Flexibility and Reputation in Repeated Prisoner's Dilemma Games

Dorothée Honhon
Jindal School of Management, University of Texas at Dallas, Dorothee.Honhon@utdallas.edu,
http://www.hyndman-honhon.com/Main/Dorothee_Honhon.html

Kyle Hyndman
Jindal School of Management, University of Texas at Dallas, KyleB.Hyndman@utdallas.edu,
http://www.hyndman-honhon.com/Main/Kyle_Hyndman.html

We study how the option to terminate a relationship affects cooperation in a repeated prisoner's dilemma. While cooperation is theoretically stainable, we show experimentally that cooperation rates are significantly lower with the option to terminate. Rather than punishing a defection, most subjects terminate the relationship, which induces more opportunistic behavior. A reputation mechanism, which signals cooperative behavior, can substantially increase cooperation rates, in some cases to a level higher than when the option to terminate the relationship is absent. Moreover, reputation mechanisms lead to stable cooperation rates over time, in contrast to the declining cooperation rates observed in their absence.

Key words: Long-term relationships; prisoner's dilemma; reputation

JEL codes: C72, C91, D83

History: **March 7, 2017**

1. Introduction

Many important business and personal relationships involve repeated interaction. For a company, maintaining good relationships with business partners is important for its long term survival. For example, Uzzi (1996) demonstrates empirically that garment firms which receive only short-term contracts were substantially more likely to go out of business than firms that have long-term contracts. Similarly, Toyota's underlying philosophy of *keiretsu* – building networks of suppliers in long-term relationships that work closely to identify cost-savings and product improvements – is one of the pillars of the Toyota Production System and has been linked to its success in the auto market (Liker and Choi 2004). In terms of personal relationships, studies show that married people are generally happier than singles (Taylor et al. 2006).

From a modeling perspective, the usual approach to studying long-lasting interactions is to assume that the parties involved are engaged in an infinitely repeated game. While many of these relationships have no fixed end date, such an approach misses one important component. Specifically, the parties engaged in an ongoing relationship typically have the option to terminate their relationship should things sour. In practice, 50% or more partnerships/joint ventures end in failure (Coopers and Lybrand 1986, Bamford et al. 2004, Anderson and Jap 2005, KPMG 2009), and about 48% of first marriages end in divorce within 20 years (Copen et al. 2012).

There are at least two reasons for a partnership to break down. First, the incentives of the parties may not be perfectly aligned, which can lead to opportunistic behavior. For example, prior to Microsoft's acquisition of Nokia the two entered into a partnership but the incentives were only partially aligned: Grundberg and Stoll (2012) note that Microsoft limited the compatibility of its then newest mobile operating system to only the newest Nokia phones, thus hurting Nokia. On the other hand, Kovach (2014) notes that, shortly before the acquisition, Nokia had announced plans to sell phones using the Android platform, which would adversely affect Microsoft. Neuville (1997) also describes an example of opportunistic behavior by an automobile parts supplier which was eventually discovered to have been cutting corners with respect to quality.

Second, the partnership simply may not live up to expectations. For example, Larsen et al. (2010) document the case of a joint venture between LEGO and Flextronics to outsource production. Despite apparent good intentions and hard work by both parties managing the joint venture, it failed after only three years due to the unanticipated complexity of coordinating the newly global supply chain and misalignments in the need for production flexibility.

In this paper, we seek to understand the implications that the ability to terminate ongoing relationships has on cooperative behavior when the parties involved can behave opportunistically, as in the classical prisoner's dilemma. On the one hand, such flexibility may be desirable because it allows players to escape from "bad" relationships. However, on the other hand, the ability to dissolve a relationship may undermine cooperation by reducing the credibility of future punishments for behaving opportunistically. In particular, a player may take an uncooperative/opportunistic action and then simply terminate the relationship rather than let it continue and accept the subsequent punishment. In fact, even a player who has suffered from an uncooperative action may prefer to terminate the relationship when meting out the punishment is too costly.

Past research (e.g., Friedman (1971)) has shown that cooperation in the indefinitely repeated prisoner's dilemma can be sustained provided that players are sufficiently patient. Kandori (1992) shows that cooperation in the prisoner's dilemma is also possible under random matching every period under certain conditions when cooperation is sustained via a community enforcement mechanism. Specifically, a defection by one player sets off a contagion towards universal defection by the population. Anticipating this, players are deterred from defecting. In the game that we consider, cooperation via community enforcement is sustainable as an equilibrium under random matching. We then use a similar logic to show that cooperation is theoretically sustainable – under the same conditions as indefinite repetition – when players have the option to dissolve relationships. In the equilibrium only groups that experience a defection dissolve relationships, creating two pools of players: cooperators and defectors. Once in the defector pool, it is optimal to continue to defect forever thereafter. Hence, the fear of entering this uncooperative pool provides the necessary incentives to sustain cooperation.

We compare the behavior of subjects as we vary the flexibility of dissolving relationships. At one extreme, we consider anonymous RM (for *random matching*) in which players are automatically rematched at the start of each period. We then consider TBA-U and TBA-M (for *Temporarily Binding Agreements – Unilateral* and *Temporarily Binding Agreements – Mutual*), in which subjects have the ability to terminate relationships and subsequently be randomly rematched from within the pool of unmatched subjects. The distinguishing feature is whether unilateral (as in TBA-U) or mutual (as in TBA-M) consent is required to terminate the relationship. At the other extreme, we consider IBA (for *indefinitely binding agreements*) in which subjects play the indefinitely repeated prisoner's dilemma and are unable to prematurely terminate the relationship.

First we establish that cooperation is theoretically possible in all four matching institutions, then we use human subjects experiments to test the empirical relevance of this claim. Our results show that cooperation is increasing in the difficulty to terminate relationships; that is, flexibility undermines cooperation. Specifically, overall, we find that the cooperation rates are 7.46%, 39.11%, 49.42% and 62.64% in the RM, TBA-U, TBA-M and IBA treatments, respectively. Moreover, the dynamics over supergames are divergent, with cooperation rates increasing under TBA-M and IBA but decreasing under RM and TBA-U. When flexibility is present, it appears that punishments lose their saliency since many relationships are dissolved following a defection but some subjects begin their next relationship by cooperating, which makes defection more attractive. This effect is partially mitigated in TBA-M, where mutual consent is required, hence we see more favorable dynamics in the level of cooperation.

Next we study how the introduction of a reputation mechanism, through which a signal of one's past cooperative behavior is given to one's match, affects cooperation when subjects have the option to unilaterally terminate the relationships. Although punishments lose credibility with the option to dissolve a relationship, subjects may be dissuaded from behaving opportunistically if doing so will send a negative signal to potential future matches. Duffy and Ochs (2009) show that providing information about the current match's previous choice increases cooperation in the prisoner's dilemma from 7.5% to 17% but that the cooperation rate is still well below that achieved under indefinite repetition. Camera and Casari (2009) also show that non-anonymous public monitoring (a form of reputation) significantly improves cooperation in the prisoner's dilemma with random matching relative to private monitoring (i.e., with no reputation mechanism).

We augment the TBA-U treatment with three different reputation treatments. In the first one, called *Objective Long-Run* (OLR), subjects saw three summary statistics: (i) the frequency of cooperation and defection[1] over all previous periods, (ii) the frequency of cooperation and defection in the first period of each new matching and (iii) the number of previous matchings. Subjects saw this information both about themselves and about their current partner. We call this treatment "long-run reputation" because this information carried over across supergames; that is, the statistics were never reset at any time during the experiment. The second mechanism, called *Objective Short-Run* (OSR), was the same as the OLR treatment except that the reputation statistics were reset at the start of each new supergame. Finally, in the *Subjective Long-Run* (SLR) treatment, subjects rated their satisfaction with their opponent on a scale of 1 to 5 each time a relationship was dissolved (either naturally or because a request was made). In each period, they saw the following summary statistics: (i) the average satisfaction score and (ii) the number of previous matchings, for both themselves and their match. Like in the OLR treatment, the satisfaction scores carried over across supergames, making it difficult to escape from a bad reputation. The comparison of

[1] Despite the fact that the frequency of cooperation is always equal to one minus the frequency of defection, we show both values to the subjects in order to avoid any framing effect.

objective and subjective measures of reputation is important as some online platforms, such as Ebay, have recently announced a shift from subjective ratings to more objective measures.[2]

We find that cooperation rates in all three reputation treatments are substantially and significantly higher than in the baseline TBA-U treatment. In fact, for the OLR treatment, cooperation rates are even significantly higher than with IBA, while for the other two reputation mechanisms, the overall cooperation rates are statistically indistinguishable from IBA. We also present some evidence suggesting that the OLR treatment leads to more cooperation than the SLR treatment, providing justification for some online platforms' move towards feedback systems based on more objective measures. Further, unlike with our TBA-U treatment which sees cooperation rates decline over time, cooperation rates in our reputation treatments are reasonably stable. Thus, we provide clear evidence that reputation mechanisms can fully counteract the adverse effects that come with the option to terminate relationships. The main reason appears to be that reputation serves as an accurate signaling device for one's cooperativeness, which players take into consideration both when deciding whether to cooperate and whether to terminate the relationship.

We also show that, in TBA-U and TBA-M subjects tend to over-use their option to dissolve relationships. Specifically, we show that matches that continue following a defection are substantially more likely to engage in mutual cooperation in the next period than are the subjects if they dissolve the relationship. When reputation mechanisms are present, high reputation subjects are both more likely to dissolve following a defection and more likely to subsequently form a mutually cooperative relationship by dissolving than by maintaining. In contrast, low reputation subjects are less likely to form a mutually cooperative relationship when the relationship is dissolved. Finally, we show that the option to dissolve relationships acts as a sorting device, allowing those in cooperative relationships to maintain productive, long-term relationships, while those who have a higher tendency to defect see their relationships terminated more frequently and are rematched from a pool of similarly uncooperative subjects. This effect is further enhanced in the reputation treatments.

The paper proceeds as follows. In the next section, we briefly summarize the related experimental literature. Section 3 introduces our experimental design, while Section 4 discusses the viability of cooperation as an equilibrium in all of our treatments. Section 5 develops our hypotheses on cooperation rates across treatments, which we subsequently test in Section 6. In Section 7 we delve more deeply into how the option to terminate relationships, and the introduction of reputation mechanisms influences cooperation. Finally, Section 8 presents some concluding remarks.

[2] See, for example, `http://pages.ebay.com/sellerinformation/news/fallupdate2015/index.html`.

2. Related Experimental Literature

There is a growing experimental literature which studies cooperation in the indefinitely repeated prisoner's dilemma. Roth and Murnighan (1978) is the first such study; they introduced the methodology of a random termination in order to mimic an infinitely repeated game. Roth and Murnighan (1978) along with Murnighan and Roth (1983) provided early evidence that cooperation can be sustained under indefinite repetition provided that it is a subgame perfect equilibrium. More recent papers, such as Dal Bó (2005) and Dal Bó and Fréchette (2011) provided more nuanced results on when cooperation can be expected to emerge. Dal Bó and Fréchette (2011) show that cooperation is further enhanced if, in addition to being subgame perfect, it is also risk dominant. Duffy and Ochs (2009) are interested in how the matching protocol affects cooperation and focus on comparing fixed matching with random matching environments. We take these as two end points on a continuum and also consider two intermediate where players have the option to dissolve relationships, either unilaterally (TBA-U) or by mutual consent (TBA-M).

A few recent papers consider endogenous dissolution or selection of groups in repeated games. First, D'Evelyn (2013) studies a marriage market where agents begin by "dating" another player in a social dilemma and subsequently decide whether to continue dating, "dump" or "marry" their partner, where marriage means that the two partners are bound thereafter until the end of the finitely repeated game. In their model, preserving flexibility by dating dominates marriage. Despite this, many subjects choose to get married. However, as predicted by the theory, effort rates subsequently go down for married versus unmarried subjects, which suggests either that subjects misunderstand the negative consequences of marriage or have non-monetary preferences for interacting with the same subject. This is different from our paper as our subjects do not choose the type of relationship they engage in; that is, they are either married (in IBA) or dating (in all other treatments) throughout the experiment.

Hyndman and Honhon (2015) consider behavior under indefinitely and temporarily (unilaterally breakable) binding agreements. Unlike the present paper, their stage game features strategic complementarities and has multiple Pareto rankable equilibria. They show that, on average, subjects earn more under indefinitely binding agreements but that there are interesting nuances. In particular, under temporarily binding agreements some groups are highly durable and benefit from a positive reinforcement process that approaches the efficient equilibrium. They show that the careful use of flexibility can be beneficial but that it can be a clear detriment when over-used. One of the key reasons why the IBA institution appears to be better on average in the strategic complementarities setting is that it promotes sophisticated learning and also prevents subjects from terminating a relationship due to bad luck, rather than bad decisions. Finally, Hyndman and Honhon (2015) show that despite the apparent advantage of IBA, most subjects, when given the choice, prefer to enter into flexible relationships and seem to earn more when the choice of institution is endogenous.

Gaudeul et al. (2015) study flexibility in a public goods game in which players can exit and contribute to a private good or maintain the relationship and contribute to a public good. In their paper, the decision to exit is reversible. They show that subjects exit excessively, leading to inefficiency; moreover, they find that introducing barriers to exiting improves welfare. Rand et al. (2011) also study cooperation when partnerships can be dissolved and new relationships can be formed according to a structured network. They show that the ability to frequently update links in the network (i.e., dissolve old links and form new links) promotes cooperation. This is consistent with our results because, in the link updating phase of their experiment, subjects observe other players' most recent action, which can be seen as a specific form of reputation.

Among the set of papers studying endogenous termination, the most closely related paper to ours is Wilson and Wu (Forthcoming). They consider an indefinitely repeated partnership game but allow for endogenous termination of relationships. However, unlike in our design, when a subject terminates a relationship, the two partners receive an exogenous termination payoff in each subsequent period until the end of the repeated game. They show that the relationship between efficiency and the termination payoff is non-monotonic. Initially, cooperation rates increase with the value of the termination payoff, which increases efficiency. However, for sufficiently high (but inefficient) termination payoffs, subjects more frequently terminate relationships and cooperation rates do not improve, leading to a decline in efficiency.

In addition to the papers on endogenous termination, we also contribute to the literature on reputation in repeated games. Duffy and Ochs (2009) and Camera and Casari (2009) also study reputation by varying the information available about subjects' past actions in repeated prisoner's dilemma games and show that they can enhance cooperation. Stahl (2013) also considers reputation in a repeated prisoner's dilemma, using a color coded system, which changes between green and purple as a function of past behavior, as a signal of reputation. We complement these papers by examining reputation mechanisms in an environment where relationships may be prematurely dissolved by the players. Finally, reputation has also been studied in other contexts. Duffy et al. (2013) examine reputation in trust games, while Bolton et al. (2004) look at reputation in online trading markets and Bolton et al. (2005) study reputation when indirect reciprocity is at play.

3. Experimental Design

In all of our experiments, subjects were matched in groups of two and played the prisoner's dilemma game depicted in Figure 1 repeatedly over a certain number of *periods*.[3] In what follows we refer to a pair of subjects as a *matching* and to a player's opponent as his or her *match*. Absent a subject requesting to be rematched, which was possible in some treatments as explained below, the probability that a matching would continue was 90% in each period; that is, with 10% chance, all matchings in the experiment would naturally break up, indicating the end of a *supergame*. In the next period, all subjects were then randomly matched to another participant and a new supergame began. This is the standard method, first used by Roth and Murnighan (1978), to implement an indefinitely repeated game in the laboratory. The experiment ended in the first period, after the 75th, in which all matchings naturally broke up. Note that subjects were not given identification numbers. Therefore they had no direct way of knowing if they started a new matching with a subject to whom they had been matched previously.[4]

Figure 1 The prisoner's Dilemma Stage Game

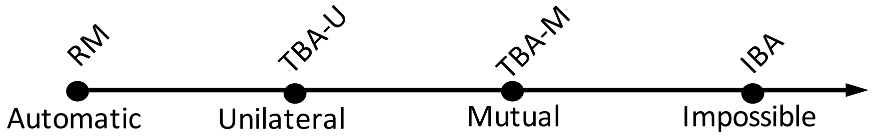
	Cooperate (C)	Defect (D)
Cooperate (C)	$40, 40$	$12, 50$
Defect (D)	$50, 12$	$25, 25$

Our first research question concerns the role of having the flexibility to dissolve a relationship has on the level and sustainability of cooperation. Figure 2 compares four possible modes of interaction between subjects in terms of how easy it is for them to dissolve a relationship. At the left-most extreme, players are engaged in a series of *one-shot* interactions and dissolutions are automatic at the end of every period. We refer to this as *Random Matching* (RM). On the right-most extreme, under *Indefinitely Binding Agreements* (IBA), players are indefinitely bound together for the entire duration of the supergame so that dissolutions are impossible. Under *Temporarily Binding Agreements* (TBA-U and TBA-M treatments), subjects have the option to request a dissolution at the end of each period. Under TBA-U, one player making such a request is enough to trigger a dissolution while under TBA-M, both players need to agree for their relationship to terminate (U stands for Unilateral and M stands for Mutual). More specifically for TBA-M, if only one of the

[3] To avoid any framing effects, we used neutral language in the experiment. In particular, the stage-game actions were labeled "Action A" and "Action B" rather than "Cooperate" and "Defect".

[4] Even in our reputation treatments, wherein subjects are shown their match's reputation score, it would be difficult for a subject to infer they are matched with someone they interacted with previously since the reputation scores dynamically varied over the course of the experiment in a non-trivial fashion.

Figure 2 **A Taxonomy of the Ease of Dissolving Relationships**



two subjects requests a dissolution of the relationship at the end of a period, their match receives a message asking them if they would consent to the breaking up the relationship. In other words, a subject who initially wants to continue the relationship but whose match requests a break up has an option to change their mind and accept the relationship dissolution. If the subject refuses to break the relationship, both subjects would continue to be matched in the next period. In this case, the subject who wanted to dissolve the relationship is explicitly informed that his request was denied by his match. In both TBA-U and TBA-M, all subjects whose relationship dissolved are then randomly matched to form new matchings.[5]

Our second research question concerns whether reputation mechanisms can restore cooperation when players have the ability to unilaterally dissolve relationships (as in TBA-U). We consider three such mechanisms: (i) Objective Long Run (OLR), (ii) Objective Short Run (OSR) and (iii) Subjective Long Run (SLR). In the objective reputation mechanisms (OLR and OSR), subjects received information about the average frequencies of cooperation and defection, overall and in the first period of a new matching, as well as the total number of matchings so far for both themselves and their match at the start of a new matching. The difference between the long run and short run mechanisms is that in the long run mechanism, the reputation carries over across supergames, while in the short run mechanism, the subjects start each supergame with a blank slate.

In the Subjective Long Run reputation treatment (SLR), each time a matching dissolved, subjects rated their level of satisfaction with their opponent on a 5-point Likert scale, which we call satisfaction score.[6] At the start of a new matching, subjects received information about the average such satisfaction score as well as the total number of matchings so far for both themselves and their match. This treatment was conducted in order to see whether a more realistic form of reputation, such as those found on eBay, Yelp or other platforms, would be effective in promoting cooperation. Implicit in the SLR treatment is the assumption that subjects rate partners who are

[5] To the extent that it was possible, the software tried to ensure that subjects would be matched with a different partner. However, if only one group dissolved their relationship, by default the same subjects would be rematched again in the next period. In this event, both subjects were told that they had been randomly rematched among the pool of subjects whose relationship had broken up in the previous period and no mention was made that it was actually the same match.

[6] In our data analysis, we rescale it to being on $[0, 1]$ to make it comparable to the other reputation treatments.

more cooperative more favorably and partners who are less cooperative less favorably so that the satisfaction scores are signals of cooperative behavior. We show that this is, in fact, the case in Section 7.2.

Table 1 Summary of Treatments

Treatment	Sessions	# of Subjects	# of Periods	Average Earnings
IBA	3	16, 16, 16	88, 98, 87	$21.74
TBA-M	3	16, 20, 18	77, 82, 81	$18.71
TBA-U	3	16, 16, 14	81, 83, 90	$18.76
RM	3	20, 18, 20	79, 75, 88	$15.95
OLR	3	20, 20, 18	75, 76, 88	$20.99
OSR	3	18, 20, 14	89, 77, 76	$19.95
SLR	3	20, 18, 20	97, 77, 94	$21.85

In Table 1, we provide a summary of the treatments conducted in the experiment. For each treatment, we ran three sessions at the experimental laboratory of a public university in the United States. The IBA and TBA-U sessions we conducted in April and November 2014, the TBA-M sessions were conducted in January and February 2016 and the RM treatments in September 2016. In total, 374 subjects participated, of which approximately 2/3 were male. All subjects were students – undergraduate and masters level – currently registered at the university. Because of the random termination rule, sessions varied in length, with the shortest session lasting for 75 periods and the longest session continuing for 98 periods. No session lasted for more than 90 minutes. The subjects' average earnings (including a $4 participation fee) for each treatment are given in Table 1. Note that there is variation in earnings both due to the treatment condition and the different number of periods each session had. The experiment was programmed in z-Tree (Fischbacher 2007). Sample instructions for the TBA-U treatment can be found in Appendix A.5.

4. Equilibrium Analysis

Consider the case of indefinitely binding agreements and let $\delta \in (0,1)$ denote the continuation probability, so that, $1-\delta$ is the probability that the supergame ends (at which point all matchings in the experiment naturally break up). Also let π_{ab} denote the payoff to a player when she chooses action $a \in \{C,D\}$ and her opponent chooses action $b \in \{C,D\}$ in the prisoner's dilemma. In order to sustain mutual cooperation as a subgame perfect equilibrium (using grim-trigger strategies), the expected total payoff from mutual cooperation until the end of the supergame, which is equal to $\frac{\pi_{CC}}{1-\delta}$, should be greater or equal than the expected total payoff from defecting in the first period, followed by mutual defection thereafter, which is equal to $\pi_{DC} + \frac{\pi_{DD}\delta}{1-\delta}$. Hence, we require:

$$\frac{\pi_{CC}}{1-\delta} \geq \pi_{DC} + \frac{\pi_{DD}\delta}{1-\delta} \Leftrightarrow \delta \geq \frac{\pi_{DC}-\pi_{CC}}{\pi_{DC}-\pi_{DD}}.$$

For the payoff matrix in Figure 1, this condition is equivalent to $\delta \geq 0.4$. Given that we set the continuation probability at 90%, mutual cooperation is a subgame perfect Nash equilibrium of our indefinitely repeated game. Moreover, cooperation is also the risk-dominant equilibrium (see, e.g., Dal Bó and Fréchette (2011)).[7]

In the Random Matching treatment, one can show that for the parameters of the experiment (payoffs, continuation probability and group size), cooperation can be sustained as an equilibrium via the community enforcement strategy of Kandori (1992).

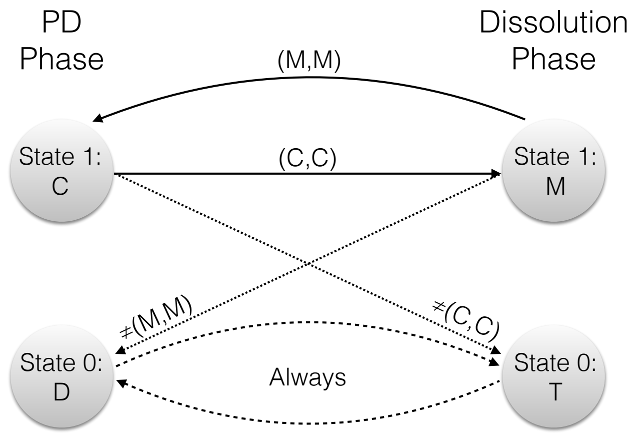
We now show that cooperation can be sustained under the same conditions on δ as in the case of indefinite repetition using simple trigger strategies under temporarily binding agreements when players can unilaterally dissolve relationships. There are two phases of the stage game under TBA-U: (i) the prisoner's dilemma phase, in which players simultaneously choose either to cooperate (C) or defect (D), and (ii) the dissolution decision phase in which subjects simultaneously choose whether to maintain (M) or terminate (T) the relationship. We characterize the equilibrium strategy profile by means of a two-state system (State 0 and State 1) as follows. Players start the supergame in State 1. Whenever subjects are in State 1, they choose to cooperate (i.e., choose C) in the prisoner's dilemma phase and maintain the relationship (i.e., choose M) in the dissolution phase of any period. The system transitions to State 0 following a defection by one or both players in a matching or following the choice to dissolve by one or both players in a matching, otherwise it remains in State 1. Once in State 0, players always choose to defect (i.e., choose D) in the prisoner's dilemma phase and to terminate (i.e., choose T) in the dissolution decision phase. Hence, State 0 is an absorbing state. To summarize, players start by cooperating and maintaining their relationship. They continue to do so as long as both players have always cooperated and maintained the relationship in every previous period, otherwise they defect and dissolve in every subsequent period. A visual depiction of the strategy is shown in Figure 3.

Now suppose that all players follow the proposed strategy profile. There are four cases to consider:

1. The player is in State 1 in the prisoner's dilemma phase: by following the proposed strategy, i.e., choosing C, she will receive a discounted payoff of $\frac{40}{1-\delta}$. If she deviates; i.e., chooses D, then her discounted payoff is $50 + \frac{25\delta}{1-\delta}$, which is less than her payoff from following the proposed strategy for $\delta \geq 0.4$.

[7] We sought a baseline cooperation rate in the IBA treatment of approximately 50% to allow for our treatment effects to go in either direction. We conducted an initial pilot using the exact parameter values (i.e., the game as in the text and a continuation probability of 0.75) as in Dal Bó and Fréchette (2011) since they reported cooperation rates of approximately 50% with these parameter values. However, in contrast to their results, in two such sessions, our subjects cooperated only about 20% of the time. We deemed this cooperation rate too low to allow for meaningful hypothesis tests; therefore, we raised the continuation probability to 90%.

Figure 3 **A Strategy That Supports Cooperation as an Equilibrium**



Note: The action to be taken according to the strategy at each state is given in the shaded circle. The rule describing transitions between states are given on the transition arrows.

2. The player is in State 1 in the dissolution decision phase: if she follows the strategy, i.e., chooses M, her continuation payoff is $\frac{40\delta}{1-\delta}$, while if she deviates, i.e., chooses T, her continuation payoff is only $\frac{25\delta}{1-\delta}$, so the optimal decision is to maintain the relationship.
3. The player is in State 0 in the prisoner's dilemma phase: the player knows that she is matched with another player in State 0 since only players whose relationship has broken up are available for rematching; hence, defection is optimal because the new match will choose action D.
4. The player is in State 0 in the dissolution decision phase: the player is indifferent between choosing M or T because, regardless of whether she chooses M or T, her opponent will choose T; hence the relationship will be terminated.

Therefore, regardless of the state she is in, the player weakly prefers to follow the proposed strategy. Hence, cooperation is sustainable as an equilibrium under temporarily binding agreements.

In the TBA-M institution, because mutual consent is required to dissolve a relationship, cooperation can be sustained in a way very similar to under IBA. Players start by cooperating and continue to cooperate until a defection, at which point they defect forever thereafter. In the dissolution phase of the game, players always choose to maintain the relationship. Off the equilibrium path, players consistently defect after one of their relationships has dissolved.

Finally, a reputation mechanism is present, it is trivial to see that if all players follow the same strategy as described above for the TBA-U institution, and simply ignore their own and their match's reputation scores, then cooperation will also be the equilibrium outcome.

5. Hypothesis Formulation

Given the previous equilibrium analysis, we see that cooperation is, at least theoretically, possible under all of our treatment conditions. Therefore, we state our **null hypothesis** as:

H_0: *Cooperation rates are identical across all treatment conditions.*

Of course, while full cooperation is *theoretically possible* in all treatments, there are multiple equilibria. Furthermore, as several authors have shown (cf. Duffy and Ochs (2009), Dal Bó and Fréchette (2011)) just because cooperation is an equilibrium does not mean that it will be achieved in an experiment. Therefore, we test this null hypothesis against the following **alternative hypothesis**:

H_a: (i) *In the absence of a reputation mechanism, cooperation rates are increasing in the difficulty to dissolve relationships. That is,* $RM < TBA-U < TBA-M < IBA$. (ii) *Introducing a reputation mechanism to the TBA-U institution will lead to higher cooperation.*

The existing literature on whether community enforcement can sustain cooperation under random matching is mixed. While Camera and Casari (2009) show that it is possible to sustain cooperation, Duffy and Ochs (2009) come to the opposite conclusion. There are three main differences between these two papers. The former paper considers cohorts of size 4, has a continuation probability of $\delta = 0.95$ and the payoff matrix is such that trigger strategies can support cooperation (in IBA) if $\delta \geq 0.25$. In contrast, the latter paper has larger groups (6 or 14), a continuation probability of $\delta = 0.9$ and can support cooperation under IBA if $\delta \geq 0.5$. Given that our set-up is closer to Duffy and Ochs (2009) (groups of size 14 or 16, $\delta = 0.9$ and critical discount factor $\delta^* = 0.4$), it is very likely that cooperation will be lower in RM than IBA. Hence, $RM < IBA$.

While the strategy that supports cooperation under temporarily binding agreements is less cognitively demanding than Kandori (1992), it still requires that subjects (correctly) anticipate that a defection will lead to subsequent defection even after the relationship dissolves and the player is rematched. To the extent that players believe they can behave opportunistically, dissolve the relationship to avoid punishment, and start a new and cooperative matching, subjects may be tempted to defect under TBA, thereby undermining the cooperative equilibrium. Thus $TBA - U < IBA$. At the same time, since a relationship has the potential to be of indefinite duration, it is likely that at least some subjects will behave cooperatively. Hence, $RM < TBA - U$.

In principle, in the $TBA - M$ treatment, since a defector cannot unilaterally dissolve the relationship, she cannot escape punishment *if* her match chooses to maintain the relationship. This may be enough to dissuade some potential defections to begin with – hence $TBA - U < TBA - M$. The question is then: are players willing to punish or do they prefer to take their chances in the rematching pool? It is natural to expect that some subjects will try to take their chances in the rematching pool rather than punish, which in turn creates an incentive to defect. Hence, $TBA - M < IBA$.

Consider now part (ii) of our alternative hypothesis. In random matching environments, Duffy and Ochs (2009) and Duffy et al. (2013) show that providing information about an opponent's past action(s) increases cooperation and trust, respectively – though the effect may be small. Camera and Casari (2009) find that cooperation rates are substantially higher under non-anonymous public monitoring than private monitoring, but anonymous public monitoring does not increase cooperation. The lesson from these papers is that reputation mechanisms can promote cooperation but they have to contain sufficiently rich information. The reason is that a rich reputation mechanism can allow for the use of *targeted strategies*; that is, if a player is matched with another who has defected a lot, then she will defect against that player, whereas if she is matched with another who has cooperated a lot, then she may choose to cooperate. Thus, our alternative hypothesis is that cooperation levels are higher when reputation information is available.

6. Results on Cooperation Rates

In this section we test our null hypothesis of equal cooperation across all treatments against our conjectured alternative hypothesis that (i) without reputation cooperation rates are increasing in the difficulty of dissolving relationships and (ii) that introducing a reputation mechanism (to the TBA-U institution) leads to higher cooperation. We also discuss briefly some different patterns that emerge within and across supergames. In the next section, we look at the finer details of behavior when relationships are only temporarily binding, and how, precisely, reputation affects subjects' behavior.

6.1. Cooperation Without Reputation as the Difficulty to Dissolve Relationships Varies

In the top portion of Table 2 we report the average frequency of cooperation, i.e., picking action C in the prisoner's dilemma phase for each treatment. The bottom portion of the table is where we explicitly test our null hypothesis of equal cooperation rates across treatments against the alternative hypothesis that cooperation is more frequent the greater the difficulty in dissolving the relationship. We report the result of a non-parametric trend test on session averages.[8] For the interested reader, we report the results of all pairwise treatment comparisons in Table A.1 in Appendix A.

The first two columns consider only the first period of each supergame. This gives an impression of subjects' intention to establish cooperative relationships and does not confound differences in

[8] We reach an identical conclusion from regressions of the form: $C = \alpha + \beta\texttt{difficulty}$, where difficulty is 0, 1, 2, or 3 for the RM, TBA-U, TBA-M or IBA treatments, respectively. For the first period of the first supergame, OLS is used while for all other treatments, we estimate a random-effects model. In all cases, standard errors are corrected for clustering at the session level. We omit these results in the interest of brevity

Table 2 Average Cooperation Rate (No Reputation)

	Treatment	First Period: First SG	First Period: All SGs	All Periods: First SG	All Periods: All SGs
Cooperation Rates	RM	43.10%	18.61%	12.01%	7.46%
	TBA-U	65.22%	36.98%	60.92%	39.11%
	TBA-M	59.26%	36.18%	54.98%	49.42%
	IBA	64.58%	63.86%	42.71%	62.64%
Non-Parametric Trend ($p-$value)		0.114	0.010	0.038	0.007

Note: The Non-Parametric Trend test takes the session average as the unit of independent observation.

treatment or supergame length. As can be seen, in the very first supergame, cooperation rates are fairly similar for TBA-U, TBA-M and IBA and only slightly lower for the RM treatment. In fact, our non-parametric trend test cannot reject the null hypothesis of equal cooperation. However, when we consider the first period of all supergames, we see a clear difference between RM, TBA-U/TBA-M and IBA. Moreover, both of the parametric and non-parametric trend tests reject the null hypothesis of equal cooperation in favor of increasing cooperation as it becomes more difficult to dissolve the relationship.

In the third and fourth columns, we report the same information but we average over all periods of the supergame. When looking at the all periods of the first supergame, cooperation deteriorated rapidly in the RM treatment (compared to the first period) but was more stable in the other treatments. On the strength of this difference, both the parametric and non-parametric trend tests reject the null hypothesis of equal cooperation in favor of our predicted alternative hypothesis. Finally, when looking at all periods of all supergames, we see the precise monotone relationship in cooperation that we predicted; namely, $RM < TBA-U < TBA-M < IBA$. Additionally, both the parametric and non-parametric tests reject the null hypothesis at the 1% level of significance. Thus, we have the following result:

Result 1 *We reject H_0 in favor of H_a (i): over the long-run, cooperation is increasing in the difficulty of dissolving relationships.*

Looking at Table 2, it suggests that the different treatments generate different dynamics within and across supergames. For example, in the RM treatment, it appears that cooperation is decreasing both within and across supergames. On the other hand, in the IBA treatment, the tendency to cooperate increases across supergames. To investigate this, in Table 3 we report the results of a linear random effects regression where the dependent variable is an indicator for whether the subject cooperated and we include the supergame index and the supergame period – separately for each treatment – as explanatory variables, as well as treatment-specific constants.

Table 3 Cooperation Across Supergames and Periods (No Reputation)

Parameter	RM		TBA-U		TBA-M		IBA	
Supergame	-0.015^{***}	(0.005)	-0.016^{**}	(0.007)	0.016	(0.012)	0.027^{***}	(0.009)
Supergame Period	-0.004^{***}	(0.001)	0.009^{**}	(0.004)	0.007^{***}	(0.002)	0.002	(0.002)
Constant (RM) / Treat. Dummy (Oth.)	0.178^{***}	(0.045)	0.255^{*}	(0.150)	0.174^{***}	(0.064)	0.454^{***}	(0.276)
Observations	17,272	R^2	0.192					

Note 1: Linear random effects regression with full set of treatment interactions. Presented as separate columns for ease of presentation.
Note 2: Robust standard errors (clustered at session level). Significance levels: *** 1%; ** 5%; * 10%.

Table 3 confirms that cooperation declines in the RM treatment both within and across supergames. For the TBA-U treatment, cooperation appears to decline across supergames, but is actually increasing within supergames. In contrast, the TBA-M treatment experiences (roughly) stable cooperation rates across supergames and increasing cooperation within supergames. The finding of increasing cooperation within supergames in the two TBA treatment is likely due to the beneficial effects of sorting which is facilitated by the ability to dissolve relationships. We will discuss sorting in more detail below. Finally, in the IBA treatment, cooperation is stable within a supergame, but cooperation increases in later supergames.

6.2. Cooperation and Reputation When Relationships Can Be Unilaterally Dissolved

We repeat a similar analysis as in the previous subsection but now focus our attention on the role of reputation in fostering cooperation. To that end, in Table 4, we report the average cooperation rate in the first period of the supergames considered (columns 1 and 2) and over all periods of the supergames considered (columns 3 and 4), and we compare the cooperation rates to the TBA-U institution, where subjects could unilaterally terminate a relationship.[9] As can be seen, cooperation rates are higher than TBA-U when a reputation is present in all four comparisons. However, the differences are not statistically significant in the first supergame and only emerge once we consider all supergames.

In Table A.2 in Appendix A, we report the results of tests comparing cooperation in either TBA-U or IBA with the three reputation treatments. The main conclusion from this exercise is that cooperation is significantly higher over the long-run in all reputation treatments than in TBA-U. Moreover it is also significantly higher than in the IBA treatment when the Objective Long Run reputation mechanism is used. Thus we have the following result:

[9] Remember that in our reputation treatments, just like in the TBA-U treatments, subjects are able to unilaterally dissolve relationships. Hence a comparison between the reputation treatments and TBA-M is not meaningful.

Table 4 **Average Cooperation Rate (Reputation)**

		First Period		All Periods	
	Treatment	First SG	All SGs	First SG	All SGs
Cooperation Rates	TBA-U	65.22%	36.98%	60.92%	39.11%
	OLR	79.31%	77.86%	81.72%	83.01%
	OSR	82.69%	70.58%	75.34%	65.72%
	SLR	75.86%	59.04%	84.03%	66.56%
Comparing TBA-U with Reputation treatments ($p-$values)	R.E. Reg.	0.372	0.004	0.290	0.000
	Rank sum test	0.459	0.034	0.309	0.034

Note 1: The row "R.E. Reg." reports the $p-$value for the test that the coefficient on a reputation dummy variable is equal to zero in a linear random effects regressions (except for the regression of first period choice in first supergame, in which there is only one observation per subject) of cooperation on an indicator variable for a reputation mechanism being present, and with standard errors clustered at the session level.
Note 2: The rank sum test takes the session average as the unit of independent observation.

Result 2 *We reject H_0 in favor of H_a (ii): over the long-run, cooperation is higher when a reputation mechanism is available compared to TBA-U.*

The fact that the differences in cooperation rates are more striking when looking at all supergames than just the first one suggests that the treatments have different dynamics across (and possibly within) supergames. Recall from Table 3 that the TBA-U treatment had cooperation rates declining across supergames but increasing within supergames. Table 5 contains the results of this exercise for the reputation treatments.

Table 5 **Cooperation Across Supergames and Periods (Reputation)**

Parameter	SLR		OSR		OLR	
Supergame	0.006	(0.004)	0.002^{**}	(0.001)	-0.001	(0.001)
Supergame Period	0.002^{***}	(0.001)	-0.003	(0.002)	0.004^{***}	(0.001)
Constant (SLR) / Treat. Dummy (Oth.)	0.625^{***}	(0.120)	0.049	(0.144)	0.167	(0.122)
Observations	14,016	R^2	0.029			

Note 1: Linear random effects regression with full set of treatment interactions. Presented as separate columns for ease of presentation.
Note 2: Robust standard errors (clustered at session level). Significance levels: *** 1%; ** 5%; * 10%.

In contrast to the TBA-U treatment, cooperation across supergames is never significantly decreasing with reputation and in the case of OSR is significantly increasing. The SLR and OLR treatments share with the TBA-U treatment that cooperation significantly increases within supergames while for the OSR treatment, the effect is negative but not significant. Thus the presence of a reputation mechanism leads to much more stable levels of cooperation over the course of the experiment. Comparing the OLR and SLR treatments, none of the coefficients are individually statistically different; however, a joint test suggests that cooperation rates are significantly higher in OLR, at least for sufficiently long supergames. Thus, an objective measure of reputation may be better at promoting cooperation than a subjective measure.

7. Behavior Under Temporarily Binding Agreements

Having shown that the patterns of cooperation are consistent with our predictions, in this section we dig deeper into behavior when subjects have the flexibility to dissolve relationships. Our goal is to investigate exactly how the flexibility to dissolve relationships, possibly combined with a reputation mechanism, leads to these results. In this section we analyze the frequency with which relationships are dissolved in our TBA-U and TBA-M treatments, as well as the underlying reasons for it. We then show that for the subjective reputation treatment, the satisfaction scores given upon dissolution of relationships are meaningful signals of cooperative behavior. We then study in more detail the role of reputation in fostering cooperation. Finally, we show that reputation facilitates sorting based on how cooperative subjects are and that it influences the continuation value to dissolving a relationship.

7.1. The Decision To Dissolve a Relationship

In Table 6, we look at how the outcome of the prisoner's dilemma phase affects the outcome of the dissolution phase for all treatments where subjects had the flexibility to dissolve relationships – both in terms of which subject(s) choose to dissolve and what this implies for whether the matching continues or is dissolved. Remember that in all treatments except the TBA-M treatment, it was enough for one subject to request a dissolution for the matching to terminate. Hence, for these treatments, the number in the row "% Matchings Dissolved" is 100 minus the frequency that neither subject requested a dissolution. In contrast, for the TBA-M treatment, a matching dissolves if and only if both subjects request it. Therefore, the "% Matchings Dissolved" is equal to the previous row (which is the frequency that both subjects choose to dissolve). Recall also that subjects in the TBA-M treatment were allowed to revise their initial decision to maintain the relationship if they found out that their match wished to dissolve. In the table we report statistics based on the final decision to dissolve; that is, cases where initially one subject wanted to continue but changed his mind after being told that his match wanted to dissolve are counted under "both subjects request dissolution". In Table A.3 in the Appendix we provide more detail for the TBA-M treatment based on which party initiated the dissolution request and whether or not it was accepted.

Panel (a) considers the case in which both subjects cooperated. As can be seen subjects virtually never choose to dissolve in the dissolution phase when both of them cooperated in the prisoner's dilemma phase. Reiterating the point made in Table 2, we see that subjects in the reputation treatments coordinated on mutual cooperation significantly more often than in either of the TBA treatments.

Table 6 Frequency of Dissolution Requests and Frequency of Matchings That Dissolve By Prisoner's Dilemma Game Outcome

(a) Both Subjects Cooperate; i.e., (C,C)

Who Requests Dissolution	OLR	OSR	SLR	TBA-U	TBA-M
Neither Subject	99.05	97.97	99.43	99.83	100.00
One Subject	0.94	2.04	0.58	0.18	0.00
Both Subjects	0.00	0.00	0.00	0.00	0.00
% Matchings Dissolved	0.94	2.04	0.58	0.18	0.00
% Outcome (C,C)	77.85	53.83	60.47	29.92	38.80

(b) One Subject Cooperates; One Subject Defects; i.e., $(C,D)/(D,C)$

Who Requests Dissolution	OLR	OSR	SLR	TBA-U	TBA-M
Neither Subject	52.94	36.60	55.84	51.82	59.35
Subject Choosing C	28.15	21.20	26.18	28.57	19.35
Subject Choosing D	13.45	26.20	12.62	12.61	6.52
Both Subjects	5.46	16.00	5.36	7.00	14.78
% Matchings Dissolved	47.06	63.40	44.16	48.18	14.78
% Outcome $(C,D)/(D,C)$	10.34	23.78	12.18	18.38	21.25

(c) Both Subjects Defect; i.e., (D,D)

Who Requests Dissolution	OLR	OSR	SLR	TBA-U	TBA-M
Neither Subject	15.07	23.57	40.45	35.86	28.21
One Subject	46.32	49.68	49.16	44.92	26.94
Both Subjects	38.60	26.75	10.39	19.22	44.86
% Matchings Dissolved	84.92	76.43	59.55	64.14	44.86
% Outcome (D,D)	11.82	22.40	27.35	51.70	39.95

Note 1: The "% Dissolve" is the percentage of matchings that dissolve. In all treatments, except TBA-M, this is 100−(frequency neither subject dissolves). In the TBA-M treatment, this is the frequency that **both subjects** dissolve.
Note 2: The "% Outcome" is the percentage of times that the given outcome (e.g., (C,C)) occurred in that treatment.

A more interesting case is panel (b), which considers the case in which one subject cooperated and the other subject defected. Here we see that many more subjects choose to dissolve the relationship. Indeed, in TBA-U, OLR and SLR, approximately 45% of matchings are dissolved when one subject defects. At the two extremes are the OSR treatment where 63.40% of matchings dissolve and the TBA-M treatment where only 14.78% of matchings dissolve. Under OSR, we note that the defectors are more likely to request a dissolution of the relationship than the cooperators (which may be indicative of a "defect-and-run" strategy), while the reverse is true in the other treatments.[10] At the same time, the fact thate – roughly 30–35% of the time – the subject who cooperated requests a dissolution indicates that many prefer to take their chances with a new match rather than enter into a punishment phase or try to build a cooperative relationship with the current match.

[10] A random effects regression with an indicator for the OSR treatment as explanatory variable indicates that defectors are significantly more likely to dissolve in the OSR treatment ($p < 0.01$) than in the other treatments.

The fact that only 14.78% of matchings dissolve after one defection in TBA-M (versus over 40% in all other treatments) speaks to the difficulty in dissolving relationships due to the mutual consent requirement.[11] Interestingly, it appears that subjects are somewhat less likely to choose to dissolve than compared to TBA-U. For example, 59.35% of the time, neither subject requests to dissolve, while this number is only 51.82% in TBA-U.[12] This could be for two non-mutually exclusive reasons: (i) subjects fear poisoning the relationship from a failed attempt at dissolving or (ii) subjects believe that they would be better off by maintaining the relationship. In Table A.3 we show two other interesting results. First, focusing on cases where only one subject defected, we see that when only one player makes an initial choice to dissolve the relationship it is, by a 3 to 1 margin, the player who cooperated. Second, we see that the subject who had initially chosen to maintain the relationship and is asked again because their match wanted to break it up declines to do so 80% of the time, regardless of whether he or she was the player who cooperated or defected.

Finally, turn next to panel (c) which considers the case in which both subjects defected. As can be seen, there is a great deal of variation in the frequency with which matchings dissolve – ranging from a low of 44.86% in the TBA-M treatment to a high of 84.92% in the OLR treatment. In terms of treatment comparisons, we note that the rate of matching dissolution is significantly lower in TBA-M than any other treatment (for TBA, $p = 0.06$; in all other cases, $p < 0.01$). Also the rate of matching dissolution is significantly lower in SLR than in the other two reputation treatments ($p < 0.033$).

The Implications of Dissolving Relationships. We next investigate subsequent outcomes conditional on relationships dissolving or being maintained. Our purpose is to gain insight into whether subjects over-use their option to dissolve relationships. Specifically, we look at the likelihood of (C, C) arising in the next period depending on the outcome in the prisoner's dilemma phase and the dissolution phase. Results can be found in Table 7.

Consider the two TBA treatments in panel (a). First, observe that when a matching plays (C, C) and the matching is maintained, then they almost always maintain mutual cooperation in the next period. In contrast, when both players defect, then they almost never experience mutual cooperation in the next period, regardless of whether the relationship is maintained or dissolved. However, when only one subject defected, subjects are more than twice as likely to play (C, C) in

[11] A random effects regression with an indicator for the TBA-M treatment as explanatory variable indicates that matchings are significantly less likely to dissolve in the TBA-M treatment ($p = 0.052$).

[12] Indeed, if we look at the initial decision to dissolve, it appears that both the subject who cooperated (32.61% versus 35.57%) and the subject who defected (11.74% versus 19.61%) are less likely to make an initial decision to dissolve than subjects in TBA-U; however neither difference is statistically significant.

Table 7 The Frequency of Observing $(\mathbf{C},\mathbf{C})$ Conditional on Outcome in the Previous Period

(a) TBA-U and TBA-M Treatments

	TBA-U		TBA-M	
Outcome Last Period	Dissolved	Maintained	Dissolved	Maintained
(C,C)	0.00%	97.76%	–	95.18%
(C,D) or (D,C)	6.21%	20.63%	8.46%	18.72%
(D,D)	0.36%	1.60%	1.87%	1.15%

(b) Reputation Treatments

	OSR		OLR		SLR	
Outcome Last Period	Dissolved	Maintained	Dissolved	Maintained	Dissolved	Maintained
(C,C)	21.43%	91.28%	44.12%	98.40%	21.43%	97.41%
(C,D) or (D,C)	18.73%	17.95%	18.75%	14.53%	11.53%	17.68%
(D,D)	9.63%	1.98%	5.45%	12.50%	4.42%	0.77%

Note: This table reports the empirical frequency that, in period t, the outcome is $(\mathbf{C},\mathbf{C})$ conditional on the outcome in period $t-1$. One could also consider a table which looks at subjects' choices conditional on outcomes in the previous period. Such a table would show the same thing: the likelihood of cooperation is higher when pairings are maintained.

the next period if they *maintain* the pairing rather than dissolve it, and this difference is significant (Wilcoxon signed-rank; $p = 0.028$). This suggests that subjects would be better off by maintaining relationships and hence that they over-use the flexibility to dissolve. It also points to a possible beneficial role of mutual consent. Whereas 48% of pairings dissolved following a period in which one player defected in TBA-U, in TBA-M the same number is only 14.8%. Therefore, by making it more difficult to dissolve relationships, TBA-M increases the chance that a matching can recover from a single defection and return to mutual cooperation. Indeed, unconditional on whether the relationship is maintained or dissolved, following a defection by one player, it is nearly 25% more likely to coordinate on mutual cooperation in the next period in TBA-M than in TBA-U.

Matters are more nuanced in our three reputation treatments. First, as in the TBA treatments, we see that, following a period of (C,C) outcome, subjects are more likely to experience another period of mutual cooperation if they maintain the relationship and in fact, dissolutions in that case are very rare as Table 6 shows. Following mutual defection, subjects are more likely (in OSR and SLR) to experience mutual cooperation if the relationship dissolves and are more likely to experience mutual cooperation (in OSR and OLR) if the relationship is dissolved following a period in which one player defected.

Comparing the mutual cooperation rates between the TBA and the reputation treatments when a matching is dissolved, we see that rates are significantly higher in the reputation treatments. Thus, another role of reputation systems is that they increase the benefit of dissolving relationships following a defection by one or both players. However, as we show in the Appendix (Table A.4), these benefits accrue primarily to subjects with above average reputation scores in their session. Specifically, following a period of either mutual defection or defection by one player, subjects with above average reputation are almost always more likely to experience mutual cooperation in the next period if the matching is dissolved than if it is maintained.[13] In contrast, subjects with below average reputation are, with two exceptions, less likely to experience mutual cooperation in the next period if the pairing is dissolved than if it is maintained.

In Table 8, we report the results of a regression analysis on the decision to terminate a relationship for our reputation treatments. The main takeaway from this analysis is that subjects at least partially recognize that dissolving matchings is more advantageous to subjects with higher reputation. This can be seen by the positive and significant coefficient on a subject's own reputation. Another interesting finding is the subjects are more forgiving (i.e., less likely to dissolve) the higher the reputation of their match. These results point to a result in Wilson and Wu (Forthcoming) who study endogenous termination decisions, but where the termination payoff is exogenously specified. In particular, they show that subjects are more likely to terminate when their exogenous termination payoff is higher. Although termination payoffs are endogenous in our experiment, Table A.4 suggests that termination payoffs are increasing in reputation and Table 8 shows that subjects with higher reputation are more likely to dissolve. However, in contrast to Wilson and Wu (Forthcoming), in our reputation treatments we observe higher termination payoffs *coexisting* with higher overall rates of cooperation.

Table 8 The Role of Reputation in the Decision to Terminate a Relationship

	OLR		OSR		SLR		Pooled	
(Own Action, Match's Action)								
(D,C)	0.242^{***}	(0.023)	0.457^{***}	(0.053)	0.288^{***}	(0.057)	0.346^{***}	(0.047)
(C,D)	0.333^{***}	(0.019)	0.365^{***}	(0.018)	0.329^{***}	(0.104)	0.339^{***}	(0.036)
(D,D)	0.594^{***}	(0.091)	0.518^{***}	(0.016)	0.461^{***}	(0.116)	0.510^{***}	(0.057)
Subject's Reputation	0.254^{***}	(0.054)	0.143^{**}	(0.059)	0.233^{***}	(0.079)	0.171^{***}	(0.043)
Match's Reputation	−0.052	(0.056)	-0.107^{***}	(0.013)	-0.140^{**}	(0.066)	-0.112^{***}	(0.029)
Subject's Number of Matchings	0.001	(0.001)	0.007^{*}	(0.004)	−0.004	(0.003)	−0.001	(0.001)
Match's Number of Matchings	0.004^{***}	(0.000)	−0.004	(0.006)	0.002^{**}	(0.001)	0.002^{**}	(0.001)
Index of Supergame	-0.006^{***}	(0.000)	0.011^{***}	(0.002)	0.003	(0.013)	0.003	(0.003)
OLR							0.056^{*}	(0.028)
OSR							0.071^{**}	(0.028)
Constant	−0.187	(0.099)	−0.082	(0.066)	−0.082	(0.050)	-0.111^{***}	(0.035)
Observations	4546		3754		4708		13008	
R^2	0.486		0.340		0.249		0.326	

Note 1: *, ** and *** denote significance at the 10, 5 and 1% levels, respectively. The table reports linear random effects models with standard errors in parentheses which have been corrected for clustering at the session level.
Note 2: The variable "Reputation" is on a continuous scale from 0 to 1, with 1 indicating a "perfect" reputation. In the OLR, reputation is the overall frequency of cooperation, in the OSR, it is the frequency of cooperation in the supergame and in the SLR, it corresponds to the satisfaction score.

[13] The one exception is the SLR treatment, following one defection.

7.2. The Subjective Reputation Score As A Proxy For Cooperative Performance

Our analysis above suggests significant differences in dissolution request rates when the reputation mechanism is objective versus subjective. In this section we further analyze the differences between these two types of reputation mechanisms. As discussed in Section 3 the objective reputation scores are statistical calculations (average frequencies of cooperation and defection, overall and in the first period of a new matching, as well as the total number of matchings so far). In contrast, the subjective reputation scores are assigned by the subjects themselves based on a 5-point Likert scale designed to report their level of satisfaction with their match following the dissolution of a relationship (either natural or caused by the subjects).[14] Hence, while a single defection affects one's reputation score in a certain and mechanistic fashion when reputation is objective, its impact is less certain when reputation is subjective as it depends on how strong a match's reaction to a defection is. Furthermore, the act of choosing to dissolve may itself be viewed negatively by one's match, who may respond with a low score. Therefore, it is interesting to understand how subjects rate their match upon the dissolution of a relationship. In particular, how accurate is the reputation score as a signal of one's cooperativeness?

Figure 4 plots a subject's average satisfaction score in the SLR treatment (normalized to $[0, 1]$) at the end of the experiment against his or her frequency of cooperation. As can be seen, the two variables are strongly positively associated with each other; that is, subjects who cooperated more frequently received higher satisfaction scores from their matches throughout the experiment. More interestingly, there appear to be two main clusters of subjects: those who cooperated frequently, and so have high satisfaction scores, and those who cooperated infrequently, and so have low satisfaction scores.

In Table 9 we report the results from a random effects ordered logit model where the dependent variable is the satisfaction score given by a subject about his/her opponent upon dissolution of the relationship. We include as explanatory variables, the fraction of times one's match defected over the course of the matching, as well as indicator variables for one's own defection in the most recent period, one's match's defection in the most recent period and whether the matching was naturally dissolved or dissolved by choice of one or both subjects in the matching. We also include both subjects' satisfaction scores (as of at the start of the last supergame), the length of the matching and the supergame index.

[14] In the SLR treatment, when a matching dissolved, subjects were taken to another screen and asked to rate their level of satisfaction with their match. On this screen, they could see the history of actions and profits for their matching. The previous average satisfaction score was not visible – though, presumably, they could recall it since this information was provided in the prisoner's dilemma phase.

Figure 4 **Satisfaction Scores and the Frequency of Cooperation (End of Experiment)**

Note: Recall that, for the analysis of results, the satisfaction scores were rescaled so that, like the OLR and OSR treatments, it is between 0 and 1.

Somewhat surprisingly, we see that the overall fraction of defections over the course of the matching by one's match does not seem to impact the satisfaction score they receive; instead, only a defection in the most recent period matters. In this way, a subjective reputation mechanism may provide stronger incentives because a player may opportunistically defect only once and still receive a low satisfaction score, while such behavior under an objective measure would only marginally lower the subject's reputation. We can also see that matchings that dissolved naturally lead to higher satisfaction scores, as did matchings that lasted longer. Interestingly, even after controlling for actions, the match's prior reputation score appears to affect the reputation score that a subject gives her: matches with higher previous average satisfaction score, receive higher satisfaction scores. Finally, it appears that subjects gave lower satisfaction scores in later supergames.

To get a better sense of the magnitude of the effects, the table also reports marginal effects for the variables. For example, if a subject's match defected in the most recent period, then she is 29.3 percentage points more likely to give a satisfaction score of 1 (lowest satisfaction) and 17.3 percentage points less likely to give a satisfaction score of 5 (highest satisfaction) than if her match had cooperated. Concerning the match's prior reputation score, it seems like the effect is largely confined to low satisfaction scores. For example, a match with a perfect reputation is 11.5 percentage points less likely to receive a score of 1 and modestly more likely to be given a satisfaction score of 2 compared with a match with a perfectly *bad* reputation.

Table 9 The Determinants of a Subject's Satisfaction With His/Her Opponent

	Model		Marginal Effect on Predicted Satisfaction Score				
			1	2	3	4	5
Fraction of Match Defections	−0.403	(0.404)	0.048	−0.005**	−0.014	−0.007	−0.021
Subject Defection (last period)	−0.257	(0.409)	0.031	−0.003	−0.009	−0.005	−0.014
Match Defection (last period)	−2.401***	(0.638)	0.293***	0.007	−0.071***	−0.056***	−0.173***
Natural Termination	1.707**	(0.830)	−0.205**	0.002	0.052***	0.035***	0.115**
Subject's Prior Reputation	0.367	(0.525)	−0.043	0.005	0.013	0.006	0.019
Match's Prior Reputation	0.975***	(0.378)	−0.115*	0.013***	0.035	0.017	0.050*
Length of Matching	0.152***	(0.029)	−0.018***	0.002**	0.005**	0.003*	0.008***
Index of Supergame	−0.051***	(0.009)	0.006***	−0.001**	−0.002**	−0.001**	−0.003***

Note: *, ** and *** denote significance at the 10, 5 and 1% levels, respectively. The table reports a random effects ordered logit model with standard errors in parentheses which have been corrected for clustering at the session level, as well as marginal effects for each variable and each possible satisfaction score.

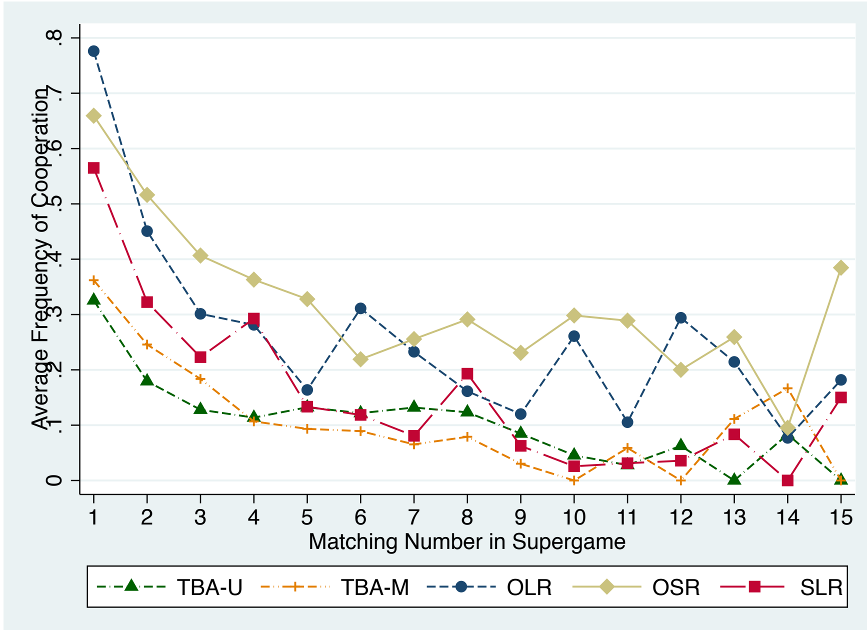
7.3. Reputation and Cooperation

In Section 7.1, we showed that most relationships are dissolved following a defection by one or both players. This represents part of the equilibrium logic that supports cooperation when players have the flexibility to dissolve their relationship. The next step in the equilibrium logic is for players to recognize that they will be rematched from within a less cooperative group; consequently the players should be more likely to defect. Indeed, it appears that subjects gradually come to this realization. To show this, in Figure 5, we plot the average cooperation rate of subjects in the first period of a matching against the index of the matching within a supergame. In this average we only consider subjects who had at least one matching dissolve during the course of a supergame. This restriction makes makes our test more conservative because subjects whose initial matching lasted until the end of the supergame generally tend to be more cooperative.[15]

As can be seen, for all treatments, the trend in cooperation is generally downwards, indicating that the pool of subjects available for rematching gets less and less cooperative as the number of matchings they have experienced since the start of the supergame increases. For all treatments, regressions confirm that this effect is significant (in all cases, $p < 0.05$). To the extent that this effect spills over across supergames, it may provide another explanation for the downward trend in cooperation seen in the TBA-U treatment. We also see that there are level effects across treatments. Specifically, subjects are uniformly more likely to cooperate in the first round of a new matching in the two objective reputation treatments, followed by the subjective reputation treatment. Subjects in the TBA-U and TBA-M treatments, where reputation is absent, are significantly less likely to cooperate in the first round of a rematching than in any of the reputation treatments. This

[15] As shown in Table 6 almost all matchings which break up saw at least one subject defect during the course of their interaction; hence it is reasonable to assume that starting with their second matching within a supergame, subjects have experienced at least one defection.

Figure 5 **The Evolution of Cooperation in the First Period of Matchings Within the Same Supergame**



Note: We only include data up to the first 15 matchings that a subject had in a supergame. This represents 96% of the relevant data.

further corroborates the results of Table 7. In other words, the existence of a reputation mechanism not only leads to fewer dissolutions but it also makes them more likely to cooperate following a rematching.

To see the role of reputation more concretely, in Table 10 we look at the determinants of cooperation in the first round of a new matching in a series of fixed effects regressions.[16] We include all relevant information that was observable to subjects at the time they made their decision. This includes their own and their match's reputation score(s), and the number of matchings that they have participated in (since the start of the experiment in OLR and SLR and since the start of the supergame in OSR).

The first thing to notice is that across all treatments, the match's reputation is the crucial determinant for whether the subject cooperates in the first period of a new matching. Interestingly, the effect of reputation is strongest in the subjective reputation treatment – a finding confirmed when we pool the data and test for treatment differences ($p < 0.01$). In particular, we see that subjects in the objective reputation treatments have a higher baseline rate of cooperation (the constant is approximately 0.23) but respond relatively weakly to their match's reputation; whereas subjects in the SLR treatment are the opposite. It is also of interest to note that while subjects only

[16] Using the `xtoverid` command suggested by Schaffer and Stillman (2010) – an alternative to the Hausman test, which can accommodate clustering – we were led to prefer the fixed effects formulation.

Table 10 The Role of Reputation in First-Period Choice to Cooperate

	OLR		OSR		SLR	
Match's Reputation (Overall)	0.149**	(0.071)	0.193***	(0.048)	0.557***	(0.020)
Match's Reputation (Period 1)	0.124	(0.103)	0.108***	(0.027)		
Match's Number of Matchings	−0.006**	(0.003)	−0.005	(0.005)	−0.003***	(0.001)
Subject's Reputation (Overall)	0.299*	(0.173)	0.029***	(0.010)	0.144	(0.154)
Subject's Reputation (Period 1)	−0.153	(0.150)	0.072	(0.071)		
Subject's Number of Matchings	0.003***	(0.001)	−0.004	(0.006)	0.002	(0.001)
Constant	0.239***	(0.073)	0.228***	(0.074)	0.011	(0.084)
Observations	974		1268		1362	
R^2	0.306		0.114		0.165	

Note 1: *, ** and *** denote significance at the 10, 5 and 1% levels, respectively. The table reports linear fixed effects models with standard errors in parentheses which have been corrected for clustering at the session level.
Note 2: The variable "Reputation" is on a scale from 0 to 1, with 1 indicating a "perfect" reputation. In the OLR, reputation is the frequency of cooperation – overall or in Period 1 of a matching; in the OSR, it is the frequency of cooperation – overall or in Period 1 – in the current supergame; and in the SLR, it corresponds to the satisfaction score.

seem to care about the overall cooperation rate in the OLR treatment, in the OSR treatment, they care both about their match's overall frequency of cooperation and their frequency of cooperation in the first period of their previous matchings. We also see that the number of previous matchings that one's match had is negatively associated with cooperation, and the effect is significant in two of the treatments. Table 10 suggests that a subject's own reputation has relatively little impact on their own decision to cooperate, though the direction is generally positive.[17]

7.4. The Option to Terminate Relationships as a Sorting Mechanism

In this section, we show that the option to terminate a relationship acts as a sorting mechanism between subjects who tend to cooperate in the Prisoner's Dilemma game and those who tend to defect and that the existence of a reputation mechanism further facilitates sorting. Also, we argue that this sorting is a large part of the reason why we see non-trivial cooperation rates in TBA and significantly higher cooperation rates in our reputation treatments.

By "sorting", we refer to the mechanism by which, by the end of a supergame, subjects who tend to cooperate more frequently, i.e., the *cooperators*, are more likely to be paired together, and subjects who defect more often, i.e., the *defectors*, are more likely to be paired together as well. To this effect, we compare two statistics: the absolute difference in cooperation rates between subjects who were matched together (i) in the first period of the first matching of a supergame and (ii) in the first period of the last matching of a supergame. In both cases, cooperation rates are measured as the frequency of choosing "Cooperate" in the Prisoner's Dilemma game up to that period.

[17] Because there is relatively little variation in a subject's own reputation score in the first period of new matchings, much of the effect of one's own reputation score is likely subsumed into the fixed effects. Indeed, when we estimate the model with random effects, one's own reputation has a significantly positive effect. However, as noted, the test of Schaffer and Stillman (2010) suggests that fixed effects is more appropriate.

Table 11 Reputation as a Sorting Device

	Absolute Difference in Cooperation Rate Between Subjects in a Matching As Of:	
Treatment	First Period of First Matching	First Period of Last Matching
IBA	0.326	0.326
TBA-M	0.341	0.286
TBA-U	0.361	0.270*
OLR	0.248	0.168***
OSR	0.314	0.194***
SLR	0.333	0.214***

Note 1: For each subject, i, we define the cooperation rate in period p as $(1/p)\sum_{t=1}^{p}\mathbf{1}[C_{it}=1]$, where C_{it} takes value 1 if subject i cooperated in period t. We then take the absolute difference between the subject's cooperation rate and her match's cooperation rate for, in the first column, the first period of **first** matching of a supergame and, in the second column, the first period of the subject's **last** matching in the supergame. Observe that for the IBA treatment, these are the same periods since rematching is not possible. Hence, as we see in the table, the numbers are the same for the IBA treatment.
Note 2: *, ** and *** denote significance at the 10, 5 and 1% levels for tests comparing the IBA and each other treatment. The test results are based on random effects regressions corrected for clustering at the session level with treatment dummies as explanatory variables. The conclusion remains the same if we run non-parametric tests on session averages; however we lose significance on the IBA/TBA comparison and the significance is only at the 5% level for the other comparisons.

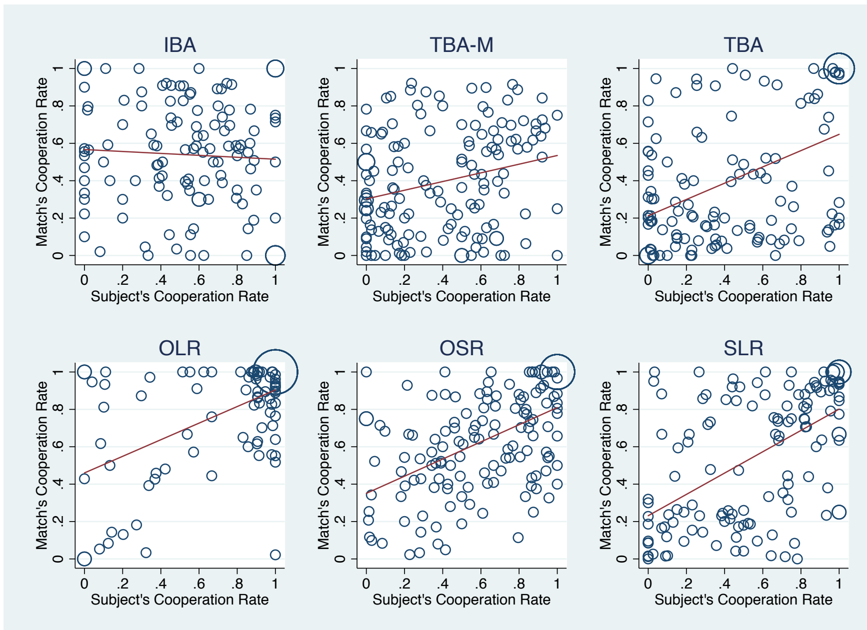
Table 11 contains these two statistics for each treatment (except RM).[18] First, note that there is no statistically significant difference between any of the treatments for the absolute difference in cooperation rates in the first period of the first matching of a supergame. This is not surprising since subjects are randomly matched into initial matchings at the beginning of each supergame in all treatments. Also, since rematching is not possible within a supergame in the IBA treatment, both statistics coincide. In contrast, in all treatments where rematching is possible, by the first period of the last matching in a supergame, the absolute difference in cooperation rates is substantially smaller. In other words, subjects are more likely to end the supergame matched with another subject who is more similar to them in terms of cooperation rate than the subject they were initially matched with at the start of the supergame. Finally, notice that for all treatments except for TBA-M, the absolute difference in cooperation rates is significantly smaller than in the IBA treatment. The reason the difference is not significantly in the TBA-M treatment is likely due fact that there are on average fewer re-matchings for each subject in a supergame due to the mutual consent requirement for a matching dissolution; in other words, there are fewer opportunities for sorting to occur.

To provide a slightly different perspective, Figure 6 represents, for each pair of subjects matched together in their last matching of a supergame, the two subject's cooperation rates up to that period. Sorting can be seen by noting the positive relationship between the two variables in all treatments except IBA. That is, consistent with Table 11, subjects who have the flexibility to dissolve matchings, are likely to end the supergame matched to someone closer to themselves in

[18] Sorting is, by definition, not possible in either the RM or IBA treatments, but we include the numbers for IBA because these treatments have similar underlying cooperation rates, making a comparison more meaningful.

terms of cooperation rates up to that period. Finally, observe that both Table 11 and Figure 6 show that sorting is stronger in the reputation treatments than in the two TBA treatments. Thus, another important role of reputation is to facilitate sorting of cooperators with cooperators and defectors with defectors.

Figure 6 **Frequency of Cooperation of Pairs up to First Period of Last Matching in a Supergame**



Note: The circles represent the raw data and larger circles indicate a larger number of observations with the given combination of subject's and match's cooperativeness. The line represents the best linear fit (ignoring the panel structure) through the data.

8. Concluding Remarks

In this paper, we experimentally study the role of flexibility to dissolve relationships on cooperative behavior in a repeated prisoner's dilemma game. At one extreme, subjects are randomly rematched in each period, while at the other extreme, subjects are matched in pairs and interact for an indefinite duration. In between these extremes, we allow players to dissolve relationships at the end of each period, and we distinguish between whether unilateral or mutual consent is required to terminate a relationship. While cooperation is theoretically sustainable in all four treatments, our results suggest that cooperation is increasing in the difficulty of dissolving relationships; that is, cooperations rates are highest in the IBA treatment, followed by TBA-M, then TBA-U and finally RM. Also, the dynamics of cooperation differ across the four treatments that we consider. Under random matching, cooperation quickly converges to a low level, while in TBA-U cooperation rates actually start at a level comparable to the IBA treatment, but declines across supergames. In the

TBA-M and IBA treatments, we observe stable or increasing cooperation rates over the duration of the experiment. Our findings on low cooperation rates in random matching provides further support to the result that cooperation under anonymous random matching is difficult, except in special circumstances (e.g., compare Duffy and Ochs (2009) with Camera and Casari (2009)). The main lesson from this analysis is that the institution in which subjects are placed plays a crucial role in the viability of cooperation, and that flexibility – especially when only unilateral consent is required – undermines cooperation in a statistically and economically meaningful way compared to a setting where the relationship cannot be broken. However, compared to the other extreme of random matching, the TBA institution leads to a meaningful increase in cooperation.

Our results suggest that subjects grasp part of the equilibrium logic under temporarily binding agreements – that is, they are very likely to dissolve following a defection, and future defections are much more likely. Yet the process is not complete, which makes opportunistic behavior a viable strategy. This, we argue, is what undermines cooperation in the TBA-U treatment. On the other hand, our results also suggest that subjects over-use the flexibility to dissolve. Specifically, we show that subjects in both the TBA-U and TBA-M treatments whose relationship continues following a single defection are significantly more likely to play mutual cooperation than they are if the relationship dissolves. Yet we find that most subjects choose to request a dissolution after experiencing a defection. A corollary to this is that in TBA-M, because the mutual consent requirement makes it more difficult to dissolve relationships, cooperation is higher than in TBA-U. Moreover, as noted, the dynamics of cooperation are more favorable.

Given that flexibility undermines cooperation, an important follow-up question is "how can we alter the institution to increase cooperation?" Reducing flexibility by requiring mutual consent is only a partial solution. We showed that providing a reputation mechanism to the TBA-U environment is much better at promoting cooperation. We find that all of the reputation mechanisms worked well and provided some evidence that an objective, long-lasting reputation mechanism may perform the best. Indeed, cooperation rates increase to a level that is significantly higher than under either TBA treatment and that equals or exceeds the cooperation rates under indefinite repetition. Moreover, we observe stable levels of cooperation over time when a reputation mechanism is available, compared to the declining rates we observed in TBA-U.

Finally, we find that temporarily binding agreements, especially when combined with a reputation mechanism, facilitate sorting between subjects so that those who frequently cooperate (and have high reputations) are more likely to be matched with each other, permitting long-lasting and cooperative relationships. The flip side is that those who frequently defect (and have low reputations) also end up being matched together and are stuck in dysfunctional relationships.

Our results nicely complement the work of Hyndman and Honhon (2015) who also compare indefinitely and temporarily (unilaterally breakable) binding agreements but use a stage game which features strategic complementarities and multiple Pareto rankable equilibria. In contrast, our prisoner's dilemma game creates incentives for the subjects to behave opportunistically. Both projects conclude that indefinitely binding agreements generally lead to better performance than temporarily binding agreements, with interesting nuances. We have shown that the introduction of a reputation mechanism can greatly improve performance when subjects can dissolve relationships. Similarly, Hyndman and Honhon (2015) show that the careful use of the flexibility to dissolve relationships can lead to stable and successful long-lasting relationships. Furthermore, they show that allowing subjects to choose whether to form a temporarily binding or indefinitely binding relationship is advantageous, possibly because it creates common knowledge of intentions. Finally, in both papers, subjects overuse the flexibility to dissolve relationships, which is detrimental.

Acknowledgments

We would like to thank an associate editor and two anonymous referees, as well as Gary Bolton and Matthew Embrey for valuable comments, and Owen Ma and Vineetha Sivadasan for their assistance in running the experiments. We gratefully acknowledge the Jindal School of Management (UT Dallas) for financial support. Kyle Hyndman also acknowledges the Institute for Research in Experimental Economics (IFREE, Small Grants Program) for additional financial support.

Appendix

A. Supplemental Results

A.1. Pairwise Comparisons of Cooperation Across Treatments

Table A.1 Pairwise Comparisons of Cooperation Across Treatments

(a) First Period; First Supergame

	RM	TBA-U	TBA-M	IBA
RM	–	–	–	–
TBA-U	0.186	–	–	–
TBA-M	0.031	0.702	–	–
IBA	0.012	0.968	0.392	–

(b) First Period; All Supergames

	RM	TBA-U	TBA-M	IBA
RM	–	–	–	–
TBA-U	0.189	–	–	–
TBA-M	0.001	0.657	–	–
IBA	0.000	0.019	0.001	–

(c) All Periods; First Supergame

	RM	TBA-U	TBA-M	IBA
RM	–	–	–	–
TBA-U	0.000	–	–	–
TBA-M	0.000	0.369	–	–
IBA	0.000	0.533	0.854	–

(d) All Periods; All Supergames

	RM	TBA-U	TBA-M	IBA
RM	–	–	–	–
TBA-U	0.000	–	–	–
TBA-M	0.000	0.367	–	–
IBA	0.000	0.003	0.146	–

The $p-$values for the pairwise tests are obtained via linear random effects regressions (except for the regression of first period choice in first supergame, in which there is only one observation per subject) of cooperation on indicator variables for each treatment, and with standard errors clustered at the session level.

Table A.2 Comparison of Cooperation Between TBA-U/IBA and Reputation Treatments

		First Period		All Periods	
		First SG	All SGs	First SG	All SGs
TBA-U vs	SLR	0.500	0.063	0.416	0.015
	OSR	0.265	0.010	0.430	0.007
	OLR	0.389	0.000	0.155	0.000
IBA vs	SLR	0.116	0.994	0.096	0.594
	OSR	0.009	0.336	0.081	0.614
	OLR	0.077	0.001	0.007	0.000

The $p-$values for the pairwise tests are obtained via linear random effects regressions (except for the regression of first period choice in first supergame, in which there is only one observation per subject) of cooperation on indicator variables for each treatment, and with standard errors clustered at the session level.

A.2. Dissolution Behavior in the TBA-M Treatment

This table provides more detail than Table 6 for the TBA-M treatment, in particular paying attention to which party requested the dissolution and whether the request was accepted or denied.

A.3. The Likelihood of (C, C) Given the Previous Round's Outcome

This table replicates Table 7(b) but breaks up subjects by above/below average reputation.

Table A.3 The Frequency of Dissolution Requests

	Outcome in Prisoner's Dilemma Phase		
Who Requests Dissolution	(C,C)	(C,D) or (D,C)	(D,D)
Neither Subject	100.00	59.35	28.21
One Subject	0.00		
Cooperator (Accepted)	n/a	4.78	
Cooperator (Denied)	n/a	19.35	
Defector (Accepted)	n/a	1.52	23.12
Defector (Denied)	n/a	6.52	26.94
Both Subjects	0.00	3.70	21.73

Table A.4 The Frequency of Observing (C,C) conditional on past outcomes

(a) Subjects With Above Average Reputation

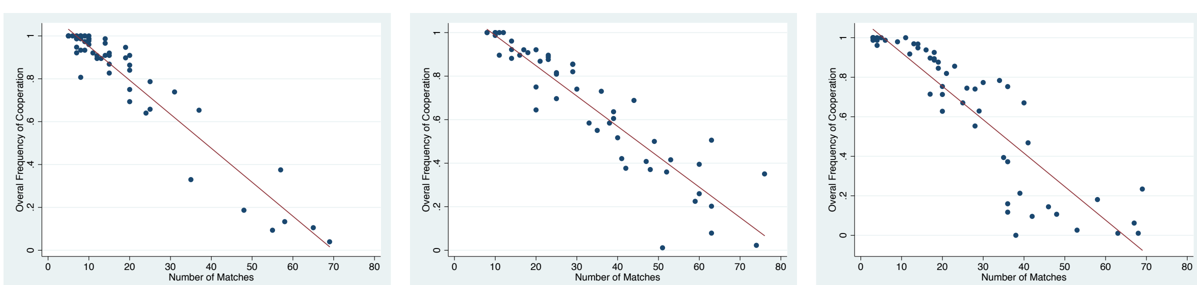
	OSR		OLR		SLR	
Outcome Last Period	Dissolved	Maintained	Dissolved	Maintained	Dissolved	Maintained
(C,C)	37.50%	92.70%	59.09%	98.81%	30.00%	97.76%
(C,D) or (D,C)	27.53%	19.17%	35.44%	15.22%	17.21%	21.55%
(D,D)	26.60%	3.70%	20.97%	0.00%	12.82%	1.89%

(b) Subjects With Below Average Reputation

	OSR		OLR		SLR	
Outcome Last Period	Dissolved	Maintained	Dissolved	Maintained	Dissolved	Maintained
(C,C)	0.00%	63.54%	16.67%	96.06%	0.00%	96.60%
(C,D) or (D,C)	12.24%	17.19%	7.08%	14.36%	6.52%	15.57%
(D,D)	6.73%	1.71%	2.78%	13.33%	2.36%	0.64%

Figure A.1 The Relationship Between Cooperation and Number of Matchings

(a) OLR (b) OSR (c) SLR



A.4. The Relationship Between Cooperation and Number of Matchings

A.5. Instructions for the Temporarily Binding Agreements (Unilateral) Treatment

Thank you for coming today. If you haven't already done so, please power off all mobile devices, tablets, computers, etc and put them in your bag or on the floor at your seat. This is an experiment on the economics of decision-making. Your earnings will depend partly on your decisions and partly on the decisions of others. By following the instructions and making careful decisions you will earn varying amounts of money, which will be paid at the end of the experiment. Details of how you will make decisions and earn money are explained below.

In this experiment, you will participate in a number of decision problems (rounds). In all rounds, you will be matched with another participant in the experiment but you will not know the identity of the other participants with whom you are matched throughout the experiment. In what follows, we will refer to the person with whom you are matched as your *match* and the two of you as a *pairing*. The experiment will last for a minimum of 75 rounds. More precise details will be given below.

Decision Problem

In each round you and your match will simultaneously choose an action A or B. The payoffs for each possible combination of actions is given in the table below:

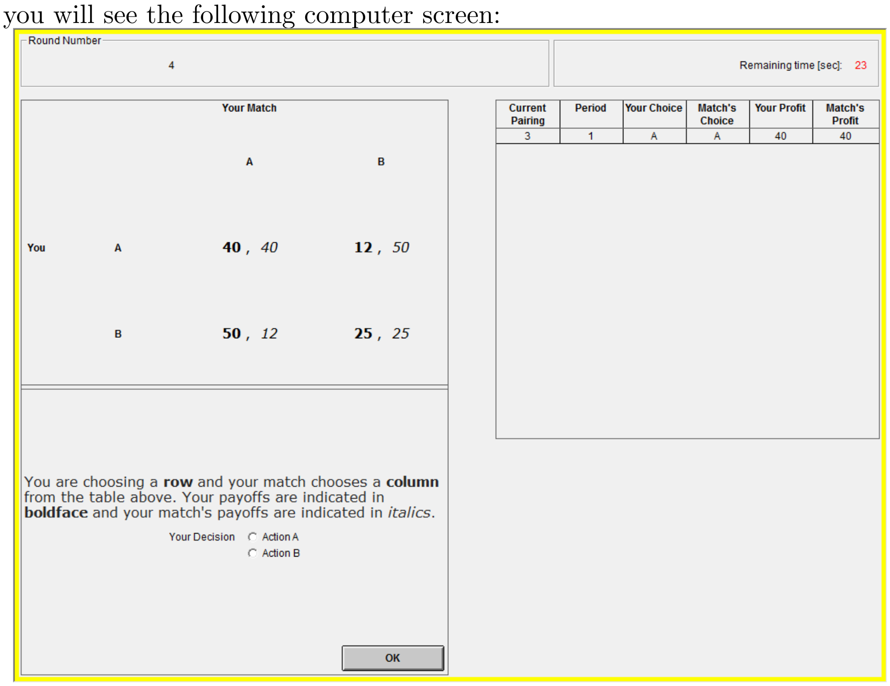
	Match's Choice	
Your Choice	A	B
A	**40**, *40*	**12**, *50*
B	**50**, *12*	**25**, *25*

The first entry in each cell (in bold) represents your payoff, while the second entry represents the payoff of the person you are matched with (in italics). As you can see, this shows the payoff associated with each choice. That is, if:

- You select **A** and your match selects **A**, you each make 40.
- You select **A** and your match selects **B**, you make 12 and your match makes 50.
- You select **B** and your match selects **A**, you make 50 and your match makes 12.
- You select **B** and your match selects **B**, you each make 25.

The Computer Screen

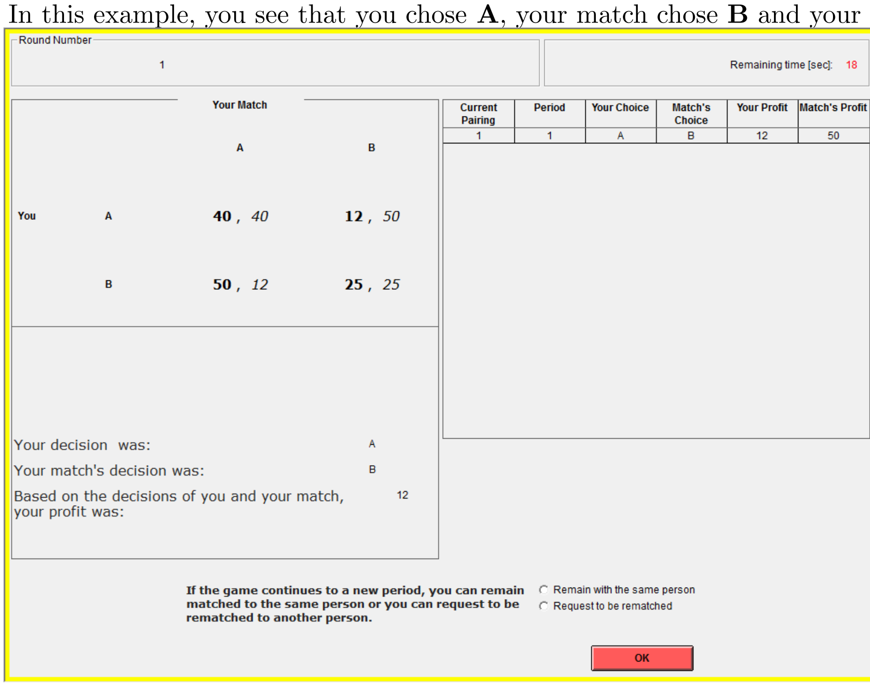
In each round, you will see the following computer screen:



On the top-left side of the screen you will see the same payoff table as depicted above. On the top-right side of the screen you can see your and your match's previous choices as well as your profits for the current pairing. That is, you will not see any information regarding choices made or outcomes from any of your previous pairings. In the example above, you see that in the first period of the third pairing, you chose **A** and your match chose **A**, which gave you a payoff of 40 and your match a payoff of 40. On the bottom-left of the screen is where you will make your decision, i.e., either action A or action B.

Payoffs

Your earnings in each round depend on your choice and on your match's choice. After both you and your match have made your choices, you will see the following screen. You see your choice, your match's choice, and your profit. In this example, you see that you chose **A**, your match chose **B** and your payoff was 12.



Pairings

At the end of every round, all participants have the option to remain matched with the same participant or to request to be rematched. If either you or your match request to be rematched, then your match for the next round will be chosen at random amongst all the participants in the experiment who either requested to be rematched or whose match requested to be rematched in the previous round.

In addition, at the end of every round, there is a 10% chance that the pairing between you and your match will **naturally** break-up. In this case, all participants will be rematched to a randomly chosen participant in the experiment and a new pairing will begin for everybody; that is, it is as if we roll a 10-sided die at the end of each round and the pairing breaks up if we roll a 1, and continues if we roll a 2, 3, . . . , 9 or 10.

Note that, absent a request for rematching, there is **always** a 90% chance that you will remain matched with the same subject in the next round. *It does not matter, for example, whether if its the first, fifth or twelfth round of your pairing; there is always a 90% chance that it will continue for one more round.*

Any time that you have been **rematched** to another participant, you will be explicitly told so; therefore, if no such announcement is made, it means you are still matched with the same participant.

End of the Experiment

The experiment will end when the first pairing **after** 75 rounds have already been played **naturally** breaks up. Note that because of the 10% chance of a natural break-up, in any round after the 75th, you can expect the experiment to continue for approximately 10 more rounds.

At the end of the experiment, we will add all your earnings in order to determine your total points. This total will be converted to a dollar amount according to the rule:

$$\$1 = 175 \text{ points.}$$

This amount will then be added to the $4.00 participation fee to give your final payment. Payments will be made in private, in cash, after the completion of the experiment.

Rules

Please do not talk with anyone during the experiment. We ask everyone to remain silent until the end of the last decision problem.

Your participation in the experiment and any information about your earnings will be kept strictly confidential. Your receipt of payment is the only place on which your name will appear. This information will be kept confidential.

If you have any questions please ask them now. If not, we will proceed to the experiment.

References

Anderson, E., S.D. Jap. 2005. The dark side of close relationships. *MIT Sloan Management Review* **46**(2) 75–82.

Bamford, James, David Ernst, David G. Fubini. 2004. Launching a world-class joint venture. *Harvard Business Review* February.

Bolton, Gary E., Elena Katok, Axel Ockenfels. 2004. How effective are electronic reputation mechanisms? an experimental investigation. *Management Science* **50**(11) 1587–1602.

Bolton, Gary E., Elena Katok, Axel Ockenfels. 2005. Cooperation among strangers with limited information about reputation. *Journal of Public Economics* **89**(5) 1457–1468.

Camera, Gabriele, Marco Casari. 2009. Cooperation among strangers under the shadow of the future. *American Economic Review* **99**(3) 979–1005.

Coopers, Lybrand. 1986. *Collaborative ventures: An emerging phenomenon in information technology*. Coopers and Lybrand, New York.

Copen, Casey E., Kimberly Daniels, Jonathan Vespa, William D. Mosher. 2012. First marriages in the united states: Data from the 2006-2010 national survey of family growth. Tech. Rep. 49, National Center for Health Statistics.

Dal Bó, Pedro. 2005. Cooperation under the shadow of the future: Experimental evidence from infinitely repeated games. *American Economic Review* **95**(5) 1591–1604.

Dal Bó, Pedro, Guillaume R. Fréchette. 2011. The evolution of cooperation in infinitely repeated games: Experimental evidence. *American Economic Review* **101**(1) 411–429.

D'Evelyn, Sean. 2013. Dump, date or marry: Endogenous group formation with varied contract length. Working Paper.

Duffy, John, Jack Ochs. 2009. Cooperative behavior and the frequency of social interaction. *Games and Economic Behavior* **66**(2) 785–812.

Duffy, John, Huan Xie, Yong-Ju Lee. 2013. Social norms, information, and trust among strangers: Theory and evidence. *Economic Theory* **52**(2) 669–708.

Fischbacher, Urs. 2007. z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* **10**(2) 171–178.

Friedman, James W. 1971. A non-cooperative equilibrium for supergames. *Review of Economic Studies* **38**(1) 1–12.

Gaudeul, Alexia, Paolo Crosetto, Gerhard Reiner. 2015. Of the stability of partnerships when individuals have outside options, or why allowing exit is inefficient. Jena Economic Research Papers #2015–001.

Grundberg, Sven, John D. Stoll. 2012. Nokia sales face more hurdles. The Wall Street Journal, `http://www.wsj.com/articles/SB10001424052702304898704577480340752531300`.

Hyndman, Kyle, Dorothée Honhon. 2015. Flexibility in long-term relationships: An experimental study. Working Paper.

Kandori, Michihiro. 1992. Social norms and community enforcement. *Review of Economic Studies* **59**(1) 63–80.

Kovach, Steve. 2014. Microsoft closes its $7.2 billion purchase of nokia. Business Insider, `http://www.businessinsider.com/microsoft-closes-nokia-acquisition-2014-4`.

KPMG. 2009. *Joint Ventures: A tool for growth during an economic downturn*. KPMG International.

Larsen, M. M., T. Pedersen, D. Stepniov. 2010. *LEGO Group: An Outsourcing Journey*. Case 9B10M094, Ivey Management Services.

Liker, Jeffrey K., Thomas Y. Choi. 2004. Building deep supplier relationships. *Harvard Business Review* December, 104–113.

Murnighan, J. Keith, Alvin E. Roth. 1983. Expecting continued play in prisoner's dilemma games: A test of several models. *Journal of Conflict Resolution* **27**(2) 279–300.

Neuville, J. 1997. La stratégie de la confiance: Le partenariat industrial observé depuis le fournisseur. *Sociologie du Travail* **20**(3) 297–319.

Rand, David G., Samuel Arbesman, Nicholas A. Christakis. 2011. Dynamic social networks promote cooperation in expeirments with humans. *Proceedings of the National Academy of Science* **108**(48) 19193–19198.

Roth, Alvin E., J. Keith Murnighan. 1978. Equilibrium behavior and repeated play of the prisoner's dilemma. *Journal of Mathematical Psychology* **17**(2) 189–198.

Schaffer, M.E., S. Stillman. 2010. xtoverid: Stata module to calculate tests of overidentifying restrictions after xtreg, xtivreg, xtivreg2 and xthtaylor. `http://ideas.repec.org/c/boc/bocode/s456779.html`.

Stahl, Dale O. 2013. An experimental test of the efficacy of simple reputation mechanisms to solve social dilemmas. *Journal of Economic Behavior & Organization* **94**(October) 116–124.

Taylor, Paul, Cary Funk, Peyton Craighill. 2006. Are we happy yet? Tech. rep., Pew Research Center.

Uzzi, B. 1996. The sources and consequence of embeddedness for the economic performance of organizations: the network effect. *American Sociological Review* **61**(4) 675–698.

Wilson, Alistair J., Hong Wu. Forthcoming. At-will relationships: How an option to walk away affects cooperation and efficiency. *Games and Economic Behavior* .