2015

Commitment Problems In Conflict Resolution

Erik O. Kimbrough
Simon Fraser University

Jared Rubin
Chapman University, jrubin@chapman.edu

Roman M. Sheremeta
Chapman University

Timothy W. Shields
Chapman University, shields@chapman.edu

Recommended Citation

Kimbrough, E., Rubin, J., Sheremeta, R., and Shields, T. (2015). "Commitment Problems in Conflict Resolution," *Journal of Economic Behavior and Organization*, 112: pp. 33-45. DOI: 10.1016/j.jebo.2015.01.005

Commitment Problems In Conflict Resolution

DOI: 10.1016/j.jebo.2015.01.005

Commitment Problems in Conflict Resolution

Erik O. Kimbrough, Simon Fraser University
Jared Rubin, Chapman University
Roman M. Sheremeta, Case Western Reserve University
Timothy W. Shields, Chapman University

October 30, 14

Abstract

Commitment problems are inherent to non-binding conflict resolution mechanisms, since an unsatisfied party can ignore the resolution and initiate conflict. We provide experimental evidence suggesting that even in the absence of binding contractual agreements individuals often avoid conflict by committing to the outcome of a conflict resolution mechanism. Commitment problems are mitigated to a greater extent for groups that opt-in to the conflict resolution mechanism, but only when opting-in is costly. Although conflict rates are higher when opting-in is costly than when it is free or exogenously imposed, commitment problems are greatly reduced among those groups who choose to opt-in.

JEL Classifications: C72, C91, D72
Keywords: conflict resolution, commitment problem, opting-in, contests, experiments

---

Corresponding author: Roman M. Sheremeta (rshereme@gmail.com)
* We received many helpful comments from the Associate Editor, two anonymous referees, Subhasish Chowdhury, Peter Leeson, Andrew McGee, Chris Muris, Paul Pecorino, Dmitry Ryvkin and seminar participants at Simon Fraser University, University of East Anglia, and Florida State University. For financial support, the authors thank the Economic Science Institute at Chapman University. The usual disclaimers apply.

1. Introduction

Since Schelling (1960) initiated the formal study of conflict, economists and others have developed numerous models highlighting the resources wasted when conflict occurs (Tullock, 1980; Konrad, 2009). Empirical studies testing these models indicate that in practice the costs of conflict are substantially higher than predicted by the theory (Sheremeta, 2013; Dechenaux et al., 2014). For this reason, there has been extensive research on mechanisms of conflict resolution that allow parties to avoid these costs, all of which depend on the ability to make credible commitments and/or binding contracts enforced by a third party (see e.g., Schelling, 1960; Williamson, 1985; Charness et al., 2007; Kimbrough and Sheremeta, 2013; Kimbrough et al., 2014). Yet, in reality, many conflict resolution efforts cannot rely on credible commitments or enforceable contracts.

The problems associated with the inability to commit are well known: in the absence of credible commitment, parties on the less favored side of any proposed resolution face incentives to ignore the resolution. Such commitment problems arise both on a macro level between countries and competing economies, as well as on a micro level between individuals. For example, a country that finds a UN resolution counter to its interests could simply exit the organization and ignore the international community. A similar commitment problem can help to explain the failures of recent WTO negotiations and the Kyoto Protocol as well as the difficulty of enforcing nuclear restrictions in North Korea and Iran.[1] Agreements between individuals often face similar problems, especially when third-party enforcement mechanisms are unavailable. Historically, weak and decentralized political rule made many contractual agreements legally unenforceable (Milgrom et al., 1990; Greif et al., 1994; Greif, 2000), and similar troubles plague

[1] See Powell (2006) for a discussion of commitment problems and armed conflict between nations.

contemporary societies where the rule of law is absent or weak. Even when the rule of law is present many contracts remain unenforceable, and commitment problems arise due to imperfect information and incomplete contracts (Harris and Raviv, 1979; Holmström, 1979; Baker, 1992; Tirole, 1999).

Given the ubiquity of commitment problems in conflict resolution, there has been surprisingly little empirical work exploring the possibility of conflict resolution in the absence of binding commitments.[2] Some conflict resolution mechanisms have been examined in the lab, including side-payment agreements (Charness et al., 2007; Kimbrough and Sheremeta, 2013, 2014), pre-commitment to an allocation proposed by a random device (Kimbrough et al., 2014), communication (Cason et al., 2013), and repeated interactions (Lacomba et al., 2014). Also, some animal behavior studies indicate that many non-human species employ non-binding resource-allocation conventions built around asymmetries between parties that are uncorrelated with their fighting abilities (Petrie et al., 1991; Maher and Lott, 2000).[3] The main difference of our paper is that we examine an environment in which parties may choose to ignore their previously made commitments to a conflict resolution mechanism. This is a crucial difference that allows us to examine commitment problems in conflict resolution.

In this paper, we explore whether (and how) non-binding conflict resolution mechanisms reduce conflict and overcome commitment problems under rather extreme conditions. We conduct an experiment where participants face potential conflict over a valuable, *indivisible*

[2] See, for example, the review of experimental literature on conflicts by Dechenaux et al. (2014).

[3] An important difference between the papers cited in this paragraph, as well as the present paper, and other papers in the literature on commitment in contests is that in these papers parties can make irreversible outlays to increase the likelihood of conflict resolution that do *not* affect the probability of winning the conflict should conflict occur. This is unlike Skaperdas and Syropoulos (2002) and Skaperdas (2006), for instance, where individuals can invest in arming themselves or in litigation, either of which increase their probability of winning should conflict occur but also make conflict less likely to happen in the first place under certain conditions. We refrain from considering such an environment because we are concerned with the possibility of conflict resolution via an independent third party.

resource. Parties may choose to allocate the resource by costly conflict (i.e., a rent-seeking contest), or they can choose to employ a conflict resolution mechanism that allocates the resource at random (i.e., a coin flip), allowing them to avoid the costs of fighting. However, for this mechanism to be effective, both participants must first opt-in to learn the proposed allocation and then mutually commit to the outcome. If either party chooses *not* to commit, the parties fight over the resource. The indivisibility of the resource is crucial here because it creates a particularly "hard case" for conflict resolution efforts: commitment requires one party to accept extreme inequality.

Clearly, standard game theoretic arguments for finitely repeated games predict that even if both parties opt-in to observe the proposed allocation, the losing party will *never* commit to the outcome. However, our treatments are designed to explore the willingness of individuals to commit to just such conflict mitigating agreements. The *Exogenous-Free* treatment, in which there is no opt-in decision and all pairs see the proposed random allocation, allows us to measure the willingness to commit in a strict test of the model. The *Endogenous-Free* and *Endogenous-Costly* treatments, in which parties must first opt-in to learn the proposed allocation and then choose whether to commit, allow us to examine alternative mechanisms for encouraging cooperation. In particular, these treatments are motivated by the idea that endogenizing the decision to employ an institution may increase that institution's effectiveness by encouraging individuals to adhere to it. This effect may operate through a number of channels: (1) endogenizing the decision may increase the "legitimacy" of the proposed allocation, (2) it may increase the expected costs of conflict by encouraging retaliation by those who receive the resource provisionally and then learn that their counterpart has rejected the proposed allocation, and (3) the act of opting-in may reveal information about a player's willingness to adhere to the

institution, particularly when opting-in is costly (i.e., in the *Endogenous-Costly* treatment, opting-in may be seen as "expensive talk").

In fact, we observe many instances in which individuals avoid conflict by committing to the outcome of a coin flip; conflict resolution rates range between 20%-54%, depending on the treatment. Commitment problems are further mitigated when groups endogenously choose to *opt-in* to the conflict resolution mechanism, but only when opting-in is costly. Although conflict rates are higher when opting-in is costly than when it is free or exogenously imposed (in many instances pairs do not opt-in), commitment problems are greatly reduced amongst those groups that opt-in.

Thus, our paper makes two key contributions. First, we provide evidence that conflict resolution efforts can overcome commitment problems, even when resolutions generate extreme payoff inequality. This suggests that the prospects for reducing the costs of conflict are not as bleak as standard theory would suggest (Tullock, 1980; Konrad, 2009). This finding assumes additional importance given the extensive experimental literature demonstrating the very high costs of conflict (Sheremeta, 2013; Dechenaux et al., 2014).[4] Second, we show that imposing explicit costs for using conflict resolution mechanisms can further mitigate commitment problems. A costly mechanism serves two purposes: it encourages reciprocity because parties both know that their counterpart paid a cost to potentially avoid conflict, and it facilitates self-selection, since individuals who are unwilling to commit are less likely to incur the cost in the

[4] Sheremeta (2013) reviews 30 conflict experiments and finds that in 28 of those experiments individuals exert significantly higher conflict expenditures than predicted. In some conflict experiments the expenditures are so high that the majority of individuals earn negative payoffs (Sheremeta, 2010a, 2010b; Price and Sheremeta, 2011, 2014; Deck and Sheremeta, 2012; Chowdhury et al., 2014).

first place.[5] Thus, we argue that designers of conflict resolution mechanisms should consider ways of ensuring that parties are invested in the conflict resolution process, and we hope that theorists can further develop models to explain the source of this effect.

2. Experimental Design and Procedures

2.1. Experimental Design

Our experiment employs a simple conflict resolution game in which participants play a finitely repeated game with a known ending, repeatedly bargaining over an indivisible resource valued at v by both players. The game, shown in Figure 1, consists of three stages. In stage 1, participants choose whether or not to use the conflict resolution device by choosing either 'Flip' or 'Don't Flip'. If at least one participant chooses Don't Flip then neither participant incurs any cost, and both proceed to stage 3. If both participants choose to Flip then they each incur a cost c, and after observing a realization of a fair coin flip, which randomly assigns the resource to one player, they proceed to stage 2. In stage 2, participants decide whether to 'Agree' to the proposed allocation. If both participants choose Agree (i.e., if they choose to *commit*) then the participant who was favored by the coin flip receives v = 100 francs, and the game is over. However, if at least one participant chooses 'Don't Agree', then both participants proceed to stage 3. In stage 3, both participants participate in a lottery contest to allocate the resource. In the lottery contest, both participants make irreversible conflict expenditures e_1, $e_2 \in [0, v]$ to increase their probabilities of receiving the prize. Specifically, for participant 1 (similarly for participant 2), the

[5] There is a large literature in biology suggesting that costly signals are more reliable than cheap signals (Zahavi, 1975, 1993; Grafen, 1990; Petrie et al., 1991). Likewise, there is a significant literature in the economics of religion suggesting that costly or stigmatizing religious rituals dramatically increase cooperation within religious "clubs" (Iannaccone, 1992, 1994; Berman, 2000; Abramitzky, 2008; Aimone et al., 2013; Carvalho, 2013).

probability of receiving is $p_1(e_1,e_2) = e_1/(e_1 + e_2)$; if both players bid 0, the allocation is determined by flipping a fair coin.

Using backward induction and assuming risk neutrality, we examine the last occurrence of the finitely repeated game. The Nash equilibrium in stage 3 is for both participants to choose expenditures of $e_1^* = e_2^* = 25$ (Tullock, 1980), receiving the expected payoff of 25 each ($v/2 - e^* = 50 - 25$). Because each participant has a positive expected value from conflict in stage 3, a participant interested in maximizing her wealth should *never* commit to the outcome after losing the coin flip. Therefore, in any subgame perfect Nash equilibrium, money-maximizing participants end up in a conflict stage 3 regardless of the cost of the coin flip c and exert wasteful expenditures competing for the resource.[6] Since conflict is certain along the equilibrium path, it follows that participants are indifferent between opting-in and not opting-in if and only if the cost of opting-in is zero (i.e., $c = 0$), and they should never opt-in when doing so is costly (i.e., $c > 0$).

To examine whether a non-binding coin flip can reduce the cost of conflict, we employ a partial two-by-two experimental design. On one dimension, we vary whether the decision to opt-in is *endogenous* or *exogenous*, and on the other dimension, we vary whether opting-in carries a *cost*. For obvious reasons, we impose the cost only when opting-in is endogenous, hence our partial design. In the *Exogenous-Free* baseline treatment, participants first observe the outcome of a coin flip. That is, there is no stage 1 decision to 'Flip' or 'Don't Flip'; instead, the 'Flip' decision is exogenously imposed. In stage 2, participants decide whether to commit to the outcome. If either participant disagrees then the game proceeds to stage 3. In stage 3, participants

[6] While we focus on the case of risk-neutrality, we can extend this argument to cases in which players have symmetric levels of risk-aversion. In such cases, there exists an equilibrium of the Tullock contest with positive expenditures and positive expected payoff, so that the subgame perfect Nash equilibrium described remains qualitatively unchanged, with players always ending up in conflict.

participate in a simple two-person contest for the resource valued at 100. In the *Endogenous-Free* treatment, participants first endogenously decide whether to observe the outcome of a coin flip for free in stage 1. Then in stage 2 they decide whether to commit to the outcome. If either participant disagrees in either stage 1 or stage 2 then the game proceeds to stage 3. Finally, the *Endogenous-Costly* treatment is similar to the *Endogenous-Free* treatment, but opting-in in stage 1 is costly, i.e., $c = 5$ francs.

2.2. Experimental Procedures

To study behavior in the proposed treatments, a total of 198 participants were recruited at random from the subject-pool consisting of graduate and undergraduate students at a private university in the United States. Participants received a participation fee for arriving to the experiment on time and received their earnings in cash privately at the end of each session. Participants were randomly assigned into treatments (44 in *Exogenous-Free*, 62 in *Endogenous-Free* and 92 in *Endogenous-Costly*) and nobody participated more than once.[7] Participants sat at, and interacted via, visually isolated computer terminals, and instructions were read aloud by the experimenter as participants followed along on paper (see the online Appendix A for instructions). The experiments were programmed using z-Tree (Fischbacher, 2007). Before the experiment began participants took a quiz (non-incentivized) to confirm their understanding of the experimental procedures. An experimenter reviewed the quiz answers and privately answered questions.

Each experimental session consisted of 30 periods of a single treatment. Participants were randomly assigned to a group of two and remained within this group for all 30 periods of the

[7] Our design was unbalanced to generate comparable sample sizes for the commitment decision across treatments.

experiment (fixed matching). We chose a fixed matching protocol because we believe that it is more realistic, as most interactions in the real world are repeated (especially macro level conflicts between countries and competing economies). In such a setting, participants seeking to avoid costly conflict might be more motivated than in one-shot games to signal their desire to other group members. Moreover, the fact that we chose the fixed matching protocol does not affect the comparison across treatments (as the same protocol was used in all three treatments).[8]

In every session, each period proceeded according to one of the treatments. In stage 1 participants chose between `Flip' and `Don't Flip' (except in the *Exogenous-Free* treatment). In stage 2, those who opted to see the coin flip chose between `Agree' and `Don't Agree' after seeing the result of the flip. In stage 3, participants chose to expend any number between 0 and 100 francs, in order to increase the probability of winning the prize of 100 francs. At the end of each period, the computer displayed individual decisions, as well as corresponding payoffs, to each participant.

At the end of each experimental session, participants completed a brief demographic survey and their total earnings from all 30 periods of the experiment were added to or subtracted from an initial endowment of 2,000 francs (which we described as a participation fee). We converted francs to USD at a rate of 100 francs = $1, and participants were paid privately in cash and dismissed from the experiment. The average experimental earnings, including the $20 participation fee, were $25.24, ranging from a low of $16.00 to a high of $46.70. Sessions lasted approximately one hour each.

[8] Baik et al. (2013) examined behavior of subjects in contests using fixed-matching and random-matching protocols and found no difference in expenditure levels. Similarly, in a meta-analysis of experimental contest studies, Sheremeta (2013) found that the matching protocol has no significant impact on contest expenditures.

3. Results

We observed early period variation that settled down after a few periods and endgame effects, both of which introduced noise into our data. Thus, we restrict our analysis to periods 5-29.[9] Table 1 reports average frequency of conflict resolution, opting-in and commitment, as well as conflict expenditures and payoffs. We also report the statistical significance of estimated treatment effects for each of the outcome variables of interest. To control for repeated measures, our ordinary least squares (OLS) regressions use each pair averaged over all periods as a single observation.[10] Rows 1, 2 and 6 of Table 1 examine averages based upon all 25 periods, and rows 3-5 examine averages conditional upon entering stage 2 or stage 3. In rows 3-5, we use probability weighting to control for the fact that each pair was observed making those decisions a different number of times (and thus has a different variance). We report two-sided *p*-values.

3.1. Conflict Resolution

In the *Exogenous-Free* treatment, pairs avoided conflict 54% of the time (i.e., they agree to the result of the Flip), while in the *Endogenous-Free* and *Endogenous-Costly* treatments, pairs avoided conflict 41% and 20% of the time, respectively. OLS regression analysis indicates that the rate of conflict resolution is significantly lower in the *Endogenous-Costly* treatment than in the *Exogenous-Free* and *Endogenous-Free* treatments (*p*-values < 0.01), while there is no significant difference between the *Exogenous-Free* and *Endogenous-Free* treatments (*p*-value = 0.22).

[9] None of our results are qualitatively altered when we vary these cutoffs, and additional regression output is available from the authors upon request. We examine end-game effects in Section 3.6.1.

[10] The results are comparable if using non-parametric analysis (i.e., Wilcoxon rank-sum test) when each pair has the same number of observations (i.e., rows 1, 2 and 6). Heterogeneity in the rates at which pairs reach stage 2 and 3 render non-parametric analysis inappropriate for rows 3-5.

Result 1: There is substantial conflict resolution in all three treatments, although pairs avoid conflict more frequently in the *Exogenous-Free* and *Endogenous-Free* treatments than in the *Endogenous-Costly* treatment.

This finding provides support for the idea that non-binding conflict resolution mechanisms can reduce the frequency of conflict, despite incentives to renege. When using the mechanism is a default, some groups are able to coordinate on a reciprocal strategy in which the losing party is willing to accept an unfavorable resolution. However, the mechanism is far from perfect. When the decision to use the mechanism is endogenous and costly, conflict frequency increased as some groups opted not to see the coin flip. Similarly, there is extensive heterogeneity in the willingness to commit to the outcome, which we discuss further in Section 3.5.

3.2. Opting-In

A major source of increased conflict in the *Endogenous-Free* and *Endogenous-Costly* treatments is failure to opt-in to the coin flip mechanism. While pairs were required to see the flip in the *Exogenous-Free* treatment, they opted to see the coin flip 71% of the time in the *Endogenous-Free* treatment and only 26% of the time in the *Endogenous-Costly* treatment. OLS regression analysis indicates that the rate of opting-in is significantly lower in the *Endogenous-Costly* treatment than in the *Endogenous-Free* treatment (p-value < 0.01). Here we exclude the *Exogenous-Free* treatment from the regression since all pairs were forced to see the coin flip.

Result 2: Subjects opted-in to the coin flip mechanism less frequently in the *Endogenous-Costly* treatment than in the *Endogenous-Free* treatment, although the likelihood was far greater than 0 (and less than 1) in both treatments.

While it is clear that the cost of observing the coin flip played an important role in reducing the frequency with which participants opted-in in the *Endogenous-Costly* treatment, the reasons are less clear in the *Endogenous-Free* treatment. In theory, participants should be indifferent between observing the coin flip and not observing the flip. However, if there was any probability that a player who lost the flip would tremble and mistakenly accept the outcome, players would strictly prefer to flip the coin. Unfortunately, our design does not allow us to test this hypothesis directly.

3.3. Commitment

The money-maximizing subgame perfect Nash equilibrium predicts that participants losing the coin flip will *never* commit to the outcome in a finitely repeated game. Comparing averages across treatments in Table 1, we found that the probability that both parties committed to the outcome (conditional on seeing the flip) were 54% in the *Exogenous-Free*, 58% in the *Endogenous-Free*, and 76% in the *Endogenous-Costly* treatment. Probability weighted OLS regression analysis indicates that the conditional probability that both parties commit to the coin flip resolution is higher in the *Endogenous-Costly* treatment than in the *Endogenous-Free* (*p*-value = 0.08) and *Exogenous-Free* (*p*-value = 0.03) treatments.

The treatment differences were almost solely driven by an increased willingness of participants who lose the coin flip to commit to its outcome. The conditional probability of

committing to the flip among those receiving unfavorable outcomes was 58% in the *Exogenous-Free*, 61% in the *Endogenous-Free*, and 78% in the *Endogenous-Costly* treatment.[11] To provide statistical support for treatment differences in the *loser*'s commitment decisions, we estimate an additional regression. The dependent variable is the probability that the loser of the coin flip in each pair committed to the coin flip, conditional on seeing it, and the independent variables are treatment dummies. The coefficient on the *Endogenous-Costly* treatment is positive and significant (p-value = 0.05), and the difference between the *Endogenous-Free* and *Endogenous-Costly* coefficients is marginally significant (p-value = 0.10). These findings are summarized in Result 3.

Result 3: The conditional probability of commitment is higher in the *Endogenous-Costly* treatment than either in the *Endogenous-Free* or *Exogenous-Free* treatments. There is no significant difference between the *Endogenous-Free* and *Exogenous-Free* treatments.

These results suggest that merely endogenizing the decision to seek third-party conflict resolution does not have an impact on the probability that parties commit to the proposed resolution – we find no difference in commitment rates in the *Endogenous-Free* and *Exogenous-Free* treatments. However, the commitment problem is significantly mitigated when both parties must incur a *cost* to learn the proposed allocation.

The relationship between opting-in and the commitment problem in the *Endogenous-Free* and *Endogenous-Costly* treatments is demonstrated in Figure 2, which plots the relative frequency of opting-in against the conditional probability of commitment for each pair with

[11] In all treatments the conditional probability of commitment for flip winners was at least 95%. Rejections are concentrated in a few individuals (44 of the 76 rejections come from 9 of 198 participants, three per treatment).

smoothing splines fit to the data by treatment.[12] Only the *Endogenous-Free* and *Endogenous-Costly* treatments are plotted, since participants in the *Exogenous-Free* treatment always saw the coin flip. Figure 2 indicates that there is a linear relationship between opting-in and committing to the outcome of the flip in the *Endogenous-Costly* treatment, but there is a convex relationship between the two in the *Endogenous-Free* treatment; indeed, there is a pair that sees the flip in every period, but *never* commits to the outcome. And, while the *Endogenous-Costly* treatment data is roughly symmetrically distributed around the 45-degree line, every *Endogenous-Free* treatment observation lies on or below the line. In the *Endogenous-Free* treatment, commitment is regularly sustained only at high rates of opting-in. This provides further evidence that the opportunity cost of opting-in in the *Endogenous-Costly* treatment facilitates the solution of the commitment problem, while the absence of such a cost in the *Endogenous-Free* treatment renders the cooperative signal less informative and is thus less effective at encouraging commitment.

3.4. Conflict Expenditures

If groups did not opt-in or failed to solve the commitment problem, the group entered the conflict stage in which they exerted expenditures to win the prize.[13] Returning to Table 1, note that the average expenditure, conditional on reaching stage 3, was greatest in the *Exogenous-Free* treatment (28.6), followed by the *Endogenous-Free* treatment (24.6), and the least in the *Endogenous-Costly* treatment (18.6).[14] OLS regression analysis indicates that conflict

[12] An alternative figure containing data from all 30 periods is in Appendix B.

[13] Figures B1-B3 in Appendix B report time series of conflict expenditures for each pair by treatment.

[14] Although we did not run a control treatment without the coin flip, our results in the *Exogenous-Free* treatment are very similar to those that have been observed in the existing experimental literature. In a meta-analysis of 30 experiments on contests, Sheremeta (2013) provides estimation results of a regression that can be used by

expenditures are significantly lower in the *Endogenous-Costly* treatment than in the *Exogenous-Free* treatment (p-value = 0.03). Expenditures are marginally lower in the *Endogenous-Costly* treatment than in the *Endogenous-Free* treatment (p-value = 0.08). There is no significant difference between the *Endogenous-Free* and *Exogenous-Free* treatments (p-value = 0.41).

Result 4: Conflict expenditures are lower in the *Endogenous-Costly* treatment than in either the *Exogenous-Free* or the *Endogenous-Free* treatment. There is no statistically significant difference between the *Endogenous-Free* and *Exogenous-Free* treatments.

3.5. Between-Pair Heterogeneity

Figure 3 displays, for each pair in each treatment, the relative frequency of opting-in, conditional frequency of commitment, and average payoff, where the pairs are sorted from left to right in each panel by opting-in rate with ties broken by commitment rate and then average payoff.

In the *Endogenous-Free* treatment, there is a significant positive correlation between a pair's observed probability of opting-in and average payoff (ρ = 0.62, p-value < 0.01) but in the *Endogenous-Costly* treatment, this relationship is not statistically significant (ρ = 0.23, p-value = 0.12). This is driven by a number of pairs in the *Endogenous-Costly* treatment that never opted-in to see the coin flip but nevertheless received very high payoffs.

It is possible to visually identify three types of pairs in the *Endogenous-Costly* treatment: (1) pairs with a high rate of both opting-in and commitment, (2) pairs with low (but positive)

researchers to compare their results to other studies. We have also used this regression to compare our results. When using the endowment equal to the prize value (endowment = 1) in a two-player contest (n = 2) with fixed matching (fixed = 1), we should expect 14% of over-expenditures in a control treatment without a coin flip. The actual over-expenditure in the *Exogenous-Free* treatment is 14.4% (28.6/25 = 1.144), which is remarkably close to the expected over-expenditure of 14%.

rates of both opting-in and commitment, and (3) pairs with zero rates of opting-in. Heuristically, we define *high flip* pairs as those with opting-in rates in the interval (0.5, 1], *low flip* pairs as those with opting-in rates in (0, 0.5], and *no flip* pairs as those with opting-in rates of exactly 0. *High flip* pairs earn on average 84 francs/period (SD = 6.77), *low flip* pairs earn 49 francs/period (SD = 30.13), and *no flip* pairs earn 73 francs/period (SD = 20.01). Interestingly, 5 of the *no flip* pairs in the *Endogenous-Costly* treatment (i.e., pairs that engaged in conflict in each period 5-29) generated average earnings greater than the maximum that could be earned per period by opting-in *and* committing in every period (i.e., greater than 90 francs). None of the *no flip* pairs in the *Endogenous-Free* treatment were able to generate average earnings greater than 90 francs.

While our data do not allow us to directly identify the source of this behavior, we conjecture the positive commitment cost generated an endogenous focal point for conflict expenditures that allowed some pairs to coordinate on expenditures far below the Nash equilibrium of 25. Specifically, we observe a number of pairs in which individuals consistently submit bids less than or equal to 5 (the cost of opting-in). This suggests that they attempt to cooperate and tacitly agree to only engage in conflict that leaves them better off in expectation than conflict resolution by the random device we have offered them. Such strategies are not always effective, but some pairs manage to sustain them, and attempts by other pairs to implement them tend to drive conflict expenditures down. Thus, Wald tests on the regression reported in row 5 of Table 1 cannot reject the null hypotheses that the average conflict expenditure for the *Exogenous-Free* or *Endogenous-Free* treatments is equal to the Nash Equilibrium value of 25 (p-value = 0.92 and p-value = 0.35, respectively). However, in the *Exogenous-Costly* treatment the average expenditure is significantly less than 25 (p-value < 0.01). When we exclude the *no flip* pairs and run a new regression that is otherwise identical, the

average expenditure in the *Endogenous-Costly* treatment rises and is not significantly different from 25 (p-value = 0.97).

3.6 Dynamics of Behavior

In this section, we highlight a number of interesting features of our data that we found after testing our main hypotheses. Specifically, we focus on behavioral dynamics, emphasizing patterns in the data that are inconsistent with the equilibrium of the finitely repeated game but nevertheless appear important for understanding behavior in our setting. Given the endogeneity issues with this analysis, we report summary statistics only, avoiding statistical tests.

3.6.1 Endgame Effects

As noted above, our main analysis excludes the final period because of the presence of endgame effects. Table 2 displays comparisons of opt-in rates, commitment rates, and conflict expenditures over periods 20-29 and period 30. Note that pairs in all treatments opt-in at roughly equal rates in periods 20-29 and period 30, but in the final period, the data reveal a sharp decrease in commitment rates. This is particularly striking and puzzling in the *Endogenous-Costly* treatment – if a player plans to opt-out, then there is no reason to incur the cost of seeing the flip in the first place. Nevertheless, the presence of an endgame effect is consistent with models in which some players are cooperative types, and others, knowing this, cooperate strategically until the penultimate period (e.g., Kreps et al., 1982). Not surprisingly, in the *Endogenous-Free* and *Endogenous-Costly* treatments, this endgame defection leads to some increase in conflict expenditures, perhaps either as retaliation or anticipation thereof.

Another interesting observation is that there are treatment differences in behavior in the conflict stage of those who were either the "victims" or the "instigators" of endgame defection. Consider only pairs that opted-in and agreed in period 29 and then also opted-in in period 30. We define victims as players who chose agree in period 30 when their partner, the instigator, chose don't agree. In both *Exogenous-Free* and *Endogenous-Free* treatments, victims respond to defection by submitting very high conflict expenditures (31 and 46, respectively), while instigators submit relatively low bids (17 and 24, respectively). In the *Endogenous-Costly* treatment, the pattern is reversed. Victims submit very small bids averaging 13 while instigators submit large bids averaging 34. Observed behavior in the *Exogenous-Free* and *Endogenous-Free* treatments can be rationalized as retaliation, and this is consistent with evidence in Kimbrough and Sheremeta (2013, 2014) in a related lottery contest environment that "reneging" leads to retaliation. The pattern in the *Endogenous-Costly* treatment is more puzzling. One possibility is that defection *after* paying to opt-in reveals information about the defector's type that deters expenditures (i.e. "this guy must be nuts").

3.6.2 Path Dependence

While path dependence does not appear to move behavior within a pair in one direction or the other *on average*, it is an important driving force. To underscore the in-group dynamics over time, Table 3 presents a transition matrix of actions. The top panel reports how a group's choice of whether to see the flip in period t affects their decision to see the flip in period $t + 1$ (note that subjects automatically see the flip in the *Exogenous-Free* treatment). There is no discernable trend in the *Endogenous-Free* treatment; subjects choose to flip at relatively consistent and high rates (0.82-0.92) regardless of their own and their partner's behavior in the previous period. In

the *Endogenous-Costly* treatment, however, it appears that there are two "types": those that frequently see the flip and those that do not, regardless of their partner's actions. Those that are willing to pay to see the flip in period t (note that they only actually pay if their partner chooses to see the flip too) choose to see the flip in period $t + 1$ with probability 0.71 if their partner chose to see the flip in period t and with probability 0.67 if their partner did not. On the other hand, those that do not choose to see the flip in period t also do not choose to see the flip in period $t + 1$ with probability 0.75 (since they do not see the flip, they do not know what their partner chose).

The second panel analyzes the effect of subjects' decisions whether to agree to the outcome of the flip (conditional on seeing the flip) on their decision to see the flip in the next period. Again, there is no discernable trend in the *Endogenous-Free* treatment. Subjects choose to see the flip at relatively consistent and high rates (0.80-0.88) regardless of their own and their partner's behavior in the previous period. In the *Endogenous-Costly* treatment, however, the decision to accept the outcome or not in period t clearly affects the (costly) choice to see the flip in the following period. When a subject and her partner agree to the outcome of the flip, she chooses to see the flip in the following period with probability 0.78. In other words, when both players cooperate, they use the mechanism that facilitated their cooperation in the following period over three-quarters of the time. However, when one subject does not agree to the outcome in period t, subjects choose to see the flip only around half of the time (0.48-0.51). This suggests that, amongst those pairs who choose to see the flip, uncooperative behavior makes cooperation much harder to sustain in the future.

Finally, the third panel reveals how a subject's willingness to agree to the outcome of the flip in period t (conditional on seeing the flip) affects their willingness to agree to the outcome of

the flip in period $t + 1$ (if there is a flip). These are most important results presented in this table, since they shed light on how cooperation is sustained over time. In the *Exogenous-Free* treatment, there is no discernable inter-period effect: both players tend to agree to the flip a little more than half of the time (0.52-0.56) regardless of actions in the previous period. In the *Endogenous-Free* treatment, an unexpected result arises: cooperation rates are higher in period $t + 1$ if both players did not agree in the previous period (0.50) than if they did (0.27). There is no obvious interpretation of this result, but it does at the very least suggest that subjects are not coordinating around the random device when viewing the outcome is free. We find substantial path dependence in the *Endogenous-Costly* treatment. Groups in which both subjects agree in period t (conditional on seeing the flip) see the flip and agree to its outcome in period $t + 1$ with probability 0.69. On the other hand, when at least one subject does not agree to the outcome of the flip in period t, groups only agree to the outcome of the flip in period $t + 1$ with probability 0.08, and they do not see the flip, which is costly to view, with probability 0.77. This suggests that the cost of seeing the flip encourages path dependence: those that agree to the outcome of the flip are more likely to do so in the future (despite the fact that seeing the flip is costly), while those that do not agree rarely pay money to flip the coin again.

4. Discussion

In all three treatments, the prediction from standard game theory is that the loser of the coin flip will never choose to commit, as the expected value from conflict is strictly positive and the game is finitely repeated. Nevertheless, our experiment shows that individuals often avoid conflict in the finitely repeated game by committing to the outcome of a randomizing conflict resolution mechanism, even in the absence of binding contractual agreements. What can explain these non-

zero levels of commitment? There are several reasons why commitment problems can be solved in the absence of external enforcement. First, if the parties expect to engage in repeated interaction, they may be able to overcome the commitment problem due to reputational concerns (Rabin, 1991; Andreoni and Miller, 1993; Dellarocas, 2006) or fear of incurring retaliation and spite (Abbink et al., 2010; Schniter et al., 2013; Lacomba et al., 2014; Mago et al., 2014). Similarly, some individuals may simply prefer to avoid conflict or be cooperative (Dorris, 1972; Molander, 1985), and given repeated interaction, they establish a reputation for cooperation that allows both parties to gain from exchange. As long as one member of the group is a cooperative type, the other member may be willing to forgo conflict to reap the gains from exchange (e.g., Kreps et al. 1982). For these reasons, it may not be that surprising that in all three treatments we observe substantial conflict resolution.

One of the main research questions was whether commitment problems could be mitigated when participants are given the option of opting-in to the conflict resolution mechanism (with and without cost). Here, the failure of participants to opt-in reduced the frequency of conflict resolution relative to the exogenous case, and more so when seeing the flip was costly. However, for participants that *did* opt-in, we observed significantly higher commitment, but *only* when opting-in was costly.

Standard models do not predict these treatment differences. What can explain these differences in commitment rates? There are several non-mutually exclusive (and non-exhaustive) possibilities. First, when individuals act collectively to propose a solution to a social dilemma, the result may be seen as more legitimate than when the proposed solution comes from “outside”, thus increasing cooperation levels. Evidence from both field studies and laboratory experiments suggests that endogenous participation in democratic institutions may encourage

both cooperation and compliance (Dal Bo et al., 2010). Although we believe that legitimacy may play a role in helping groups to resolve commitment problems, our data do not provide support for this hypothesis. Specifically, the legitimacy hypothesis would suggest that since participants choose to opt-in to the conflict resolution mechanism, the coin flip outcome will acquire additional legitimacy in both participants' eyes and will thus facilitate solving the commitment problem. Hence, we would expect commitment rates to be greater in both the *Endogenous-Free* and *Endogenous-Costly* treatments than in the *Exogenous-Free* treatment. However, we find no significant difference in commitment rates between the *Endogenous-Free* and *Exogenous-Free* treatments.

A second possible explanation of our observations is that fears of retaliation and spite may discourage participants from reneging; i.e., opting-in and then not abiding by the outcome may encourage high bids by angered flip winners (Andreoni and Miller, 1993; Dellarocas, 2006; Schniter et al., 2013; Lacomba et al., 2014). While we cannot rule out this explanation, our data do not reveal notable differences in the conflict expenditures of flip winners and losers. Additional unreported regressions (available upon request) indicate that the only significant difference is that flip losers bid more than winners in the *Endogenous-Costly* treatment, which is opposite of the direction of the retaliation hypothesis.

A third explanation is that individuals who pay to see the coin flip may be subject to a sunk cost fallacy; they might justify committing to the coin flip in order to avoid the regret of wasting money on unused advice (Arkes and Blumer, 1985; Gino, 2008). We cannot rule out this explanation – though it seems implausible since the expected value of conflict is positive and greater than the sunk cost of seeing the flip.

A final explanation receives the most substantial support from our data: the act of opting-in may serve as a signal of participant's willingness to cooperate. That is, commitment is encouraged when both parties know that the other has deliberately sent a signal of willingness to cooperate, but this signal is more than cheap talk only when it involves incurring an opportunity cost.[15] Under this interpretation, the willingness to incur a cost of opting-in indicates a willingness to accept the outcome of the coin flip, and agents are able to infer cooperative intentions (to commit) more readily from costly opting-in than from free opting-in. In this sense, the mechanism encourages self-selection among cooperative types. In the repeated game setting, non-cooperative types are going to be found out early on unless they feign being a cooperative type. It may therefore be worthwhile to opt-in to see whether one is paired with a cooperative type, since they can always avoid the costly conflict resolution mechanism in the future after they have been cheated once. These considerations may explain why commitment rates are greater in treatments where opting-in is costly than when it is exogenous or costless. They may also explain why we find significant path dependence in the treatment where opting-in is costly but not in the other treatments: once the players have sent each other costly signals in the past, coordinating around the cooperative outcome is easier in the future.

5. Conclusion

This paper explores the effectiveness of non-binding institutions for conflict resolution and the effect of endogenous institutional choice on their success. Empirically, the main problem with such mechanisms is: how do groups commit to abiding by the outcome of a conflict resolution mechanism? This paper suggests that one way of mitigating this problem is to make the choice of

[15] Indeed, there is evidence for this in other contexts where incurred opportunity costs facilitate reciprocity (McCabe et al., 2000, 2003; Schniter et al., 2013).

opting-in to the mechanism costly. Although the mere act of opting-in has no significant impact on commitment rates, costly opting-in substantially increases commitment for groups who pay the price.

Our results have important implications for the sustainability of cooperation and the mitigation of conflict in environments with repeated interaction and weak contract enforcement. First, using a coin flip as a conflict resolution mechanism has numerous benefits for studying the commitment problems inherent in conflict resolution: it is *transparent* and easy to understand, it provides an *unbiased* means of allocating indivisible resources (consistent with egalitarian norms), and its "all-or-nothing" nature exacerbates the commitment problem, since losing participants must commit to receiving nothing.[16] Second, endogenizing and imposing a cost on the decision to opt-in can encourage reciprocity – by requiring individuals to undertake a cost to seek conflict resolution, the decision substantially increases the rate at which parties commit to conflict resolution. These and similar behavioral mechanisms are clearly important drivers of economic behavior and thus have an important role in any institutional design framework. When developing institutions to reduce conflict and facilitate mutually beneficial exchange, designers should consider not only standard incentive-compatibility issues but also the interaction between institutions and norms of cooperation and reciprocity. In particular, when the absence of exogenous enforcement mechanisms renders incentive-compatible arrangements infeasible, one way to reduce the costs of conflict may involve institutions designed to encourage reciprocity-driven pro-sociality.

[16] Using a random device as a conflict resolution mechanism has a long history. Iannaccone et al. (2011) argue that the Oracle at Delphi promoted cooperation amongst the Greek city-states despite providing random outcomes. Likewise, Leeson (2014) argues that the historical purpose of oracles was to resolve conflicts by correlating conflicting parties' strategy randomization. Another well-known historical example of commitment to a random device is the battlefield agreement to settle a conflict by single combat between two renowned warriors. More familiar examples of conflict resolution via random device are abundant. Rock-paper-scissors, drawing straws, and throwing dice settle many friendly disputes.

References

Abbink, K., Brandts, J., Herrmann, B. & Orzen, H. (2010). Inter-group conflict and intra-group punishment in an experimental contest game. American Economic Review, 100, 420-447.

Abramitzky, R. (2008). The limits of equality: Insights from the Israeli Kibbutz. Quarterly Journal of Economics, 123, 1111-1159.

Aimone, J., L.R. Iannaccone, M. Makowsky, & J. Rubin. (2013). Endogenous group formation via unproductive costs. Review of Economic Studies, 80(4), 1215-1236.

Andreoni, J.A., & Miller, J.H. (1993). Rational cooperation in the finitely repeated Prisoner's Dilemma: Experimental evidence. Economic Journal, 103, 570-585.

Arkes, H.R., & Blumer, C. (1985). The psychology of sunk cost. Organizational Behavior and Human Decision Processes, 35, 124–140.

Baik, K.H., Chowdhury, S.M., & Ramalingam, A. (2013). Matching Protocol in Contest Experiments. Working Paper.

Baker, G.P. (1992). Incentive contracts and performance measurement. Journal of Political Economy, 100, 598-614.

Bardhan, P. (2000). Irrigation and cooperation: An empirical analysis of 48 irrigation communities in south India. Economic Development and Cultural Change, 48, 847-865.

Berman, E. (2000). Sect, subsidy, and sacrifice: An economist's view of ultra-orthodox Jews. Quarterly Journal of Economics, 65, 905-953.

Bernstein, P. (1996). Against the Gods: The remarkable story of risk. New York: John Wiley & Sons.

Carvalho, J-P. (2013). Veiling. Quarterly Journal of Economics, 128, 337-370.

Cason, T.N., Sheremeta, R.M. & Zhang, J. (2012). Communication and efficiency in competitive coordination games. Games and Economic Behavior, 76, 26-43.

Charness, G., Frechette, G.R., & Qin, C.Z. (2007). Endogenous transfers in the Prisoner's Dilemma game: An experimental test of cooperation and coordination. Games and Economic Behavior, 60, 287-306.

Chowdhury, S.M., Sheremeta, R.M., & Turocy, T. (2014). Overbidding and overspreading in rent-seeking experiments: Cost structure and prize allocation rules. Games and Economic Behavior, 87, 224-238.

Dal Bó, P., Foster, A., & Putterman, L. (2010). Institutions and behavior: Experimental evidence on the effects of democracy. American Economic Review, 100, 2205-2229.

Dechenaux, E., Kovenock, D., & Sheremeta, R.M. (2014). A survey of experimental research on contests, all-pay auctions and tournaments. Experimental Economics, forthcoming.

Deck, C. & Sheremeta, R.M. (2012). Fight or flight? Defending against sequential attacks in the Game of Siege. Journal of Conflict Resolution, 56, 1069-1088.

Dellarocas, C. (2006). Reputation mechanisms. Handbooks in Information Systems, Vol. I, Handbook on Economics and Information Systems, edited by Terrence J. Hendershott, Elsevier, 629-660.

Dorris, J. W. (1972). Reactions to unconditional cooperation: A field study emphasizing variables neglected in laboratory research. Journal of Personality and Social Psychology, 22, 387-397.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. Experimental Economics, 10, 171-178.

Gino, F. (2008). Do we listen to advice just because we paid for it? The impact of advice cost on its use. Organizational Behavior and Human Decision Processes, 107, 234-245.
Grafen, A. (1990). Biological signals as handicaps. Journal of Theoretical Biology, 144, 517-546.
Greif, A. (2000). The fundamental problem of exchange. Review of European Economic History, 4, 251-284.
Greif, A., P. Milgrom, & B. Weingast (1994). Coordination, commitment, and enforcement: The case of the merchant guild. Journal of Political Economy, 102, 745-776.
Harris, M., & Raviv, A. (1979). Optimal incentive contracts with imperfect information. Journal of Economic Theory, 20, 231-259.
Holmström, B. (1979). Moral hazard and observability. Bell Journal of Economics, 10, 74-91.
Iannaccone, L.R. (1992). Sacrifice and stigma: Reducing free-riding in cults, communes, and other collectives. Journal of Political Economy, 100, 271-291.
Iannaccone, L.R. (1994). Why strict churches are strong. American Journal of Sociology, 99, 1180-1211.
Iannaccone, L.R., Haight, C.E., & Rubin, J. (2011). Lessons from Delphi: Religious markets and spiritual capitals. Journal of Economic Behavior and Organization, 77, 326-338.
Kimbrough, E.O. & Sheremeta, R.M. (2013). Side-payments and the costs of conflict. International Journal of Industrial Organization, 31, 278-286.
Kimbrough, E.O. & Sheremeta, R.M. (2014). Why can't we be friends? Entitlements and the costs of conflict. Journal of Peace Research, 51, 487-500.
Kimbrough, E.O., Sheremeta, R.M. & Shields, T. (2014). When parity promotes peace: Resolving conflict between asymmetric agents. Journal of Economic Behavior and Organization, 99, 96-108.
Konrad, K. A. (2009). Strategy and dynamics in contests. New York, NY: Oxford University Press.
Lacomba, J. A., Lagos, F., Reuben, E., & Van Winden, F. (2014). On the escalation and de-escalation of conflict. Games and Economic Behavior, 86, 40-57.
Leeson, P. (2014). Oracles. Rationality and Society, 26, 141-169.
Lissau, R. (2011). Coin flips choose winners in Island Lake, Green Oaks. Daily Herald. http://www.dailyherald.com
Mago, S.D., Savikhin, A.C. & Sheremeta, R.M. (2014). Facing your opponents: Social identification and information feedback in contests. Journal of Conflict Resolution, forthcoming.
Maher, C.R., & Lott, D.F. (2000) A review of ecological determinants of territoriality within vertebrate species. American Midland Naturalist, 143, 1-29.
McCabe, K.A., Rigdon, M., & Smith, V.L. (2003). Positive reciprocity and intentions in trust games. Journal of Economic Behavior and Organization, 52, 267-275.
McCabe, K.A., Smith, V.L., & Lepore, M. (2000). Intentionality detection and 'mindreading': Why does game form matter? Proceedings of the National Academy of Sciences, 97, 4404-4409.
Milgrom, P.R., North, D.C., & Weingast, B.R. (1990). The role of institutions in the revival of trade: The law merchant, private judges, and the champagne fairs. Economics and Politics, 2, 1-23.
Molander, P. (1985). The optimal level of generosity in a selfish, uncertain environment. Journal of Conflict Resolution, 29, 611-618.

Petrie, M., Halliday, T., & Sanders, C. (1991). Peahens prefer peacocks with elaborate trains. Animal Behavior, 41, 323-331.

Powell, R. (2006). War as a commitment problem. International Organization, 60, 169-203.

Price, C.R., & Sheremeta, R.M. (2011). Endowment effects in contests. Economics Letters, 111, 217-219.

Price, C.R., & Sheremeta, R.M. (2014). Endowment origin, demographic effects and individual preferences in contests. Journal of Economics and Management Strategy, forthcoming.

Rabin, M. (1991). Reneging and renegotiating. Working Paper.

Schelling, T. (1960). The strategy of conflict. Cambridge, MA: Harvard University Press.

Schniter, E., Sheremeta, R.M., & Sznycer, D. (2013). Building and rebuilding trust with promises and apologies. Journal of Economic Behavior and Organization, 94, 242-256.

Sheremeta, R.M. (2010a). Expenditures and information disclosure in two-stage political contests. Journal of Conflict Resolution, 54, 771-798.

Sheremeta, R.M. (2010b). Experimental comparison of multi-stage and one-stage contests. Games and Economic Behavior, 68, 731-747.

Sheremeta, R.M. (2013). Overbidding and heterogeneous behavior in contest experiments. Journal of Economic Surveys, 27, 491-514.

Skaperdas, S. (2006) Bargaining versus fighting. Defence and Peace Economics, 17, 657-676.

Skaperdas, S., & Syropoulos, C. (2002). Insecure property and the efficiency of exchange. Economic Journal, 112, 133-146.

Tirole, J. (1999). Incomplete contracts: Where do we stand? Econometrica, 67, 741-781.

Tullock, G. (1980). Efficient rent seeking. In Toward a Theory of the Rent-seeking Society, ed. James M. Buchanan, Robert D. Tollison, and Gordon Tullock. 97-112. College Station, TX: Texas A&M University Press.

Williamson, O. (1985). Credible commitments: Using hostages to support exchange. American Economic Review, 73, 519-540.

Zahavi, A. (1975). Mate selection: a selection for a handicap. Journal of Theoretical Biology, 53, 205-214.

Zahavi, A. (1993). The fallacy of conventional signaling. Philosophical Transactions of the Royal Society London, Series B, 340, 227-230.

Table 1: Average Conflict Resolution, Opting-in, Commitment, Conflict Expenditures and Payoffs (Periods 5-29)

	Treatment			
	Exogenous-Free	*Endogenous-Free*	*Endogenous-Costly*	Joint Significance of Treatment Variables
Conflict Resolution	0.54 (0.50) *N* = 550	0.41 (0.49) *N* = 775	0.20 (0.40) *# *N* = 1,150	F(2,96) = 6.85 *p*-value < 0.01
Both Opt-In	1.00 *N* = 550	0.71 (0.46) * *N* = 775	0.26 (0.44) *# *N* = 1,150	F(1,75) = 28.17 *p*-value < 0.01
Both Commit (if Pair Reaches Stage 2)	0.54 (0.50) *N* = 550	0.58 (0.49) *N* = 547	0.76 (0.43) *@ *N* = 296	F(2,73) = 28.17 *p*-value = 0.06
Flip-Loser Commits (if Pair Reaches Stage 2)	0.58 (0.49) *N* = 550	0.61 (0.49) *N* = 547	0.78 (0.42) ^ *N* = 296	F(2,73) = 2.39 *p*-value = 0.09
Conflict Expenditure (if Pair Reaches Stage 3)	28.56 (18.93) *N* = 502	24.62 (20.08) *N* = 912	18.58 (19.27)*@ *N* = 1,848	F(2,83) = 3.08 *p*-value = 0.05
Payoff	36.97 (52.54) *N* = 1,100	35.52 (51.59) *N* = 1,550	33.53 (51.19) *N* = 2,300	F(2,96) = 0.51 *p*-value = 0.60

Standard deviation reported in parentheses. Rows 1 through 4 report frequency, and rows 5 and 6 report the experimental currency francs, where 100 francs equals one US $1. Each reported variable was averaged over all periods per pair and the average was regressed against treatment dummy variables. For rows 3 through 5, we used probability-weighted regressions, as each pair did not always have the same number of observations (e.g., reached stage 2 or stage 3). Amounts noted with * signify the treatment coefficient was significantly different from the *Exogenous-Free* coefficient, *p*-value < 0.05. Amounts noted with # signify the treatment coefficient was significantly different from the *Endogenous-Free* coefficient, *p*-value < 0.05. Amounts noted with ^ signify the treatment coefficient was significantly different from the *Exogenous-Free* coefficient, *p*-value < 0.10. Amounts noted with @ signify the treatment coefficient was significantly different from the *Endogenous-Free* coefficient, *p*-value < 0.10. For row 2, * signifies the coefficient is significantly different from 1, *p*-value < 0.01.

Table 2: Average Endgame Effects, by Treatment

		Periods	
Treatment	Variable	20-29	30
Exogenous-Free	Both Opt-In	1.00	1.00
	Both Commit	0.51	0.36
	Conflict Expenditure	28.8	24.9
Endogenous-Free	Both Opt-In	0.67	0.71
	Both Commit	0.60	0.23
	Conflict Expenditure	23.7	29.1
Endogenous-Costly	Both Opt-In	0.25	0.24
	Both Commit	0.85	0.45
	Conflict Expenditure	17.6	19.4

Table 3: Transition Matrix of Actions, Periods 5-29

Period t		*Exogenous-Free*			*Endogenous-Free*				*Endogenous-Costly*			
		Period t+1			Period t+1				Period t+1			
Player i	Player j	Player i Flip	Player i Don't Flip	N	Player i Flip	Player i Don't Flip	N		Player i Flip	Player i Don't Flip	N	
Flip	Flip				0.824	0.176	1098		0.709	0.291	598	
Flip	Don't Flip	N/A			0.924	0.076	211		0.670	0.330	448	
Don't Flip	Either				0.863	0.137	241		0.247	0.753	1254	
Period t		Period t+1			Period t+1				Period t+1			
Player i	Player j	Player i Flip	Player i Don't Flip	N	Player i Flip	Player i Don't Flip	N		Player i Flip	Player i Don't Flip	N	
Agrees	Agrees				0.799	0.201	636		0.784	0.216	444	
Agrees	Don't Agree	N/A			0.876	0.124	225		0.514	0.486	72	
Don't Agree	Either				0.844	0.156	237		0.476	0.524	82	
		Period t+1			Period t+1				Period t+1			
Period t		Both choose Agree	At least one chooses Don't Agree	N	Both choose Agree	At least one chooses Don't Agree	Don't Flip	N	Both choose Agree	At least one chooses Don't Agree	Don't Flip	N
Both choose Agree		0.560	0.440	600	0.267	0.352	0.381	636	0.685	0.027	0.288	444
At least one chooses Don't Agree		0.524	0.476	500	0.502	0.229	0.268	462	0.078	0.156	0.766	154

Figure 1: Game Tree

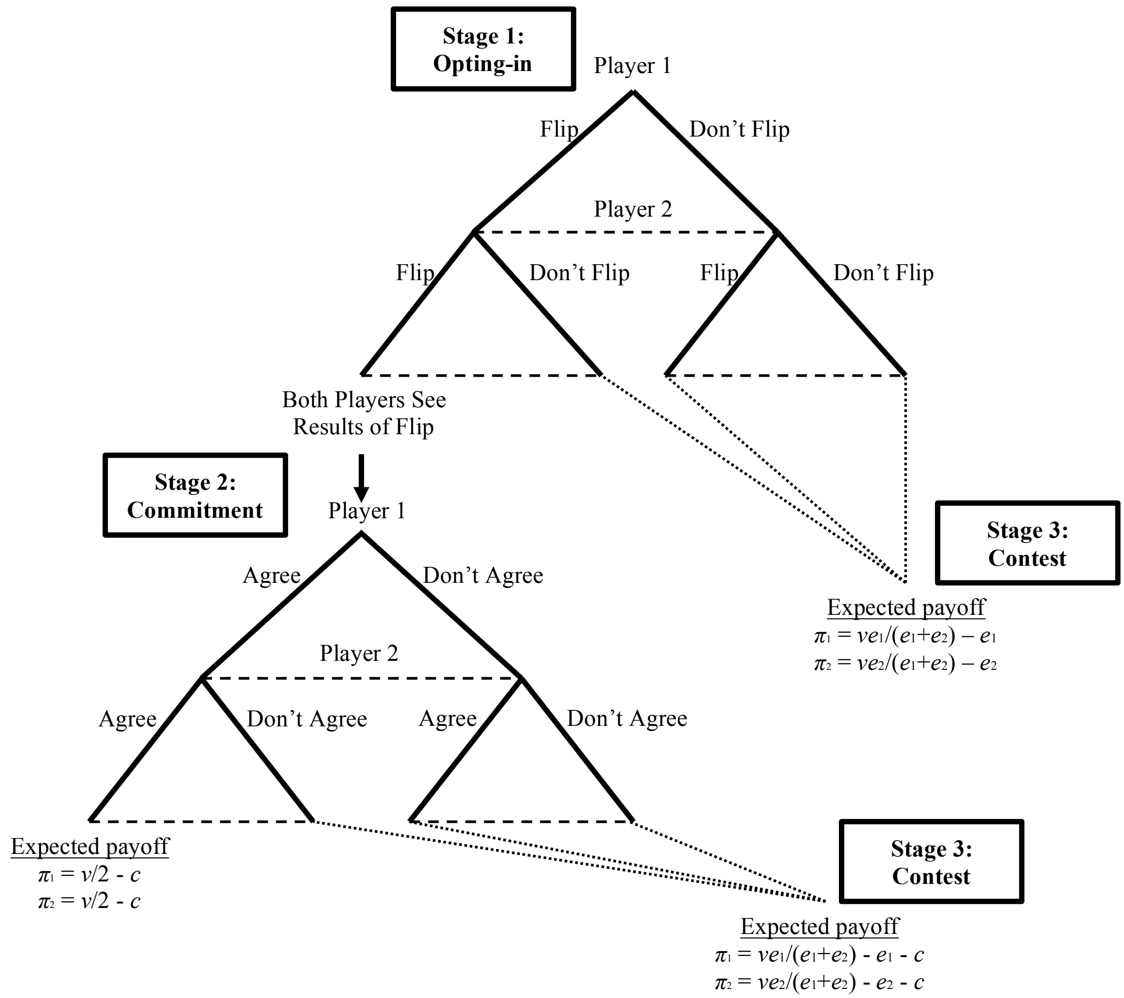
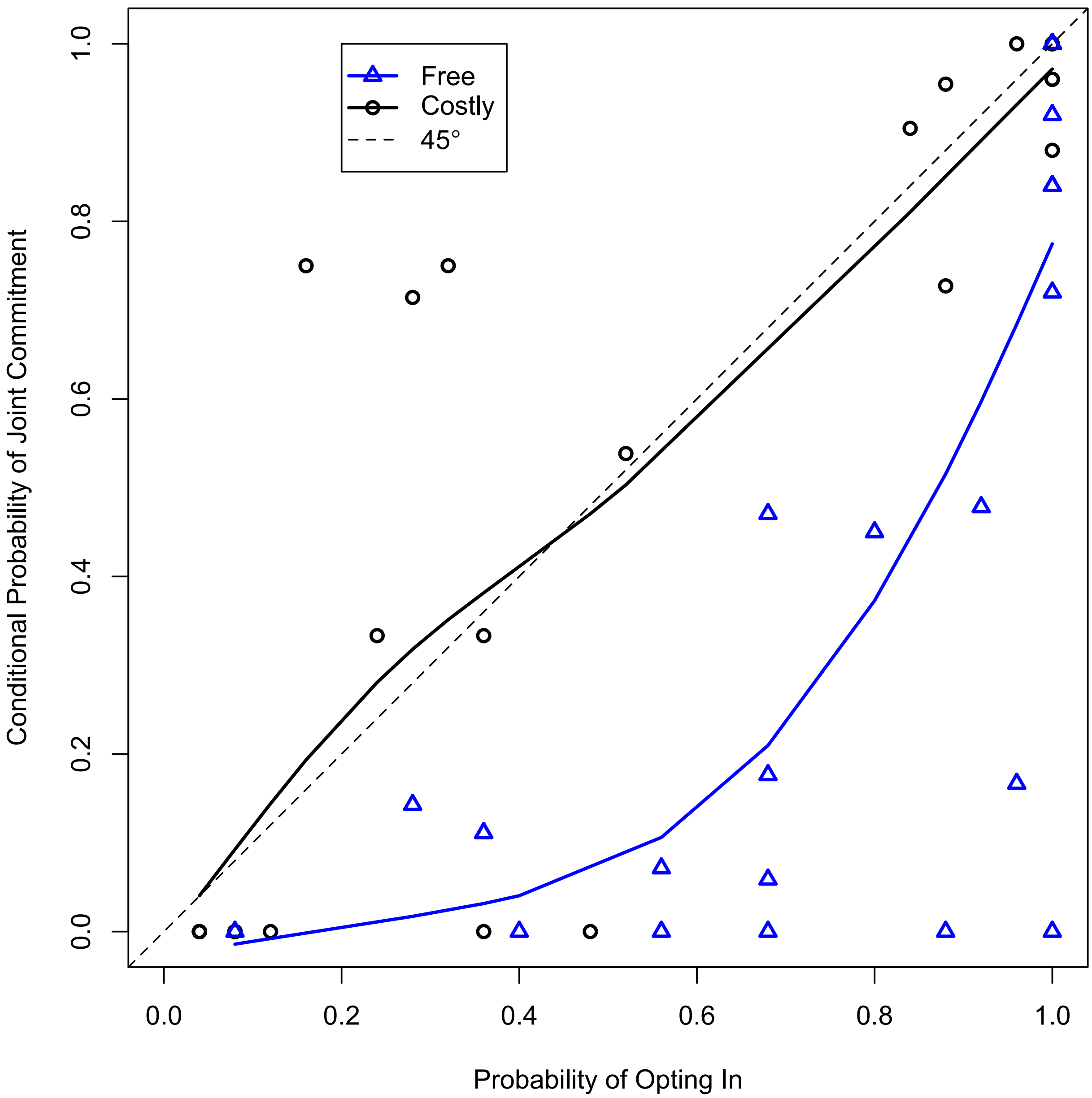
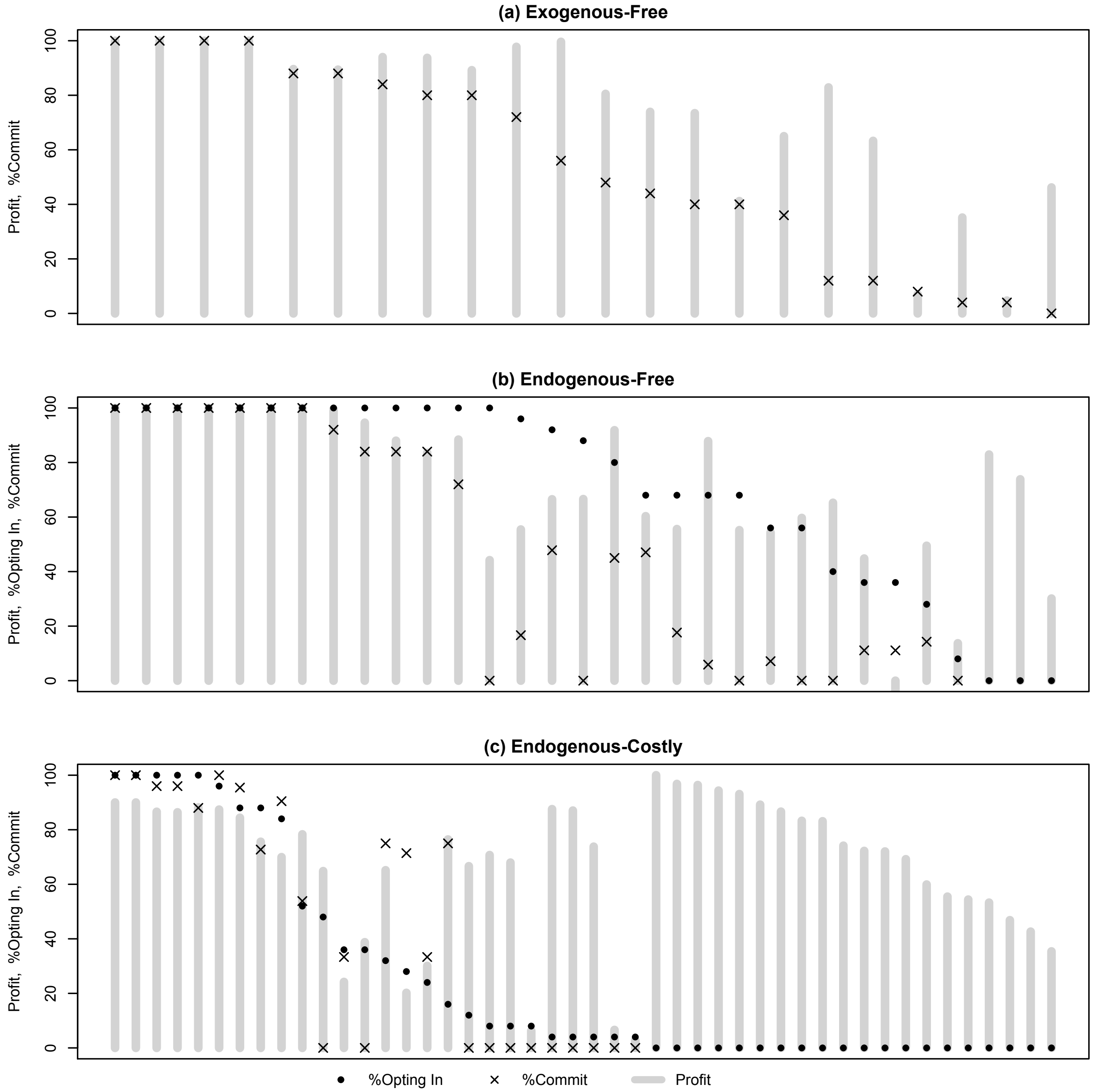


Figure 2: Probability of Both Opting-in vs. Conditional Probability of Both Committing (*Endogenous-Free* and *Endogenous-Costly* Treatments, Periods 5-29)



Note: Each point represents the probability of opting-in and the conditional probability of commitment for a single pair over periods 5-29; the lines plot smoothing spline fits to the data with 3 DF.

Figure 3: Average Profits, Opting-in, and Commitment by Pair and Treatment (Periods 5-29)



Note: Each column displays the average of three variables for a single pair over periods 5-29. Colored bars represent profits, black dots represent the opt-in rate, and black x's represent the commitment rate, conditional on opting-in. The pairs are sorted by opt-in rate with ties broken by commitment rate, and then profit to allow visual identification of relationships between the variables. Note, that there is a strong relationship between commitment and profit only in the *Exogenous-Free* and *Endogenous-Free* treatments.

Appendix A (For Online Publication) – Instructions

Instructions for the *Exogenous-Free* Treatment

GENERAL INSTRUCTIONS

This is an experiment in the economics of strategic decision-making. Various research agencies have provided funds for this research. If you follow the instructions closely and make appropriate decisions, you can earn an appreciable amount of money, which will be paid to you in cash.

The currency used in the experiment is francs. Francs will be converted to U.S. Dollars at a rate of **_100_** francs to **_1_** dollar. You have already earned **a $20.00 participation fee** (this includes the $7 show up fee). The experiment will consist of **30** periods and at the end of the experiment we sum your total earnings for all 30 periods and convert them to a U.S. dollar payment.

It is very important that you remain silent and do not look at others' decisions (screens). If you have any questions, or need assistance of any kind, please raise your hand and an experimenter will come to you. If you talk, laugh, exclaim out loud, etc., you will be asked to leave and you will not be paid. We expect and appreciate you following the laboratory's rules. The remainder of the instructions will describe the decisions you may face in each period.

The participants in today's experiment will be randomly assigned into two-participant groups. You and the other participant in your group will make choices that will determine your payoffs. The experiment contains 30 periods. You will remain within the same two-participant group for all 30 periods.

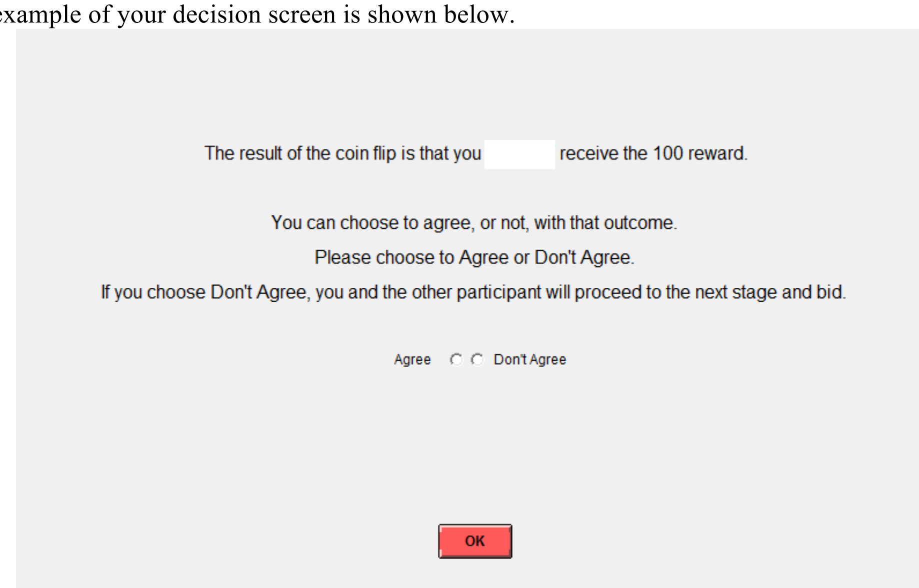
In each period of the experiment one of the two participants in your group will receive the **reward**. The reward is worth **100** francs. Each period of the experiment consists of as many as **two decision stages**.

YOUR DECISION IN STAGE 1

Before making a decision in **Stage 1**, the computer will **flip a coin**. There is a 50% chance the coin lands heads, and 50% chance the coin lands tails. If the computer coin lands heads one participant will receive the reward, if it lands tails, the other participant receives the reward. The flip outcome determines who receives the reward. So, there are two possible payoffs:

If You Receive the Reward	Earnings = 100
If The Other Participant Receives the Reward	Earnings = 0

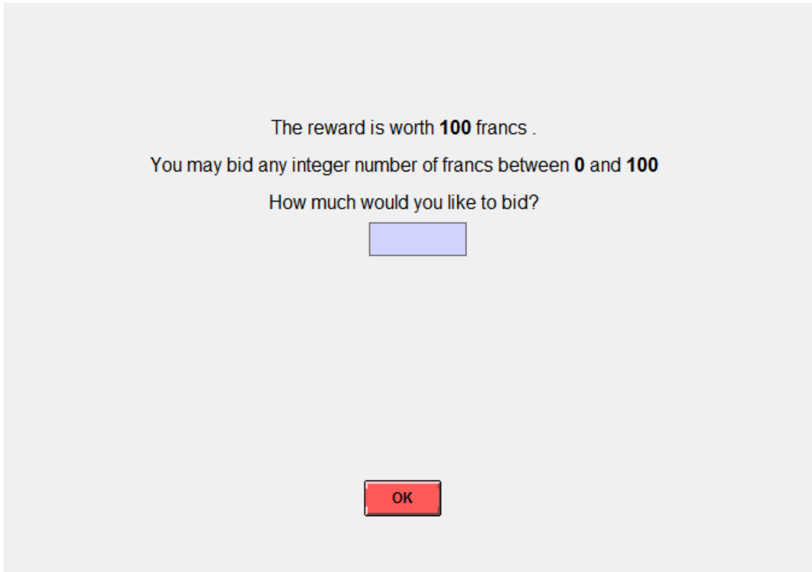
In **Stage 1**, both participants will have to choose whether they want to **Agree** to the outcome of a computer coin flip. An example of your decision screen is shown below.



If you want to agree, check **Agree**. If **both** participants choose to **Agree**, the computer will assign earnings to participants according to the coin flip and the period is over, meaning that neither participant enters **Stage 2**. However, if **either** participant chooses **Don't Agree**, then both participants will enter **Stage 2**.

YOUR DECISION IN STAGE 2

If either participant checked **Don't Agree** in **Stage 1**, each participant enters **Stage 2**. In this stage, each participant may **bid for the 100 franc reward**. You may bid any integer number of francs between **0** and **100**. An example of your decision screen is shown below.



The more you bid, the more likely you are to receive the reward. The more the other participant bids, the less likely you are to receive the reward. Specifically, for **each franc** you bid, you will receive **1 lottery ticket** and for **each franc** the other participant bids, the other participant will receives **1 lottery ticket**. After both participants make their bids, the computer will **draw randomly** one ticket among all the tickets purchased by **you and the other participant**. The owner of the winning ticket receives the reward of 100 francs. Each ticket has an equal chance of winning. So your chance of receiving the reward is given by the number of tickets you buy divided by the total number of tickets bought by you and the other participant.

$$\text{Chance of Receiving the Reward} = \frac{\text{Your Total Lottery Tickets}}{\text{Sum of Your and Other Participant's Lottery Tickets}}$$

If both participants bid *zero* the reward is randomly assigned to one of the two participants.

After both participants make their bids, your earnings for the period are calculated. Regardless of who receives the reward, *both* participants will have to pay their bids. So your earnings will be calculated in the following way:

If either participant decided **Don't Agree** in Stage 1:

If you receive the reward: Earnings = 100 – Your Bid

If you do not receive the reward: Earnings = 0 – Your Bid

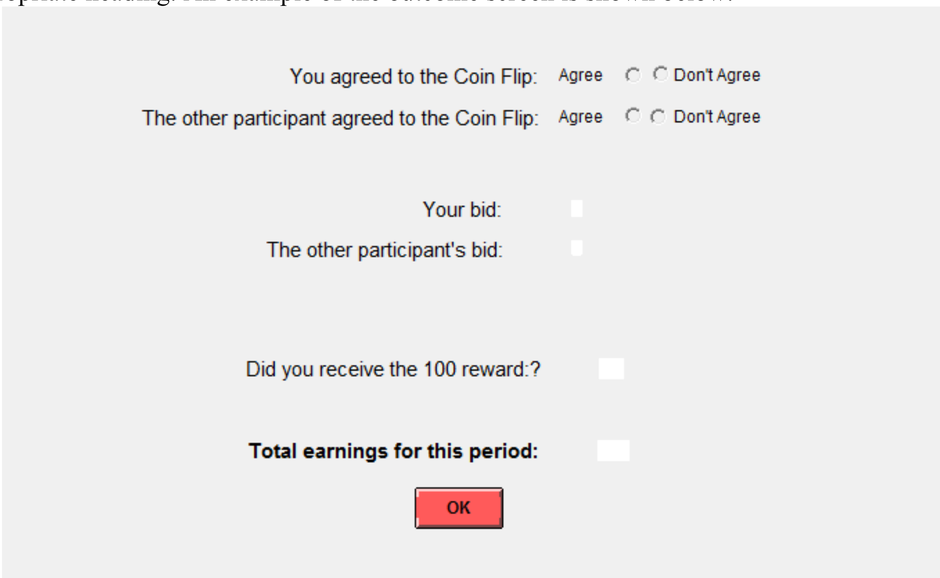
Remember you have already earned a $20.00 participation fee (equivalent to 2,000 francs). In any period, you may receive either positive or negative earnings. At the end of the experiment we will sum the total earnings for all 30 periods of the experiment and convert them to a U.S. dollar payment. If the summed earnings are negative, we will subtract them from your participation fee. If the summed earnings are positive, we will add them to your participation fee.

An Example

This is a hypothetical example used to illustrate how the computer determines who wins the reward of 100 francs. If you bid 30 francs and the other participant bids 20 francs, then you receive 30 lottery tickets and the other participant receives 20 lottery tickets. Then the computer randomly draws one lottery ticket out of 50 (30 + 20). As you can see, you have a higher chance of receiving the reward, 0.60 = 30/50. The other participant has a 0.40 = 20/50 chance of receiving the reward.

After both participants bid, the computer will make a random draw that will determine who receives the reward. Then the computer will calculate your period earnings based on your bid and whether you received the reward or not.

At the end of each period, the computer will display all decision in all three stages on the outcome screen. Once the outcome screen is displayed you should record your results for the period on your **Personal Record Sheet** under the appropriate heading. An example of the outcome screen is shown below.



IMPORTANT NOTES

In each period, you and another participant have the opportunity to receive a **reward** worth 100 francs. Before making a decision in **Stage 1**, the computer will **flip a coin**. In **Stage 1** the outcome of the coin flip is revealed, and both participants choose either **Agree** or **Don't Agree**. If both participants choose **Agree**, the outcome of the coin flip is made final and the payoffs are computed for each participant. On the other hand, if either participant chooses **Don't Agree** in **Stage 1**, both participants enter **Stage 2**. Then, each participant submits a bid for the reward, and the reward is allocated by a random computerized draw. Are there any questions?

Instructions for the *Endogenous-Free* Treatment

GENERAL INSTRUCTIONS

This is an experiment in the economics of strategic decision-making. Various research agencies have provided funds for this research. If you follow the instructions closely and make appropriate decisions, you can earn an appreciable amount of money, which will be paid to you in cash.

The currency used in the experiment is francs. Francs will be converted to U.S. Dollars at a rate of **_100_** francs to **_1_** dollar. You have already earned **a $20.00 participation fee** (this includes the $7 show up fee). The experiment will consist of **30** periods and at the end of the experiment we sum your total earnings for all 30 periods and convert them to a U.S. dollar payment.

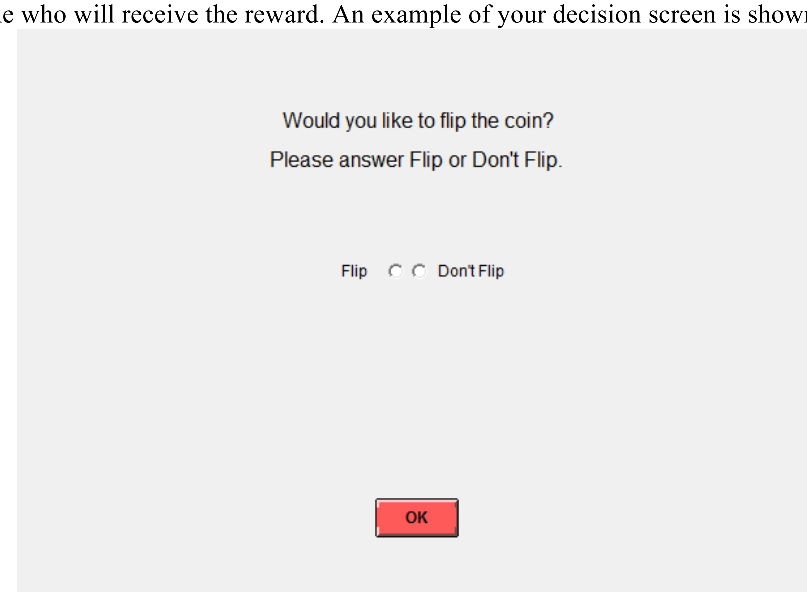
It is very important that you remain silent and do not look at others' decisions (screens). If you have any questions, or need assistance of any kind, please raise your hand and an experimenter will come to you. If you talk, laugh, exclaim out loud, etc., you will be asked to leave and you will not be paid. We expect and appreciate you following the laboratory's rules. The remainder of the instructions will describe the decisions you may face in each period.

The participants in today's experiment will be randomly assigned into two-participant groups. You and the other participant in your group will make choices that will determine your payoffs. The experiment contains 30 periods. You will remain within the same two-participant group for all 30 periods.

In each period of the experiment one of the two participants in your group will receive the **reward**. The reward is worth **100** francs. Each period of the experiment consists of as many as **three decision stages**.

YOUR DECISION IN STAGE 1

In **Stage 1**, both participants will have the opportunity to choose whether they want to **Flip** a computer coin in order to determine who will receive the reward. An example of your decision screen is shown below.



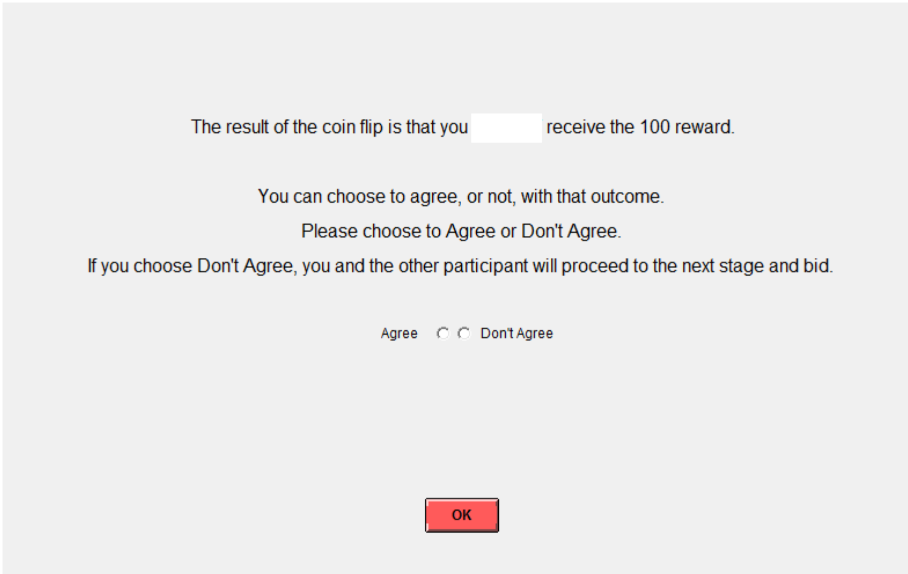
If you want to flip the coin, check **Flip**. If **both** participants choose **Flip**, then both participants will enter **Stage 2**. However, if **either** participant chooses **Don't Flip,** then both participants will skip **Stage 2** and enter **Stage 3**.

YOUR DECISION IN STAGE 2

If both participants choose **Flip** in **Stage 1**, the computer will flip a coin. There is a 50% chance the coin lands heads, and 50% chance the coin lands tails. If the computer coin lands heads one participant will receive the reward, if it lands tails, the other participant receives the reward. The flip outcome determines who receives the reward. So, there are two possible payoffs:

If You Receive the Reward	Earnings = 100
If The Other Participant Receives the Reward	Earnings = 0

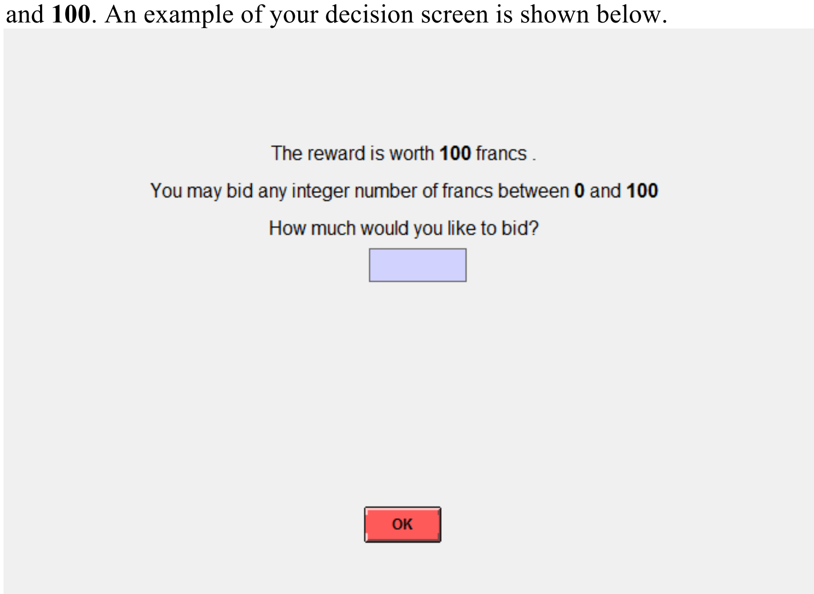
In **Stage 2**, both participants will have to choose whether they want to **Agree** to the outcome of a computer coin flip. An example of your decision screen is shown below.



If you want to agree, check **Agree**. If **both** participants choose to **Agree**, the computer will assign earnings to participants according to the coin flip and the period is over, meaning that neither participant enters **Stage 3**. However, if **either** participant chooses **Don't Agree**, then both participants will enter **Stage 3**.

YOUR DECISION IN STAGE 3

If either participant checked **Don't Flip** in **Stage 1** or checked **Don't Agree** in **Stage 2**, each participant enters **Stage 3**. In this stage, each participant may **bid for the 100 franc reward**. You may bid any integer number of francs between **0** and **100**. An example of your decision screen is shown below.



The more you bid, the more likely you are to receive the reward. The more the other participant bids, the less likely you are to receive the reward. Specifically, for **each franc** you bid, you will receive **1 lottery ticket** and for **each franc** the other participant bids, the other participant will receives **1 lottery ticket**. After both participants make their bids, the computer will **draw randomly** one ticket among all the tickets purchased by **you and the other participant**. The owner of the winning ticket receives the reward of 100 francs. Each ticket has an equal chance of winning. So your chance of receiving the reward is given by the number of tickets you buy divided by the total number of tickets bought by you and the other participant.

$$\text{Chance of Receiving the Reward} = \frac{\text{Your Total Lottery Tickets}}{\text{Sum of Your and Other Participant's Lottery Tickets}}$$

If both participants bid *zero* the reward is randomly assigned to one of the two participants.

After both participants make their bids, your earnings for the period are calculated. Regardless of who receives the reward, *both* participants will have to pay their bids. So your earnings will be calculated in the following way:

If either participant decided **Don't Flip** in Stage 1:

If you receive the reward:	Earnings = 100 – Your Bid
If you do not receive the reward:	Earnings = 0 – Your Bid

If both participants decided **Flip** in Stage 1, but either participant decided **Don't Agree** in Stage 2:

If you receive the reward:	Earnings = 100 – Your Bid
If you do not receive the reward:	Earnings = 0 – Your Bid

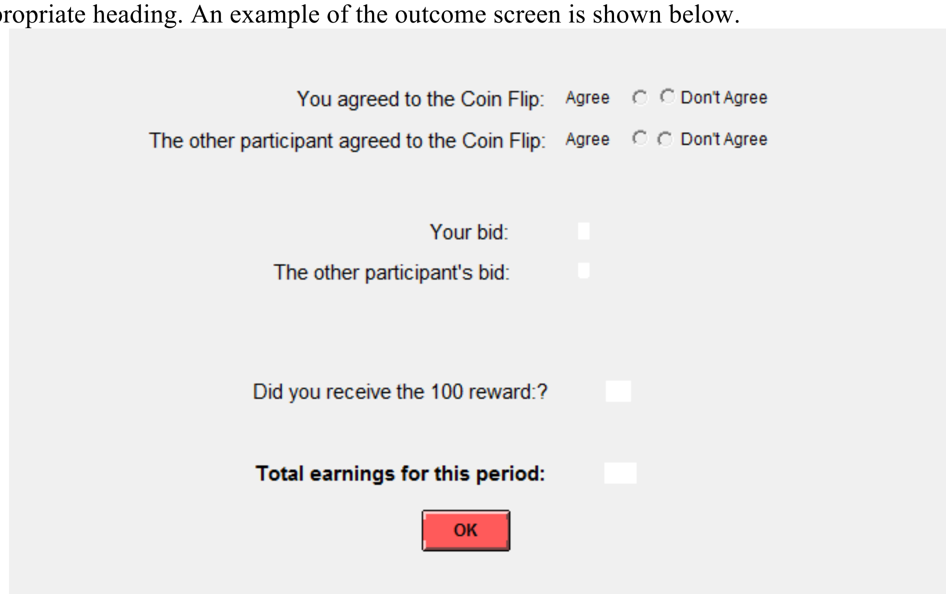
Remember you have already earned a $20.00 participation fee (equivalent to 2,000 francs). In any period, you may receive either positive or negative earnings. At the end of the experiment we will sum the total earnings for all 30 periods of the experiment and convert them to a U.S. dollar payment. If the summed earnings are negative, we will subtract them from your participation fee. If the summed earnings are positive, we will add them to your participation fee.

<u>An Example</u>

This is a hypothetical example used to illustrate how the computer determines who wins the reward of 100 francs. If you bid 30 francs and the other participant bids 20 francs, then you receive 30 lottery tickets and the other participant receives 20 lottery tickets. Then the computer randomly draws one lottery ticket out of 50 (30 + 20). As you can see, you have a higher chance of receiving the reward, 0.60 = 30/50. The other participant has a 0.40 = 20/50 chance of receiving the reward.

After both participants bid, the computer will make a random draw that will determine who receives the reward. Then the computer will calculate your period earnings based on your bid and whether you received the reward or not.

At the end of each period, the computer will display all decision in all three stages on the outcome screen. Once the outcome screen is displayed you should record your results for the period on your **Personal Record Sheet** under the appropriate heading. An example of the outcome screen is shown below.



IMPORTANT NOTES

In each period, you and another participant have the opportunity to receive a **reward** worth 100 francs. In **Stage 1** each participant chooses either **Flip** or **Don't Flip**. If both participants choose **Flip**, then the experiment moves to **Stage 2**. In **Stage 2** the outcome of the coin flip is revealed, and both participants choose either **Agree** or **Don't Agree**. If both participants choose **Agree**, the outcome of the coin flip is made final and the payoffs are computed for each participant. On the other hand, if either participant chooses **Don't Flip** in **Stage 1** or if either participant chooses **Don't Agree** in **Stage 2**, both participants enter **Stage 3**. Then, each participant submits a bid for the reward, and the reward is allocated by a random computerized draw. Are there any questions?

Instructions for the Endogenous-Costly Treatment

GENERAL INSTRUCTIONS

This is an experiment in the economics of strategic decision-making. Various research agencies have provided funds for this research. If you follow the instructions closely and make appropriate decisions, you can earn an appreciable amount of money, which will be paid to you in cash.

The currency used in the experiment is francs. Francs will be converted to U.S. Dollars at a rate of **_100_** francs to **_1_** dollar. You have already earned **a $20.00 participation fee** (this includes the $7 show up fee). The experiment will consist of **30** periods and at the end of the experiment we sum your total earnings for all 30 periods and convert them to a U.S. dollar payment.

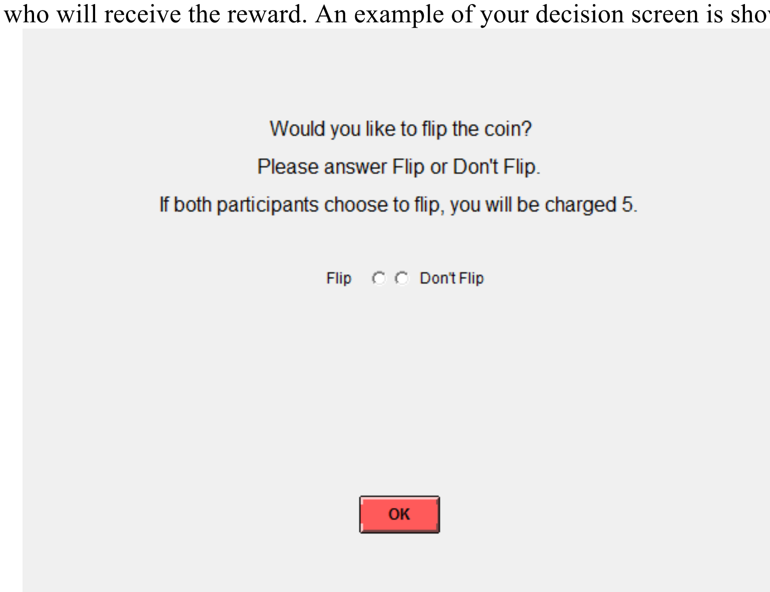
It is very important that you remain silent and do not look at others' decisions (screens). If you have any questions, or need assistance of any kind, please raise your hand and an experimenter will come to you. If you talk, laugh, exclaim out loud, etc., you will be asked to leave and you will not be paid. We expect and appreciate you following the laboratory's rules. The remainder of the instructions will describe the decisions you may face in each period.

The participants in today's experiment will be randomly assigned into two-participant groups. You and the other participant in your group will make choices that will determine your payoffs. The experiment contains 30 periods. You will remain within the same two-participant group for all 30 periods.

In each period of the experiment one of the two participants in your group will receive the **reward**. The reward is worth **100** francs. Each period of the experiment consists of as many as **three decision stages**.

YOUR DECISION IN STAGE 1

In **Stage 1**, both participants will have the opportunity to choose whether they want to **Flip** a computer coin in order to determine who will receive the reward. An example of your decision screen is shown below.



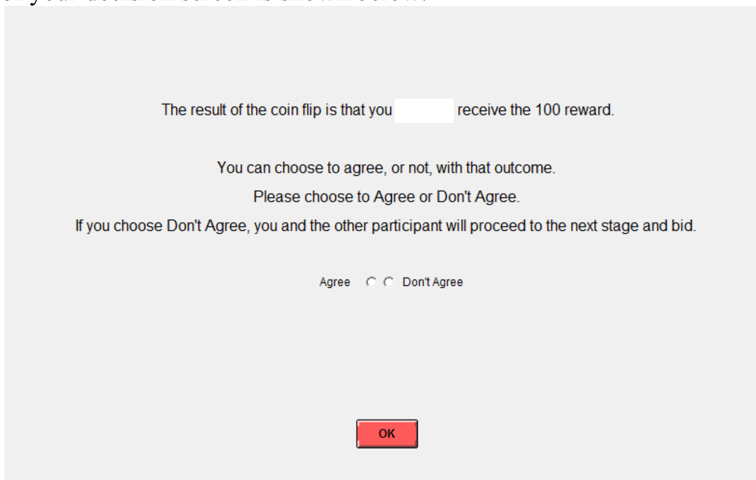
If you want to flip the coin, check **Flip**. If **both** participants choose **Flip**, the computer will subtract **5 francs** from both participants' earnings for the period, and both participants will enter **Stage 2**. However, if **either** participant chooses **Don't Flip,** then both participants will skip **Stage 2** and enter **Stage 3**, and the 5 francs will not be subtracted from either participant.

YOUR DECISION IN STAGE 2

If both participants choose **Flip** in **Stage 1**, the computer will flip a coin. There is a 50% chance the coin lands heads, and 50% chance the coin lands tails. If the computer coin lands heads one participant will receive the reward, if it lands tails, the other participant receives the reward. The flip outcome determines who receives the reward. So, there are two possible payoffs:

If You Receive the Reward	Earnings = 100 – 5 francs
If The Other Participant Receives the Reward	Earnings = 0 – 5 francs

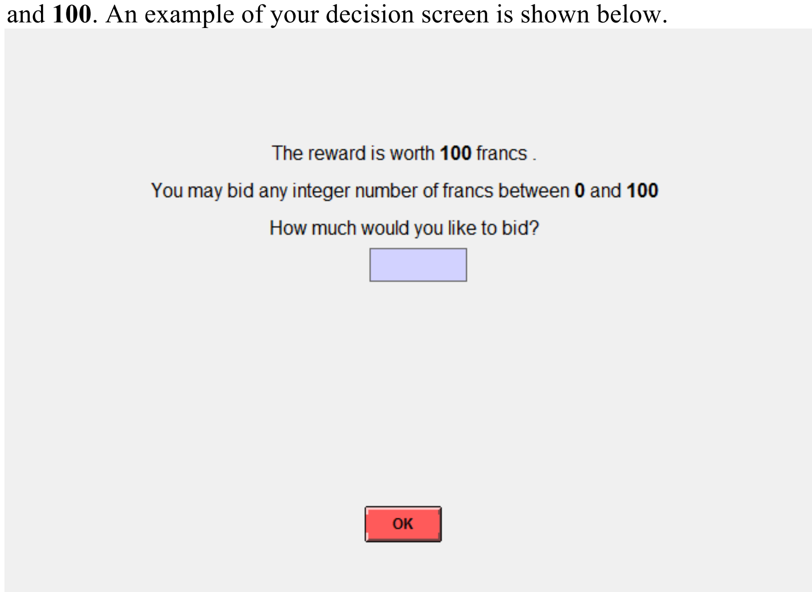
In **Stage 2**, both participants will have to choose whether they want to **Agree** to the outcome of a computer coin flip. An example of your decision screen is shown below.



If you want to agree, check **Agree**. If **both** participants choose to **Agree**, the computer will assign earnings to participants according to the coin flip and the period is over, meaning that neither participant enters **Stage 3**. However, if **either** participant chooses **Don't Agree**, then both participants will enter **Stage 3**, but the 5 francs will be subtracted from both participants since they have agreed in **Stage 1** to flip a coin.

YOUR DECISION IN STAGE 3

If either participant checked **Don't Flip** in **Stage 1** or checked **Don't Agree** in **Stage 2**, each participant enters **Stage 3**. In this stage, each participant may **bid for the 100 franc reward**. You may bid any integer number of francs between **0** and **100**. An example of your decision screen is shown below.



The more you bid, the more likely you are to receive the reward. The more the other participant bids, the less likely you are to receive the reward. Specifically, for **each franc** you bid, you will receive **1 lottery ticket** and for **each franc** the other participant bids, the other participant will receives **1 lottery ticket**. After both participants make their bids, the computer will **draw randomly** one ticket among all the tickets purchased by **you and the other participant**. The owner of the winning ticket receives the reward of 100 francs. Each ticket has an equal chance of winning. So your chance of receiving the reward is given by the number of tickets you buy divided by the total number of tickets bought by you and the other participant.

$$\text{Chance of Receiving the Reward} = \frac{\text{Your Total Lottery Tickets}}{\text{Sum of Your and Other Participant's Lottery Tickets}}$$

If both participants bid *zero* the reward is randomly assigned to one of the two participants.

After both participants make their bids, your earnings for the period are calculated. Regardless of who receives the reward, *both* participants will have to pay their bids. So your earnings will be calculated in the following way:

If either participant decided **Don't Flip** in Stage 1:

If you receive the reward:	Earnings = 100 – Your Bid
If you do not receive the reward:	Earnings = 0 – Your Bid

If both participants decided **Flip** in Stage 1, but either participant decided **Don't Agree** in Stage 2:

If you receive the reward:	Earnings = 100 – Your Bid – 5 francs
If you do not receive the reward:	Earnings = 0 – Your Bid – 5 francs

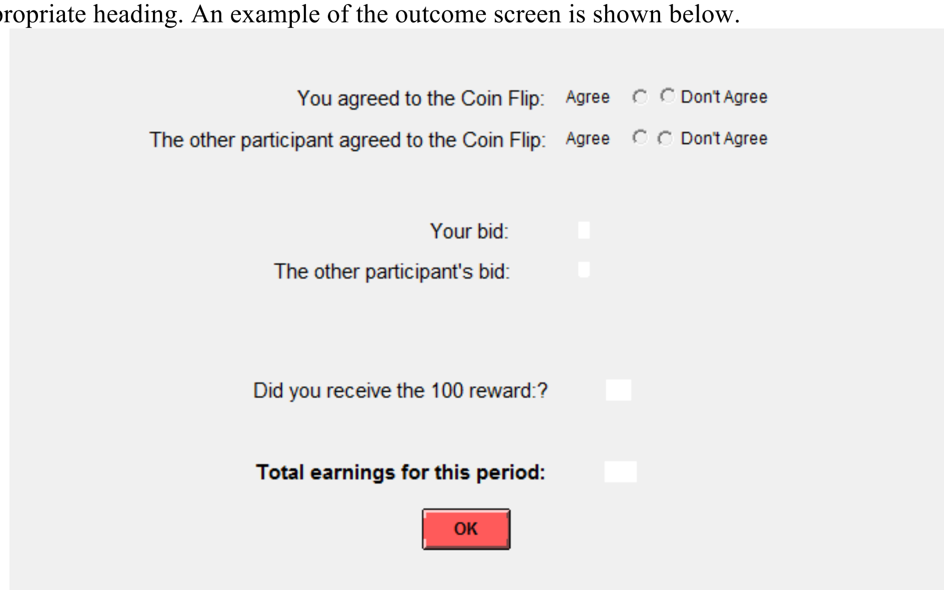
Remember you have already earned a $20.00 participation fee (equivalent to 2,000 francs). In any period, you may receive either positive or negative earnings. At the end of the experiment we will sum the total earnings for all 30 periods of the experiment and convert them to a U.S. dollar payment. If the summed earnings are negative, we will subtract them from your participation fee. If the summed earnings are positive, we will add them to your participation fee.

<u>An Example</u>

This is a hypothetical example used to illustrate how the computer determines who wins the reward of 100 francs. If you bid 30 francs and the other participant bids 20 francs, then you receive 30 lottery tickets and the other participant receives 20 lottery tickets. Then the computer randomly draws one lottery ticket out of 50 (30 + 20). As you can see, you have a higher chance of receiving the reward, 0.60 = 30/50. The other participant has a 0.40 = 20/50 chance of receiving the reward.

After both participants bid, the computer will make a random draw that will determine who receives the reward. Then the computer will calculate your period earnings based on your bid and whether you received the reward or not.

At the end of each period, the computer will display all decision in all three stages on the outcome screen. Once the outcome screen is displayed you should record your results for the period on your **Personal Record Sheet** under the appropriate heading. An example of the outcome screen is shown below.

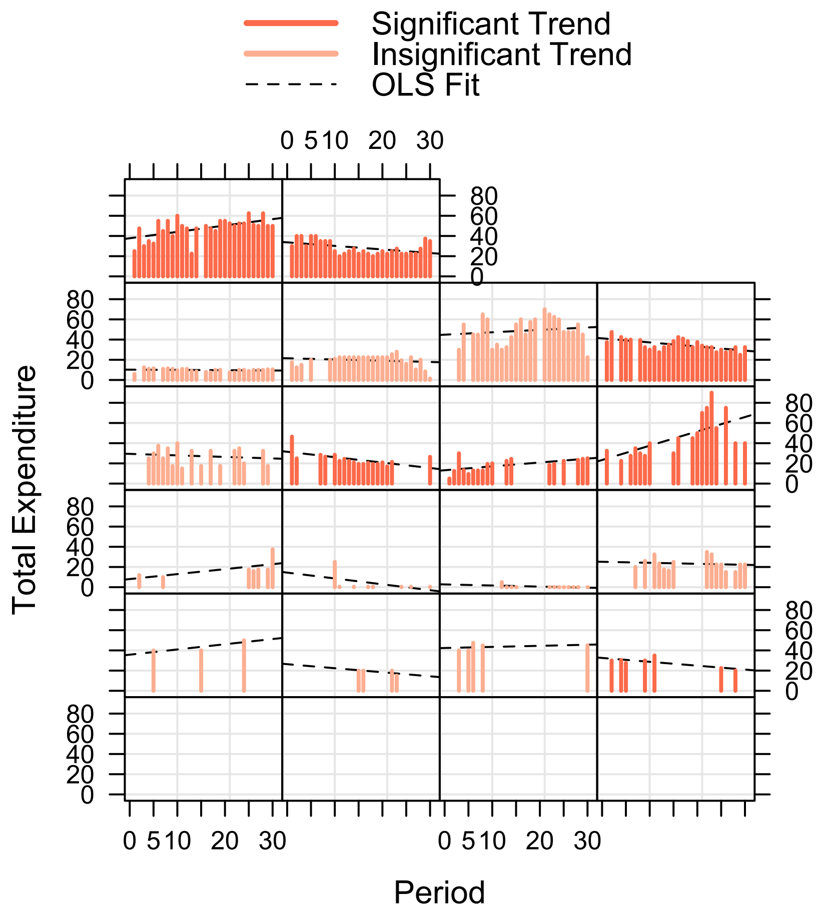


IMPORTANT NOTES

In each period, you and another participant have the opportunity to receive a **reward** worth 100 francs. In **Stage 1** each participant chooses either **Flip** or **Don't Flip**. If both participants choose **Flip**, then both participants pay 5 francs and the experiment moves to **Stage 2**. In **Stage 2** the outcome of the coin flip is revealed, and both participants choose either **Agree** or **Don't Agree**. If both participants choose **Agree**, the outcome of the coin flip is made final and the payoffs are computed for each participant. On the other hand, if either participant chooses **Don't Flip** in **Stage 1** or if either participant chooses **Don't Agree** in **Stage 2**, both participants enter **Stage 3**. Then, each participant submits a bid for the reward, and the reward is allocated by a random computerized draw. Are there any questions?

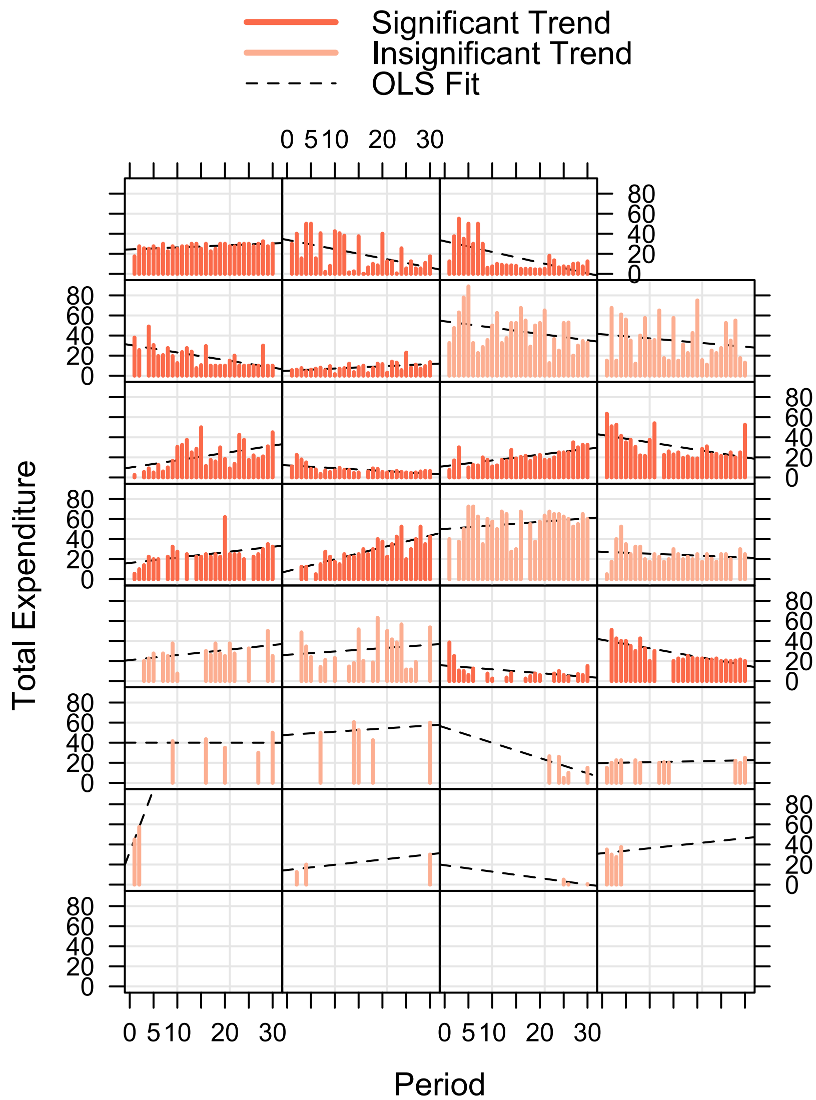
Appendix B (For Online Publication) – Additional Figures

Figure B1: Time Series of Conflict Expenditures by Pair, *Exogenous-Free* Treatment



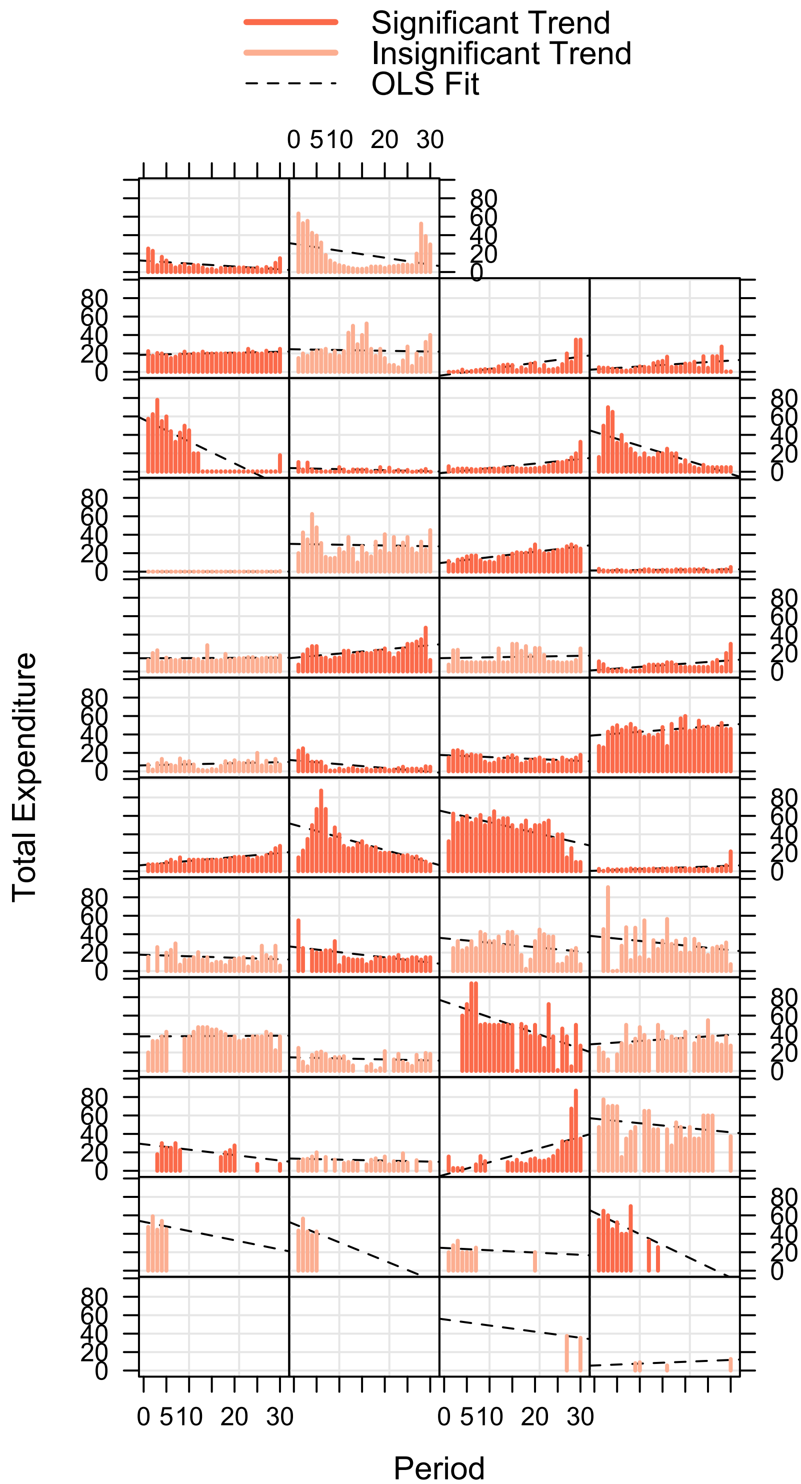
Note: Panels ordered from the top by probability of conflict.

Figure B2: Time Series of Conflict Expenditures by Pair, *Endogenous-Free* Treatment



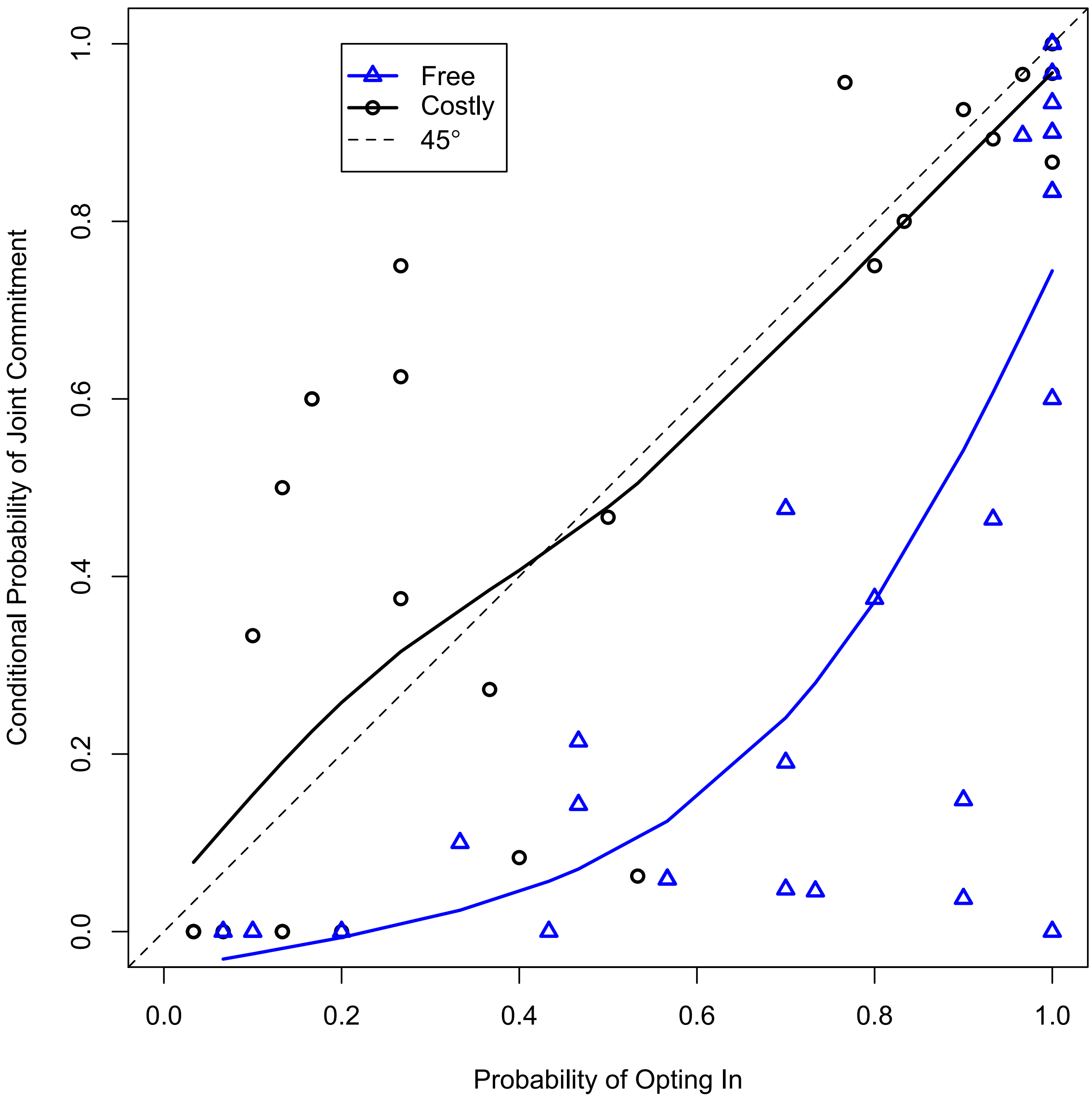
Note: Panels ordered from the top by probability of conflict.

Figure B3: Time Series of Conflict Expenditures by Pair, *Endogenous-Costly* Treatment



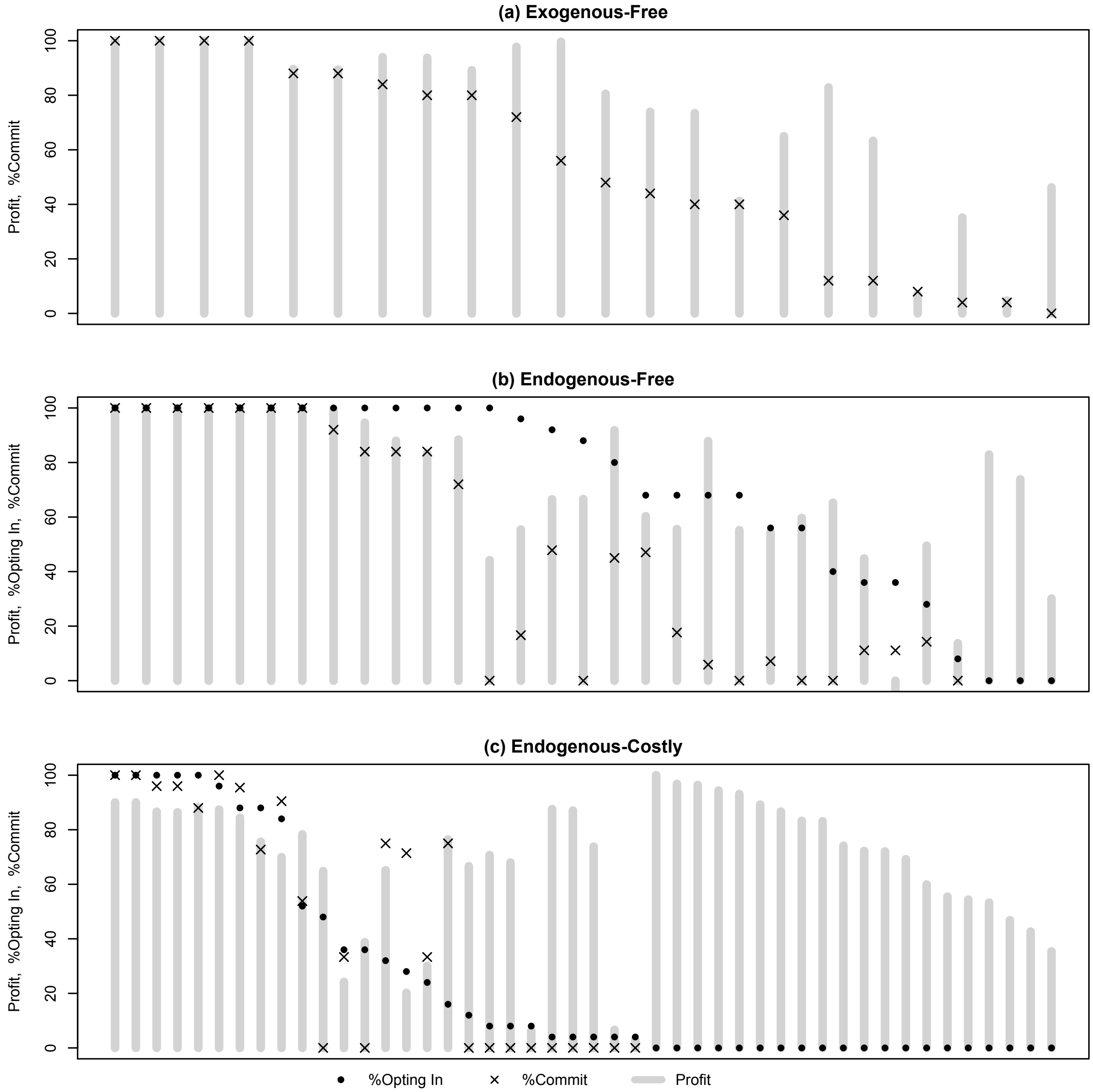
Note: Panels ordered from the top by probability of conflict.

Figure B4: Probability of Both Opting-in vs. Conditional Probability of Both Committing (*Endogenous-Free* and *Endogenous-Costly* Treatments, All Periods)



Note: Each point represents the probability of opting-in and the conditional probability of commitment for a single pair over all 30 periods; the lines plot smoothing spline fits to the data with 3 DF.

Figure B5: Average Profits, Opting-in and Commitment by Pair and Treatment (All Periods)



Note: Each column displays the average of three variables for a single pair over periods 1-30. Colored bars represent profits, black dots represent the opt-in rate, and black x's represent the commitment rate, conditional on opting-in. The pairs are sorted by opt-in rate with ties broken by commitment rate, and then profit to allow visual identification of relationships between the variables. Note, e.g., that there is a strong relationship between commitment and profit only in the *Exogenous-Free* and *Endogenous-Free* treatments.