

**DEPARTMENT OF ECONOMICS AND FINANCE
COLLEGE OF BUSINESS AND ECONOMICS
UNIVERSITY OF CANTERBURY
CHRISTCHURCH, NEW ZEALAND**

ABC on Deals

Martin Dufwenberg, Maroš Servátka and Radovan Vadovič

WORKING PAPER

No. 14/2012

**Department of Economics and Finance
College of Business and Economics
University of Canterbury
Private Bag 4800, Christchurch
New Zealand**

ABC on Deals

Martin Dufwenberg, Maroš Servátka & Radovan Vadovič *

August 26, 2012

Abstract

We develop, and experimentally test, a behavioral model of deal-making which includes binding contracts and informal agreements as distinct but related special cases. The key assumptions: people are mostly honest; they suffer costs of overcoming temptation to renege; and they tend to split gains down the middle.

*MD: University of Arizona, University of Gothenburg, and the CE-Sifo Network; martind@eller.arizona.edu. MS: NZEEL, University of Canterbury; maros.servatka@canterbury.ac.nz. RV: University of Texas at Dallas; rxv120030@utdallas.edu. We thank Jim Andreoni, Rachel Croson, Nick Feltovich, Uri Gneezy, Joel Sobel, Johan Stennek, Fernando Vega-Redondo, and several participants in seminars for helpful comments and discussion. For financial support, MD thanks the NSF and RV thanks the Asociación Mexicana de Cultura.

1 Introduction

Bargaining theory, e.g. building on Nash (1950, 1953), focuses on binding contracts. Less attention has been given to informal (non-binding) agreements. A likely reason is that if people maximize own income, a common assumption, then there is limited scope for informal agreements to have impact.¹ A selfish agent would simply renege if this were in his interest.

Humans have tendencies that curb such opportunism. Successful entrepreneur Karl Eller, for example, wrote his book *Integrity Is All You've Got* (2005) in which that message is clear. One can justify a preference for honesty with reference to repetition or reputation, but that cannot be the whole story. Eller writes about “the happiness that comes with knowing you’ll never be ashamed to face yourself in the mirror” (p. 103). Indeed, experiments indicate that honesty matters even in non-repeat settings with anonymity guaranteed. For example, Kessler & Leider (2012) find that subjects who were offered an opportunity to enter an informal agreement often did so and then delivered although they could have profitably reneged.²

We propose that people are mainly honest, and offer a theory of deal-making which treats informal agreements and binding contracts as special cases. Honest folk have much to gain by striking informal agreements. Binding contracts may be infeasible (e.g. in developing countries with unreliable courts), illegal (e.g. for cartelists), or costly (e.g. nuptials). This begs questions regarding the shape and impact of informal agreements, and how they compare with binding contracts. Our theory delivers predictions.

Assuming that honesty is prevalent does not imply that informal agreements and binding contracts look the same. We highlight two reasons. First,

¹Informal agreements may allow Pareto improvements in games with multiple equilibria; see e.g. MacLeod & Malcolmson (1989), McCutcheon (1997), Levin (2003).

²Irlenbusch (2004), Ben-Ner & Putterman (2009), and (from social psychology) Malhotra & Murnighan (2002) present related results. Several studies have to boot shown that people may have preferences to not deceive or to honor promises. See for example, Kerr & Kauffman-Gilliland (1994), Ellingsen & Johannesson (2004), Gneezy (2005), Charness & Dufwenberg (2006), Vanberg (2008), and Servátka, Tucker & Vadovič (2011).

lots of research indicates that humans suffer cognitive costs when fighting temptations.³ An honest person fulfils the terms he agrees to, and yet he may not be immune to costs incurred when overcoming temptation to renege. The anticipation may affect his evaluation of what deals are worth striking. Second, while we assume most people are honest, we do not assume all are. When deciding whether to enter an informal agreement, a party takes account of the costs associated with opportunistic counter-part behavior.

Apart from responding to these costs we propose that informal agreements look like binding contracts. So how do binding contracts look? Classical bargaining theory provides answers. However, we are skeptical of the relevance of this scholarship for describing the behavior of honest people. Much of classical theory assumes players know each others' utilities. As an empirical proposition this was questioned early on – Luce & Raiffa (1957, p. 134) find it “extremely doubtful” – but over time scholars seem to have not worried too much. Again this may be because economists often accept the idea that utility equals (or is monotone in) own money, making the issue largely moot. However, for honest deal-makers own monetary gain is not all that matters. More generally, modern research on social preferences shows that people often care not only about their own monetary rewards. Experimental research has revealed many insights regarding what people care about instead; nevertheless there is a lot of debate and little consensus.⁴

If human motivation is complicated, rich, and hard-to-pin-down in mathematical formulae, perhaps one should not be surprised if humans strike deals that refer to readily observable data like dollars earned? Indeed, many agreements often simply divide dollar gains equally.⁵ We assume that binding contracts have this property. We then propose that informal agreements

³See Benartzi & Thaler (2004), Brown, Chua & Camerer (2009), and Martinsson, Myrseth & Wollbrant (2012) for experiments and Thaler & Shefrin (1981), Gul & Pesendorfer (2001), Loewenstein & O'Donoghue (2005), Fudenberg & Levine (2006), Salant, Silverman & Ozdenoren (2011), and Lipman & Pesendorfer (2011) for theory.

⁴For surveys, see Camerer (2003) and Fehr & Schmidt (2006).

⁵See, e.g., Andreoni & Bernheim (2010) or Binmore, Shaked & Sutton (1989).

work the same way except that new participation constraints enter the picture to take into account costs of temptation and co-player opportunism.

The framework in which we formulate these statements precisely is game-theoretic and structured as follows:

1. An agreement, be it binding or informal, may be reached by a group of players about to play a game.
2. The object of negotiation concerns which strategy profile to play.
3. If the agreement is binding then the players have to play accordingly.
4. If the agreement is informal then the players still have to play the game.
5. If no agreement is reached then the players still have to play the game.

A key feature here is the explicit anchoring on an “underlying game” that has to be played. This allows us to be explicit about the nature of the economic situation in which an agreement may be struck, and it allows informal agreements to be truly non-binding, since post-agreement the game must still be played. The approach is, however, not idiosyncratic to informal agreements; as seen in 1-3 & 5, it allows for binding contracts as well.

The underlying game does not describe the players’ haggling. Agreement-formation is instead captured implicitly, through a solution-concept: Let S be the set of strategy profiles in the underlying game. We select $a, b, c \in S$ such that a is the agreement, b is the behavior following the agreement, and c the conflict outcome that would happen if the players failed to agree. Our predictions are formulated in terms of restrictions on a, b , and c .

In envisaging agreements as strategy profiles we connect to a classic tradition. Von Neumann & Morgenstern (1944) approach all games other than two-player zero-sum ones with this outlook (see e.g. pp. 223-4). Nash (1953) assumes players strike binding contracts regarding which strategy profile to play in some underlying game, and before negotiations start they announce “threats” (strategies) which an “umpire” forces them to implement if they

subsequently fail to reach an agreement.⁶ We share the outlook that strategy profiles are objects of negotiation, but neither limit attention to binding contracts nor presuppose access to an umpire.⁷

The contribution of our paper has two parts. Developing the theory is #1. What's next? One option would be to explore applications, say prenups, cartel collusion, auction bidding rings, political logrolling, or diplomacy. However, it seems premature to spend such effort before one has direct evidence that the theory's key features make empirical sense. Our contribution part #2 reports on an experiment designed to provide a relevant testbed.

Sections 2, 3, & 4 present theory, experiment, & concluding remarks.

2 Theory

This section describes and interprets the game-theoretic framework (2.1) and develops our solution concept (2.2).

2.1 General Framework

Our analysis of a bargaining scenario starts with a two-player extensive game form Γ with given material/dollar payoffs specified at end nodes. Let S_i be player i 's set of strategies (taken to be singleton if i owns no information set), and $S = S_1 \times S_2$. Let $m_i : S \rightarrow \mathbb{R}$ be i 's dollar-payoff-from-strategy-profile function, derived from Γ . We'll write $m_i(s \setminus s'_j)$ to indicate i 's dollar payoff if j unilaterally deviates from $s = (s_i, s_j)$ to use s'_j instead of s_j .

⁶See Kalai (1977) and Kalai & Tauman Kalai (2010) for more work in this vein.

⁷The cheap talk literature (see e.g. Crawford & Sobel 1982, Farrell & Rabin 1996) also studies the effect of communication in games. Unlike our approach, cheap talk is modeled as explicit choices and, most importantly, presumed not to affect preferences (over strategy profiles) in the underlying game. By contrast, in our approach players have a preference for playing as they agree, so talk is not cheap. There is also the game-theoretic literature on communication equilibria (e.g. Forges 1986, Myerson 1986), which (like us) captures the effect of messages through solution concepts but (like the cheap talk literature) assumes communication does not affect preferences over strategy profiles.

This underlying game Γ describes the strategic structure of a situation where two persons just “met” and face some opportunity of collaboration for mutual gain. The payoffs represent dollar increments relative to whatever wealth the two had before. If someone gets a payoff of 0 when the underlying game is played, we interpret this to mean that his overall dollar wealth remains the same as if he had never met the other player.

We assume that the underlying game allows enough opportunities for the players to transfer money between each other that they can achieve the following: Let H denote the highest sum of the players’ dollar payoffs at any endnode in Γ . We assume that $H > 0$ and that if $\$1, \$2 \geq 0$ and $\$1 + \$2 = H$ then Γ admits some endnode with dollar payoffs $(\$1, \$2)$. This implies, among other things, that Γ allows infinite choice sets and that there is a way to play Γ that results in equal dollar payoffs for each player.

We assume that Γ is a multi-stage game form with observed actions (see Fudenberg & Tirole 1991, Chapter 3), so that all instances of imperfect information concern simultaneous choices. This restriction simplifies the key definitions below by allowing us to refer to subgames in a useful way, without essentially compromising the scope of the model since most applied and experimental work is concerned with such games.

The players haggle about which strategy profile to play. Γ does not give an explicit description of this process. Rather, we capture its effect through a solution concept with a special structure. An agreement, which may be formal or informal, takes the form of a strategy profile which we denote $a \in S$. Of course, a also describes off-path play. The interpretation differs depending on whether a is an informal agreement or a binding contract. In the former case the off-path part of a reflects the players’ agreed upon understanding, presumably obtained through the un-modeled haggling process, of what would happen following any deviation. In the latter case the off-path part of a should be neglected since the involved choices are inconceivable (think of a as representative of the equivalence class of strategy profiles that generate the same path).

When the negotiation phase is over the players have to play the game. What will they do? Our framework asks for specification of two more strategy profiles $b \in S$ and $c \in S$. b describes what actually happens in the game after agreement a is struck and c describes what would happen if negotiation failed. b and c also describe off-path play, again interpreted as reflecting the players' agreed upon understanding of what would happen following any possible deviation.⁸

Several clarifying comments are warranted. First, we theorize only about what happens when negotiations generate agreements. c should thus be understood counter-factually; it describes what would happen had a not occurred, given that a does occur. As we shall see below it is possible that in some game no deal (a, b, c) exists which satisfies the postulated properties. The interpretation is that no agreement would be reached in that case. We offer no explicit prediction for play following such non-counterfactual negotiation-breakdown.

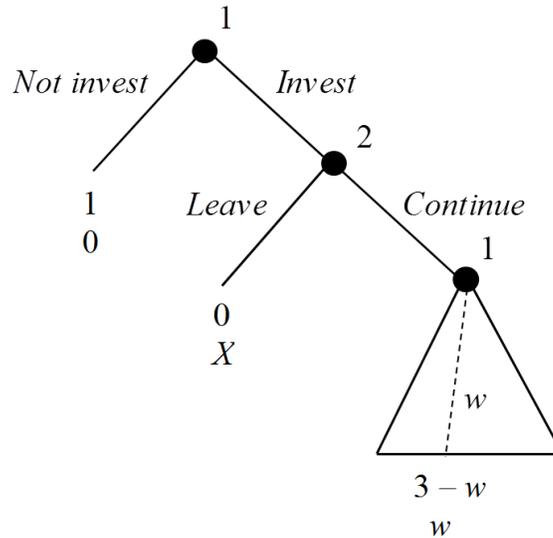
Second, with binding contracts $b = a$ by definition, but with non-binding agreements it is conceivable that $b \neq a$. Third, as regards c , one could imagine a richer structure such that c depend on how negotiations stranded (e.g., which player caused the break down). We abstract away from such nuances. Fourth, we interpret off-path parts of a, b, c as reflecting common understanding by all players regarding off-path behavior. One could imagine alternatives, e.g., as in a self-confirming equilibrium (Fudenberg & Levine 1993; cf. Greenberg 2000). We do not do so in this paper.

Fifth, we elucidate why we do not explicitly model the strategic structure of the pre-play negotiation. Consider the example in Figure 1, where $1 < X < 3$, which comes with a story:

Player 1 is hospital and player 2 is an employed radiologist. A new radiography technique has been made available, and at the root 1 decides whether to *Invest* or *Not invest* in costly training for 2 to learn the new methods.

⁸With binding agreements $b = a$ and off-path play is by definition not possible so (as with a) the off-path part of b should be neglected in this case.

Figure 1: Hospital-Doctor game



In the former case 2 becomes more productive but also more attractive to other hospitals; choice *Leave* with subsequent payoffs reflects what happens if 2 resigns and takes employment at Johns Hopkins. That would be bad for 1 who stands to gain if 2 instead *Continues* at the current job. In that case, 1 can choose what wage $w \in [0, 3]$ to pay 2, thereby determining 1's life-time income.⁹

The point of this example is that the game, like many others, can describe a *somewhat meaningful* situation where deals can be made. However, that situation also seems unrealistically barren, as it incorporates no opportunities for wage negotiation, promises, threats, etc. A *more meaningful* situation would arguably be the same one augmented so that the players can meet and strike agreements regarding whether player 1 should pay for the training, and what player 2's pension should be.

How should one model such considerations? How should one make pre-

⁹The implicit assumption is that later in 2's life he has fewer outside opportunities and is therefore vulnerable to hold-up.

dictions? One possibility may be to change the game, to include every offer, every counter-offer, every promise, every threat, every signing on a dotted line, etc., as explicit choices in a larger game. But that game is likely to get unmanageable. The players' strategy sets would be very complex. It may be intractable to apply an adequate solution concept. It is against this backdrop that we propose our approach. It is reminiscent of cooperative game theory in that we do not model every aspect of the negotiation, yet it is non-cooperative in that the strategic structure of the situation to which we add a negotiation possibility is allowed to influence our analysis.

What agreements may one expect to occur? That depends on many things, including the nature of the players' motivation, whether the agreement is binding or non-binding, and game details such as X . We propose a generally applicable theory next.

2.2 The Model

What psychological and economic principles determine a deal $(a, b, c) \in S^3$? Many answers are conceivable. We propose a particular one, as follows:

$$c \in S : \textit{Selfish SPE}$$

Before we describe the agreement ($a \in S$) and post-agreement behavior ($b \in S$), let us ponder what would have happened should negotiations fail ($c \in S$). We assume that players would behave selfishly in the sense that c is a subgame perfect equilibrium (SPE) of Γ using the dollar payoffs.

This assumption is not obvious, as it is known from many studies that players are often not selfish (cf. footnote 4). However, we justify the assumption as follows: On the one hand, in many game forms (e.g. prisoners' dilemmas or trust games or public goods games) that exhibit a tension between individual and collective dollar-payoff-maximization, subjects manage to reach efficient outcomes. This suggests that players often appreciate the well-being of others. On the other hand, it does not seem unreasonable to

suppose that players who do not manage to agree would end up being irritated with one another. All in all, it seems unclear whether $c \in S$ should reflect that players would be friendly or hostile towards one another. Our assumption that players would act selfishly takes a middle road.¹⁰

b ≈ a: Pacta sunt servatka

With a binding contract, $b = a$ by definition. Informal agreements, on the other hand, are not explicitly enforced. We assume that if an informal agreement $a \in S$ is struck, most players i honor the agreement and choose $b_i = a_i$. Some economists may find this assumption extreme. Our experiences from life as well as readings of related literature (cited in the introduction, including footnote 2) tell us it actually makes sense. It is indeed a stark antidote to the more typical economists' assumption that people are selfish. But ultimately it is an empirical question whether most people are honest, one than can be experimentally tested and potentially rejected.

Honest behavior can be actually justified with reference to a variety of social preferences. For example, if one thinks of informal agreements as embodying a process where players make promises, then several authors have suggested that players may prefer to keep promises or (more generally) not to have lied.¹¹ One may alternatively imagine that players simply obey some social norm that says that one should honor agreements.¹² Or players could be guilt averse, inclined not to hurt others relative to their expectations;¹³ negotiation may then foster common and self-fulfilling beliefs that an informal agreement will be honored. Honesty may also be backed up by reciprocity, if the agreed on strategy profile plus correct beliefs imply that the players

¹⁰Note that there is scant data to guide our modeling choice: existing data on the relevance of social preferences typically neither concern games played after negotiations break down nor concern what would happen after counter-factual negotiation breakdown.

¹¹See the references in footnote 2 (except Malhotra & Murnighan) plus Demichelis & Weibull (2008) and Kartik (2009).

¹²See Malhotra & Murnighan, Kessler & Leider, and Miettinen (2011).

¹³See e.g. Charness & Dufwenberg (2006), Battigalli & Dufwenberg (2007).

are kind to one another.¹⁴ Or players may be uncertain about the nature of their own preferences and stick with the agreement to avoid negative self-signalling.¹⁵ While it is an interesting issue to determine which motivational story may be most empirically relevant, in this paper we just assume that agreements are mostly honored.

What choices should one expect from those who renege? Are they perhaps spiteful and minimize others payoffs? Do they trade off their own gain against the loss of their co-players? While these possibilities do not seem unreasonable, we shall assume that renegers are selfish and maximize their own payoff. This way we obtain fairly easy to state definitions and a clear benchmark to test with data.

a ∈ S : Equal gains

Recall our remarks from the introduction regarding classical bargaining theory: While that scholarship is conceptually related to our (a, b, c) -approach, it is of questionable applied relevance as regards the behavior of honest individuals. Most classical models presume that utilities are commonly known, but modern research on social preferences has in our view made it clear that utilities are nowhere near commonly known. Is it not reasonable, then, that bargainers give scant reference to such complicated, rich, and hard-to-pin-down notions? Is it not likely that they instead strike deals that, most of the time, refer to readily observable data like dollars gained?

Even if dollar gains is the object of negotiation it is not obvious which split is most likely. Different contexts may trigger different thinking. In an intriguing recent paper, Isoni, Poulsen, Sugden & Tsutsui (2011) discuss Schelling's (1960) idea that outcomes under tacit bargaining (where communication is incomplete or impossible) may depend on focal points which in turn may depend on cues such as object proximity, existing location of bargaining parties, salience of geographical boundaries (e.g. a river), precedence

¹⁴See Rabin (1993), Dufwenberg & Kirchsteiger (2004), Falk & Fischbacher (2006).

¹⁵See Benabou & Tirole (2002).

of supply chains, or a historical consumer base. Isoni *et al* experimentally test Schelling’s theory and find support. We suggest that these ideas may naturally extend beyond tacit bargaining, to explicit haggling, which is our focus. That said, while it would seem an exciting long run goal to merge the ideas of Schelling + Isoni *et al* with our framework, in this paper we focus on a simpler norm which may nevertheless be very relevant in many contexts (that perhaps lack salient locations, rivers, or historical antecedents): 50/50 splits.

Andreoni & Bernheim (2010, p. 1607) reference a variety of studies documenting prevalence of equal splits of dollar gains (e.g. joint ventures between corporations, share tenancy in agriculture, bequests to children, negotiation and arbitration, business partners splitting earnings from joint projects, or friends splitting tabs). Some experiments point in the same direction. Binmore, Shaked & Sutton (1989) let subjects bargain over sums of money. If one of the parties breaks the negotiations both get prespecified outside option payoffs. It turns out that these do not influence the agreements, except as imposing lower bounds on what each party will get. The money is simply split equally, as long as each party at least gets his outside option. Binmore *et al* call this a “deal-me-out solution.”¹⁶

Our solution may be viewed as a version of deal-me-out. It modifies or adds to Binmore *et al*’s notion in two ways: First, we assume that the object of the negotiation is a strategy profile ($a \in S$) rather than a sum of money. Second, we do not restrict attention to binding contracts but also consider informal agreements.

We propose that deals generate equal dollar payoffs, as long as each player i is made no worse off than when play proceeds according to $c \in S$. If the deal involves a binding contract then what this means is straightforward: i is made no worse off if $m_i(a) \geq m_i(c)$. (Think of $m_i(c)$ as the counterpart in our theory to Binmore *et al*’s outside option for i .)

¹⁶For further evidence supporting deal-me-out, see Binmore, Proulx, Samuelson & Swierzbinski (1998) and (much less conclusively) Feltovich & Swierzbinski (2011).

If the deal involves an informal agreement then it is not as obvious whether or not i is made no worse off relative to $c \in S$. There are (at least) two reasons. First, there are the dishonest individuals who renege. Suppose that player i believes that with probability $\varepsilon_i > 0$ his co-player j will renege and choose selfishly; let the relevant strategy be $\hat{s}_j \in \arg \max_{s_j \in S_j} m_j(a \setminus s_j)$.¹⁷ Then i 's expected dollar payoff under an informal agreement will be lower than with the same agreement under a binding contract. If the difference is big enough, i may accept an equal-split inducing $a \in S$ as a binding contract but not as an informal agreement.

A second factor complicating how a player evaluates whether or not he is made no worse off concerns the temptations to renege that he has to overcome if he honors an informal agreement. There is a sizable literature on human tendency to resist temptations.¹⁸ It is often argued that humans can overcome temptation, but that this comes at a cost. If player i considers such costs when evaluating an informal agreement, then his subjective dollar gain (i.e., net of the temptation cost) under an informal agreement will be lower than with the same strategy profile as a binding contract. Again, if the difference is big enough, i may accept $a \in S$ as a binding contract but not as an informal agreement.

How should one calculate costs of overcoming temptation to renege? Are they linear or perhaps convex in how much a player may gain (cf. Fudenberg & Levine 2006 who discuss both versions)? Are they stochastic (cf. Dekel & Lipman 2011)? Do they depend on how many times along a path a player is tempted (cf. Salant, Silverman & Ozdenoren 2011), or does only the

¹⁷This formulation assumes a selfish player j does not take into account the possibility that his co-player reneges too, a feature we could easily change with little conceptual gain, much added complexity, and no importance in many games incl. those of Sect. 4.

¹⁸See the references in footnote 3. The literature mainly focuses on single decision maker settings (Loewenstein & O'Donoghue's section VI is an exception), not temptation to renege and hurt a co-player as we have, but that extension seems very plausible to us. Indeed, Martinsson *et al* reports support for "the proposition that individuals may experience a self-control conflict between the temptation to act selfishly and the better judgment to act pro-socially."

maximum temptation along the path matter? Are they moderated to the extent that renegeing hurts others (cf. Gneezy 2005), or even via some nuanced notion of “empathy” (O’Donoghue & Loewenstein 2005)? The answers are by no means obvious. We go with the following straightforward version that allows us to state definitions fairly easily. The cost of overcoming temptation is proportional to the maximum foregone dollar gains from renegeing; let the proportional factor of player i be $\gamma_i > 0$.¹⁹ With an informal agreement, i ’s cost-of-temptation associated with $a \in S$ equals

$$\gamma_i \times \left(\max_{s_i \in S_i} m_i(a \setminus s_i) - m_i(a) \right).$$

With a binding contract, by contrast, temptation-costs equal zero; there cannot be any renegeing on a binding contract, hence no temptation-cost.

We next give definitions that formally pin down, respectively, binding contracts and informal agreements. We then offer further comments on interpretation and predicted differences.

- A binding contract (BC) $a \in S$ satisfies two conditions:

BC(i) for $i = 1, 2$ and some $M > 0$ we have

$$m_i(a) = \max\{M, m_i(c)\},$$

BC(ii) M is the largest number for which BC(i) holds.

- An informal agreement (IA) $a \in S$ satisfies three conditions:

¹⁹We adopt notation γ_i from Fudenberg & Levine. Their Assumption 5’ makes a comparable linearity assumption.

IA(i) for $i = 1, 2$ and some $M > 0$ we have

$$m_i(a) = \max\{M, m_i(c) + \gamma_i[\max_{s_i \in S_i} m_i(a \setminus s_i) - m_i(a)] + \varepsilon_i[m_i(a) - m_i(a \setminus \hat{s}_j)]\},$$

where $\gamma_i, \varepsilon_i > 0$ and $\hat{s}_j \in \arg \max_{s_j \in S_j} m_j(a \setminus s_j)$,

IA(ii) M is the largest number for which IA(i) holds,

IA(iii) off its induced path, a prescribes SPE-play using the dollar payoffs.

Condition BC(i) is our take on deal-me-out for binding contracts; players simply split dollar payoffs equally, unless someone would then get less than he would at c . BC(ii) is an efficiency condition such that no money is left on the table; obviously $M = \max_{a \in S} [m_1(a) + m_2(a)]$.

Condition IA(i) captures the deal-me-out idea for informal agreements. Comparing with BC(i), one sees that i 's "participation constraint" becomes tighter; he now needs to get not only $m_i(c)$ but also be compensated for the costs of overcoming temptation and of co-player opportunism.²⁰ Note that if $\gamma_i, \varepsilon_i \rightarrow 0$ then IA(i) approaches BC(i). In this sense, relative to informal agreements, binding contracts "switch off" the costs of temptation and opportunism. IA(ii) is again an efficiency condition, but this time it may be that $M < \max_{a \in S} [m_1(a) + m_2(a)]$ (the reader may verify this by constructing an example). IA(iii) is an assumption about play after a player reneges,²¹ analogous to what lead us to assume that $c \in S$ is a subgame perfect equilibrium using the dollar payoffs.

²⁰IA(i) reflects the idea that player i factors in temptation costs evaluated with respect to a fully honest opponent, independently of ε_i . This makes sense if one assumes that any temptation cost is borne (only) at the point of agreement, before any actions are taken. This modeling has the advantage of allowing a straightforward mathematical formulation, and it is in line with the modeling choices of Fudenberg & Levine (2006).

²¹There is no need for a counterpart condition BC(iii), since off-path play is inconceivable under a binding contract.

The informal agreement definition implicitly assumes that γ_i and ε_i are commonly known by the parties. We also assume that γ_i and ε_i differ across players, so that the nature of a deal will vary across negotiation instances. There is an obvious tension between those two sentences! How can this be reconciled? We interpret the solution to presume that the parties reveal γ_i and ε_i to each other in the course of negotiations. After all, they are mainly honest.

We admit that this assumption may be quite strong. Sticking to an agreement may be one thing, revealing private information about γ_i and ε_i quite another.²² Nevertheless, the assumption is at least consistent with the idea that players are honest, and it allows us to close the model and to generate predictions for general games. Extreme as the assumption may be, we feel it may be a useful benchmark for exploring the deal-making of honest people.

Note that the predicted agreement in many cases does not change with γ_i and ε_i . For example, if γ_i and ε_i are “small enough” and there is a “large enough” distance between $m_i(a)$ and $m_i(c)$, for all i , then the players will simply go for the equal split in all cases.

a, b, c ∈ S: Square dealing

It is time to sweep everything together in our key definition. According to dictionaries, “square” can mean “straightforward and honest.” It can also mean (in math) that all sides are equal. Since we assume that most players are honest and that many deals involve straightforward equal splits, we find it appropriate to call our predictions “square deals”:

Definition: The triple $(a, b, c) \in S^3$ is a *square deal* if 1-3 hold:

1. Depending on whether a is a binding contract or an informal agreement,

²²In addition, in line with Loewenstein’s (1996) contention that “people underestimate the impact of visceral factors on their own future behavior” (his Proposition 5), we note that assuming that i knows γ_i is not entirely innocuous.

a satisfies conditions BC(i) & BC(ii) or IA(i) & IA(ii) & IA(iii).

2. If a is a binding contract or if i is honest, then $b_i = a_i$. If a is an informal agreement and i is dishonest then he chooses $\hat{s}_i \in \arg \max_{s_i \in S_i} m_i(a \setminus s_i)$,
3. c is an SPE of Γ using the dollar payoffs.

It follows from the restrictions we imposed on the underlying game Γ that if we consider binding contracts then a square deal always exists. However, if we consider informal agreements then existence is an issue. The interpretation is natural: taking costs of overcoming temptation and of co-player opportunism into account, there is no square deal available that can make both players feel at least as well off as they would be at c . We illustrate non-existence further in the next section as we analyze in detail the games we use in our experiment.

3 An Experiment

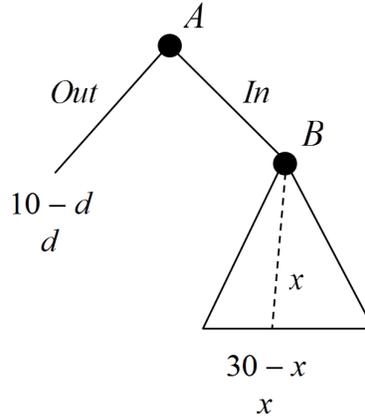
Is the theory empirically relevant? We shed light on this issue through an experiment. Next we introduce the games and derive predictions (3.1); then present the design (3.2) and report results (3.3–3.4).

3.1 Games and Predictions

We use the lost wallet game (Dufwenberg & Gneezy, 2000) presented in Figure 2, which may be viewed as a simplified version of the doctor-hospital game of section 2.1. The lost wallet game is appropriate in that it possesses several key qualities: It is easy to explain to subjects and implement in the lab, yet it is rich enough to allow a deal with equal payoffs. The theory is simple to apply and generates sharp comparative statics predictions across treatments (binding contracts vs. informal agreements, and a payoff variation). As regards $c \in S$, the game has a unique subgame perfect equilibrium

using the dollar payoffs. The two players, A and B (named as in the experiment), only have to consider ε_A and γ_B (and can ignore ε_B and γ_A), which substantially reduces complexity regarding unobservables and predictions.²³

Figure 2: The Lost Wallet game



Player A chooses In or Out . If he chooses Out , then A and B receive their respective payoffs $10 - d$ and d , where $d \in \{0, 5\}$ is a parameter (which in the experiment varies by treatment). Alternatively, if A chooses In then B divides \$30 between them. B keeps $x \in [0, 30]$ and transfers $30 - x$ to A . The subgame perfect equilibrium for selfish players is $(Out, x = 30)$, with resulting dollar payoffs $(5, 5)$ or $(10, 0)$ depending on d .

In the rest of section 3 we refer to y as the agreed-upon x and to z as the post-agreement choice of x . Note that with a binding contract $z = y$ by definition, while with an informal agreement $z \neq y$ is possible.

Now suppose that before playing the game the parties negotiate about what strategy profile to play. We consider the case where negotiations can lead to an informal agreement as well as the case when they can lead to a binding contract. With $d \in \{0, 5\}$ the dollar payoffs following Out are $(5, 5)$ and $(10, 0)$. We thus get four cases, which we accordingly label IA[5,5],

²³For instance, if we had instead used the hospital-doctor game, then ε_B and γ_A would also have entered the analysis and made the solution more complex.

IA[10,0], BC[5,5], and BC[10,0]:

Binding contracts: Reneging is ruled out by definition. This immediately takes costs of temptations as well as anticipated opportunism out of the picture. Since $30/2 = 15 > m_i(c)$ for $i \in \{A, B\}$, the theory predicts that players will agree on (and then play to) an equal split of 15 for each player (cf. BC(i)). That is, $a = (In, y = 15)$ and $m_A(a) = m_B(a) = 15$.

Informal agreements: For sufficiently low ε_A and γ_B we again obtain $a = (In, y = 15)$ as the prediction. To see this clearly, glance at condition IA(i) and note that, since $m_i(a) > m_i(c)$, for small enough ε_i and γ_i we get $m_i(a) > m_i(c) + \gamma_i[\max_{s_i \in S_i} m_i(a \setminus s_i) - m_i(a)] + \varepsilon_i[m_i(a) - m_i(a \setminus \hat{s}_j)]$. Hence, we can set $M = 15 > m_i(c)$. However, if either ε_A or γ_B is large enough we get either a compensated deal where $y \neq 15$ or no deal at all. IA(i) implies that a square deal must compensate player A so that $y < 15$ when

$$15 - \varepsilon_A(15 - 0) - \gamma_A \times 0 < 10 - d \iff \varepsilon_A > (5 + d)/15. \quad (1)$$

Suppose (1) holds. Using IA(i), we see that the compensated deal must allocate A an amount $30 - y$ such that

$$(30 - y) - \varepsilon_A[(30 - y) - 0] - \gamma_A \times 0 = 10 - d \iff 30 - y = \frac{10 - d}{1 - \varepsilon_A}.$$

Notice, however, that there is a limit on how much A can get in a square deal. Because compensation is a direct transfer from B to A , a too high demand may not be feasible in the sense that B would eventually reject it, i.e., conditions IA(i) and IA(ii) would fail to hold for B . Specifically, player B is willing to compensate player A if

$$(30 - \frac{10 - d}{1 - \varepsilon_A}) - \varepsilon_A \times 0 - \gamma_B(\frac{10 - d}{1 - \varepsilon_A}) \geq d \iff \gamma_B \leq \frac{20 - \varepsilon_A(30 - d)}{10 - d}. \quad (2)$$

If this inequality is violated, then there is no informal agreement that qualifies as a square deal. This exemplifies well the following key feature of the theory: An agreement which is a square deal under binding contractual

arrangement need not be a square deal when it is informal and not explicitly enforced. With a binding contract, each player perceives that he is better off than under the conflict outcome (since $m_i(a) > m_i(c), i \in \{A, B\}$). With an informal agreement being better off means not only comparing $m_i(a)$ and $m_i(c)$ but also factoring in the costs of overcoming temptation and of the belief about the other player's opportunism. If ε_A and γ_B are large enough it is not possible to compensate both parties, so no informal agreement is viable as a square deal.

To complete the characterization of square deals for this game similar calculations need to be made for player B . Suppose the agreement is for him to choose y . IA(i) implies that a square deal must compensate B so that $y > 15$ when

$$15 - \varepsilon_A \times 0 - \gamma_B \times 15 < d \Leftrightarrow \gamma_B > (15 - d)/15, \quad (3)$$

in which case he must get

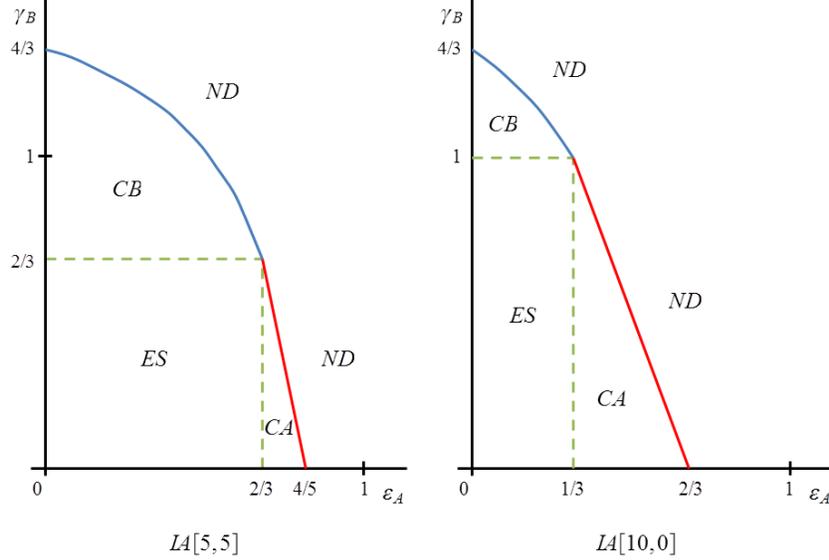
$$y = d + \gamma_B \times 15.$$

The deal can of course be accepted by player A only if

$$\begin{aligned} [30 - (d + \gamma_B 15)] - \varepsilon_A \times ([30 - (d + \gamma_B 15)] - 0) - \gamma_A \times 0 &\geq 10 - d \\ \Leftrightarrow \gamma_B &\leq \frac{20 + \varepsilon_A(30 - d)}{15(1 - \varepsilon_A)}. \end{aligned} \quad (4)$$

Figure 3 gives a graphical overview of these predictions. In the left panel $d = 5$; in the right panel $d = 0$. Each point corresponds to a player A with ε_A and a player B with γ_B . For each case we divide the parameter space into four different regions depending on the predicted deals: ES is the region with an equal split (bounded by (1) and (3)); CA is the region where the deal compensates player A (between (1) and (2)); CB is the region where the deal compensates player B (between (3) and (4)); ND denotes all $(\varepsilon_A, \gamma_B)$ pairs that do not admit any square deal.

Figure 3: Predictions



The corresponding regions in the two treatments vary in size and position. Moreover regions ES and regions ND have different shapes and nontrivial overlap. Because of this, predictions will depend on the distribution of types in our subject population. Ex-ante we have no way of knowing what the distribution of types might be. However, many distributional assumptions would admit qualitative conclusions based simply on a comparison of the relevant region-areas in Figure 3. This would be the case for example for any distribution that has strictly positive density across $[0, 4/5] \times [0, 4/3]$ which is non-increasing in directions away from the origin. For any such distribution, letting $ES_{IA[5,5]}$ denote the region ES in the IA[5,5] panel of Figure 3 etc, we get the following predictions:

1. **Deals compensating B:** The region $CB_{IA[5,5]}$ is a strict superset of (the closure of) the region $CB_{IA[10,0]}$ (in $(\varepsilon_A, \gamma_B)$ -space). It follows that $CB_{IA[5,5]}$ is more likely than $CB_{IA[10,0]}$, so deals compensating A should

be more frequent in $IA[5, 5]$ than in $IA[10, 0]$.

2. **Deals compensating A:** The area of region $CA_{IA[10,0]}$ is larger than area of region $CA_{IA[5,5]}$; moreover $CA_{IA[5,5]}$ is located east of $CA_{IA[10,0]}$ (in $(\varepsilon_A, \gamma_B)$ -space). It follows that $CA_{IA[10,0]}$ is more likely than $CA_{IA[5,5]}$. Hence, deals compensating A should be more frequent in $IA[10, 0]$ than in $IA[5, 5]$.
3. **No deals:** $CA_{IA[5,5]} \cup CB_{IA[5,5]} \cup ES_{IA[5,5]}$ is a strict superset of (the closure of) $CA_{IA[10,0]} \cup CB_{IA[10,0]} \cup ES_{IA[10,0]}$, which implies that $ND_{IA[10,0]}$ is more likely than $ND_{IA[5,5]}$. Hence, an outcome with no deal should be more frequent in $IA[10, 0]$ than in $IA[5, 5]$.
4. **Equal split deals:** The area of $CA_{IA[5,5]}$ is smaller than the area of $ES_{IA[5,5]}$; moreover, $CA_{IA[5,5]}$ is located east of $ES_{IA[5,5]}$. It follows that $ES_{IA[5,5]}$ is more likely than $CA_{IA[5,5]}$. By similar reasoning, $ES_{IA[5,5]}$ is more likely than $CB_{IA[5,5]}$, and $ES_{IA[10,0]}$ is more likely than either of $CA_{IA[10,0]}$ and $CB_{IA[10,0]}$. Hence, in each treatment, equal split deals should be more frequent than any particular kind of deals compensating one of the parties.²⁴

3.2 Design & Procedures

The experiment was computerized and conducted at the University of Arizona's Economic Science Laboratory. The software was written in Visual Basic 6. In total, 204 undergraduate students participated as subjects; the sessions and participation is summarized in Table 7 in Appendix A. Subjects played one game – no repetitions – and were then privately paid.

Once all subjects were seated at computer terminals separated by privacy dividers, hard copies of instructions were handed out (see Appendix B) and subjects were given 10-15 minutes to read them. When everyone

²⁴The nontrivial overlap of the ES regions does not yield clear-cut predictions for between treatment comparison of equal split deals.

had finished reading, the instructions were also read out loud. After this, the experimenters answered any questions individually. The software then started up with a set of comprehension questions. Every subject had to get all answers correct before the experiment proceeded further.

Our theory presumes pre-play negotiation but leaves the strategic details of this process implicit, reflected only through the solution concept. In the lab, however, one has to offer some specific format for the haggling. We chose an alternating-offer structure. After being acquainted with game details, and learning their respective roles as player *A* or player *B*, the subjects could send proposals back and forth and agree on how to play. One person from each pair was randomly selected to make an opening proposal. Each proposal specified whether player *A* would choose *In* or *Out*, and, conditional on *In*, the amount, y , that player *B* would keep. The party who received a proposal could accept it, make a counter-proposal, or disagree and quit negotiating. Acceptance of a proposal led to an agreement. This ended the negotiations and a message saying either “Player *A* chooses OUT” or, for, e.g., $y = 18$, “Player *A* chooses IN and Player *B* keeps \$18 and gives \$12 to Player *A*” appeared on the pair’s computer screens. A counter-proposal reversed the negotiation roles while a disagreement terminated the negotiation process. There was no limit imposed neither on the length of negotiations nor on the time within which a subject had to submit his decision.

We implemented a 2×2 design, varying the type of agreement, to be either a binding contract or an informal agreement, and the payoff following option *Out* to be either $(10, 0)$ or $(5, 5)$. We refer to the treatments as IA[5,5], IA[10,0], BC[5,5], and BC[10,0], in accordance with the cases of section 3.1. If an informal agreement was made, or if no agreement was made, the paired subjects entered the game stage where they were free to make any decisions. In the BC-treatments the agreements from the negotiation stage were automatically implemented.

On average, a session lasted about 50 minutes. The average final payment was \$19.90, including a \$5 show-up fee.

3.3 Main Results

In this section we present a series of results regarding specific tests of our theory. This is then followed (in section 3.4) by a few other observations of striking data patterns for which our theory has no direct implications.

Table 1 presents raw data on negotiated agreements (y) and path of play.²⁵ In the IA-treatments, the path of play is described by the amount kept (z) by player B , implying that A chose *In*, or by indicating that A chose *Out* (and hence implying that B had no decision to make in the game). In the BC-treatments, $z = y$ by definition.

Agreement formation

Table 1 shows that apart from two cases in BC[5,5] all other pairs of subjects reached an agreement. This is in line with the theory in the BC-treatments. All binding contracts involved player A choosing *In*.

In the IA-treatments temptations or high suspicion of dishonesty could have led to disagreements, but this did not happen. 100% of our subject-pairs formed an agreement. From the vantage point of our theory, this would suggest that the subjects' γ_B 's and ε_A 's are relatively low.²⁶ Finally, all but one informal agreements involved player A choosing *In*.²⁷

Do players honor agreements?

Table 2, distilled from Table 1, provides a summary of reached agreements and subsequent behavior. The first column, *Obs*, denotes the number of subject pairs who participated in a given treatment. The second column, *Agr*, provides the count of reached agreements which we further split (in subsequent columns) into what these agreements prescribe that players A and B will do. For A 's we compare the number of subjects who agreed on a

²⁵Descriptive statistics could be found in Appendix A.

²⁶Due to lack of disagreements in the data we cannot perform a meaningful test of our prediction involving no deals.

²⁷One pair in IA[5,5] agreed on *Out*, then Player A chose *In* and Player B kept 15.

Table 1: Raw data on agreements and path of play

| IA[10,0] | | IA[5,5] | | BC[10,0] | BC[5,5] |
|-----------|------------|------------|-----------|-----------|----------------|
| y | z | y | z | $y = z$ | $y = z$ |
| 0 | 15 | 15 | 15 | 14 | 14 |
| 10 | 20 | 15 | 15 | 15 | 15 |
| 13 | 18 | 15 | 15 | 15 | 15 |
| 15 | 15 | 15 | 15 | 15 | 15 |
| 15 | 15 | 15 | 15 | 15 | 15 |
| 15 | 15 | 15 | 15 | 15 | 15 |
| 15 | 15 | 15 | 15 | 15 | 15 |
| 15 | 15 | 15 | 15 | 15 | 15 |
| 15 | 15 | 15 | 15 | 15 | 15 |
| 15 | 15 | 15 | 15 | 15 | 15 |
| 15 | 15 | 15 | 15 | 15 | 15 |
| 15 | 15 | 15 | 15 | 15 | 15 |
| 15 | 15 | 15 | 16 | 15 | 15 |
| 15 | 15 | 15 | 17 | 15 | 15 |
| 15 | 15 | 15 | 25 | 15 | 15 |
| 15 | 15 | 15 | 20 | 15 | 15 |
| 15 | 15 | 15 | 30 | 15 | 15 |
| 15 | 15 | 15 | 30 | 15 | 15 |
| 15 | 15 | 15 | 30 | 15 | 15 |
| 15 | 15 | 16 | 16 | 15 | 15 |
| 15 | 15 | 16 | 20 | 15 | 15 |
| 15 | 20 | 17 | 15 | 15 | 15 |
| 15 | 30 | 18 | 18 | 15 | 16 |
| 15 | 30 | 18 | 18 | 15 | 17 |
| 15 | Out | 20 | 20 | 20 | 20 |
| | | 20 | 20 | | 20 |
| | | 24 | 22 | | Disagr. |
| | | Out | 15 | | Disagr. |

Note: y refers to the agreed-upon amount that player B would keep and z to the amount B actually kept.

Highlighted in bold are all observations that differ from 15.

certain deal which involved them choosing In (see column “Agreed to In ”)

Table 2: Agreements and honesty

| Treat. | Obs | Agr | Player A | | Player B | | |
|----------|-----|-----|---------------------|-----------------|----------|---------|---------|
| | | | Agreed to <i>In</i> | Chose <i>In</i> | $z < y$ | $z = y$ | $z > y$ |
| IA[10,0] | 24 | 24 | 24 | 23 | 0 | 17 | 6 |
| IA[5,5] | 27 | 27 | 26 | 27 | 2 | 16 | 8 |
| BC[10,0] | 24 | 24 | | 24 | | 24 | |
| BC[5,5] | 27 | 25 | | 25 | | 25 | |

Note: In IA[5,5] one pair has agreed on player *A* choosing *Out*. Following this player *A* chose *In* and player *B* kept \$15.

with the number who indeed chose *In* (column “Chose *In*”). For example, in the IA[10,0] treatment, presented in the first row, twenty-four *A*’s agreed to choose *In*, out of which twenty-three subsequently honored the agreement and chose *In*. In IA[5,5] twenty-six *A*’s agreed on *In*, all twenty-six indeed chose *In*. The twenty-seventh observation is the subjects who agreed to choose *Out* but in the end chose *In*. Finally, in the rightmost part of the table we list the number of *B*’s for whom the amount they kept (z) was smaller than, equal, or greater than the agreed upon amount (y). Naturally, in the BC-treatments, presented in the two bottom rows, there is no variation between the agreement and observed behavior of either player.

Table 2 shows that a majority of agreements were indeed honored. In all cases where player *A* agreed to choose *In*, he indeed did so. *B*’s face a direct temptation to renege. Nevertheless, the proportion of *B*’s who honored the agreement is quite high. In IA[10,0] 74% of *B*’s did exactly as they agreed and in IA[5,5] the proportion was slightly lower at 64%.²⁸

What about the subjects who reneged? Do they represent the selfish-fringe (captured by ε_i) as assumed by our theory? Recall that a player *B* who reneges on the agreement should act in his best self-interest by keeping

²⁸Unlike Player *A*’s, Player *B*’s cost themselves a lot of money by not reneging. In the IA[10,0] treatment Player *B*’s who lived up to their agreements earned on average \$15 but could have earned as much as \$30! In treatment IA[5,5] Player *B*’s earned on average \$16.55 which means they left on average \$14.35 on the table.

all \$30. Out of sixteen subjects who reneged (about 10% of all player B 's), five kept everything (i.e. 31% for those who reneged). The remaining eleven gave their paired player A 's a non-zero amount.²⁹

Equal splits & effect of d

The theory makes the following predictions: First, in the BC-treatments subjects will agree on equal splits and there will be no sensitivity to d . Second, in IA-treatments equal splits will be more frequent than deals compensating player A , and also more frequent than deals compensating player B . Third, in IA[5,5], deals compensating B will be more frequent and deals compensating A less frequent than in IA[10,0].

Table 3: Agreements

| Treat. | $y < 15$ | $y = 15$ | $y > 15$ |
|----------|----------|----------|----------|
| BC[10,0] | 1 | 22 | 1 |
| BC[5,5] | 1 | 20 | 4 |
| IA[10,0] | 3 | 21 | 0 |
| IA[5,5] | 0 | 18 | 8 |

Table 3 categorizes data on individual agreements with respect to equal splits. In line with the first two predictions, in all treatments (BC and IA), equal splits ($y = 15$) are the most frequent agreements. We observe 86% pairs in the BC-treatments agree on the equal split and 78% pairs in the IA-treatments. There is a relatively small number of data points on compensated deals. Nevertheless, we find statistical support for the remaining predictions regarding treatment comparisons: First, in the BC-treatments we observe only a few deals that are not equal splits and find that binding contracts are not sensitive to variation in payoffs following *Out*, i.e., there is no statistical

²⁹We discuss these subjects further in a subsequent section.

difference in frequencies presented in the first two rows of Table 3 (two-sided Fisher’s exact test has p -value = 0.58). Second, when it comes to compensated deals ($y \neq 15$), there is a distinct pattern: In IA[10,0] all pairs agree on player B keeping at least \$15 (three pairs or 12.5% of the sample agreed on $y < 15$) while in IA[5,5] all pairs agree on B keeping at most \$15 (eight pairs or 30.7% of the sample agreed on $y > 15$). The equality of distributions is rejected at 1% level (two-sided Fisher’s exact test has p -value = 0.001).³⁰ Informal agreements thus show sensitivity in the predicted direction: A ’s (B ’s) get compensated more (less) often when their payoff following *Out* increases (decreases).

3.4 Additional Observations

As seen in section 3.3, the theory holds up rather well. However, the data exhibits a few seemingly systematic patterns which are either at odds with the theory, or not explicitly predicted. This section highlights three such occurrences that we find intuitively striking. We allow ourselves some leeway commenting on what may be going on.

Reneging and semi-honesty

Our first comment belongs to subjects who reneged but did not keep all \$30. Such behavior lies outside the tight boundaries of our simple model. It is not easy to judge whether these subjects acted in an honest or a dishonest manner. On one hand, they did break the agreement; on the other hand, they still showed concern for their respective player A ’s by sending them some money. Perhaps one might refer to them as semi-honest.

What should we think of their behavior? In IA[10,0] three pairs negotiated compensated deals, according to which B ’s were supposed to keep $y = 0, 10$, and 13. Following each of those deals, B ’s reneged by “shading”

³⁰If one runs a test on just the observations for which $y \neq 15$, the results are virtually the same. For BC distributions the p -value is 0.524, for IA distributions it is 0.006.

the agreed-on amount by some fraction and keeping $z = 15, 20$, and 18 , respectively. It seems as if these semi-honest B 's had different deals in mind – ones that compensate them instead. The remaining data are in line with this story. In IA[5,5] we observe eight compensated deals ($y = 16, 17$, and 24) but this time only three of the B 's reneged ($z = 20, 15$, and 22 , respectively). Each of the three subjects gave his matched player A a positive amount. In two of those cases B actually gave A more money than what they agreed on!

The selfish fringe

Next let us examine the behavior of player B subjects who have kept all \$30 for themselves. Only five subjects fall into this category. It might nevertheless be interesting to look at the negotiation patterns of these subjects. If their behavior was planned then they knew at the point of the agreement that they were going to renege. One would think that their main objective would then be to maximize the chances that their paired player A chooses In . What is the most likely behavior to do the job?

Looking at the data all five B 's in question ended up agreeing on an equal split. Three of them accepted the opening equal split proposals made by their respective A 's. One of them proposed an equal split which was accepted. And the last one initially proposed \$25 for himself but that was rejected and countered with an equal split. This B accepted.

Beware of people who do not goof around! The selfish fringe in our data set hide among the subjects who strike 50/50 deals. We find it intriguing that there thus seem to be some degree of conformity in the community of confidence tricksters. An analogous finding, for a different strategic setting with asymmetric information, is reported by Charness & Dufwenberg (2011; see Section III.C)

Bargaining delay

Our next remark concerns the following systematic pattern of bargaining delay. Namely, most of the time the parties agree quickly. However, in almost

all instances where the negotiating proceeds several rounds, this happens in the BC-treatments and involves some player who demands more than \$15. To illustrate this, in Table 4 we list the sequences of proposal exchange for all deals that gave player B more money in treatments BC[5,5] and IA[5,5]. Notice that in BC[5,5] compensated deals involved a struggle between the paired subjects, with one pair negotiating for fifteen periods. In contrast, the compensated deals in IA[5,5] were negotiated in early rounds. We find the same pattern for other departures from equal split in BC[10,0] and IA[10,0] (see Table 5 which presents data on the length of negotiations broken down by final agreement y). Agreements that depart from equal splits in the BC-treatments were hard bargains while this is not the case in the IA-treatments.

Why do we observe this bargaining delay with informal agreements but not with binding contracts? Our intuition is as follows. In the BC-treatments a square deal is for the players to split the \$30 right down the middle. On the other hand, in the IA-treatments any y could be part of a possible square deal. If we suppose that the only legitimate proposals are those that could become a part of a square deal, then we might expect these offers to be accepted more easily than others. These data patterns suggest an auxiliary observation that square deals form smoothly and without much delay. On the other hand, agreements that are not square deals must have involved illegitimate proposals and have led to conflict and longer offer-exchanges.

There are multiple ways of evaluating this conjecture. One is in terms of the length of negotiations. Table 5 provides clear support for the argument suggested above. The average length of negotiations for illegitimate proposals (in the BC-treatments when $y \neq 15$) is distinctly longer, 4.5-7 rounds, than for legitimate proposals, 1.048-1.818.³¹

Another way of looking at the same issue is by comparing acceptance rates for the initial proposals.³² Table 6 gives the summary of the data. One can see that the acceptance rate for initial proposals that are equal splits (top

³¹The difference is statistically significant; $p=0.0005$ on an Epps-Singleton test.

³²Only for the initial offers we are guaranteed to have independent observations.

Table 4: Sequences of proposals

| | IA[5,5] | | | | | | | | BC[5,5] | | | |
|------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|------------|----------|----------|
| Obs: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 1 | 2 | 3 | 4 |
| Op. Prop.: | <i>B</i> | <i>B</i> | <i>A</i> | <i>A</i> | <i>B</i> | <i>B</i> | <i>A</i> | <i>A</i> | <i>B</i> | <i>B</i> | <i>B</i> | <i>B</i> |
| Rnd: 1 | 16 | 20 | 17 | 18 | 16 | 20 | 20 | 24 | 20 | 20 | 20 | 20 |
| 2 | | 15 | | | 18 | | | | 15 | 15 | 15 | |
| 3 | | 18 | | | | | | | 17 | 18 | 20 | |
| 4 | | 16 | | | | | | | 16 | 16 | 10 | |
| 5 | | | | | | | | | | 18 | 20 | |
| 6 | | | | | | | | | | 17 | 15 | |
| 7 | | | | | | | | | | 18 | 20 | |
| 8 | | | | | | | | | | 15 | 15 | |
| 9 | | | | | | | | | | 18 | 20 | |
| 10 | | | | | | | | | | 16 | 15 | |
| 11 | | | | | | | | | | 18 | 20 | |
| 12 | | | | | | | | | | 17 | 15 | |
| 13 | | | | | | | | | | 18 | 20 | |
| 14 | | | | | | | | | | <i>Out</i> | | |
| 15 | | | | | | | | | | 17 | | |
| Agreem.: | 16 | 16 | 17 | 18 | 18 | 20 | 20 | 24 | 16 | 17 | 20 | 20 |

Note: Row “Op. Prop.,” lists the player (*A* or *B*) who opened the negotiations.

Sequences of proposal-exchange run from top to bottom. E.g., sequence 2 in IA[5,5] reads as follows: player *B* made the first proposal to keep 20; player *A* countered with 15; player *B* rejected this and suggested he keeps 18; then player *A* went up to 16; and this was accepted by player *B*.

row) is higher than for rest of the proposals (bottom row). However, even proposals that are not equal splits could be legitimate in the IA-treatments because they may according to the theory reflect subjects’ ε_A ’s and γ_B ’s. The acceptance rate for these is still substantially higher (50%) than for their counterparts in the BC-treatments (0-10%). All in all, the difference in the acceptance rate between illegitimate proposals (found only in treatments BC[10,0] and BC[5,5]) and legitimate proposals is statistically significant

Table 5: Average length of negotiations

| | IA[10,0] | IA[5,5] | BC[10,0] | BC[5,5] |
|-------------|-----------------------|----------------------|----------------------|----------------------|
| $y = 15$ | 1.048 (0.218),{21} | 1.5 (0.985),{18} | 1.818 (1.79),{22} | 1.55 (1.791),{20} |
| $y \neq 15$ | 1.667 (1.155),{3} | 1.556 (1.014),{9} | 4.5 (2.121),{2} | 7 (6.52),{5} |

Note: Standard deviations are in parentheses; number of observations are in braces; two cases in BC[5,5] where subjects failed to reach an agreement are excluded.

Table 6: Acceptance rates of initial proposals

| | IA[10,0] | IA[5,5] | BC[10,0] | BC[5,5] |
|------------------------|---------------|----------------|----------------|----------------|
| Op. proposal = 15 | 20/20 100% | 13/15 86.7% | 14/15 93.3% | 16/17 94.1% |
| Op. proposal \neq 15 | 2/4 50% | 6/12 50% | 0/9 0% | 1/10 10% |

(Fisher’s exact test has p -value = 0.000). This evidence also suggests that the departures from equal splits in the BC-treatments are typically found unjustified and become hard bargains. The resulting deals are then likely driven by an imbalance in subjects’ patience and obstinacy.

Informal agreements affect play

Our experimental goal is to test our theory, and since it concerns deal-making the design always includes opportunities for informal agreements or binding contracts. However, when presenting the paper we were sometimes asked whether B ’s sharing behavior resembles what would happen without pre-play deal-making. A comparison with preceding Lost Wallet game studies (without pre-play negotiation) indicates this is not the case. Servátka & Vadovič (2009) (S&V) have treatments with $d = 0$ and $d = 5$ and we focus

on their data.³³ The comparison is not perfect as the available money for the second-mover to split was \$20, not \$30, but with that caveat we can record the following differences most of which are statistically significant (we give proportions only but number of observations and formal test results are available on request):

Our *B*'s choose more equal splits than S&V's (78% vs 35% for the [10,0]-treatments); 50% vs 31% for the [5,5]-treatments). Our *B*'s give zero ($z = 0$) less often than S&V's (9% vs 38% for [10,0]; 12% vs 31% for [5,5]). Our *B*'s give more than S&V's (the average amounts given, measured as percentages of the available pie, are 44% vs 31% for [10,0]; 39% vs 26% for [5,5]).

4 Discussion

Informal agreements have been given rather scant attention in economic theory. Are they economically unimportant? Couldn't agents simply always rely on binding contracts to achieve good partnership outcomes? We do not think so for several reasons.

First, binding contracts may be infeasible. Consider two impatient fishermen who live in a developing country where neither courts nor policemen are reliable. It may be impossible to draft a binding contract which regulates the fishermen's access to a nearby lake. Does this mean that they are destined to inefficient excessive depletion of the fish stock? Even if the interaction is repeated, classical theory would say yes (because of the impatience). However, according to our theory, the answer may be no, if the fishermen rely on an informal agreement.

Second, binding contracts may be illegal. Think of collusion in a one-shot government procurement auction in industrial countries. Courts exist and police can be relied on, yet bid rigging would not be legal. Does that imply that the outcome will conform with standard auction theory, where

³³Similar conclusions would be reached if we considered the data of Dufwenberg & Gneezy (2000), Charness, Haruvy & Sonsino (2007), or Cox, Servátka & Vadovič (2010).

the competing firms maximize profit taking each others' actions as given? Perhaps not. Suppose the firm representatives meet in a bar, have a pint, shake hands, and agree to collude. Will this stick? If the representatives take pride in standing by their word, the answer may be yes.

Third, even if binding contracts are feasible in principle (as they perhaps usually are) they may be costly. A guy meets a girl and they play the (one-shot, sequential) game of life with decisions on kids, who works, divorce, alimony, etc. A binding contract may involve significant costs ranging from lawyers' fees to unforeseen contingencies to awkward feelings regarding legal chit-chat during courtship. Perhaps, instead, the couple will shun the formalities, look one another in the eye, and promise to be faithful forever?

Just how compelling are these examples? Under which circumstances will informal agreements work? How will the terms be structured? How will the deals shape subsequent play? How do the outcomes compare to those that would obtain under a binding contract? Answers require theory.

We have proposed a theory of deal-making which covers informal agreements and binding contracts as special cases. We suggest that people – who deep down may be motivated in complicated, rich, and hard-to-pin-down-in-mathematical-formulae ways – do not engage in any elaborate utility calculus when striking deals. They rather focus on easy-to-observe data, namely dollars gained. Inspired by Binmore et al's so-called deal-me-out solution, we assume that people agree to split dollar gains equally, as long as each person is thereby made better off than if the negotiations stranded.

A key assumption is that most people are honest. Informal agreements therefore work in similar ways as binding contracts. However, the two are not interchangeable. With informal agreements players incur costs of overcoming temptation and of co-player opportunism. This influences players' perception of whether a deal makes them better off. These concerns are irrelevant with binding contracts. Therefore, our theory generates systematic testable predicted differences between the shapes and impact of informal agreements and binding contracts. We ran an experiment to evaluate the empirical relevance

of these predictions, and found considerable support.

We hope this will inspire economists to apply our model, to study various economic settings where people strike deals, be they informal or binding. That said, we conclude by noting a few issues we did not address so far. First, in context where both informal agreements and binding contracts are feasible, one may wish to endogenize the choice between the two. So far we assumed the feasible deals to involve one or the other. Second, one may imagine deals that regulate some but not all choices, and so would be intermediate to the polar cases of an informal agreement and a binding contract. Our framework may be adequate for exploring such “incomplete contracts”. Third, in many contexts material costs and revenues are not as readily observable as our above account (with given m_i functions) may suggest. For example, how should considerations of unobserved cost-of-effort or consumer surplus be dealt with? Fourth, (as discussed in section 2.2) even when dollar payoffs are given, 50/50 splits may not be focal in all setting and a refined theory may consider alternatives. Fifth, we limited attention to games with two players, but many situations involve multiple bargaining parties. We plan to explore these extentions in future work.

References

- [1] Andreoni, J. & B.D. Bernheim (2009): “Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects,” *Econometrica* 77, 1607–1636.
- [2] Battigalli, P. & M. Dufwenberg (2007): “Guilt in Games,” *American Economic Review*, Papers and Proceedings, 97, 170–176.
- [3] Benabou, R. & J. Tirole (2002): “Self-confidence and Personal Motivation,” *Quarterly Journal of Economics* 117, 871–915.

- [4] Benartzi, S. & R. Thaler (2004): “Save More Tomorrow: Using Behavioral Economics to Increase Employee Savings,” *Journal of Political Economy* 112, 164–187.
- [5] Ben-Ner, A. & L. Putterman (2009): "Trust, Communication and Contracts: An Experiment," *Journal of Economic Behavior & Organization* 70, 106–121.
- [6] Binmore, K., A. Shaked & J. Sutton (1989): “An Outside Option Experiment,” *Quarterly Journal of Economics* 104, 753–770.
- [7] Binmore, K., C. Proulx, L. Samuelson & J. Swierzbinski (1998): “Hard Bargains and Lost Opportunities,” *Economic Journal* 111, 1279–1298.
- [8] Brown, A. L., Z. E. Chua & C. F. Camerer (2009): “Learning and Visceral Temptation in Dynamic Savings Experiments,” *Quarterly Journal of Economics* 124, 197–231.
- [9] Camerer, C. (2003): *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton University Press.
- [10] Charness, G. & M. Dufwenberg (2006): “Promises & Partnership,” *Econometrica* 74, 1579–1601.
- [11] – (2011) : “Participation,” *American Economic Review* 101, 1213–39.
- [12] Charness, G., E. Haruvy & D. Sonsino (2007): "Social distance and reciprocity: an internet experiment," *Journal of Economic Behavior & Organization*, 63(1), 88–103.
- [13] Cox, J., M. Servátka & R. Vadovič (2010): “Saliency of Outside Options in the Lost Wallet Game,” *Experimental Economics*, 13(1), 66–74.
- [14] Crawford, V. & J. Sobel (1982): “Strategic Information Transmission,” *Econometrica* 50, 1431–1451.

- [15] Dekel, E. & B. Lipman (2011): “Costly Self Control and Random Self Indulgence,” unpublished manuscript.
- [16] Demichelis, S. & J. Weibull (2008): “Language, Meaning and Games – A Model of Communication, Coordination and Evolution,” *American Economic Review* 98, 1292–1311.
- [17] Dufwenberg, M. & U. Gneezy (2000): “Measuring Beliefs in an Experimental Lost Wallet Game,” *Games & Economic Behavior* 30, 163–182.
- [18] Dufwenberg, M. & G. Kirchsteiger (2004): “A Theory of Sequential Reciprocity,” *Games & Economic Behavior* 47, 268–298.
- [19] Eller, K. (2004): *Integrity Is All You’ve Got*, McGraw-Hill.
- [20] Ellingsen, T. & M. Johannesson (2004): “Promises, Threats and Fairness,” *Economic Journal* 114, 397–420.
- [21] Falk, A. & U. Fischbacher (2006): “A Theory of Reciprocity,” *Games & Economic Behavior* 54, 293–315.
- [22] Farrell, J. & M. Rabin (1996): “Cheap Talk,” *Journal of Economic Perspectives* 10, 103–118.
- [23] Fehr, E. & K. Schmidt (2003): “Theories of Fairness & Reciprocity – Evidence & Economic Applications,” in: M. Dewatripont, L. Hansen, S. Turnovsky (eds.), *Advances in Economics & Econometrics*, 8th World Congress Econometric Society Monographs, Cambridge Univ Press.
- [24] Feltovich, N. & J. Swierzbinski (2011): “The Role of Strategic Uncertainty in Games: An Experimental Study of Cheap Talk and Contracts in the Nash Demand Game,” *European Economic Review* 55, 554–574.
- [25] Forges, F. (1986): “An Approach to Communication Equilibria,” *Econometrica* 54, 1375–1385.

- [26] Fudenberg, D. & D.K. Levine (1993): "Self-Confirming Equilibrium," *Econometrica* 61(3), 523–545.
- [27] – (2006): "A Dual-Self Model of Impulse Control," *American Economic Review* 96, 1449–1476.
- [28] Fudenberg, D. & J. Tirole (1991): *Game Theory*, MIT Press.
- [29] Geanakoplos, J., D. Pearce & E. Stacchetti (1989): "Psychological Games and Sequential Rationality," *Games & Economic Behavior* 1, 60–79.
- [30] Gneezy, U. (2005): "Deception: The Role of Consequences," *American Economic Review* 95, 384–394.
- [31] Greenberg, J. (2000), "The Right to Remain Silent," *Theory & Decision* 48, 193-204.
- [32] Gul, F. & W. Pesendorfer (2001): "Temptation and Self-Control," *Econometrica* 69, 1403–1435.
- [33] Irlenbusch, B. (2004): "Relying on a Man's Word? An Experimental Study on Non-Binding Contracts," *International Review of Law & Economics* 24, 299-332.
- [34] Isoni, A., A. Poulsen, R. Sugden & K. Tsutsui (2011): "Focal Points in Tacit Bargaining," CBESS Discussion Paper 11-13, University of East Anglia.
- [35] Kalai, E. (1977): "Proportional Solutions to Bargaining Situations: Inter-Personal Utility Comparisons," *Econometrica* 45, 1623–1630.
- [36] Kalai, A., & E. Tauman Kalai (2010): "A Cooperative Value for Bayesian Games," Working paper 1512, Northwestern University.

- [37] Kartik, N. (2009): “Strategic Communication with Lying Costs,” *Review of Economic Studies* 76, 1359–1395.
- [38] Kerr, N.L. & C.M. Kauffman-Gilliland (1994): “Communication, Commitment, and Cooperation in Social Dilemmas,” *Journal of Personality & Social Psychology* 66, 513–529.
- [39] Kessler, J. & S. Leider (2012): “Norms & Contracting,” *Management Science* 58, 62–77.
- [40] Levin, J. (2003): “Relational Incentive Contracts,” *American Economic Review* 93, 835–57.
- [41] Lipman, B. & W. Pesendorfer (2011): “Temptation,” unpublished manuscript.
- [42] Loewenstein, G. (1996): “Out of Control: Visceral Influences on Behavior,” *Organizational Behavior & Human Decision Processes* 65, 272–292.
- [43] Loewenstein, G. & T. O’Donoghue (2005): “Animal Spirits: Affective and Deliberative Processes in Economic Behavior,” unpublished manuscript.
- [44] MacLeod, B. & J. Malcolmson (1989): “Implicit Contracts, Incentive Compatibility and Involuntary Unemployment,” *Econometrica* 56, 447–480.
- [45] Malhotra, D. & J. Murnighan (2002): “The effects of contracts on interpersonal trust,” *Administrative Science Quarterly* 47, 534–559.
- [46] Martinsson, P., K. Myrseth & C. Wollbrant (2012): "Reconciling Pro-Social vs. Selfish Behavior: On the Role of Self-Control," *Judgment & Decision Making* 7, 304-315.
- [47] Miettinen T. (2008): “Contracts & Promises – An Approach to Pre-play Agreements,” unpublished manuscript.

- [48] Myerson, R. (1986): “Multi-Stage Games with Communication,” *Econometrica* 54, 323–358.
- [49] Nash, J. (1950): “The Bargaining Problem,” *Econometrica* 18, 155–162.
- [50] – (1953): “Two-person Cooperative Games,” *Econometrica*, 21, 128–40.
- [51] Ozdenoren, E., S. Salant & D. Silverman (2011): “Willpower and the Optimal Control of Visceral Urges,” *Journal of the European Economic Association*, forthcoming.
- [52] Rabin, M. (1993): “Incorporating Fairness into Game Theory and Economics,” *American Economic Review* 83, 1281–1302.
- [53] Schelling, T. (1960): *The Strategy of Conflict*, Harvard University Press.
- [54] Servátka, M., S. Tucker & R. Vadovič (2011): “Words Speak Louder Than Money,” *Journal of Economic Psychology* 32, 700–709.
- [55] Servátka, M. & R. Vadovič (2009): “Unequal Outside Options in the Lost Wallet Game,” *Economics Bulletin*, 29(4), 2870-2883.
- [56] Thaler, R. & H. Shefrin (1981): “An Economic Theory of self-Control,” *Journal of Political Economy* 89, 392–410.
- [57] Vanberg, C. (2008): “Why Do People Keep Their Promises? An Experimental Test of Two Explanations,” *Econometrica* 76, 1467–1480.

Appendix

A Session overview

Table 7: Treatments

| Tr. Name | Type of Agreement | Outside Opt. | Session | # of subj. |
|----------|-------------------|--------------|-------------|------------|
| IA[10,0] | Informal | [10,0] | Sess. 1 | 20 |
| IA[10,0] | Informal | [10,0] | Sess. 2 | 28 |
| IA[5,5] | Informal | [5,5] | Sess. 1 | 28 |
| IA[5,5] | Informal | [5,5] | Sess. 2 | 26 |
| BC[10,0] | Binding | [10,0] | Sess. 1 & 2 | 10 |
| BC[10,0] | Binding | [10,0] | Sess. 3 | 28 |
| BC[5,5] | Binding | [5,5] | Sess. 1 | 30 |
| BC[5,5] | Binding | [5,5] | Sess. 2 | 24 |

Table 8: Descriptive statistics

| Tr. | Obs. | Agreement | | | Decisions | |
|----------|------|-------------|----------------------|----------------------------|----------------------|----------------------------|
| | | Agreed (%) | Pl. A: <i>In</i> (%) | Pl. B: <i>y</i> (st. dev.) | Pl. A: <i>In</i> (%) | Pl. B: <i>y</i> (st. dev.) |
| IA[10,0] | 24 | 24 (100) | 24 (100) | 14.08 (3.19) | 23 (96) | 16.13 (5.65) |
| IA[5,5] | 26 | 26 (100) | 25 (96) | 16.11 (2.321) | 26 (100) | 18.54 (4.99) |
| BC[10,0] | 24 | 24 (100) | 24 (100) | 15.17 (1.05) | 24 (100) | 15.17 (1.05) |
| BC[5,5] | 27 | 25 (93) | 25 (100) | 15.48 (1.45) | 25 (100) | 15.48 (1.45) |

Note: In IA[10,0] one pair has agreed that player *A* chooses *Out*. Following this player *A* chose *In* and player *B* kept \$15. In only two instances, both in BC[5,5], subjects have disagreed. In both cases player *A*'s chose *In*, player *B*'s kept \$20 and \$30 respectively.

B Instructions

In what follows we present the universal version of the instructions in which {... *or* ...} always contains two different versions of the text that was used appropriately in different treatments.

Now that the experiment has begun, we ask that you do not talk with each other for the duration of the experiment. If you have a question after we finish reading the instructions, please raise your hand and the experimenter will approach you and answer your question in private.

You will receive \$5 for participating in this experiment. You may also receive additional money, depending on the choices made (as described below). Your earnings will be paid to you in cash individually and privately.

During the session, you will be paired with another person. However, no participant will ever know the identity of the person he or she is paired with.

In the experiment, one person from each pair will be randomly selected to be Player A and the other to be Player B. The players will interact in two stages: 1. The Negotiation Stage and 2. The Game. In the negotiation stage the players can form an agreement about how to play the game. Any agreement reached in the negotiation stage { *will or will not* } be enforced and the players {will have to play according to the agreement *or* be free to make any decisions} in the game that follows. The decisions in the negotiation stage will determine how much each of the players earns in the experiment.

We next describe first the game and then the negotiation stage that precedes it.

The Game

Player A moves first and chooses either IN or OUT by clicking a button labeled either “IN” or “OUT.”

Player B moves second:

- If Player A chose OUT, then the game ends. *Player A receives { \$5 or \$10} and Player B receives { \$5 or \$0} .*
- If Player A chose IN, then *Player B splits \$30 between the two of them: Player B keeps \$x and gives \$30-x to Player A, choosing x such that $0 \leq x \leq 30$.*

The Negotiation Stage

Before the game is played the players can form an agreement about how to play the game. One player from each pair will be randomly selected to make the first proposal and the other player will be asked to respond to it. A proposal describes the choices of Player A and Player B in the game.

It could be:

Player A chooses OUT

or it could be

Player A chooses IN and
Player B keeps \$x and gives \$30 - x to Player A.

The proposal is sent to the other player by clicking on the “Submit” button. The responding player observes the proposal and chooses one of the following three options:

- *Agree with the proposal* by clicking on the button “Agree.” In this case an agreement is formed and { *will or will not* } be enforced.
- *Make a counter-proposal* by clicking on the button “Make a counter-proposal.” This reverses the roles of the players in the negotiation. Now, the player who clicked this button makes a new proposal and sends it to the other player. The other player will then have the chance to respond by either agreeing with the proposal, or making a counter-proposal, or disagreeing.

- *Disagree and quit negotiating* by clicking on the button “Disagree and quit negotiating.” In this case no agreement is reached and negotiations terminate. Both players proceed to play the game.