. For the remaining Bs who receive informal contracts, $b \cdot c$ percent of Bs also send informal contracts, in this case, $p = 1$, and $20 - l \leq 5 < 10$, the best response is *Out*; $(1 - b) \cdot c$ percent of Bs do not send informal contracts but receive one from C, $p = 1$ and $10 < 20p$, their best response is instead *In*. In sum, $_{Double}^{B}\varphi_{Out} = 1 - (1 - b) \cdot c$.

3) In C-B treatment, if B is rational and send informal contracts, he will only choose *In* if $20p > 10$, i.e., $p > 0.5$. If $l < 15$, *Out* is B's best response since $p = 0$. If $l \geq 15$, c percent of Bs receive informal contracts from C ($p = 1$), *In* is the best response. For the remaining $1 - c$ percent of Bs who do not receive informal contracts from C, *Out* is the best response. And $_{CB}^{B}\varphi_{Out} = 1 - c$. Since $c > c(1 - b) > 0$, we have $_{CB}^{B}\varphi_{Out} < _{Double}^{B}\varphi_{Out} < _{BA}^{B}\varphi_{Out}$.

4) If C is rational and send informal contracts, he/she will only choose *Left* if $25 > 40 - l$. In both C-B and Double treatment, the decision problem for C is the same. Therefore, we expect no differences between the two treatments. And the rate of C choosing *Right* in both C-B and Double treatment will be smaller than B-A treatment where there is no incentive for C to choose *Left* at all. ∎

Combine both Proposition 1 and 2, we have:

Proposition 3 *(EJ model)* *If $l < 15$, $_{i}^{p}\varphi_{s} = _{j}^{p}\varphi_{s}$, where i, $j \in t$ and $i \neq j$; If $l \geq 15$, **for player B**, $_{Base}^{B}\varphi_{Out} = _{BA}^{B}\varphi_{Out} = _{Double}^{B}\varphi_{Out} < _{CB}^{B}\varphi_{Out}$; **for player C**, $_{Base}^{C}\varphi_{R} = _{BA}^{C}\varphi_{R} = _{Double}^{C}\varphi_{R} > _{CB}^{C}\varphi_{R}$.*

In the EJ model, everyone suffers the same cost from lying; therefore, if the benefit of lying outweighs the cost, any communication is futile since words said will never be kept. However, if the cost of lying outweighs the benefit, we expect people to keep their contracts (if they send one). Thus, in both the B-A and Double treatments where player B may send a contract , we hypothesize a higher rate of *In* than in the C-B (C-A) treatments. Similarly, we expect to see a higher rate of cooperative action from player C choosing *Left* in both the C-B (C-A) and Double treatments than in any other treatments.

Heterogeneous cost to lying model

Gibson et al. (2013) propose and test a heterogeneous preference for honesty model (we call GTF model). The GTF model is very similar to the EJ model in that the cost of lying is independent from any type of belief. The difference is that instead of a fixed l for all players, the GTF model assumes that each individual might have a different cost l_i associated with lying. The utility function is as follows:

$$u_i = \begin{cases} m_i - l_i & \text{if player } i \text{ lies} \\ m_i & \text{otherwise.} \end{cases} \quad (2)$$

where l_i indicates the utility loss player i endures when he(or she breaks the promise or lies and $l_i \in [0, +\infty]$.

Game G_5, G_6, G_7 in Figure 5, 6, 7 incorporate those. And we have Proposition 4 below.

Proposition 4 *(GTF model)* **For player B**, ${}^B_{Base}\varphi_{Out} = {}^B_{BA}\varphi_{Out} > {}^B_{Double}\varphi_{Out} > {}^B_{CB}\varphi_{Out} = {}^B_{CA}\varphi_{Out}$; **for player C**, ${}^C_{Base}\varphi_R = {}^C_{BA}\varphi_R > {}^C_{Double}\varphi_R = {}^C_{CB}\varphi_R = {}^C_{CA}\varphi_R$.
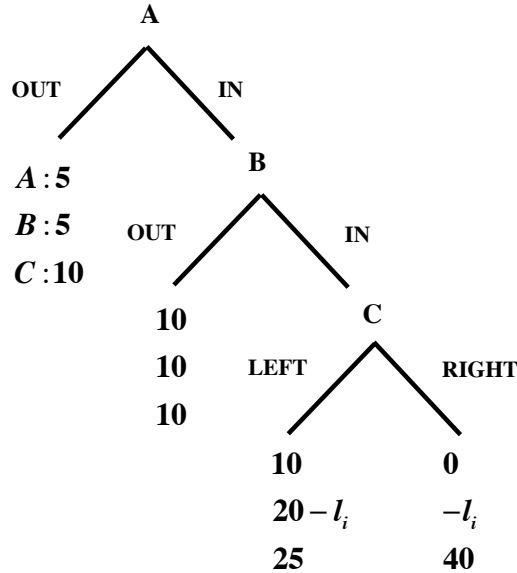

Figure 5: G_5 - B-A treatment

Proof. 1) Similarly, in the B-A treatment, all Bs choose *Out* given that *Right* is the dominant strategy for Cs.

2) In the Double treatment, for the $1 - c$ percent of Bs who do not receive informal contracts, *Out* is the best response. For the $b \cdot c$ percent of Bs who both send and receive informal contracts from C, they will only choose *In* if $p(20 - l_i) - (1 - p)l_i \geq 10$. $p = 1$ if
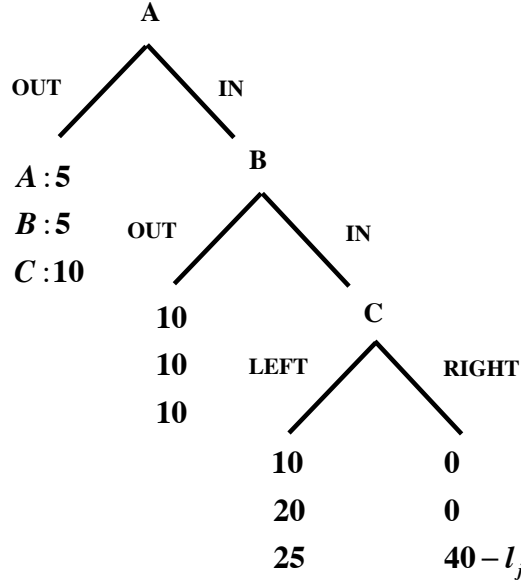
Figure 6: G_6 - C-B(C-A) treatment

$l_j \leq 15$, and $p = 0$ otherwise. Assume that $Pr(l_j \geq 15) = w$ and $Pr(l_i \leq 10) = k$, only $w \cdot k \cdot b \cdot c$ percent of B (whose cost to default on informal contracts is small enough, while his/her partners cost is big enough) choose *In* and $(1 - w \cdot k)b \cdot c$ choose *In*. For the remaining $(1 - b)c$ percent of Bs who do not send but receive informal contracts from C, they will only choose *In*, if $20p \geq 10$. We have $w \cdot (1 - b)c$ percent of Bs choosing *In*, and $(1 - v)(1 - b)c$ percent of Bs choosing *Out*. In total, $w \cdot k \cdot b \cdot c + w \cdot (1 - b)c$ percent of Bs choose *In* and the rest choose *Out*.

3) In the C-B treatment, similar to the Double treatment, the $1 - c$ percent of Bs who do not receive informal contracts choose *Out*. The rest will choose *In* only if $20p \geq 10$. Since $p = 1$ with probability w, we have $w \cdot c$ percent of Bs choosing *In* and $1 - w \cdot c$ percent of Bs choosing *Out*. Since $^B_{Double}\varphi_{In} = w \cdot k \cdot b \cdot c + w \cdot (1 - b)c$ is smaller than $^B_{CB}\varphi_{In} = w \cdot c$, we should expect a higher frequency of Bs choosing *In* from C-B to the Double treatment. ∎

For player B's behavior, the GTF model hypothesizes that B may choose *In* more frequently in the C-B treatment than in the Double treatment and also more frequently in Double treatment than B-A and Baseline treatments. The reasons are as follows. In the B-A and baseline treatments, it is always the best response for B to choose *Out*, anticipating C to prefer *Right*. In the Double treatment, however, choosing *In* can become the best response for some message-receiving Bs when there is increased probability of C choosing *Left* (for those Cs with $l_j > 15$). In the C-B treatment,
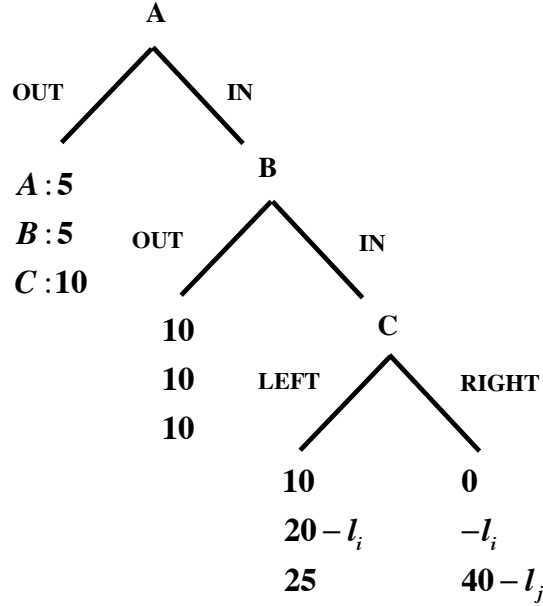
12

Figure 7: G_7 - Double treatment

choosing *In* can be the best response for ALL message-sending Bs when their matched Cs are with $l_j > 15$.

2.3.2 Consequence-based Model

Charness and Dufwenberg (2006) proposed a guilt from blame model built on the psychological game theory framework developed by Geanakoplos et al. (1989), furthered by Dufwenberg and Kirchsteiger (2004) and formalized in Battigalli and Dufwenberg (2007). In this model, the cost of lying comes from the guilt of letting someone down, and the level of guilt a player suffers depends on the level of harm he imposes on others relative to what the others believe they will suffer (i.e., the difference between the players actual action and the action the player believes others believe he would take). In a sense, this model is a different take on social preference models, where the degree one cares about others also depends on the belief one holds about others belief about him/her.

According to the Guilt Aversion Model (we call CD model from now onwards), player i has the following modified utility:

$$u_i = m_i - \gamma_i \cdot \tau_i \cdot \Delta m_j, \ i \neq j \tag{3}$$

where γ_i denotes player i's sensitivity to guilt, and it is independent from τ_i; $\gamma_i \in$

$[0, +\infty)$; τ_i denotes player is belief about js belief about i, $\tau_i = E(\mu_j)$, where μ_j is the probability player j assigns to is move, $\mu_j \in [0, 1]$; Δm_j denotes j's monetary loss between what i thinks j thinks that i would do and what i actually does. In light of our game with B-A treatment where B sends a contract to A, Δm_A in this case would be \$10, which is what A would get given B's informal contracts (this is also the monetary payoff B would expect A to believe B would be able to give him/her) minus 0 (if, instead of choosing *Out*, B deviates from his/her contract and chooses *In*. since the best strategy for C is to choose *Right*, A would get 0 given B and Cs (*In, Right*) choices) equals 10.

G_8, G_9, G_{10} in Figure 8, 9, and 10, respectively, incorporate the CD model for all treatments. Notice that in the B-A treatment, player B doesn't suffer from guilt, when he/she chooses *In* (violating informal contracts) and player C chooses Left. The reason is that when player C plays *Left*, A receives \$10, which is the same if B has chosen *Out*. In other words, violating the contract has no monetary consequences to A if C chooses *Left*, thus B doesnt feel any letting down guilt. Similarly, in the Double treatment, B doesnt suffer utility loss from guilt choosing *In* and breaking informal contracts as long as C chooses *Left*.
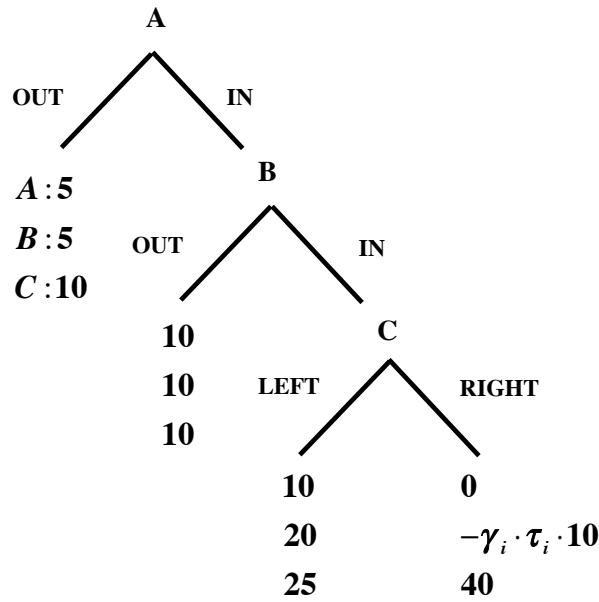
Figure 8: G_8 - B-A treatment

Proposition 5 *(CD model)* **For player B,** $^B_{Base}\varphi_{Out} = {}^B_{BA}\varphi_{Out} > {}^B_{CB}\varphi_{Out} = {}^B_{CA}\varphi_{Out} = {}^B_{Double}\varphi_{Out}$*; for player C,* $^C_{Base}\varphi_R = {}^C_{BA}\varphi_R = {}^C_{Double}\varphi_R > {}^C_{CB}\varphi_R = {}^C_{CA}\varphi_R$.
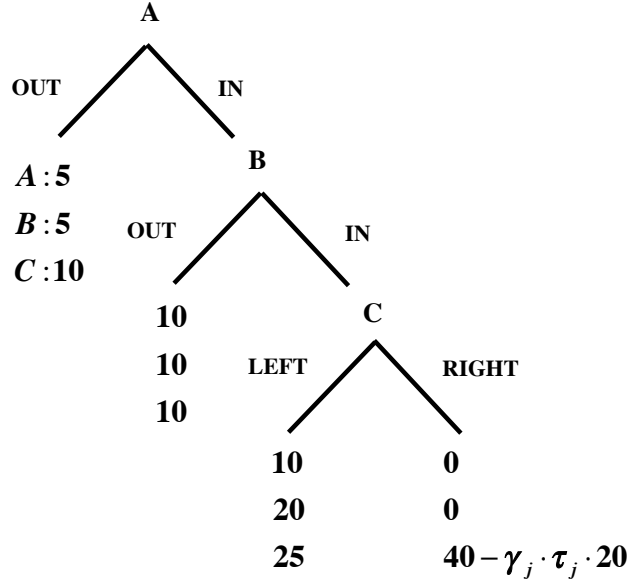
Figure 9: G_9 - C-B(C-A) treatment

Proof. 1) Similar to the analysis for the EJ model, b percent of Bs in B-A treatment will only choose *In*, if $20p - \gamma_i \cdot \tau_i \cdot 10 \cdot (1-p) \geq 10$. Anticipating that *Right* is Cs dominant strategy, $p = 0$, *Out* is the best response for all Bs.

2) In the Double treatment, for the $1 - c$ percent of Bs who do not receive informal contracts, *Out* is the best response. For the $b \cdot c$ percent of Bs who both send and receive informal contracts from C, they will only choose *In* if $20p - \gamma_i \cdot \tau_i \cdot 10(1-p) \geq 10$. $p = 1$ if $25 \geq 40 - \gamma_j \cdot \tau_j \cdot 20$, and $p = 0$ otherwise. Assume that $Pr(25 \geq 40 - \gamma_j \cdot \tau_j \cdot 20) = v$, we have $v \cdot b \cdot c$ percent of B choose *In* and $(1-v)b \cdot c$ choose *Out*. For the remaining $(1-b)c$ percent of Bs who do not send but receive informal contracts from C, they will only choose *In*, if $20p \geq 10$. We have $v \cdot (1-b)c$ percent of Bs choosing *In*, and $(1-v)(1-b)c$ percent of Bs choosing *Out*. In total, $v \cdot c$ percent of Bs choose *In* and $1 - v \cdot c$ percent of Bs choose *Out*, i.e., $^B_{Double}\varphi_{Out} = 1 - v \cdot c$.

3) In the C-B treatment, similar to the Double treatment, the $1 - c$ percent of Bs who do not receive informal contracts choose *Out*. The rest will choose *In* only if $20p \geq 10$. Since $p = 1$ with probability v, we have $v \cdot c$ percent of Bs choosing *In* and $1 - v \cdot c$ percent of Bs choosing *Out*, which is the same as in Double treatment. ∎

The CD model offers the same predictions on player Cs behavior as EJ and GTF models; likewise, the intuition (i.e., lying is costly, although in the CD model defaulting on informal contracts is costly due to the fact that lying always lowers a partners payoff) is similar among all three models. For player B, the CD model differs from the
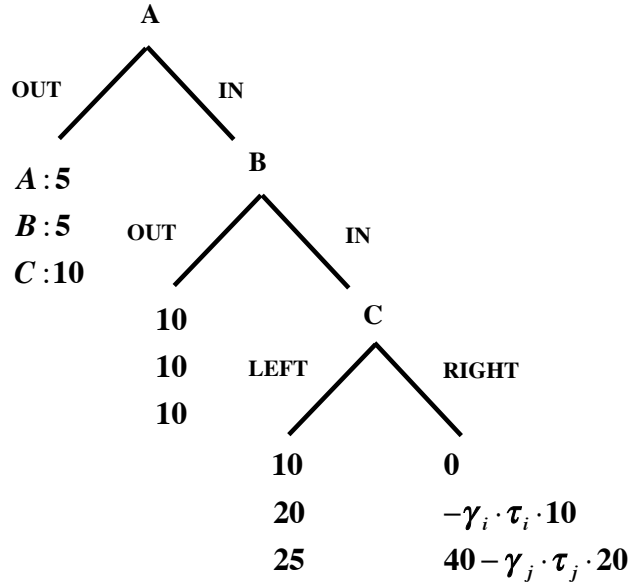
Figure 10: G_{10} - Double treatment

EJ model in that each player may have varying costs from lying, and it is different from GTF model in that the different costs are from "letting-down" resulted guilt. Also, the level of guilt depends on guilt sensitivity and potential harm that may incur to others while in GTF model, the cost of lying is innate and independent from consequences. The key prediction difference between the CD and EJ models (when $l > 15$) is player B's behavior in the Double treatment: in the EJ model, breaking informal contracts is costly, and consequently player B in the Double treatment behaves the same way in the B-A treatment; in the CD model, however, breaking informal contracts can be costless for player B as long as there is no foreseeable harm to player A (which is clearly the case if C chooses *Left*); as a result, we should observe an increased rate of B choosing *In* for the Double treatment versus the B-A treatment.

2.3.3 Simple Type Model

Hurkens and Kartik (2009) put forward a simple type model that can make sense of the observations in Gneezy (2005) (we call it the HK model). The HK model assumes that there are two types of people, one with infinite cost of lying (honest type), and the other with zero cost of lying (economic type).

$$u_i = m_i - i \cdot L \tag{4}$$

16

where i indicates players' type, $i \in \{0, 1\}$, if $i = 0$, player i is economic type with no cost to lying, while if $i = 1$, i is honest type with infinite cost to lying; L denotes an enormous cost to breaking informal contracts and $L \to +\infty$. Assume that in the population ρ percent are honest types, and $1 - \rho$ percent are economic types. The implications of the HK model for our game are detailed in Proposition 6.

Proposition 6 *(HK model)* **For player B**, $^B_{Base}\varphi_{Out} = {}^B_{BA}\varphi_{Out} > {}^B_{Double}\varphi_{Out} = {}^B_{CB}\varphi_{Out} = {}^B_{CA}\varphi_{Out};$ **for player C**, $^C_{Base}\varphi_R = {}^C_{BA}\varphi_R > {}^C_{Double}\varphi_R = {}^C_{CB}\varphi_R = {}^C_{CA}\varphi_R.$

Proof. 1) As with all other theories, the HK model predicts that all Bs choose *Out* given that *Right* is the dominant strategy for Cs in the B-A treatment.

2) In the Double treatment, for the $1 - c$ percent of Bs who do not receive informal contracts, *Out* is the best response. For the $b \cdot c$ percent of Bs who both send and receive informal contracts from C, $\rho \cdot b \cdot c$ of them are the honest type, and will always choose *Out* as they state in the informal contracts. Among the remaining $(1 - \rho)b \cdot c$ economic types, they will only choose *In* if $20p \geq 10$. $p = 1$ if they receive informal contracts from an honest type, which happens with probability of ρ, and $p = 0$ otherwise. Therefore, we have $\rho(1 - \rho)b \cdot c$ percent of Bs choosing *In* and $(1 - \rho)^2 b \cdot c$ choosing *Out*. For the remaining $(1 - b)c$ percent of Bs who do not send but receive informal contracts from C, they will only choose *In*, if $20p \geq 10$. We have $\rho \cdot (1 - b)c$ percent of Bs choosing *In*, and $(1 - \rho)(1 - b)c$ percent of Bs choosing *Out*. In total, $\rho \cdot c(1 - b \cdot \rho)$ percent of Bs choose *In* and the rest choose *Out*.

3) In the C-B treatment, similar to the Double treatment, $1 - c$ percent of Bs who do not receive informal contracts choose *Out*. The rest will choose *In* only if $20p \geq 10$. Since $p = 1$ with probability ρ, we have $\rho \cdot c$ percent of Bs choosing *In* and $1 - \rho c$ percent of Bs choosing *Out*, which is the greater than Double treatment. ∎

The predictions under the HK model are indistinguishable from the GTF model in our game settings, and the reasoning behind the hypotheses is quite similar. In the GTF model, an honest type never lies, whereas an economic type has no cost of lying if the outcome from lying is preferred. Communication thus only has an effect on the behaviors of an honest type who chooses to send a message. As for player C, the changes in aggregate behavior only come from honest types who communicate. For player B, anticipating that some Cs would switch and choose Left, all Bs from the C-B treatment, all economic-type Bs, and some honest-type Bs who choose not to send a message from the Double treatment have the incentive to switch to *In*. And in both B-A and baseline treatment, there is no incentive for any type of Bs to deviate from choosing *Out*.

3   Key Hypotheses

Following the analysis in section 2, Table 1 and 2 below summarizes hypothesis under different existing theories. The tables can be read as follows: the inequality (equal) sign represents the comparison outcome between the row treatment and the column treatment. For example, the equal sign on row 3 column 2 implies that the frequency of Bs choosing Out is expected to be the same between Baseline and B-A treatments. Notice that all theories except for EJ model with $l < 15$ predict the same treatment effects compared with Baseline for both players B and C. However, the predictions differ when we compare between treatments. Hypothesis 1–4 investigates the treatment effects compared with Baseline, and hypothesis 5–6 focuses on the between-treatment differences under which existing theories offer different predictions.

Table 1: Frequency of Bs Choosing Out: $_t^B \varphi_{Out}$

	EJ model ($l < 15$)				EJ model ($l \geq 15$)/GTF model/HK model				CD model			
	B-A	C-B	C-A	Double	B-A	C-B	C-A	Double	B-A	C-B	C-A	Double
Baseline	=	=	=	=	=	>	>	>	=	>	>	>
B-A		=	=	=		>	>	>		>	>	>
C-B			=	=			=	>			=	=
C-A				=				>				=

Note: the inequality (equal) sign represents the comparison outcome between the row treatment and the column treatment.

Table 2: Frequency of Cs Choosing $Right$: $_t^C \varphi_R$

	EJ model ($l < 15$)				EJ model ($l \geq 15$)/GTF model/HK model/CD model			
	B-A	C-B	C-A	Double	B-A	C-B	C-A	Double
Baseline	=	=	=	=	=	>	>	>
B-A		=	=	=		>	>	>
C-B			=	=			=	=
C-A				=				>

Note: the inequality (equal) sign represents the comparison outcome between the row treatment and the column treatment.

Hypothesis 1 *(B-A vs. Baseline)* *the proportion of Bs choosing Out is the same in both the baseline and B-A treatments, $_{Base}^{B}\varphi_{Out} = {_{BA}^{B}}\varphi_{Out}$; the proportion of Cs choosing Right is the same in both the baseline and B-A treatments, $_{Base}^{C}\varphi_{R} = {_{BA}^{C}}\varphi_{R}$.*

All the models offer the same prediction with regard to the B-A and baseline treatments. If this hypothesis is rejected, we can invalidate all models under consideration. Otherwise, all models are possibly valid.

Hypothesis 2 *(C-B vs. Baseline)* *there is a lower percentage of Bs choosing Out in the C-B treatment than in the Baseline treatment, $_{CB}^{B}\varphi_{Out} < {_{Base}^{B}}\varphi_{Out}$; and a lower percentage of Cs choosing Right in the C-B treatment than in the Baseline treatment, $_{CB}^{C}\varphi_{R} < {_{Base}^{C}}\varphi_{R}$.*

If this hypothesis is rejected, we can invalidate the EJ model with $l \geq 15$, CD and the GTF/HK model. Otherwise, we can invalidate the EJ model with $l < 15$.

Hypothesis 3 *(Double vs. Baseline)* *there is a lower percentage of Bs choosing Out in the Double treatment than in Baseline treatment, $_{Double}^{B}\varphi_{Out} < {_{Base}^{B}}\varphi_{Out}$; and a lower percentage of Cs choosing Right in the Double treatment than in the Baseline treatment, $_{Double}^{C}\varphi_{R} < {_{Base}^{C}}\varphi_{R}$.*

If this hypothesis is rejected, we can invalidate CD and GTF/HK model. Otherwise, we can invalidate both EJ models.

Hypothesis 4 *(C-A vs. Baseline)* *there is a lower percentage of Bs choosing Out in the C-A treatment than in the Baseline treatment, $_{CA}^{B}\varphi_{Out} < {_{Base}^{B}}\varphi_{Out}$; and a lower percentage of Cs choosing Right in the C-A treatment than in the Baseline treatment, $_{CA}^{C}\varphi_{R} < {_{Base}^{C}}\varphi_{R}$.*

Similar to Hypothesis 2, if this hypothesis is rejected, we can invalidate the EJ model with $l \geq 15$, the CD model, and the GTF/HK model. Otherwise, we can invalidate the EJ model with $l < 15$.

Hypothesis 5 *(C-A vs. C-B)* *the percentage of Bs choosing Out is the same in both the C-A and C-B treatments, $_{CA}^{B}\varphi_{Out} < {_{CB}^{B}}\varphi_{Out}$; and the percentage of Cs choosing Right is the same in both the C-A and C-B treatments, $_{CA}^{C}\varphi_{R} < {_{CB}^{C}}\varphi_{R}$.*

All the existing theories treat informal contracts/commitments the same regardless of the decision relevance of the receiver of the promise. A promise from C to A should be treated the same as a promise to B. If this hypothesis is rejected, we can invalidate all models.

Hypothesis 6 *(C-B vs. Double) the proportion of Bs choosing Out is the same in both the C-B and Double treatments, $_{CB}^{B}\varphi_{Out} < {}_{Double}^{B}\varphi_{Out}$; the proportion of Cs choosing Right is the same in both the C-B and Double treatments, $_{CB}^{C}\varphi_{R} < {}_{Double}^{C}\varphi_{R}$.*

If the hypothesis for player B is rejected, we can invalidate the CD model. Otherwise, we can invalidate all other models. And if the hypothesis for player C is rejected, we can invalidate all models.

4   Experimental Procedure

The experimental sessions were conducted in the experimental laboratory of the Interdisciplinary Center for Economic Science at George Mason University. Participants were seated at spaced intervals. We conducted 17 sessions with 9-12 participants per session (total of 273 subjects). Participants could only participate in one session. Average earnings were $17; sessions lasted approximately 90 minutes.

We used within design in our experiment. In each session, subjects played three games (treatments) in random order. Subjects were fully aware that for each game they were matched with complete strangers with whom they had never previously interacted before. Only one of the games was randomly selected at the end of the experiment to be a paid game. The session and corresponding game played are shown in Table 3[8]. During each session, participants were referred to as A or B or C (as in the games in section 2). Participants roles in the experiment were randomly determined at the beginning of the experiment when subjects privately drew from a stack of cards with a letter ("A", "B" or "C") and a number written on them. The letter indicated the participant's role in the session. Participants' roles stayed constant throughout the experiment. Note that the Double treatment was never run together with either the B-A or C-B treatments; the reason is that the Double treatment effectively combines the B-A and C-B treatment. If had run the Double treatment together with the B-A and(or or C-B treatments, participants would have had to write messages to the same role twice, which could have potentially confounded treatment differences (such as different contents in the messages).

We adopted the strategy method in conducting our experiment, as did Charness and Dufwenberg (2006)[9]. In the Baseline treatment, no messages were allowed. In all other

---

[8]Notice that in Table 3 ,A-C treatment isnt discussed in this paper since it is less relevant for the purpose of this paper.

[9]This is an effort to make the results more comparable for theory testing purposes. Also Amdur and Schmick (2012) suggest that there is no behavioral difference between the use of strategy method and direct response for our type of game with communication.

Table 3: Session – Game information

Session No.	Game Played
1-7	Baseline, B-A treatment, C-B treatment
7-13	Baseline, Double treatment, C-A treatment
14-17	Baseline, Double treatment, A-C treatment

treatments with communication, each potential message-sender had the option to send a nonbinding messages to their matched partner prior to their partners' decision; they were given a sheet of paper, but could decline to send a message by circling the letter (A, B or C) that indicated their role in the experiment at the top of the otherwise-blank sheet. Then messages were transmitted to the respective partners. Upon completion of the message transmission, participants played the game; B made his/ her choice of *In* or *Out* without knowing A's actual choice of *In* or *Out* (similarly, C made his/her decision without know the actual decision of B), but the instruction explained that B's choice would be immaterial if A chose *Out*. We therefore obtained an observation for every B and C.

5 Results

We present a summary of communication and detail players behavioral patterns in Section 5.1 and 5.2 respectively. In Section 5.3, we test the hypothesis from Section 3.

5.1 Communication Summary - Messages and Their Contents

What messages were sent? Free-form messages can potentially be classified in various ways. To simplify the analysis, we assume that a player can make a statement regarding his/her planned action or stay silent. For instance, Player B can make a statement regarding his(or her planned action (*In* or *Out*) and Player C can make a statement (*Left* or *Right*). Staying silent indicates two things: no messages are transmitted or the message shows no indication of the players planned action. Here, we denote a statement of planned action as Informal Contract/Commitment, and staying silent as Silence. From the messages that we collected, informal contracts from B always involve a statement indicating the action *Out*, while all the informal contracts from C involve a statement indicating the action *Left*.

Table 4 below summarizes the frequency of messages (communication) and informal contract in each of the treatments. The difference between the two indicates the

percentage of messages that belongs to Silence. The data generally support our assumption that both the rate of communication and the frequency of informal contract are constant for the same role as long as there is no contingency that must occur before the contract is realized.

As shown in Table 4, Player B sends messages about 88 percent of the time in both the B-A and Double treatments ($z = 0.11, p = 0.92$). Player C sends messages about 80 percent of the time in both the C-B and Double treatments ($z = 0.83, p = 0.40$); however, in the C-A treatment, player Cs are significantly less likely to send messages ($z = 2.60, p = 0.01$); less than half of them sent messages to As.

Table 4: Communication Summary by Treatment

	All Messages		Informal Contracts	
Treatment	Player B	Player C	Player B	Player C
B-A	21/24		13/24	
	(88%)		(54%)	
C-B		20/24		16/24
		(83%)		(67%)
C-A		13/27		7/27
		(48%)		(26%)
Double	38/43	32/43	22/43	29/43
	(88%)	(74%)	(51%)	(67%)

For informal contracts/commitments, player B sends informal contracts/commitments) about 53 percent of time in both the B-A and Double treatments ($z = 0.23, p = 0.81$). And player C sends informal contracts commitments around 67 percent of the time in both the C-B and Double treatments ($z = 0.06, p = 0.95$). In the C-A treatment, consistent with the pattern for messages, Cs are significantly less likely to send informal contracts/commitments to As compared with C-B and Double ($z = 2.89, p = 0.00; z = 3.36, p = 0.00$, respectively); around a quarter of them sent informal contracts/commitments to As. However, we can reject the null hypotheses ($p = 0.00$) that there are no informal contracts/commitments exchanged in the C-A treatment as compared to Baseline.

5.2 Data Summary

5.2.1 Behavioral Patterns Across Treatments

Figure 11 summarizes the choices of players A, B and C for the various treatments. In Baseline, where there is no communication opportunity, 81% of As chose Out, 76% of Bs chose Out, and 73% of Cs chose the defective option Right. The behaviors we observed are well described by the unique backward induction Nash equilibrium strategy profile (*In, Out, Right*).
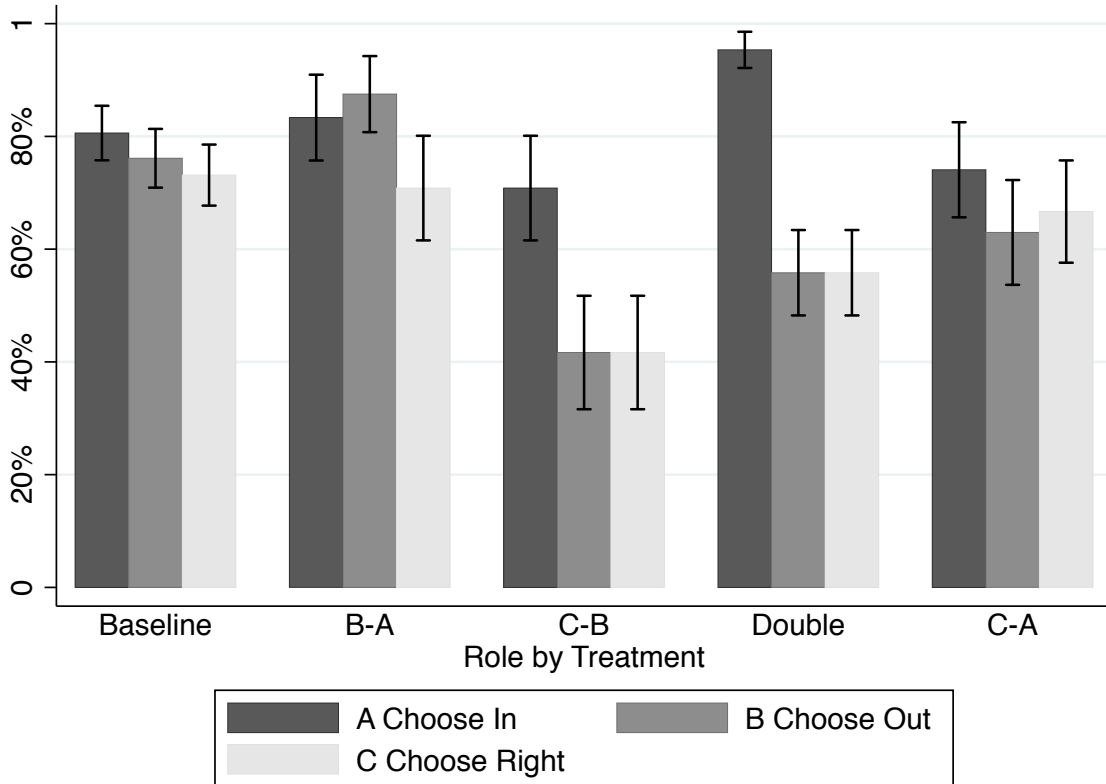

Figure 11: Choices By Role Across All Treatments

In the B-A treatment, we observe a similar percentage of As choosing *In* (83%), even more Bs - 87% - choosing *Out*, and a similar percentage of Cs choosing *Right* (71%). Compared with Baseline, more Bs chose *Out*, although the percentage is not statistically significant. The informal contracts from Bs reduce Bs willingness to explore potential Pareto improving opportunities.

In the C-B treatment, 71% of As chose In, half of Bs switched and chose *Out* (42%), and more than half of Cs (58%) chose the cooperative action - *Right*. When B

23

receives informal contracts from C, not only does B choose to trust C, C also recip-
rocates. The informal contracts from Cs are effective in binding Cs behaviors despite
the misalignment of monetary interest between B and C.

In the Double treatment, 95% of As chose *In*, only 56% of Bs chose *Out*, and
56% of Cs chose *Right*. B sends informal contracts to A indicating the willingness to
choose *Out*; however, when new opportunities arise (C sends commitment conveying
the willingness to cooperate), significantly less Bs ($p = 0.00$) chose *Out* (an action
that is consistent with their informal contracts). In this environment, informal con-
tracts/commitments are less effective in binding peoples behaviors.

When C could send a message to A (C-A treatment), 74% of As chose *In*, 63% of
Bs chose Out, and 67% of Cs choose Right. The behaviors in this treatment resemble
those in Baseline, which suggests that communication in the form of informal contracts
is ineffective in promoting trust and reciprocity.

5.2.2 Comparison of Nash Play and Pareto Improving Play Across Treat-ments

We use the bootstrap method to compare the frequency of Nash equilibrium strat-
egy profile (*In, Out, Right*) and Pareto efficient strategy profile (*In, In, Left*) among
different treatments. The distributions of frequencies for Nash strategy profile (*In,
Out, Right*) and Pareto Efficient profile (*In, In, Left*) across treatments are shown in
figure 12 and 13, respectively.

We find that the B-A treatment has the highest frequency of Nash strategy profile
($p = .26$ compared with Baseline, $p = .00$ with C-B, $p = .06$ with Double, $p = .00$
with C-A), and on average, this strategy profile is played about half of the time. B-A
is followed by Baseline ($p = .12$ compared with Double treatment, $p = .00$ with C-B,
$p = .03$ with C-A) and the Double treatment ($p = .21$ compared with C-A treatment,
$p = .00$ with C-B), while the C-B treatment has the lowest frequency of Nash play
where the Nash profile is played around 4% of the time on average.

As for the frequency of Pareto efficient strategy profile, we observe that the C-
B treatment has the highest frequency for (*In, In, Left*) strategy profile ($p = .00$
compared with Baseline, $p = .00$ with B-A, $p = .08$, with Double, $p = .02$ with C-A),
and the Pareto efficient strategy is played 29% of the time on average. (*In, In, Left*) is
played in the Double treatment about 14% of the time ($p = .19$, compared with C-A
treatment, $p = .01$ compared with Baseline, $p = .00$ with B-A). The B-A treatment has
the lowest frequency of Pareto efficient strategy profile, and it is almost never played.
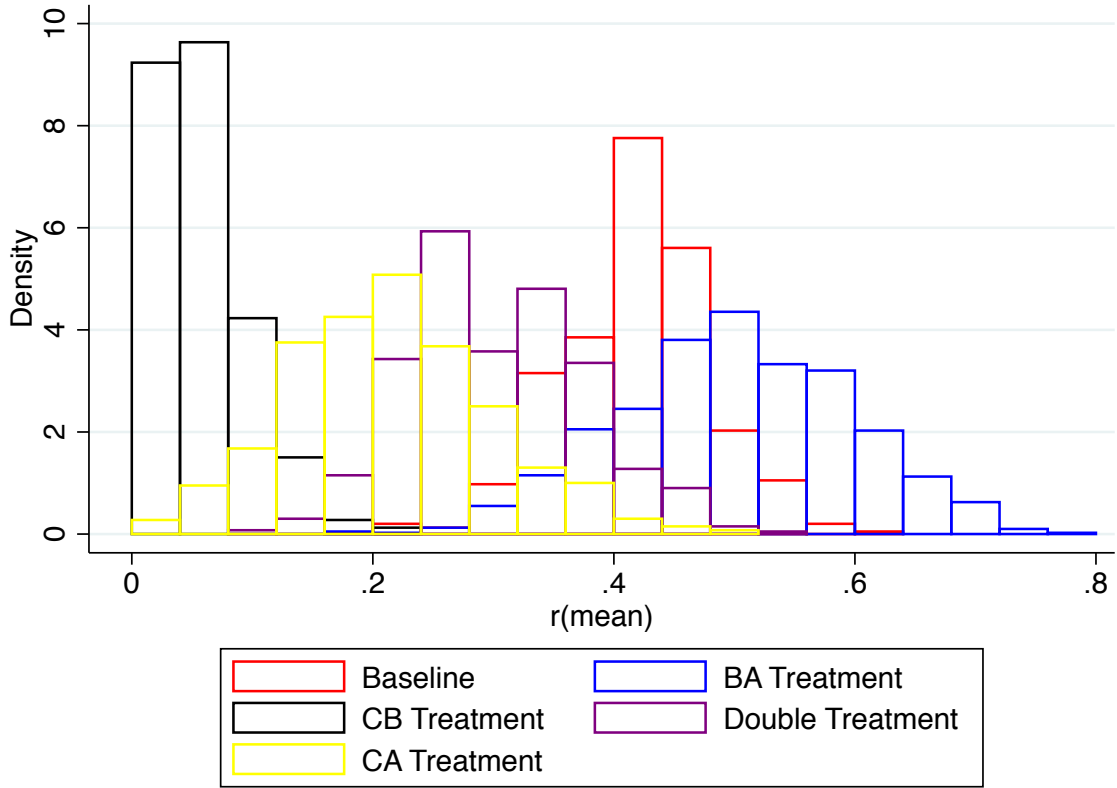
Figure 12: Bootstrap distribution of frequency of Nash strategy profile (*In, Out, Right*) by treatments (Resample $N = 999$)

5.3 Hypothesis Testing

In summary, our experimental data support Hypotheses 1 –3 and Hypothesis 6, while failing to support Hypotheses 4 and 5.

Result 1 *Treatment effects compared with Baseline.*

Tables 5 and 6 below present the Wilcoxon-Mann-Whitney two-sided test results with the null hypotheses that $^{B}_{Base}\varphi_{Out} = \ ^{B}_{t'}\varphi_{Out}$, where $t' \in \{BA, CB, Double\}$. As predicted by all theories, behaviors in the B-A treatment are statistically indistinguishable from Baseline. We observe significantly fewer Bs and Cs choosing Out(or Right in both the C-B and Double treatments than in Baseline; thus, we can invalidate the EJ theory with $l < 15$. However, we cannot find support for Hypothesis 4; instead we did not observe any treatment differences between the C-A treatment and Baseline. On the contrary, all existing theories predict negative treatment effects for both B and C, and we can therefore invalidate all theories under consideration.
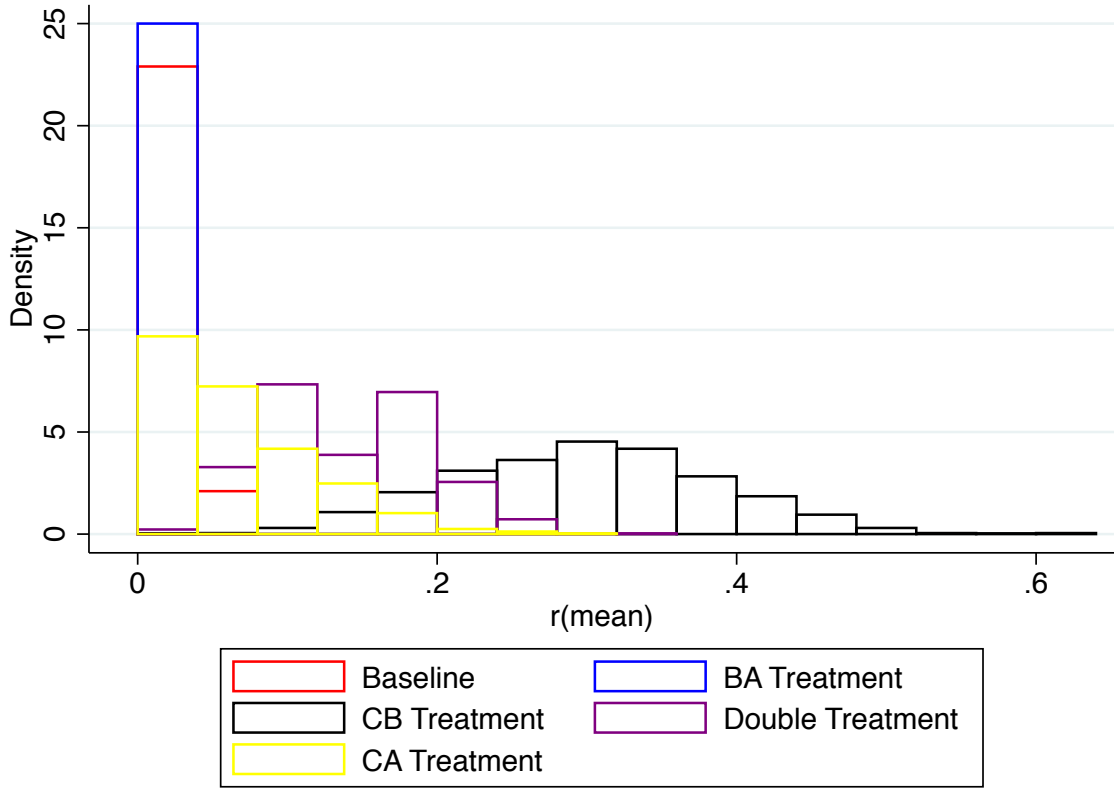
Figure 13: Bootstrap distribution of frequency of Pareto improving strategy profile (*In, In, Left*) by treatments (resample $N = 999$)

Result 2 *Within treatment comparison.*

Table 7 presents the Wilcoxon-Mann-Whitney two-sided test results with the null hypotheses that $^B_{CB}\varphi_R = {}^B_{t"}\varphi_R$, where $t" \in \{CA, Double\}$. There is significant difference in Cs behavior between the C-A and C-B treatments, which is contrary to what all theories predicted. However, this difference may result from the fact that there are significantly fewer informal contracts/commitments exchanged in the C-A treatment than in the C-B treatment. Additionally, we conducted another test to determine whether there are behavioral differences between the two treatments among those who sent informal contracts/commitments. We fail to reject the null hypothesis ($p = 0.40$) that there are behavioral differences, although this failure to reject may also arise from the small number of observations ($n = 16$ for C-B treatment and $n = 7$ for C-A treatment). Comparing C-B with the Double treatment, we cannot reject the null hypothesis that $^C_{CB}\varphi_R = {}^C_{Double}\varphi_R$ and the evidence is more in support of the CD model.

26

Table 5: Wilcoxon-Mann-Whitney two-sided test results for Player B

Frequency of Bs Choosing **Out**: $_t^B \varphi_{Out}$				
Baseline	B-A	C-B	C-A	Double
51/67 (76%)	21/24 (87%) $p = 0.24$	10/24 (42%)*** $p = 0.00$	17/27 (63%) $p = 0.20$	24/43 (56%)** $p = 0.03$

Note: *,**, and *** indicate $p < 0.10$, 0.05 and 0.01, respectively, two-sided tests.

Table 6: Wilcoxon-Mann-Whitney two-sided test results for Player C

Frequency of Cs Choosing **Right**: $_t^C \varphi_R$				
Baseline	B-A	C-B	C-A	Double
49/67 (73%)	17/24 (71%) $p = 0.83$	10/24 (42%)*** $p = 0.01$	18/27 (67%) $p = 0.53$	24/43 (56%)* $p = 0.06$

Note: *,**, and *** indicate $p < 0.10$, 0.05 and 0.01, respectively, two-sided tests.

To summarize, we find only limited support for the existing theories. For a detailed comparison between the theory predictions and observed behaviors, please refer to Tables 8 and 9 below.

Table 7: Wilcoxon-Mann-Whitney Two-sided Test Results For Player B and C (Within Treatment)

Frequency of Bs Choosing **_Out_**: $_t^B\varphi_R$			Frequency of Cs Choosing **_Right_**: $_t^C\varphi_R$		
C-B	C-A	Double	**C-B**	C-A	Double
10/24	17/27	24/43	10/24	18/27	24/43
(42%)	(63%)	(56%)	(42%)	(67%)*	(56%)
	$p = 0.13$	$p = 0.27$		$p = 0.08$	$p = 0.27$

Note: *,**, and *** indicate $p < 0.10$, 0.05 and 0.01, respectively, two-sided tests.

Table 8: Predictions *vs.* Observed For Player B

Frequency of Bs Choosing *Out*: $^B_t \varphi_{Out}$

	EJ model ($l < 15$)				EJ model ($l \geq 15$)/GTF model/HK model				CD model				Observed Behaviors			
	B-A	C-B	C-A	Double	B-A	C-B	C-A	Double	B-A	C-B	C-A	Double	B-A	C-B	C-A	Double
Baseline	=	=	=	=	=	>	>	>	=	>	>	>	=	>***	=	>**
B-A		=	=	=		>	>	>		>	>	>		>**	>***	>***
C-B			=	=			=	>			=	=			=	=
C-A				=				>				=				=

Note: *, **, and *** indicate $p < 0.10$, 0.05 and 0.01, respectively, one-sided tests.

Table 9: Predictions vs. Observed For Player C

Frequency of Cs Choosing *Right*: $_t^C \varphi_R$

	EJ model ($l < 15$)				EJ model ($l \geq 15$)/GTF model/HK model/CD model				Observed Behaviors			
	B-A	C-B	C-A	Double	B-A	C-B	C-A	Double	B-A	C-B	C-A	Double
Baseline	=	=	=	=	=	>	>	>	=	>***	=	>***
B-A		=	=	=		>	>	>		>***	=	=
C-B			=	=			=	=			<***	=
C-A				=				>				=

Note: *, **, and *** indicate $p < 0.10$, 0.05 and 0.01, respectively, one-sided tests.

6   Discussion

We find, as have many others, that communicationthe formation of informal contractsimpacts behavior in the game we study. One of the important new observations from this paper is that the extent to which a person feels behaviorally bound by the contract they form depends on whether that contract is formed with a person with whom they directly interact. We call this a "contingency effect". Specifically, the likelihood with which people will follow an informal contract (or plan) declines according to the number of contingencies that must occur in order for the contract to be realized. To our knowledge, this pattern is unaccounted-for by existing theories, and has not been observed empirically due to the fact that previous studies have focused on the effect of direct communication between two individuals (or groups). Below, we offer several potential explanations for the contingency effect.

Charness (2000) proposed a responsibility alleviation effect to explain the increased generosity from the subjects in a gift-exchange game when wages are determined by a random process rather than assigned by a third party. The responsibility alleviation effect states that people's innate pro-sociality is moderated when they can shift the responsibility of the final outcome. Similarly, Ellman and Pezanis-Christou (2010) demonstrate that vertical decision-making structure sharply diffuses each individuals sense of personal responsibility, thereby reducing pro-social behaviors. In our game, C's decision is only partially responsible for the outcome of A, as B has the option to choose *Out* and shares the responsibility of the final outcome of the game. Consistent with our data, this predicts that C may behave in a more self-interested way and become less likely to choose the option indicated on the informal contract.

Another possibility is that players follow descriptive norms that emerge during the game (see, Bicchieri and Xiao, 2009). In our case, Player C's decision is realized only when B chooses *In*; however, in doing so, B indicates that it is appropriate to choose a selfish option. Taking this as the relevant norm, C may be more likely to conform and choose according to his/her own self-interest, even though the informal contract mandates otherwise.

7   Conclusion

Previous research indicates that unenforceable informal contracts/commitments promote trust and reciprocity. While such contracts can benefit existing exchange, they may, in dynamic environments, decrease ones willingness to explore Pareto effi-

cient opportunities. This issue arises in both business and social contexts, including industry non-compete agreements and personal relationship commitment decisions. We study an environment that differs from previous environments in three important ways: (1) we consider communication among three parties, thereby enriching the communication environment; (2) we study cases where informal contracts/commitments bind people to Pareto-inferior outcomes; and (3) we allow multiple contingencies necessary for these plans to be realized.

Our results indicate that when, between the point of commitment and the point of decision, no new information is received regarding the expected value of commitment, then people overwhelmingly decide in accord with their informal contracts and avoid exploring potentially Pareto improving opportunities. However, when new information arises that reduces the relative value of commitment, and when this occurs following the commitment but before the decision, then people are significantly more likely to deviate from their informal contract and pursue a Pareto improvement. And lastly, we find that the extent to which a person feels behaviorally bound by the contract they form depends on whether that contract is formed with a person with whom they directly interact – the contingency effect. When contingency of such informal contracts increases (even by one level), the binding power of such contracts drastically decreases.

We also present predictions from three types of existing relevant behavioral models based on intrinsic preference, consequence, and types, respectively. When communication is allowed, all theories offer some degree of trust and cooperation, although through different mechanisms. Regarding treatment differences, simple type models share predictions with intrinsic preference-based models, though the mechanism differs. While intrinsic preference-based models make predictions that allow little behavioral deviation from informal contract across different treatments, models based on consequences, *e.g.*, the CD model, allow some degree of deviation.

As for the behavioral pattern we observed across treatments, each of the three types of theories could capture these patterns to some degree. For example, all theories predict treatment differences from baseline that are consistent with our observations; the difference we observe between the Double and C-B treatments is more consistent with consequence–based models. However, the contingency effect we uncovered was not considered by any of the existing theories, suggesting that future behavioral models that permit communication to foster trust and cooperation may also need to take contingency into account to better capture the observed behaviors.

References

David Amdur and Ethan Schmick. Does the direct-response method induce guilt aversion in a trust game? 2012.

P Battigalli and Martin Dufwenberg. Guilt in games. *The American economic review*, 97(2):170–176, 2007. URL http://www.jstor.org/stable/10.2307/30034441.

Pierpaolo Battigalli, Gary Charness, and Martin Dufwenberg. Deception: The role of guilt. *Journal of Economic Behavior & Organization*, pages 1–6, April 2013. ISSN 01672681. doi: 10.1016/j.jebo.2013.03.033. URL http://linkinghub.elsevier.com/retrieve/pii/S0167268113000784.

Cristina Bicchieri and Erte Xiao. Do the right thing: but only if others do so. *Journal of Behavioral Decision Making*, 22(2):191–208, 2009. ISSN 08943257. doi: 10.1002/bdm.621. URL http://doi.wiley.com/10.1002/bdm.621.

M Bigoni, S Bortolotti, and Marco Casari. Trustworthy by Convention. *Working Paper*, 2012. URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2057607.

Alessandra Cassar and Mary Rigdon. Trust and trustworthiness in networked exchange. *Games and Economic Behavior*, 71(2):282–303, March 2011. ISSN 08998256. doi: 10.1016/j.geb.2010.04.003. URL http://linkinghub.elsevier.com/retrieve/pii/S0899825610000643.

G. Charness and M. Dufwenberg. Promises and partnership. *Econometrica*, 74(6):1579–1601, 2006. URL http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0262.2006.00719.x/abstract.

Gary Charness. Responsibility and effort in an experimental labor market. *Journal of Economic Behavior & Organization*, 42(3):375–384, July 2000. ISSN 01672681. doi: 10.1016/S0167-2681(00)00096-2. URL http://linkinghub.elsevier.com/retrieve/pii/S0167268100000962.

Gary Charness and Martin Dufwenberg. Participation. *The American Economic Review*, 101(June):1211–1237, 2011.

Gary Charness, Ramón Cobo-Reyes, Natalia Jiménez, Juan a Lacomba, and Francisco Lagos. The Hidden Advantage of Delegation: Pareto Improvements in a Gift Exchange Game. *American Economic Review*, 102(5):2358–2379, August 2012. ISSN

0002-8282. doi: 10.1257/aer.102.5.2358. URL http://pubs.aeaweb.org/doi/abs/10.1257/aer.102.5.2358.

Gary Charness, Ninghua Du, Chun-lei Yang, and Lan Yao. Promises in Contract Design. 2013.

Martin Dufwenberg and Georg Kirchsteiger. A theory of sequential reciprocity. *Games and Economic Behavior*, 47(2):268–298, May 2004. ISSN 08998256. doi: 10.1016/j.geb.2003.06.003. URL http://linkinghub.elsevier.com/retrieve/pii/S0899825603001908.

Tore Ellingsen and Magnus Johannesson. PROMISES , THREATS AND FAIRNESS *. *The Economic Journal*, 114:397–420, 2004.

M Ellman and P Pezanis-Christou. Organizational Structure, Communication and Group Ethics. *The American Economic Review*, 100(5):2478–2491, 2010. URL http://www.jstor.org/stable/10.2307/41038770.

Ernst Fehr and Klaus M Schmidt. A Theory Of Fairness, Competition, and Cooperation. *Quarterly Journal of Economics*, 114(3):817–868, 1999. ISSN 00335533. doi: 10.1162/003355399556151. URL http://www.mitpressjournals.org/doi/abs/10.1162/003355399556151.

J Geanakoplos, David Pearce, and E Stacchetti. Psychological games and sequential rationality. *Games and Economic Behavior*, 1(1):60–79, 1989. URL http://www.sciencedirect.com/science/article/pii/0899825689900055.

Rajna Gibson, Carmen Tanner, and Alexander F Wagner. Preferences for Truthfulness: Heterogeneity among and within Individuals. *American Economic Review*, 103(1):532–548, February 2013. ISSN 0002-8282. doi: 10.1257/aer.103.1.532. URL http://pubs.aeaweb.org/doi/abs/10.1257/aer.103.1.532.

U. Gneezy. Deception: The Role of Consequences. *The American Economic Review*, 95(1):384–394, 2005. URL http://www.jstor.org/stable/10.2307/4132685.

Sjaak Hurkens and Navin Kartik. Would I lie to you? On social preferences and lying aversion. *Experimental Economics*, 12(2):180–192, 2009. ISSN 13864157. doi: 10.1007/s10683-008-9208-2. URL http://www.springerlink.com/index/10.1007/s10683-008-9208-2http://www.springerlink.com/index/h8078018898516w0.pdf.

NL Kerr and Cynthia M Kaufman-Gilliland. Communication , Commitment , and Cooperation in Social Dilemmas. *Journal of Personality and Social*, 66(3):513–529, 1994. URL http://psycnet.apa.org/journals/psp/66/3/513/.

James L. Loomis. Communication, the development of trust, and cooperative behavior. *Human Relations*, 12:305–315, 1959.

Tobias Lundquist, Tore Ellingsen, Erik Gribbe, and Magnus Johannesson. The aversion to lying. *Journal of Economic Behavior & Organization*, 70(1-2):81–92, May 2009. ISSN 01672681. doi: 10.1016/j.jebo.2009.02.010. URL http://linkinghub.elsevier.com/retrieve/pii/S0167268109000523.

T. Miettinen and S. Suetens. Communication and Guilt in a Prisoner's Dilemma. *Journal of Conflict Resolution*, 52(6):945–960, July 2008. ISSN 0022-0027. doi: 10.1177/0022002708322325. URL http://jcr.sagepub.com/cgi/doi/10.1177/0022002708322325.

Thomas a. Rietz, Roman M. Sheremeta, Timothy W. Shields, and Vernon L. Smith. Transparency, efficiency and the distribution of economic welfare in pass-through investment trust games. *Journal of Economic Behavior & Organization*, October 2012. ISSN 01672681. doi: 10.1016/j.jebo.2012.09.019. URL http://linkinghub.elsevier.com/retrieve/pii/S0167268112002004.

D. Sally. Conversation and Cooperation in Social Dilemmas: A Meta-Analysis of Experiments from 1958 to 1992. *Rationality and Society*, 7(1):58–92, January 1995. ISSN 1043-4631. doi: 10.1177/1043463195007001004. URL http://rss.sagepub.com/cgi/doi/10.1177/1043463195007001004.

Roman M. Sheremeta and Jingjing Zhang. Three-Player Trust Game With Insider Communication. *Economic Inquiry*, pages no–no, March 2013. ISSN 00952583. doi: 10.1111/ecin.12018. URL http://doi.wiley.com/10.1111/ecin.12018.

Christoph Vanberg. Why Do People Keep Their Promises? An Experimental Test of Two Explanations. *Econometrica*, 76(6):1467–1480, 2008. ISSN 0012-9682. doi: 10.3982/ECTA7673. URL http://doi.wiley.com/10.3982/ECTA7673.