

# Social Learning and Coordination in High-Stakes Games

## Evidence from *Friend or Foe*

Felix Oberholzer-Gee  
Harvard Business School

Joel Waldfogel  
Matthew White  
The Wharton School  
University of Pennsylvania and NBER

June 5, 2003

### Abstract

We analyze the behavior of game-show contestants who play a one-shot game called *Friend or Foe*. While it is a weakly dominant strategy not to cooperate, almost half the contestants on the show choose to play “friend.” Remarkably, the behavior of contestants remains unchanged even when stakes are very high, ranging from \$200 to more than \$10,000. We conclude that the frequent cooperation observed in one-shot social dilemma games is not an artefact of the low stakes typically used in laboratory experiments. Strategic decisions on *Friend or Foe* change markedly if players can observe previous episodes. We show that these contestants play “friend” if they have reason to expect their opponent to play “friend,” and they play “foe” otherwise. The observed decisions are consistent with recent fairness theories that characterize individuals as conditional cooperators. Using information about past play, some groups (e.g., pairs of women) manage to stabilize cooperation in this high-stakes environment. For most others, improved coordination implies a drastic decline in monetary winnings. Prior to playing the social dilemma game, contestants “produce” their endowment by answering trivia questions. We find some evidence for reciprocal behavior: Players who produce fewer correct answers for their team are more likely to cooperate in the social dilemma game.

JEL: H41, K42, A13, C93

We are grateful to Mary Benner and Hannah Waldfogel for introducing us to *Friend or Foe* and to Sarah Waldfogel for spirited research assistance. Rachel Croson, Dan Kessler, Peter Zemsky, and seminar participants in the Wharton Summer Applied Economics workshop provided useful and thought-provoking comments.

#### IV. Introduction

In June 2002, the Game Show Network began airing a cable television show in which contestants play a high-stakes, one-shot game called *Friend or Foe*.<sup>1</sup> In this game, each of two players simultaneously chooses whether to play *Friend* or *Foe*. Each player's payoff depends on the action chosen by the other contestant in the following way:

Figure 1

		Player 2	
		Friend	Foe
Player 1	Friend	$(x/2, x/2)$	$(0, x)$
	Foe	$(x, 0)$	$(0, 0)$

This game is a variant of the classic Prisoner's Dilemma, and is similar in structure to games analyzed previously in experimental studies.<sup>2</sup> Here playing "foe" is a weakly-dominant strategy for each player. In contrast to the prior literature, however, the stakes in *Friend or Foe* are quite high.<sup>3</sup> The payoff  $x$  varies between \$200 and \$16,400, with an average of \$2,800 at stake per game. In 315 games, the 630 contestants take home over \$700,000.

The data from *Friend or Foe* reveal two striking facts about play. First, individuals choose "friend" at a remarkably high rate, even at very high stakes. The fraction of players choosing "friend" when the stakes are under \$3,000 is 55 percent, and it is 54 percent when the stakes exceed \$3000 and 55 percent when the stakes exceed \$5,000. Second, the fraction of games with split outcomes—when one player chooses "friend" and the other "foe"—declines substantially over the

---

<sup>1</sup> We are not the first to use data generated by television game shows. Gertner (1993) examines attitudes toward risk on *Card Sharks*, Metrick (1995) studies betting behavior on *Jeopardy!*, and Berk, Hughston, and Vandevande (1996) examine learning and bounded rationality on *The Price is Right*. Nor are we the first to analyze play on *Friend or Foe*. List (2003) uses data on the first 40 episodes to draw inferences about discrimination.

<sup>2</sup> Social dilemma games are among the most commonly studied experimental situations (for surveys, see Ledyard, 1995 and Laury and Holt, forthcoming).

<sup>3</sup> For reasons of cost, most high-stakes experiments have been undertaken in low-income countries. The evidence in this paper is (to our knowledge) the first to study subjects from an advanced Western economy play such a high-stakes one-shot game (for a survey on the role of stakes, see Camerer and Hogarth, 1999).

course of the show. That is, players evidently learn to “coordinate” on outcomes along the main diagonal of the payoff matrix above, despite the simultaneous-move nature of the game. Moreover, these outcomes become correlated with players’ observable characteristics (e.g., gender, race, and age) over time.

A third interesting feature arises because the show’s payoff stakes  $x$  is determined in pre-game play by the pairs’ answers to trivia questions. The relative productivity of the two players is predictive of whether each chooses to share (i.e., play “friend”) or steal (play “foe”) the fruit of their efforts, at rates which are again stakes-invariant.

This paper documents these features of play, providing reinforcement for some prior theories of behavior but pointing to gaps in the literature with respect to other empirical regularities. The observation that individuals play “friend” at a high rate is consistent with various empirical results from laboratory experiments. These findings can be rationalized by supposing that at least some players care about the fairness of outcomes in this game (Rabin, 1993). Fehr and Schmidt (1999) argue that individuals have preferences over outcomes that depend in part on the *difference* in monetary payoffs between players, and Bolton and Ockenfels (2000) indicate that individuals dislike taking a disproportional *share* of the winnings. Nevertheless, in previous papers a number of authors expressed a concern that the findings from laboratory studies with fairly small stakes would not generalize to contexts with larger economic consequences (Binswanger, 1980; Kachelmeier and Shehata, 1992; Fehr, Fischbacher and Tougareva, 1995; Cameron, 1995; Slonim and Roth, 1998).<sup>4</sup> The results from *Friend or Foe* allay some concerns about generalizability. The overwhelming evidence from this game is that players’ cooperative tendencies are remarkably *stakes invariant*.

---

<sup>4</sup> Moreover, the literature has expressed concerns that experiments performed in relatively poor countries to simulate high stakes might not be generalizable to Western economies, given substantial uncontrolled differences in cultural norms (see Roth et. al., 1991).

The second fact about play is less readily explained with the existing literature on behavior in games. Our evidence on this point is derived from a useful feature of the production history of the show. The show was filmed in two “seasons,” with the first 40 episodes produced before the show’s on-air debut in June 2002. The remaining 65 episodes were taped in late summer 2002, after the airing of the first season. Players on the first season therefore had little show-related basis for forming beliefs about opponent play, while players in the second season could observe the play of 240 prior contestants. That is, *Friend or Foe* is an intergenerational game with a sequence of non-overlapping “generations” of players. In many experiments, players learn through personal experience by repeatedly playing the same game. *Friend or Foe* is an example of social learning, where contestants learn by observing the decisions of other contestants, as in Schotter and Sopher (2003). While these authors find that relying on history alone does not allow players to coordinate their actions, we find strong evidence that players learn to coordinate on symmetric “friend-friend” or “foe-foe” outcomes. The evidence is particularly strong for pairs of players with common observable characteristics; among pairs of women, for example, the fraction of games with split outcomes falls from 58 percent in the first season (N=26) to 40 percent (N=40) in the second season.

This observation raises two questions: *How* do later generations of players manage to achieve symmetric outcomes, and *why* do they prefer to do this? A tantalizing piece of evidence on the former is that during the first season, a player’s choice in the game is statistically independent of his or her opponent’s, but this choice *is* correlated with the player’s observable characteristics such as age, race and gender. In the second season play depends on upon both the individual’s *and* his or her opponents’ observable characteristics. This suggests that later generations of players have learned to condition their strategies on their opponents’ observable characteristics.

Explaining *why* play would evolve toward this behavior – when “foe” is always a weakly dominant strategy in monetary terms – requires appeal to both theories of fairness and “social” (intergenerational) learning. Building on the literature, we suggest that some fraction of the population are “conditional cooperators” who prefer to play “friend” as a best response to an opponent playing “friend,” but prefer “foe” against “foe.” The rest of the population may have less-intense preferences regarding fairness, electing “foe” regardless. If these unobservable aspects of preferences are correlated during early rounds of the game with players’ observable characteristics, then updating by conditional cooperators after the first season would explain the evolution of intergenerational play observed in the data.

This explanation has two noteworthy implications. The first is how powerful preferences for “fair” outcomes must be, given the evidence from *Friend or Foe*. Specifically, because of the high stakes for which *Friend or Foe* is played, the implied money-metric value of a “fair” (i.e., cooperative) outcome must be quite large for conditional cooperators – on the order of hundreds or thousands of dollars in this simple game. The second implication is that players with demographic attributes associated with conditionally-cooperative preferences will come to fare better (monetarily) over time, relative to players without observable characteristics initially associated with cooperation. In the data, changes in take-home winnings over time are consistent with this prediction. In essence players ‘stereotyped’ by observable characteristics associated with uncooperative play in early generations do progressively worse monetarily, as opponents become likely to play foe against them.

The final feature of *Friend or Foe* we examine arises because players determine the stakes of the game by initially answering trivia questions. Before playing the game shown in Figure 1, pairs of players answer up to 18 questions. For each pair, the number of right answers determines

the game's stakes,  $x$ . We directly observe how each pair comes to choose a particular answer, allowing us to record the number of correct and incorrect answers for which each player is responsible. We use this information to examine theories of equity and reciprocity in subsequent (non-cooperative) play. For example, if a player's partner has contributed a larger number of correct answers, does the player show "gratitude" with a greater tendency to play "friend"? The answer is affirmative, at least for the first season. Contestants who make smaller contributions than their partners are more likely to play "friend." Interestingly, this effect disappears in the second season, and is quite insensitive to the monetary payoffs at stake.

The paper proceeds in four sections. Section 2 describes the game context that generates the data. Section 3 discusses the theoretical background with reference to the relevant literature. Section 4 describes the data, and Section 5 presents the results with respect to stakes, learning and coordination, and pre-game contribution activity.

## **V. The Quasi-Experimental Context**

### **A. The Game**

*Friend or Foe* is a television game show that was aired on the Game Show Network beginning June 3<sup>rd</sup>, 2002. The game has two components, a *production phase*, in which player pairs jointly contribute to answering questions, and a *distribution phase*, in which contestants play a social dilemma game to determine how the pie they have produced will be divided. The Game Show Network (2003) provides the following description of the game:

"The show consists of six strangers who pair up at the start of the show to form three teams of two. Each team is separated into isolation chambers where all trivia rounds will be played. The newly formed teams have to work together and agree on answers to trivia questions, in order to build a bank account. At the end of each round (there are 3 rounds total) the lowest scoring team is eliminated. Before the team is dismissed, they enter the "Trust Box" where they decide how their winnings up to that point are divided."

The first round has four trivia questions, worth \$500 each. The second round has four questions worth \$1000 each. In the third round, the remaining team answers up to ten \$500 “right or wrong” questions. If the third-round team answers all ten questions correctly within 60 seconds, their entire score is doubled. The winnings to be divided can therefore range between \$200 and \$22,000. Some trivia questions are “absurdly easy” (TV Guide Online, 2003), many others are impossible to know.<sup>5</sup> For each question in rounds 1 and 2, the host provides four possible answers. The two players in a team need to agree on an answer before they can respond. The show airs the pair trying to arrive at consensus, and it is generally possible to observe which of the pair first suggested the answer they ultimately choose. By coding this information for every question, we can determine the total number of positive and negative contributions that each contestant makes to her team’s score. After each round, the lowest-scoring team plays the game in Figure 1 to decide how the money in their account is to be split, then exits the show.<sup>6</sup>

The show, which was taped in Santa Monica, CA, aired in two “seasons.” The first season consisted of 40 episodes taped prior to the show’s on-air premiere on June 3, 2002. These episodes aired twice daily on weekdays (at 4 and 10pm), then rerun on weekends. A second season of 65 new episodes was taped in late summer 2002. These were aired beginning October 1, 2002. Contestants on the show during the second season therefore can have seen the play from the first season, while contestants during the first season could not have seen any episodes on television prior to taping.<sup>7</sup>

---

<sup>5</sup> A typical example for the latter category is: “According to a Blockbuster survey, which celebrity do people most want to clone: Britney Spears, Sharon Stone, Brad Pitt, or John Travolta?” (Answer: Britney Spears).

<sup>6</sup> Daniel R. Coleridge, writing in TV Guide Online, June 3, 2002, describes this part of the show as follows: “*Friend or Foe*’s rather Satanic premise has contestant duos answer questions, then pits them against each other in a nasty mind game: Can they trust each other to split their jackpot equally, or will one lie and steal all the cash from the other?”

<sup>7</sup> There were a few minor changes in the show between the first and second seasons. During the first season each team was given \$200 for their account prior to the start of the game (so that a team answering no questions correctly would have something to split). The top possible score in the first season is \$22,200. In the second season the show started

Partner assignment is not random; instead, selection worked as follows (again, according to show producers):

“Prior to the taping of each episode the six game players will be gathered together backstage. There, the three contestants and the three potential players will be introduced to one another via a producer. The producer will first expose the three contestants’ positive and negative attributes. Once all of the contestant’s information has been shared, the producer will then disclose all three potential partners’ positive and negative attributes. After each potential partner has been introduced, the producer will ask the three contestants to select one person they would like to partner with. These choices, made in security, will be written down and then displayed one by one. If all three have selected different partners, the producer will identify each team as official partners. If two or even three contestants have selected the same partner, then the choice falls to this selected partner. After the selected partner has chosen, the remaining contestants select their second choice partner. This process will continue until three teams of two have been formed.”<sup>8</sup>

## **B. TV Shows and Laboratory Experiments**

Our context has some clear advantages and disadvantages. The first and foremost advantage is that *Friend or Foe* allows us to observe decisions in a social dilemma situation with very high stakes. Balanced against this advantage are a few features that distinguish our context from standard laboratory experiments.

The game show differs from many experiments in that players interact in person. While face-to-face one-shot interaction is not obviously less realistic than double-blind exchanges – many business and social situations constitute one-shot games where people countenance their opponents

---

teams with nothing but gave teams answering no questions correctly \$200 to split in the dilemma game. The top possible score in the second season is \$22,000.

<sup>8</sup> Producer Melissa Rudman, email communication with authors, April 8, 2003. During the first season the show aired the partner selection process, while this part was not aired during the second season. Hence, one cannot know whether partners chose each other in the latter 65 episodes.

– personal interactions reduce the degree of control in the experiment because it is difficult to empirically assess whether looks, hugs, and smiles influence the observed strategic decisions. Second, our contestants play on a televised show where play is not anonymous. While it is unlikely that acquaintances of the contestants happen to see the show by accident – only 0.6 percent of cable television households watch *Friend or Foe* (Greco, 2002) – it is possible that some players informed friends or family that they would be on TV. While this visibility diverges from standard laboratory experiments, we do not view it as clearly bad. In real life, only very rarefied examples of one-shot interaction have no chance of being observed by third parties.

These issues aside, our context has two challenges. The first challenge is that events in the production phase of the game may affect behavior in the distribution phase. Contestants’ “friend” or “foe” decisions may be affected by their relative contributions to their teams’ score. For example, contestants contributing less to the kitty may feel gratitude and may be more likely to play friend. Our strategy for dealing with this is to examine whether players’ choices in the *Friend or Foe* game in Figure 1 are robust to players’ contributions during the production phase. We examine this issue directly in section IV.C.

A related concern is that the size of the stakes is correlated with the history of the production phase. More successful teams who play for higher stakes have a longer “production” history, and they get to observe the strategic decisions of the less successful contestants. Our approach for dealing with this problem is to analyze the relationship between stakes and the foe tendency both within and between rounds of the game.

Finally, as with subjects in most controlled experiments, our subjects are not a random sample of the general population, which may affect our ability to generalize our results.

### III. Theoretical Issues Associated with the Game

The distribution phase of *Friend or Foe* is the simple game in Figure 1. In this game, playing “foe” is a weakly dominant strategy. Only if a player believes his or her opponent will play “foe” with probability one is it even rational for an income-maximizing contestant to play “friend.” Given the frequency of friendly play—over the course of the show 45% of contestants choose “friend” —the belief that a partner will play “foe” for certain seems implausible. Furthermore, because of the simplicity of the payoff structure it strains credulity to think that cognitive shortcomings kept contestants from choosing “foe.” Similarly, repeated interaction cannot explain cooperative behavior as it is common knowledge that contestants play a one-shot game.

A more promising avenue for understanding the behavior on the show is to consider additional ‘fairness’ payoffs not captured in Figure 1. Rabin (1993) argues that people want to be nice to those who treat them fairly, and they want to punish those who hurt them. The fairness of actions depends on the players’ intentions, which can be inferred from the distribution of payoffs that these actions induce. One implication of Rabin’s theory is that contestants who expect their opponents to choose “foe” prefer to punish their partner by destroying the entire endowment. There is much evidence on *Friend or Foe* that supports this view. The “friend”-playing partner in a “friend-foe” or “foe-friend” pair often displays sadness or anger. Insults are common, and “friend” players’ remarks sometimes need to be bleeped over. In the context of a public good game, Fehr and Gächter (2000) provide systematic evidence showing that people are willing to punish those who do not cooperate, even if punishment is costly.

While Rabin’s (1993) model builds on players’ intentions, other fairness theories assume that relative material payoffs directly enter the utility functions of individuals. In their theory of

equity, reciprocity and competition, Bolton and Ockenfels (2000) posit that a player's *relative share* of the total pecuniary payout affects utility. An alternative model is due to Fehr and Schmidt (1999), who argue that it is the *difference* between players' payoffs that is relevant. While these fairness theories differ in their specific assumptions and predictions, they all assume that the complete payoff matrix includes both monetary and non-monetary considerations of the following form:

Figure 2

		Player $j$	
		Friend	Foe
Player $i$	Friend	$(x/2, x/2)$	$(-S_i, x-G_j)$
	Foe	$(x-G_i, -S_j)$	$(0, 0)$

The term  $S_i > 0$ , for *sucker's dismay*, motivates a player to destroy the entire endowment if he or she believes the opponent will try to grab the entire pie. The term  $G_i$ , or *guilt*, captures feelings of guilt or shame for having played "foe" when the partner played "friend." On the show, this type of embarrassment is frequently observed. Many apologize for having chosen "foe" when the other contestant was friendly. Some explain they really needed the money, while others say they chose "foe" only because they thought—incorrectly—that the other player would choose "foe."

The non-monetary terms in this augmented game,  $S_i$  and  $G_i$ , reflect unobservable preferences. We assume that these attributes vary across players. Strategies in this augmented game then depend on these values in the following way. If  $p_i$  denotes player  $i$ 's belief that his or her opponent will play "foe," then player  $i$  prefers to play "friend" if and only if

$$G_i - S_i \cdot p_i > x/2$$

where  $\theta_i = p_i / (1-p_i)$ . That is, a player chooses “friend” if her guilt from stealing the entire pie exceeds  $x/2$  by a multiple of her sucker’s dismay should her opponent do so; the potential dismay weighs more heavily in the decision as the prior on an opponent playing “foe” becomes larger.

In this setting, it is useful to distinguish between two ‘types’ of players. Given a game at stakes  $x$ , players with  $G_i > x / 2$  are *conditional cooperators*. For such types there exists a set of beliefs about the likelihood of an opponent choosing “foe” for which it is optimal to also play “foe.” Below a critical threshold of  $p_i$ , however,  $i$  will choose “friend.” That is, a conditional cooperator prefers to play “friend” against an opponent she believes is (sufficiently) likely to also play “friend,” but prefers to meet “foe” with “foe” to avoid sucker’s dismay. The other type of player that it is useful to distinguish here has  $G_i < x / 2$ , or a lower level of guilt from taking the entire pie than a conditional cooperator would have. Such players have a dominant strategy of playing “foe” in the augmented game, assuming all players have a positive value of  $S_i$  (that is, no one likes being played the sucker to at least *some* degree).

Before proceeding further, it is worth noting what conditionally cooperative types have at stake empirically in *Friend or Foe*. Although  $G_i$  and  $S_i$  are not observable directly, informative bounds on the former can be inferred for a sizeable share of the population based on observed play. Overall, approximately 45% of players choose “friend;” for these players, the median stakes  $x$  is \$2700. Thus for nearly half of the 630 players, the money-metric “cost” of playing “foe” against a possibly “friend”-playing partner—a cost we interpret as guilt or shame—must be upwards of \$1,350. This strikes us as a remarkably large sum, especially given the truly end-game nature of players’ *Friend or Foe* dilemma and their quite brief pre-game interactions as show contestants. There is, however, no way around the facts of how people play for the stakes in the data or the structure of the game. Since the magnitude of this lower bound on  $G_i$  (for half the players) does not

depend on the players' (unknown) prior beliefs, we infer that such 'fairness' considerations must be quite powerful even in one-shot social interactions.

### **A. Learning and Coordination**

Fairness theories posit that some individuals—but not everyone—reward friendly behavior with cooperative play. In *Friend or Foe*, one can extend the inference that  $G_i$  exceeds thousands of dollars for many people to the supposition that players may have a high level of motivation to discern attributes of opponents that correlate with these unobserved aspects of preferences. In Season 1, players have little game-related experience to form beliefs about opponents' play. In contrast, players in Season 2 can use Season 1 outcomes to update beliefs about  $p_i$ . For example, if older people in Season 1 were observed to play "friend" at higher-than-average rates—as we will see was the case—then one might infer that fairness considerations of the sort described above are greater for older players.

More generally, we postulate that players' unobservable preferences over the fairness of outcomes may be correlated with players' observable attributes. Such correlations might then be deduced by later generations of players from observation of earlier players' actions. Although it is not essential to our argument, it is reasonable to expect Season 1 players to approach the game with (comparatively) uninformed priors over opponents' play; we know they had little show-related basis for forming such priors given the show's production history. If so, then during Season 1 only players with exceptionally large values of  $G_i$  would choose "friend." Over time, however, the relative frequency of "friend" play can be determined for different groups of people based upon their observable (and non-imitable) characteristics. Since Season 2 players had access to all

(potentially) of the results of Season 1's play, it is natural for the later players to update accordingly.

To see how learning will affect outcomes, imagine two contexts in which players know nothing and everything, respectively, about whether their opponents are conditional cooperators. In each case, suppose it is common knowledge that 50 percent of individuals are conditional cooperators that prefer to play "friend" with a prior of  $p=0.5$ , while the other 50 percent are "money players" who care only about monetary payoffs. In the case without further information, we would expect the individual foe rate to be 50 percent, half of the games to end in off-diagonal (split) outcomes, and a quarter each to end with "friend-friend" and "foe-foe" outcomes.

How do outcomes change when players can perfectly predict an opponent's type? Then three-quarters of pairings should end "foe-foe," while the remaining quarter is "friend-friend." Half of the time, conditional cooperators are paired with money players, and in those cases both play "foe." In the quarter of cases when money players face money players, both play "foe." Only in the remaining quarter of cases when conditional cooperators face other conditional cooperators does either party wish to play friend. Greater player knowledge thus has two effects. First, it raises the individual foe rate, in this case from 50 to 75 percent. Second, it also increases the degree of coordination, or situations in which both players choose the same action and avoid asymmetric outcomes.

In reality – and presumably in the data as well – players cannot perfectly predict opponent types, but in Season 2 they have more information than in Season 1. Then a player can update his or her prior over an opponent's play (correctly) based on observable characteristics, and we should see greater 'coordination' – *i.e.* more play that avoids the non-monetary costs of asymmetric "friend-foe" outcomes. Consider an interaction of two persons whose observable characteristics

are similar to earlier players that typically chose “friend.” Each player might expect his or her opponent to care about the fairness of outcomes because of the correlation observed among earlier players. For our purposes it does not matter whether such players individually update toward a high opponent  $G_j$ , a low “foe” prior value of  $p_i$ , or both; either way, the implication is that both players are more likely to play “friend” than if either faced a player whose observable characteristics were not historically correlated with “friendly” play. We would similarly expect a pairing of two players with observable characteristics that predicted uncooperative behavior in the first season to yield increasingly high “foe-foe” rates in the second season, as players update away from the possibility of high  $G_j$ ’s and ensuing friendly play.

Among conditional cooperators, uncooperative behavior is contagious (Fehr and Schmidt, 1999). When a sufficiently high fraction of persons does not care about fairness, increases in “foe”-playing behavior over time emerges as a positive prediction. Social learning—that is, updating by later generations based on the correlation of historical play with players’ observable characteristics—has the implication that player behavior should depend on both own and opponent characteristics in the second season.

## **B. Gratitude vs. Resentment**

Fairness theories rely on a reference outcome to describe ‘fair’ and ‘unfair’ distributions. In many laboratory experiments, an equal split appears to be a natural reference point. Subjects do not know each other, they are most often students at the same university, and their roles in the experiments are randomly assigned. On *Friend or Foe*, however, players can contribute differently to their teams’ success. For example, if one team member contributes many correct answers while the other contributes few, the former player has, in effect, made a large gift to the joint kitty. A

large body of experimental work in psychology and economics asks how recipients respond to gifts, and in particular whether they reciprocate. In our context this question reduces to the following: Does the player who contributes fewer answers show *gratitude* by playing “friend,” or does she show *resentment* by playing “foe”?

#### **IV. Data**

A total of 105 *Friend or Foe* episodes were produced, with 6 players on each episode, for a total of 630 individual contestants in 315 game pairs. For each contestant we observe gender, age, occupation, race, her team’s score (the value of  $x$  in Figures 1 and 2, or the “stakes”), number of positive and negative contributions to her team’s score (i.e. correct and incorrect answers she contributes), and the amount she – individually – takes home (the “take”).

Our data come from two sources. First, incomplete data on all 105 episodes – episode airdate, player names, “friend” or foe decision, and each player’s take – were obtained via the web.<sup>9</sup> Player name generally allows inference about gender, so the gender variable is available for 621 observations. The stakes were unavailable at this source for cases where pairs played “foe-foe” (and therefore the takes were 0). Second, we obtained complete data on 100 episodes - and 600 players – by taping and coding outcomes and player date directly from the tapes.

Table 1 summarizes characteristics of the sample. About half of the contestants are male, and nearly a sixth of the contestants are black (nearly one per episode). Players tend to be young adults: the median contestant age is 27. The 25<sup>th</sup> and 75<sup>th</sup> percentile age players are 23 and 33, respectively. Dividing the players’ hometowns into four Census divisions and “foreign,” about half

---

<sup>9</sup> <http://gameshowfavorites.classictvfavorites.com/FriendorFoe/episodeguide.html> (accessed May 8, 2003)

of players are from the West, 17 percent are from the Midwest, about 15 percent are from each of the South and the Northeast, and the remaining 6 percent are foreign.<sup>10</sup>

Table 2a shows how game outcomes – the scores, the tendency to play “foe,” and individuals’ winnings – vary by player gender, age, and race. Table 2b shows how play varies with both own and opponent characteristics. We discuss these tables below.

## **V. Results**

In this section we examine 1) how learning affects coordination, 2) how stakes affect play, and 3) how contributions affect play.

### **A. Learning to Coordinate**

#### **1. Learning from Season One**

In Season 1, the tendency to play cooperatively differs across player demographics. Men play foe more often than women (52.5 percent vs. 46.3 percent), although this difference is not statistically significant (two-sided p-value=0.34). Players under the median age ( $\leq 27$ ) choose “foe” more often than players over the median (64.6 vs. 38.9 percent, two-sided p-value $<0.01$ ), and blacks play foe more often than whites (57.5 vs. 48.1 percent, two-sided p-value=0.33). Column (3) of table 3 reports a bivariate probit of both players’ tendencies to play “foe” on their own characteristics in Season 1. The pairwise differences appear here as well, again with low levels of statistical significance. Yet, we resoundingly reject the hypothesis that a players’ own characteristics are unrelated to his or her play (p-value=0.0005).

One thing that Season 1 play does not depend on is an opponent’s characteristics. When we include both own and opponent characteristics in the bivariate probit for Season 1 play – see

---

<sup>10</sup> One can infer from context – show banter – that “hometown” refers to upbringing rather than current residence.

column (5) – we continue to reject the hypothesis that own characteristics do not matter (p-value=0.0005), but we cannot reject the hypothesis that opponent characteristics do not matter (p-value=0.96).

These results support our assumption that we can draw inferences about types' ( $G, S$ ) distributions from Season 1 play. A player familiar with Season 1 would come away with the expectation that women have more “conscience” than men – that is, a ( $G, S$ ) distribution that, for a given set of beliefs about opponent play inclines them to play “friend” with higher frequency – and that older players have more such “conscience” than younger players.<sup>11</sup> We – and Season 2 players – can use these inferences to make predictions about play in Season 2.

## **2. Coordination in Season Two**

What sort of Season 2 play do we expect to see in our data? First, the overall “foe” rate should increase. Second, cooperative players should play “foe” at lower rates against other cooperative players than they do against less cooperative players. As a result, we should see as a third implication that the amount of coordinated play should increase.

Our data bear out these predictions. First, as table 2a shows, the individual “foe” rate jumps from 49 to 58 percent between Season 1 and Season 2 (two-sided p-value=0.03). The “foe” rate does not increase across the board, however. Instead, we see the clear emergence of conditional cooperation: predictably cooperative types are less likely to play “foe” against other predictably cooperative types than against predictably less-cooperative opponents. We can examine this by gender, age, and race.

---

<sup>11</sup> Blacks play foe more often than whites, but this has to be understood in light of the fact blacks never play against other blacks. Hence, black foe rates could simply reflect a high expectation that whites will play foe against them.

In Season 1 men play “foe” 52 percent of the time when playing against men and a statistically indistinguishable 54 percent of the time when playing against women. Similarly, in Season 1, women play statistically indistinguishable “foe” rates against men (48 percent) and women (44 percent). In Season 2, however, “foe” rates depend on both own and partner gender. Men and women both play “foe” over 60 percent against men, and men play “foe” 58 percent of the time against women. Yet, women play “foe” only 45 percent of the time against women. We cannot reject the hypothesis that men play the same against men and women in Season 2, but we can reject the hypothesis that women play the same against both men and women (p-value=0.007, two-sided). See figure 3a.

Results by race combination are similar. In Season 1 blacks play “foe” 58 percent of the time, while whites play “foe” 48 percent, and the “foe” rate for whites differs indistinguishably in Season 1 depending on whether they are playing against whites (51 percent) or blacks (47 percent). In Season 2, by contrast, the “foe” rate is roughly three quarters for whites playing against blacks and for blacks playing against whites, but it is only 53 percent – and statistically significantly lower – for whites playing against whites (p-value=0.03, two sided). See figure 3b.

There is additional evidence along these lines by age. Young players play “foe” nearly two thirds of the time in Season 1, while older players play “foe” only 39 percent of the time, and this difference is significant. In Season 2, the “foe” rate if either player is young averages 60 percent, and one cannot reject the hypothesis that it is equal across young-young, old-young, and young-old pairs. Compared to this, the “foe” rate for older players against older players is only 50 percent, although it is not statistically significantly lower than their rate against younger players in Season 2 (p-value=0.20, two-sided).

Coordinated play increases in Season 2. As table 4 shows, the fraction of games with outcomes on the main diagonal increases from 49 to 57 percent (two-sided p-value=0.15). We can also document this in the context of bivariate probit models of the players' "foe" decisions, as shown in table 3. The first two columns include only constant terms. The  $\rho$  parameter, the correlation of players' foe-playing tendencies not explained by observables, reflects *de facto* coordination and is insignificantly different from 0 in Season 1. In the second Season its estimated value is 0.22 (p-value=0.057), reflecting greater coordination. Columns (3) and (4) include specifications with players' *own* observable characteristics. Players' own characteristics affect their play in both seasons. With players' own observables included,  $\rho$  remains zero in Season 1 and is estimated to be 0.17 (s.e.=0.12) in Season 2.

We suspect that increased coordination in Season 2, if it is based on updating of beliefs based on Season 1, is driven by both variables that we can observe (and record) as well as factors that the players observe but are not in the data. Accordingly, we can ask how much of the increased coordination of play is attributable to our simple characterization of observables. We examine this by including own and opponent observables in the bivariate probit models. Columns (5) and (6) report bivariate probits including both own and opponent characteristics as determinants of play. The  $\rho$  parameter remains essentially zero in Season 1, and there is no evidence that opponent characteristics affect own play in Season 1. In Season 2, however, there is evidence that own play is related to opponent characteristics. With own and opponent observables included, the  $\rho$  parameter estimate declines somewhat to 0.15 (s.e.=0.13), suggesting that our additive observables explain some of the increase in coordination between seasons.

We can revisit the question of how conditional cooperation emerges by including variables indicating whether *both* players have observable characteristics associated with conditional

cooperation. The last two columns include variables for whether players are both female, both old, or both white. “Both female” and “both white” are negative and significant in Season 2 but not in Season 1, affirming the emergence of conditional cooperation in Season 2 suggested in figure 3.

### 3. Coordination and Winnings

From the standpoint of money winnings, coordination is not necessarily beneficial. If half of players are cooperative by nature, but they have the uninformed beliefs about opponents described earlier (c.f. section III.A.), then players will arrive at “foe-foe” – with no winnings for either player – only one quarter of the time. With perfect information, by contrast, there are no winnings for either player three quarters of the time. In the context of this game, information only helps those who are expected to be conditionally cooperative, and then only when they are paired with expected conditional cooperators. If we denote the probability that a contestant plays “foe” as  $\pi$ , and we normalize stakes to 1, then total expected winnings when play is independent are  $1-\pi^2$ , while expected winnings with fully informed players are  $(1-\pi)^2$ . Information reduces winnings regardless of  $\pi$ , and the decline is larger as  $\pi$  is larger.

What happens between seasons? Because the “foe-foe” rate increases, winnings decline even in relation to an overall decline in game stakes between seasons. As table 2a shows, in Season 1 players score an average of \$3718 and take home an average of 39 percent (\$1463). In Season 2 scores average \$3062, and players take home an average of 30 percent.<sup>12</sup>

The ratio declines differently across groups. In particular, it declines less for pairs of expected conditional cooperators. For pairs of women, whites, and older players collectively, the average winnings ratio declines from 39 to 32 percent. For the remainder of the sample, it declines

---

<sup>12</sup> One can alternatively calculate the average of each individual’s take ratio (winnings/score) to arrive at 38.3 in season 1 and 31.9 in season 2.

from 43 to 22 percent. Two-women teams appear to solve the coordination problem differently than others. In Season 2 they have the highest “friend-friend” rate and the lowest “foe-foe” rate. As a result of the latter, women playing women have the highest “take” rate among all groups in Season 2. In essence, women have established an “island of cooperation” in Season 2, at least relative to other players.

### **B. Do Stakes Matter?**

While it is generally prohibitively expensive to run high-stakes experiments, a number of investigators have run experiments in low-income countries, where it is possible to give high stakes by local standards (for a survey, see Camerer and Hogarth, 1999). Translated to current US dollars, the highest stakes in these games seem to be about \$500. The conclusion of this literature is that stake variation over this range does not affect behavior, allaying some of the concern that prior results for low-stakes does not generalize. Yet, two worries remain. First, experiments run in different countries tend to give different results, with stakes constant (see Roth, et al. (1991)). Hence, the high-stakes behavior of Slovak or Indonesian subjects may not generalize to US persons. Second, even though behavior does not change as stakes vary from a few dollars to, say, \$500, behavior may change if stakes rise by another order of magnitude.

The data in our study allow us to revisit the question of how stakes affect play. The team score – based on the number of questions they answer correctly – determines the stakes of the *Friend or Foe* game they play.

Figure 4 shows the relationship between stakes and the tendency to play foe. Each circle in the diagram shows the relative frequency of “foe” play for all players with a particular score. The horizontal axis shows the logarithm of scores. Circle size reflects the number of persons with each

score. The leftmost circle indicates a 39 percent “foe” rate for the 36 individuals who play with \$200 at stake. Other than a depressed “foe” rate with \$200 at stake, there is no apparent relationship between stakes and the individual tendency to play foe.

Figure 5 reports the “foe”-stakes relationship for players eliminated in the first round of each episode. These scores can vary between \$200 and \$2200, depending on whether the team gets 0, 1, 2, 3, or 4 questions right. In order to exit the game in the first round a team must have the lowest score, and the highest score observed for these teams is \$1700.<sup>13</sup> Figure 2 vividly indicates that the “foe” rate is lower for \$200 games (for losers in round 1); the figure also shows an oscillating pattern (low for \$200, high for \$500 and \$700 stakes, lower for \$1000 and \$1200, then higher again for \$1500 and \$1700).<sup>14</sup>

Figure 6 breaks down the individual foe-stakes relationship by season. Although the graphs reveal no unconditional relationships between stakes and the individual tendency to play “foe,” we can explore the relationship further with regressions that account for players’ characteristics along with the game’s stakes. Table 5 presents four groups of probit estimates of the individual “foe” rate on stakes, with and without explanatory covariates. Each group of three uses data from a) the full sample, b) Season 1 only, and c) Season 2 only. The first group of four includes play from all rounds of the game and no covariates. The second group of three includes only the play of teams losing in the first round. These groups repeat in columns 7-12, with covariates for age, gender, race, and region. Standard errors in all regressions are clustered on the pair to account for the possible dependence of opponents’ play.

---

<sup>13</sup> Recall that teams are spotted \$200 prior to the game in the first season, so that first season scores are \$200, 700, 1200, etc, while second season scores are \$200, 500, 1000, etc.

<sup>14</sup> The oscillation suggests a possible explanatory role for relative contributions. When \$200 is at stake, neither player has contributed a correct answer, whereas when \$500 or \$700 is at stake, only one of the two players has contributed a correct answer (unless their answers were truly simultaneous). We explore this further below.

The foe-stakes relationship is not statistically significant in any of the 12 specifications in table 5, although it has borderline significance in the Season 2 regressions using round 1 play (columns 6 and 12). Some observables are generally significant: age bears a generally negative relationship with the tendency to play “foe” (particularly in Season 1 regressions), and westerners are more likely to play foe.

In addition to the relationship between stakes and individual “foe” rates, we can also explore the relationship between stakes and the tendency for cooperative – and coordinated – pair outcomes. Table 6 reports probit estimates of 3 mutually exclusive pair outcomes (foe-foe, friend-friend, and split) on log stakes along with controls for the demographic composition of the pairs, by season. Stakes are not significant in any of the equations.

Except for the teams facing stakes of \$200 – who are slightly more likely to play “friend” – there is no evidence that stakes affect cooperativeness, either within or across rounds of the game. The difference between the “foe” rate at \$200 stakes and higher stakes suggests a possibility that stakes matter only at that lower bound, although this difference could also be due to an effect of answering no questions correctly, independent of stakes.

These issues aside, what is remarkable about our data is that even with very large sums of money at stake, the “foe” rate remains stable and does not increase above 55 percent. If we are correct to conjecture that players understand the game, then these deviations from the (weakly) dominant strategy in the money payoff game are not due to cognitive limitations. Rather, the observed patterns of play must reflect non-monetary payoffs that differ substantially from the money payoffs. In particular, since a player chooses “foe” if and only if  $G_i - S_i \cdot \pi_i > x/2$ , and this occurs with a 55 percent frequency regardless of  $x$ , it must be the case that the 55<sup>th</sup> percentile of

$G_i - S_i \cdot \beta_i$  is proportional to  $x$ . Non-monetary concerns appear to scale up roughly linearly with the monetary gains at stake.

### **C. Do Contributions Affect Play?**

Each team member contributes differently to the kitty during the production phase of the game, and the tendency to play “foe” may depend the number of questions correctly answered. We record the number of questions that each player answers – correctly and incorrectly – in each round, and we use this to create measures of the number of correct and incorrect answers contributed by each team member along with their dollar-value analogues. For the dollar analogues, we weight questions by \$500 for rounds 1 and round 3, and by \$1000 for round 2.

Table 7 presents a first look at these data. The top panel divides the sample according to whether the difference between the net contribution of player 1 less the net contribution of player 2 is above or below the median for this gap. Each player’s net contribution in questions is the number of correct answers less the number of incorrect answers that he contributes. The average above-median net question contribution is about 2.5, while the average below-median net question contribution is about 0. In Season 1, players above the median in terms of net question contribution are more likely to play “foe” (two-sided p-value = 0.091). There is no statistically discernible difference in Season 2.

The second panel of the table revisits the question with dollar-weighted net contributions, and the results are similar. The average above-median net contribution is about \$1400, while the average below-median net contribution is zero or negative. Again, players with above-median dollar-weighted contributions are more likely than others to play “foe” in Season 1 but not in Season 2.

Table 8 examines this question in a series of regressions. Each regression includes team fixed effects. The first two columns include the numbers of correct and incorrect answers separately, by season. Columns (3) and (4) include the net questions (correct – incorrect). Columns (5)-(8) revisit the question with dollar-weighted answers. The only individually significant variable in the table is the (raw or dollar-weighted) number of correct answers in Season 1 (see columns (1) and (5)). That is, there is evidence of reciprocity in Season 1: the player contributing fewer correct answers to the team is more likely than his opponent to play friend. When we divide the sample into high and low stakes groups, we cannot reject the hypothesis that contributions have the same effect on playing “foe” across stakes.

The results beg two questions. First, why do negative contributions have no effect on play? Based on watching the show, we speculate that negative contributions do not affect play because many wrong answers do not harm the team. Most negative contributions occur when neither player knows the answer. Second, why is there no evidence of reciprocity in Season 2? Given our evidence on learning, players in Season 2 have much better predictions of their opponents’ likely play than they did in Season 1. Perhaps what they have learned swamps reciprocity.

We note here that the inclusion of contribution variables in unreported foe-rate regressions does not change the stake invariance reported above.

## **VI. Conclusion**

Our study has a number of interesting findings. First, we have documented that players learn from the history experienced by other players. In particular, they learn how players of various types are expected to play, and what they learn facilitates conditional cooperation. In Season 2, cooperative individuals play “friend” at elevated rates but only when paired with other

cooperative types. While better-informed players tend to bring about smaller dollar winnings, more cooperative types of players, such as women and older contestants, can create “island of cooperation” where winnings decline considerably less.

Second, we document that even very high stakes don’t induce Nash play. This indicates that the non-monetary payoffs associated with disutility of asymmetric (or inequitable) outcomes are large and proportional to dollar stakes. If these sorts of payoffs, which are well-documented in small-stakes laboratory experiments, generalize to other contexts, then one may infer that cooperation can arise without the usual suspects of repeat play or extra-game sanctions.

Finally, we document reciprocity, at least in the first season of play. Players contributing fewer correct answers are more likely than their opponents to play friend. This behavior disappears in the second season, perhaps because player’s ability to predict their opponents’ play creates incentives that swamp the desire to reciprocate.

## Literature

- Berk, Jonathan B., Eric Hughson and Kirk Vandezande (1996). The Price Is Right, but Are the Bids? An Investigation of Rational Decision Theory. *American Economic Review* 86(4): 954-970.
- Binswanger, Hans P. (1980) Attitudes toward Risk: Experimental Measurement in Rural India. *American Journal of Agricultural Economics* 62(3): 395-407.
- Bolton, Gary E. and Axel Ockenfels (2000). ERC: A Theory of Equity, Reciprocity, and Competition. *American Economic Review* 90(1): 166-93.
- Camerer, Colin F. and Robin M. Hogarth (1999). The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework. *Journal of Risk & Uncertainty* 19(1-3): 7-42.
- Cameron, Lisa A. (1995). Raising the Stakes in the Ultimatum Game: Experimental Evidence from Indonesia. *Economic Inquiry* 37(1): 47-59.
- Fehr, Ernst and Simon Gächter (2000). Cooperation and Punishment in Public Goods Experiments. *American Economic Review* 90(4): 980-94.
- Fehr, Ernst and Klaus M. Schmidt (1999). A Theory of Fairness, Competition, and Cooperation. *The Quarterly Journal of Economics* 114(3): 817-68.
- Fehr, Ernst, Urs Fischbacher and Elena Tougareva (2002). Do High Stakes and Competition Undermine Fairness? Evidence from Russia. Institute for Empirical Economics. University of Zurich, Working Paper No. 120.
- Game Show Network (2003). Friend or Foe. At <http://www.gameshownetwork.com/index.html>, accessed April 15, 2003
- Gertner, Robert (1993). Game Shows and Economic Behavior: Risk-Taking on "Card Sharks." *The Quarterly Journal of Economics* 108(2): 507-21.
- Greco, Melissa (2002). Gamer net game for more 'Foe,' 'Russian'. *Daily Variety*, August 12: 15.
- Kachelmeier, Steven J. and Mohamed Shehata (1994). Examining Risk Preferences under High Monetary Incentives: Reply. *American Economic Review* 84(4): 1105-06.
- Laury, Susan K. and Charles A. Holt (forthcoming). Voluntary Provision of Public Goods: Experimental Results with Interior Nash Equilibria. In: C.R. Plott and V. Smith (eds.), *Handbook of Experimental Economic Results*. North Holland, Amsterdam.
- Ledyard, John O. (1995). Public Goods: A Survey of Experimental Research. In: Kagel, John H. and Roth, Alvin E. (eds.), *The Handbook of Experimental Economics*. Princeton: Princeton University Press, 111-194.
- List, John A. (2003) "Friend or Foe: A Natural Experiment of the Prisoner's Dilemma." Mimeo, University of Maryland.
- Metrick, Andrew (1995). A Natural Experiment in "Jeopardy!" *American Economic Review* 85(1): 240-53.

Rabin, Matthew (1993). Incorporating Fairness into Game Theory and Economics. *American Economic Review* 83: 1281–1302.

Roth, Alvin E., Vesna Prasnikar, Masahiro Okuno-Fujiwara, and Shmuel Zamir (1991). Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study. *American Economic Review* 81: 1068-1095.

Slonim, Robert and Alvin E. Roth (1998). Learning in High Stakes Ultimatum Games: An Experiment in the Slovak Republic. *Econometrica* 66(3): 569-96.

TV Guide Online (2003). MTV's Kennedy: Game Show Foe. At <http://www.tvguide.com/newsgossip/insider/020603c.asp>, accessed April 17, 2003.

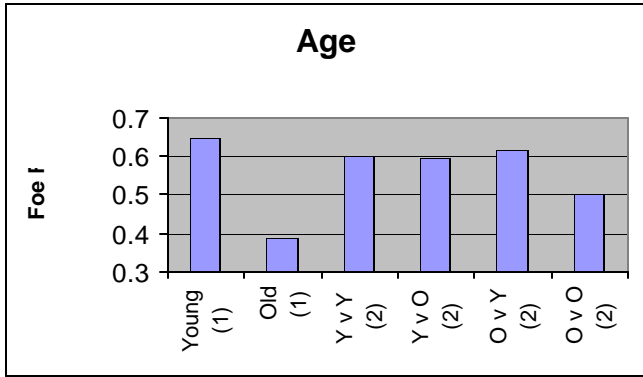


Figure 3a

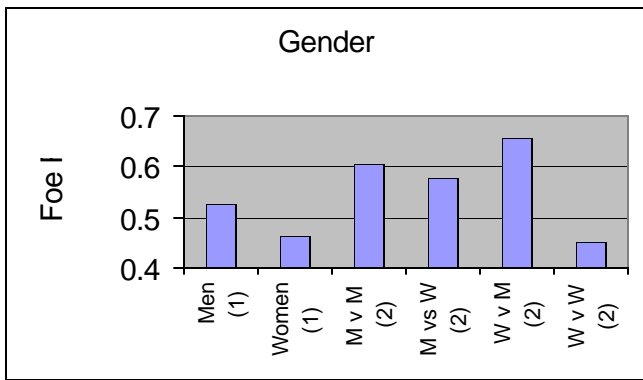


Figure 3b

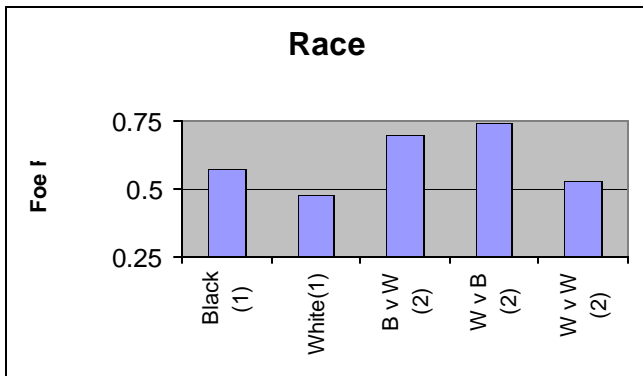


Figure 3c

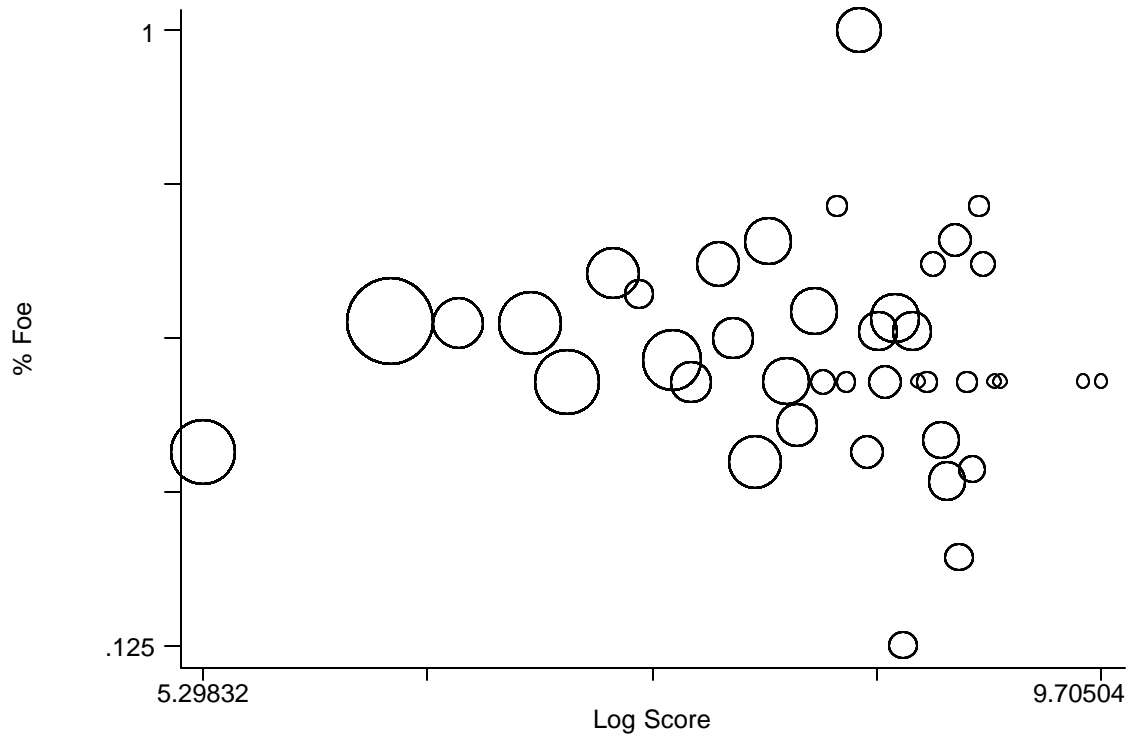


Figure 4: Stakes and the Individual “foe” Rate

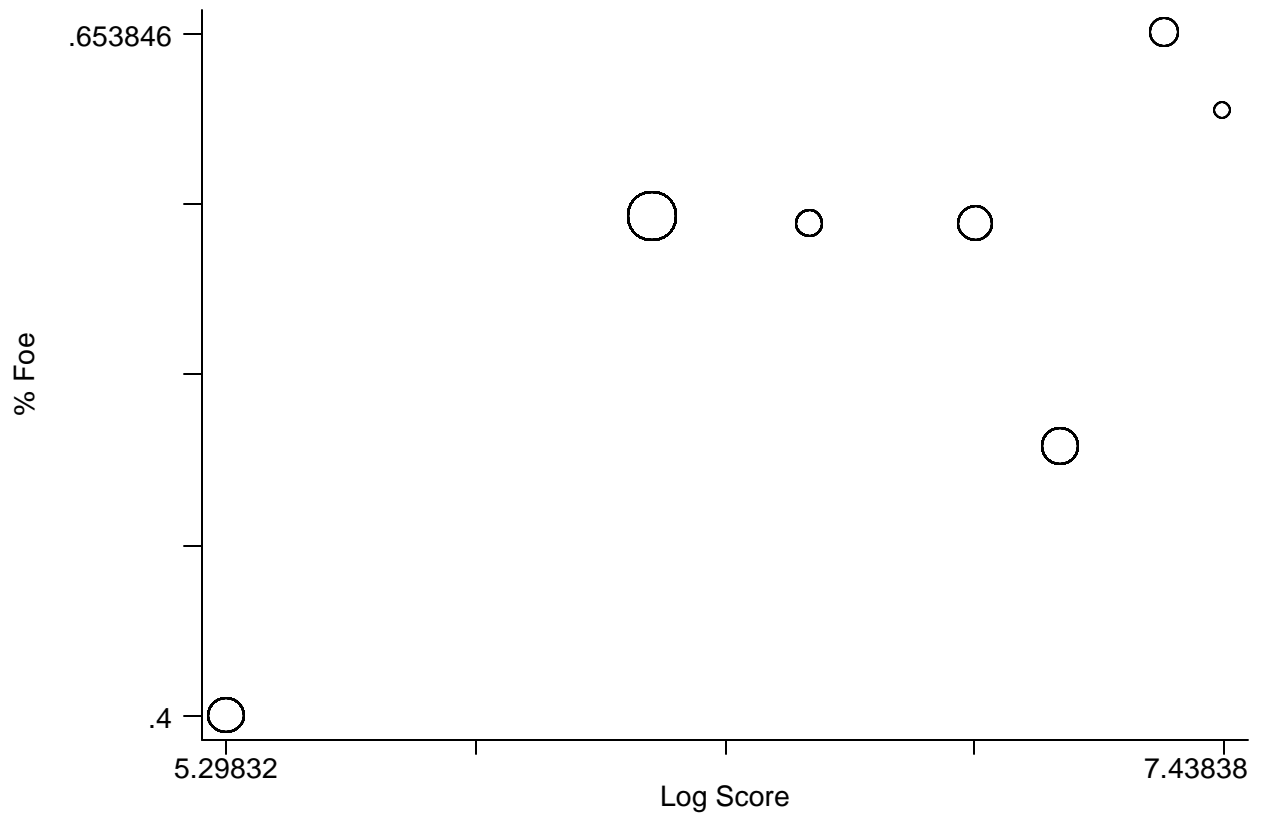
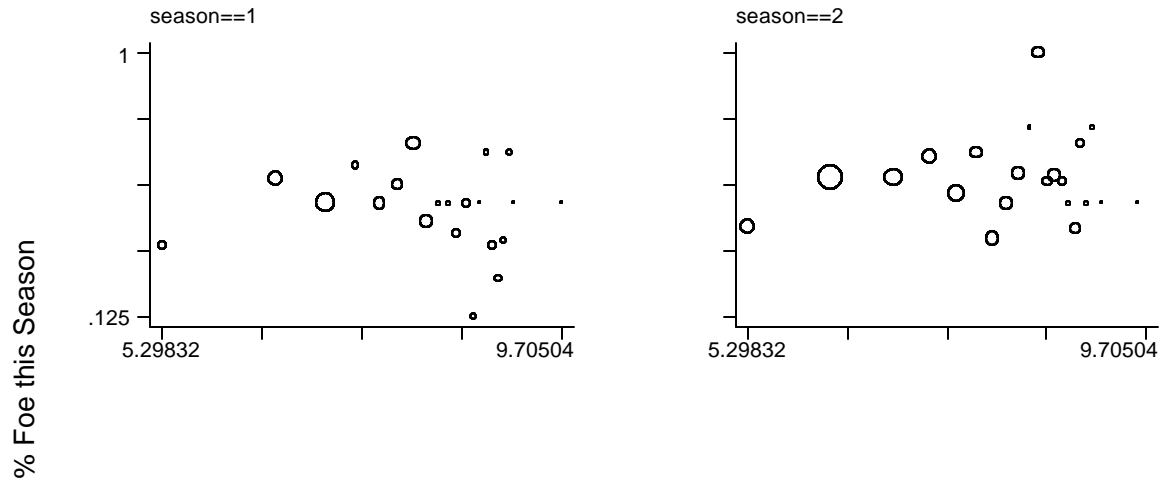


Figure 5: Stakes and the Individual “foe” Rate for Games Ending in Round 1



Log Score  
**Graphs by season**

Figure 6: Stakes and the Individual “foe” Rate by Season

**Table 1: Contestant Characteristics**

group	Number	percent
All	630	100
Male	314	50.1
Female	313	49.9
Foreign	38	6.3
North Central	103	17.2
Northeast	80	13.3
South	92	15.3
West	287	47.8
Black	95	15.8
White	470	78.3
Other	35	5.8
Mean age	29.1	
25 <sup>th</sup> pct	23	
median	27	
75 <sup>th</sup> pct	33	

**Table 2a: Individual Outcomes by Group and Season**

Group	season	Foe	N	Stakes	Winnings	Mean Winnings/
						Mean Stakes
All	Total	54.8%	630	3314.1	1130.6	34.12%
	1	49.2%	240	3718.3	1462.9	39.34%
	2	58.2%	390	3062.7	926.2	30.24%
Men	Total	56.5%	313	3637.8	1207.5	33.2%
	1	52.5%	118	4330.5	1847.5	42.7%
	2	59.0%	195	3216.5	820.3	25.5%
Women	Total	53.2%	314	2990.1	1054.9	35.3%
	1	46.3%	121	3100.8	1074.4	34.6%
	2	57.5%	193	2919.9	1042.7	35.7%
White	Total	54.3%	470	3501.1	1201.1	34.3%
	1	48.1%	183	3915.3	1388.5	35.5%
	2	58.2%	287	3236.9	1081.5	33.4%
Black	Total	65.3%	95	2840.0	856.8	30.2%
	1	57.5%	40	2912.5	1528.8	52.5%
	2	70.9%	55	2787.3	368.2	13.2%
Other Race	Total	48.6%	35	2620.0	931.4	35.6%
	1	45.5%	11	3563.6	2340.9	65.7%
	2	50.0%	24	2187.5	285.4	13.0%
Young (<=27)	Total	61.3%	310	3444.8	1205.0	35.0%
	1	64.6%	96	3686.5	1665.1	45.2%
	2	59.8%	214	3336.4	998.6	29.9%
Old (> 27)	Total	48.4%	320	3185.8	1058.6	33.2%
	1	38.9%	144	3739.6	1328.1	35.5%
	2	56.3%	176	2722.1	838.1	30.8%

**Table 2b: Individual Outcomes, by Own and Partner Characteristics**

	season	Foe	N	Stakes	Winnings	Mean Winnings/ Mean Stakes
Male vs Male	1	52.1%	48	5541.7	2291.7	41.4%
	2	60.7%	84	3576.2	1234.5	34.5%
	Total	57.6%	132	4290.9	1618.9	37.7%
Male vs Female	1	53.6%	69	3460.9	1520.3	43.9%
	2	57.7%	111	2941.8	506.8	17.2%
	Total	56.1%	180	3141.9	895.3	28.5%
Female vs Male	1	47.8%	69	3460.9	1023.2	29.6%
	2	65.8%	111	2941.8	963.5	32.8%
	Total	58.9%	180	3141.9	986.4	31.4%
Female vs Female	1	44.2%	52	2623.1	1142.3	43.5%
	2	45.0%	80	2920.0	1172.5	40.2%
	Total	44.7%	132	2803.0	1160.6	41.4%
White vs. White	1	51.5%	136	4225.0	1542.6	36.5%
	2	53.1%	226	3314.2	1115.0	33.6%
	Total	52.5%	362	3656.4	1275.7	34.9%
White vs. Black	1	45.9%	37	2929.7	833.8	28.5%
	2	74.5%	47	2859.6	786.2	27.5%
	Total	61.9%	84	2890.5	807.1	27.9%
Black vs White	1	56.8%	37	2929.7	1652.7	56.4%
	2	70.2%	47	2859.6	350.0	12.2%
	Total	64.3%	84	2890.5	923.8	32.0%
Young vs Young	1	65.6%	32	3512.5	940.6	26.8%
	2	60.0%	120	3671.7	1163.3	31.7%
	Total	61.2%	152	3638.2	1116.4	30.7%
Young vs Old	1	64.1%	64	3773.4	2027.3	53.7%
	2	59.6%	94	2908.5	788.3	27.1%
	Total	61.4%	158	3258.9	1290.2	39.6%
Old vs Young	1	37.5%	64	3773.4	1041.4	27.6%
	2	61.7%	94	2908.5	862.8	29.7%
	Total	51.9%	158	3258.9	935.1	28.7%
Old vs Old	1	40.0%	80	3712.5	1557.5	42.0%
	2	50.0%	82	2497.4	809.8	32.4%
	Total	45.1%	162	3112.7	1179.0	37.9%

**Table 3: Play and Own and Opponent Characteristics**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Season:	1	2	1	2	1	2	1	2
Constant	-0.0209 (0.0795)	0.1961 (0.0679)	1.3649 (0.7766)	-0.4491 (0.5906)	1.4897 (0.9031)	-0.8429 (0.6852)	2.0848 (0.9996)	-0.1590 (0.7786)
Log Score			-0.0494 (0.0842)	0.0555 (0.0647)	-0.0539 (0.0878)	0.0567 (0.0659)	-0.0656 (0.0880)	0.0587 (0.0659)
Age			-0.0404 (0.0104)	-0.0009 (0.0090)	-0.0406 (0.0104)	-0.0009 (0.0091)	-0.0453 (0.0122)	0.0014 (0.0106)
Black			0.3311 (0.2299)	0.3293 (0.1932)	0.3335 (0.2321)	0.4165 (0.1980)		
Male			0.2156 (0.1723)	0.0671 (0.1357)	0.2079 (0.1767)	0.1264 (0.1397)	0.1628 (0.2510)	-0.1635 (0.1987)
West			0.2277 (0.1752)	0.4505 (0.1372)	0.2309 (0.1750)	0.4481 (0.1394)	0.1945 (0.1742)	0.4592 (0.1400)
Opponent Age					0.0001 (0.0099)	0.0058 (0.0091)	-0.0056 (0.0114)	0.0083 (0.0106)
Opponent Black					-0.0001 (0.2314)	0.3968 (0.1999)		
Opponent Male					-0.0542 (0.1766)	0.2686 (0.1393)	-0.1147 (0.2517)	-0.0137 (0.1955)
Opponent West					-0.1286 (0.1747)	-0.0238 (0.1388)	-0.1244 (0.1741)	-0.0165 (0.1392)
Both female							-0.1108 (0.3441)	-0.5739 (0.2883)
Both Older							0.2140 (0.2443)	-0.1734 (0.2752)
Both white							-0.1298 (0.1826)	-0.4084 (0.1590)
$\rho$	-0.0528 (0.1421)	0.1872 (0.1117)	-0.0182 (0.1531)	0.1654 (0.1224)	-0.0168 (0.1532)	0.1538 (0.1226)	-0.0283 (0.1534)	0.1349 (0.1240)
H0: Own Characteristics don't matter (p-val)			20.20 (0.0005)	14.12 (0.0069)	20.19 (0.0005)	15.51 (0.0038)		
H0: Opponent characteristics don't matter (p-val)					0.64 (0.9586)	8.69 (0.0694)		
H0: Interaction of pair characteristics don't matter (p-val)							1.37 (0.7129)	11.10 (0.0112)
N (pairs)	120	195	117	183	117	183	117	183

Note: bivariate probit models of player pairs' "foe" decisions.

**Table 4: Pair Outcomes**

	Season	Coordination			Friend-Friend/Total		N
		Foe-Foe	Friend-Friend	Total	Split	Coord.	
All	1	23.9%	24.8%	48.7%	51.3%	50.9%	117
	2	36.4%	20.7%	57.1%	42.9%	36.3%	198
	Total	31.7%	22.2%	54.0%	46.0%	41.2%	315
2 Female	1	15.4%	26.9%	42.3%	57.7%	63.6%	26
	2	25.0%	35.0%	60.0%	40.0%	58.3%	40
	Total	21.2%	31.8%	53.0%	47.0%	60.0%	66
1 Male, 1 Female	1	28.4%	26.9%	55.2%	44.8%	48.6%	67
	2	41.6%	18.6%	60.2%	39.8%	30.9%	113
	Total	36.7%	21.7%	58.3%	41.7%	37.1%	180
2 Male	1	20.8%	16.7%	37.5%	62.5%	44.4%	24
	2	33.3%	11.9%	45.2%	54.8%	26.3%	42
	Total	28.8%	13.6%	42.4%	57.6%	32.1%	66
2 Young	1	50.0%	18.8%	68.8%	31.3%	27.3%	16
	2	40.0%	20.0%	60.0%	40.0%	33.3%	60
	Total	42.1%	19.7%	61.8%	38.2%	31.9%	76
1 Young, 1 Old	1	20.3%	18.8%	39.1%	60.9%	48.0%	64
	2	40.4%	19.1%	59.6%	40.4%	32.1%	94
	Total	32.3%	19.0%	51.3%	48.7%	37.0%	158
2 Old	1	18.9%	37.8%	56.8%	43.2%	66.7%	37
	2	27.6%	17.2%	44.8%	55.2%	38.5%	29
	Total	22.7%	28.8%	51.5%	48.5%	55.9%	66
2 White	1	25.6%	29.5%	55.1%	44.9%	53.5%	78
	2	31.9%	25.7%	57.6%	42.4%	44.6%	144
	Total	29.7%	27.0%	56.8%	43.2%	47.6%	222
1 White, 1 Black	1	20.5%	15.4%	35.9%	64.1%	42.9%	39
	2	48.1%	7.4%	55.6%	44.4%	13.3%	54
	Total	36.6%	10.8%	47.3%	52.7%	22.7%	93

**Table 5: Individual Foe Rates and Stakes**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	full	Season 1	Season 2	full	Season 1	Season 2	full	Season 1	Season 2	full	Season 1	Season 2
Log Score	0.0015 (0.0487)	-0.0335 (0.0748)	0.0426 (0.0637)	0.1172 (0.1310)	0.1780 (0.1727)	0.1717 (0.1880)	0.0049 (0.0500)	-0.0492 (0.0820)	0.0554 (0.0647)	0.1624 (0.1330)	0.2208 (0.1751)	0.2094 (0.1904)
Age							-0.0207 (0.0065)**	-0.0404 (0.0113)**	-0.0003 (0.0082)	-0.0184 (0.0102)	-0.0249 (0.0180)	-0.0130 (0.0123)
Black							0.3468 (0.1514)*	0.3312 (0.2438)	0.3698 (0.2012)	0.3112 (0.2455)	0.4306 (0.3961)	0.3105 (0.3248)
Male							0.1642 (0.0992)	0.2163 (0.1569)	0.0938 (0.1331)	0.1349 (0.1864)	-0.1400 (0.2949)	0.2395 (0.2459)
West							0.3474 (0.1052)**	0.2294 (0.1828)	0.4497 (0.1279)**	0.2832 (0.1815)	0.2604 (0.3025)	0.3405 (0.2279)
Constant	0.1313 (0.3738)	0.2510 (0.5862)	-0.0798 (0.4819)	-0.6251 (0.8306)	-1.1815 (1.1004)	-0.8682 (1.1713)	0.4099 (0.4416)	1.3610 (0.7386)	-0.4854 (0.5647)	-0.6215 (0.8886)	-0.8368 (1.1974)	-1.0636 (1.2427)
Observations	600	234	366	200	78	122	600	234	366	200	78	122

Notes: Robust standard errors in parentheses.\* significant at 5%; \*\* significant at 1%.

**Table 6: Pair Outcomes and Stakes**

	(1)	(2)	(3)	(4)	(5)	(6)
	Friend-Friend Season 1	Foe –Foe Season 1	Split Outcome Season 1	Friend-Friend Season 2	Foe –Foe Season 2	Split Outcome Season 2
Log Score	0.0301 (0.0408)	-0.0037 (0.0397)	-0.0265 (0.0459)	0.0033 (0.0266)	0.0320 (0.0331)	-0.0353 (0.0341)
Man and Woman	0.0305 (0.1017)	0.0992 (0.0989)	-0.1297 (0.1144)	-0.1823 (0.0736)*	0.1677 (0.0917)	0.0146 (0.0943)
2 Men	-0.1015 (0.1279)	0.0540 (0.1243)	0.0475 (0.1438)	-0.2228 (0.0888)*	0.0356 (0.1107)	0.1872 (0.1139)
Black and White	-0.1335 (0.0863)	-0.1033 (0.0839)	0.2369 (0.0970)*	-0.1639 (0.0634)*	0.1412 (0.0790)	0.0226 (0.0813)
2 Blacks	-0.2507 (0.4366)	0.7421 (0.4243)	-0.4914 (0.4908)	-0.1876 (0.3931)	0.6731 (0.4902)	-0.4855 (0.5040)
Old and Young	0.0067 (0.1218)	-0.3081 (0.1184)*	0.3014 (0.1369)*	0.0177 (0.0644)	-0.0084 (0.0803)	-0.0093 (0.0826)
2 Old	0.1824 (0.1302)	-0.3065 (0.1266)*	0.1242 (0.1464)	-0.0268 (0.0879)	-0.1226 (0.1096)	0.1494 (0.1127)
Constant	-0.0001 (0.3388)	0.4926 (0.3293)	0.5075 (0.3809)	0.3636 (0.2144)	0.0164 (0.2673)	0.6200 (0.2749)*
Observations	117	117	117	183	183	183
R-squared	0.08	0.11	0.13	0.08	0.06	0.04

Notes: Probit estimates, with standard errors in parentheses. \* significant at 5%; \*\* significant at 1%

**Table 7: Relative Contributions and the Foe Rate**

A. Questions

<u>Season 1</u>	Player 2 Makes Larger Net Contribution (Questions)		Player 1 Makes Larger Net Contribution (Questions)	
	Contribution	Foe rate	Contribution	Foe rate
Player 1	0.05	43.3%	2.46	60.9%
Player 2	2.45	50.7%	0.04	46.0%
N	67		50	
<u>Season 2</u>				
Player 1	-0.65	62.9%	1.94	68.7%
Player 2	1.14	48.3%	-0.96	64.2%
N	116		67	

B. Dollars

<u>Season 1</u>	Player 2 Makes Larger Net Contribution (Dollars)		Player 1 Makes Larger Net Contribution (Dollars)	
	Contribution	Foe rate	Contribution	Foe rate
Player 1	56.34	45.1%	1358.7	58.7%
Player 2	1563.4	54.9%	-250.0	39.1%
N	71		46	
<u>Season 2</u>				
Player 1	-509.1	62.7%	1068.5	68.5%
Player 2	781.8	48.2%	-787.7	63.0%
N				

**Table 8: Contributions and the Tendency to Play Foe, by Season**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Foe	Foe	Foe	Foe	Foe	Foe	Foe	Foe
	Season 1	Season 2	Season 1	Season 2	Season 1	Season 2	Season 1	Season 2
# Correct Answers	0.0466 (0.0232)*	-0.0170 (0.0192)						
# Incorrect Answers	-0.0085 (0.0328)	0.0149 (0.0245)						
# Net Correct Answers			0.0358 (0.0210)	-0.0163 (0.0163)				
\$ Correct Answers					0.06836 (0.03542)	-0.03635 (0.02984)		
\$ Incorrect Answers					0.00353 (0.05063)	0.02072 (0.003313)		
\$ Net Correct Answers							0.04781 (0.03179)	-0.02948 (0.002360)
Constant	0.3627 (0.0917)**	0.6064 (0.0736)**	0.4510 (0.0420)**	0.6011 (0.0247)**	0.3526 (0.0929)**	0.6248 (0.0710)**	0.4618 (0.0399)**	0.59970 (0.0243)**
Observations	234	366	234	366	234	366	234	366
Number of pair	117	183	117	183	117	183	117	183
R-squared	0.03	0.01	0.02	0.01	0.03	0.01	0.02	0.01

Notes: Dependent variable is foe dummy. All regressions include pair fixed effects. Standard errors in parentheses. \* significant at 5%; \*\* significant at 1%.