

Split or Steal? Cooperative Behavior When the Stakes Are Large

Martijn J. van den Assem

Department of Business Economics, Erasmus School of Economics, Erasmus University Rotterdam,
3062 PA Rotterdam, The Netherlands, vandenasse@ese.eur.nl

Dennie van Dolder

Department of Applied Economics, Erasmus School of Economics, Erasmus University Rotterdam,
3062 PA Rotterdam, The Netherlands, vandolder@ese.eur.nl

Richard H. Thaler

Booth School of Business, University of Chicago, Chicago, Illinois 60637, richard.thaler@chicagobooth.edu

We examine cooperative behavior when large sums of money are at stake, using data from the television game show *Golden Balls*. At the end of each episode, contestants play a variant on the classic prisoner's dilemma for large and widely ranging stakes averaging over \$20,000. Cooperation is surprisingly high for amounts that would normally be considered consequential but look tiny in their current context, what we call a "big peanuts" phenomenon. Utilizing the prior interaction among contestants, we find evidence that people have reciprocal preferences. Surprisingly, there is little support for conditional cooperation in our sample. That is, players do not seem to be more likely to cooperate if their opponent might be expected to cooperate. Further, we replicate earlier findings that males are less cooperative than females, but this gender effect reverses for older contestants because men become increasingly cooperative as their age increases.

Key words: natural experiment; game show; prisoner's dilemma; cooperation; cooperative behavior; social behavior; social preferences; reciprocity; reciprocal behavior; context effects; anchoring

History: Received July 14, 2010; accepted April 17, 2011, by Brad Barber, Teck Ho, and Terrance Odean, special issue editors. Published online in *Articles in Advance* October 7, 2011.

1. Introduction

Cooperation is vital for the functioning of society, and the organizations and communities that form its fabric. Not surprisingly, cooperative behavior is the focus of many studies across a wide range of scientific disciplines, including psychology (Dawes 1980, Dawes and Messick 2000), sociology (Marwell and Ames 1979, 1980; Raub and Snijders 1997), economics (Ledyard 1995, Fehr and Gächter 2000a, Fischbacher and Gächter 2010), political science (Ostrom et al. 1992), and biology (Gardner and West 2004, West et al. 2007). The key question in this literature is why humans cooperate even in situations in which doing so is not in line with their material self-interest.

Although cooperation is ubiquitous in social life and an important topic for all kinds of economic interaction, field data rarely allow for a clean discrimination among competing theories. Because carefully designed laboratory experiments do allow for such rigorous comparisons, laboratory experiments have provided numerous important insights into cooperative behavior, and the resulting rich literature forms the basis of most of our knowledge on human cooperation. Still, laboratory settings inevitably have

limitations that some argue may hinder the generalization of findings to situations beyond the context of the lab (Levitt and List 2007, 2008). Subjects are often volunteering students who thus constitute a nonrandom sample of the population at large. Also, they generally have less familiarity with decision tasks in the laboratory than with those in everyday life, no opportunity to seek advice from friends or experts, and they know that their behavior is examined in detail.

From an economic perspective, another obvious drawback to lab studies is that the financial stakes employed tend to be relatively small. Even those experiments that utilize relatively large payoffs do not involve amounts in excess of a few hundred dollars (e.g., Hoffman et al. 1996a, List and Cherry 2000, Carpenter et al. 2005), giving rise to the question to what extent findings will generalize to situations of significant economic importance. One solution is to perform experiments in low-income countries, where small nominal amounts carry a larger value. In the domain of social interaction, such experiments are, for example, employed by Slonim and Roth (1998), Cameron (1999), Fehr et al. (2002), Munier

and Zaharia (2002), Johansson-Stenman et al. (2005), and Kocher et al. (2008). Although this might appear an ideal approach, it has its own drawbacks. Culture, for example, has been shown to play an important role in social interaction (Henrich et al. 2001, 2004; Herrmann et al. 2008), making it difficult to generalize findings from low-income countries.¹ And although the stakes in these experiments are larger than commonly employed, they still rarely exceed a few months' wages.

In the current paper, we study cooperative behavior using another source of data, namely the behavior of contestants on the British television (TV) game show *Golden Balls*. Although the game show setting is an unusual environment, it has the benefit of employing large and varying stakes. Furthermore, game shows are markedly different from laboratory experiments in terms of participant selection, scrutiny, and familiarity of participants with the decision task. Combined with the strict and well-defined rules, game shows can therefore provide unique opportunities to investigate the robustness of existing laboratory findings.

Because game shows are often competitive in nature and ask contestants to make risky or strategic choices, is it not surprising that they have mostly been used to study decision making under risk (e.g., Gertner 1993, Metrick 1995, Beetsma and Schotman 2001, Post et al. 2008) or strategic reasoning (e.g., Bennett and Hickman 1993, Berk et al. 1996, Tenorio and Cason 2002). More recently, however, game shows have also been used to study social interaction, in particular discrimination (Levitt 2004, Antonovics et al. 2005) and cooperative behavior (List 2004, 2006; Belot et al. 2010; Oberholzer-Gee et al. 2010). The current paper is in the latter category.

In the final stage of *Golden Balls*, contestants make a choice on whether or not to cooperate in a variant of the famous prisoner's dilemma. In particular, the two final contestants independently have to decide whether they want to "split" or "steal" the jackpot. If both contestants choose split, they share the jackpot equally. If one chooses split and the other chooses steal, the one who steals takes the jackpot and the other gets nothing. If they both steal, both go home empty-handed. On average, the jackpot is over \$20,000. The variation is large: from a few dollars to about \$175,000.

If we assume that each player only cares about maximizing her immediate financial payoff, the choice problem in *Golden Balls* can be labeled as a

"weak" form of the prisoner's dilemma.² Where in the classic form of the prisoner's dilemma defecting strictly dominates cooperating, here defecting only weakly dominates cooperating: choosing steal always does at least as well, and sometimes better than choosing split. Of course, contestants may consider other factors aside from their own monetary payoff when deciding which choice to make. Much experimental research suggests that people have social preferences in the sense that the payoffs to others enter their utility functions. For discussions, see, for example, Fehr and Gächter (1998, 2000b), Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Charness and Rabin (2002), Camerer (2003), Fehr and Gintis (2007), and Cooper and Kagel (2009).

The fact that the show is aired on TV of course creates another set of rather special circumstances that could affect our results, although there is little existing theory to suggest what the effect of a large TV audience would be. One might argue that players would not want to be seen as a "jerk" on national television and so would be more likely to cooperate, but one can also argue that a player would not want to be seen as a "sucker" (or someone who cannot detect the weakly dominant solution to a simple game) in public.³ The public nature of the choice could also magnify subtle features created by the fact that the game is a weak form of the prisoner's dilemma. Specifically, if a player thinks that the other player will steal, she might decide to split on the grounds that it costs her nothing to appear "nice" on TV. These complications do not render our results uninteresting, but do need to be incorporated in any attempt to evaluate how our results should be interpreted in the context of existing theories and experimental findings on cooperation.

Although *Golden Balls* is unique in its format, the show shares the prisoner's dilemma element with a few game shows from other countries including *Friend or Foe* (United States, 2002–2003) and *Deelt ie 't of deelt ie 't niet?* (English translation: *Will He Share or Not?*; the Netherlands 2002). These two shows have been studied in four different papers. List (2004, 2006) and Oberholzer-Gee et al. (2010) analyze data from *Friend or Foe*, and Belot et al. (2010) use the Dutch show. List (2004, 2006) focuses on the effects of demographic variables such as gender, race, and age.

² Rapoport (1988) introduced this terminology. For the sake of brevity, we will simply use the term prisoner's dilemma to refer to the game studied here.

³ Most studies related to the issue of observability indicate that people display more other-regarding behavior when they are or feel more subject to public scrutiny (see, for example, Hoffman et al. 1996b, Rege and Telle 2004, Haley and Fessler 2005), but there is also contradictory evidence (Dufwenberg and Muren 2006). Kerr (1999) suggests that the effect depends on conditions related to social expectations and sanctions.

¹ Interestingly, though not generally acknowledged, this argument at the same time questions the universal applicability of the many findings from higher-income countries, including ours. We refer to Henrich et al. (2010) for a discussion on this issue.

Studying the same game show, Oberholzer-Gee et al. (2010) compare the behavior in the first season of the show with later seasons in which the contestants have had a chance to observe prior episodes. Finally, Belot et al. (2010) find that making a promise to cooperate prior to the decision is positively related to cooperation if the promise was voluntary, but not if it has been elicited by the host.

In this paper we replicate many of the earlier investigations but also undertake several novel analyses that are possible because of some unique features in the format of *Golden Balls*. The way the stakes are determined and the very wide range they cover provide the basis for new insights into the effect of stakes and context. The dynamic setting of the show enables us to look at reciprocity in cooperative behavior and also at the effect of earlier deceitful behavior.

In our sample, individual players on average cooperate 53% of the time. Although this rate is similar to earlier findings from the experimental literature (Dawes and Thaler 1988, Sally 1995), direct comparisons are hampered by systematic differences in the stakes, the visibility of decisions, characteristics of the subjects, and preceding opportunities for communication or other social interaction.

We find only limited support for the notion that cooperation will decrease if the stakes get significant. The cooperation rate is unusually high when the stakes lie in the low range of our sample, perhaps because contestants think that for so little money (relatively speaking) they might as well cooperate in public. Cooperation does decline with the stakes for stakes below the median but plateaus at around 45% for medium to large amounts.

The high cooperation rate for relatively small stakes suggests that context can convert money amounts that would normally be considered consequential or “big” into amounts that are perceived to be small, just “peanuts.” This idea is supported by our finding that cooperation is not only based on the actual stakes but also on what the jackpot potentially could have been. This effect is especially pronounced for those who appeared in the earlier episodes of the show and had no or little opportunity to watch the show on TV and learn what sizes are large or small in the context of this game.

A special property of *Golden Balls* is the interaction that occurs among contestants prior to the final. Utilizing the dynamic setting, we find evidence that contestants show some tendency toward reciprocity. Among contestants whose final opponent has attempted to vote them off the show, the propensity to cooperate is significantly lower. Contestants do not appear to reciprocate against opponents who have lied earlier in the game. Lying seems to be accepted here, similar to bluffing in poker. A possible reason

for this is that, in contrast to a vote cast against someone, lying is a defensive act that is not aimed at anyone in particular.

Surprisingly, we find little evidence that contestants’ propensity to cooperate depends positively on the likelihood that their opponent will cooperate. Although an opponent’s promise to cooperate is a strong predictor of her actual choice, contestants appear not to be more likely to cooperate if their opponent might be expected to cooperate. Our final main result is that young males cooperate less than young females. This difference decreases and even reverses as age increases and men become increasingly cooperative.

The paper proceeds as follows. Section 2 describes the game show in more detail, discusses our data, and presents descriptive statistics. Sections 3–7 cover the various possible factors behind cooperative behavior included in our analysis. Each of these sections provides related literature and other background, explains the variables that we use, and discusses the results of our probit regression analyses. Section 8 concludes.

2. Game Show and Data

2.1. Description of *Golden Balls*

The TV game show *Golden Balls* was developed by the Dutch production company Endemol. Its debut was on the ITV network in the United Kingdom in June 2007 and the show ran until December 2009. Each episode consists of four rounds and starts with four contestants, usually two men and two women.

In round 1, twelve golden balls are randomly drawn from the “golden bank,” a lottery machine containing one hundred “golden” balls. Each of these balls has a hidden cash amount inside, ranging from a minimum of £10 to a maximum of £75,000.⁴ Contestants know that this is the range for the amounts in the balls, but they do not know the precise distribution (though this becomes clearer over time as the show is aired). At a later stage of the game, a subset of the cash balls drawn will contribute to the final jackpot. Also, four balls hiding the word “killer” inside are mixed with the twelve cash balls. Killer balls are undesirable in a way we will explain below. From the 16 balls, each contestant receives four balls at random. For each contestant, two are placed on the front row with their contents—either a cash amount or the word killer—openly displayed; the other two are placed on the back row and their contents are known by the particular contestant alone. (Poker players can think of

⁴ Values in British pounds can be translated into U.S. dollars using a rate of \$1.75 per pound, an approximate average of the exchange rate during the period in which the show ran.

this as two “up” cards and two “down” cards.) The contestants now have to decide by vote which player will be kicked off the show. Because the balls of voted-off contestants are removed from the game and the remaining contestants’ balls matter for the ultimate jackpot, there is a strong incentive to retain players with high value balls and kick off players with low value balls or killer balls.

Before the voting starts, each contestant publicly announces the contents of the balls on her back row (knowing that these values will subsequently be revealed, but only after the vote). Then, the four contestants together have an open discussion in which they can voice their evaluation of other players’ statements and their opinion of who should be voted off. Each player then anonymously casts a vote against one specific opponent. After the votes are tallied, the player who received the most votes leaves the game.⁵ Lastly, all the players reveal the values of their back row balls, and differences between the actual values and the previous claims are noted.

In round 2, two additional cash balls from the lottery machine and one extra killer ball are added to the twelve remaining balls from round 1. The 15 balls are then randomly allocated to the three contestants. Each of them receives two balls on her front row and three on her (hidden) back row. Similar to round 1, contestants make (cheap talk) statements on the balls on their back row, a round of banter follows, votes are cast anonymously and tallied, the player who receives the most votes leaves the game, and all hidden ball values are revealed.⁶ Two players and their 10 balls proceed to round 3.

Round 3 determines the size of the final jackpot. First, one additional killer ball is mixed with the 10 balls from round 2. Then five of the balls are selected sequentially at random. If a ball selected is a cash ball, its face value is added to the jackpot. If a killer ball is drawn, the current cumulative jackpot is divided by ten. For example, if the first two balls were £50,000 and £1,000 and the third is a killer ball, the level of the jackpot is reduced from £51,000 to £5,100. A killer ball does not affect the jackpot contribution of cash balls drawn thereafter. If the fourth and fifth ball in our example were another killer ball and £25,000, respectively, then the actual jackpot would be £25,510.

⁵ If two contestants receive two votes each, their opponents openly discuss who they want to keep in the show. If they cannot decide, a decision is made at random. If all four contestants receive one vote each, contestants openly attempt to form a coalition against one specific contestant. Again, if they cannot decide, a decision is made at random.

⁶ The procedure in the case of a tie is similar to that in round 1. Tie-breaking occurs by discussion or by random draw if no agreement is reached.

Note that this round is a completely stochastic process, and that contestants have full information on the balls that are in play. Before the five balls are drawn, special attention is always paid to the highest possible jackpot (that is, the sum of the five largest cash values). This value and the number of killer balls are explicitly stressed by the game show host.

After round 3 determines the jackpot, in the fourth and final round the contestants play a variant of the prisoner’s dilemma. Each contestant receives two golden balls. One of the balls says “split” and the other says “steal” on the inside. The contestants then simultaneously have to decide which ball they want to play. If both choose split, they share the jackpot equally. If one chooses split and the other chooses steal, the contestant who steals takes the whole jackpot and the other gets nothing. If they both choose steal both go home empty-handed. Before each contestant makes her actual decision, a brief time period is reserved for a discussion between the players in which they can make nonbinding promises, ask about intentions, or attempt to get assurances of cooperative behavior. This is the final round of cheap talk. Importantly, the contestants have not met before the game starts and have no opportunity before or during the show to make any kind of collusive agreement.

A relevant question is how the contestants are selected. A spokeswoman of Endemol informed us that anyone can apply to be on *Golden Balls* by submitting a detailed application form. Shortlisted contestants are then invited to an audition in order to determine their skills at playing the game, their character, and their suitability to appear on a TV show such as *Golden Balls*. Producers watch tapings of these auditions and put together shows such that, according to the producers, “a good mix of characters” is represented on each show. Thus, although the contestants are not a random sample of society, the selection process does not seem to create any obvious confounds with the analyses we conduct here.

2.2. Data and Descriptive Game Characteristics

We examine the split and steal decisions of 574 final contestants appearing in 287 episodes aired between June 2007 (when the show was introduced) and December 2009. During this period, 288 episodes were aired, and, at the time of writing, no further episodes were aired thereafter. Recordings from the show and additional information such as recording and airing dates were kindly provided by Endemol’s local production company, Endemol UK. The one missing episode could not be supplied because it was not present in their archives.⁷

⁷ Sixteen episodes in our data set feature returning contestants. In 12 of these, players who previously had lost in the final (opponents

Table 1 Selected Game Show Characteristics

	<i>N</i>	Mean	Std. dev.	Min	Median	Max
<i>Cash ball (overall)</i>	4,018	5,653.92	10,478.49	10.00	1,500.00	75,000.00
<i>Cash ball (round 3)</i>	2,257	6,775.15	12,204.39	10.00	1,600.00	75,000.00
<i>No. of killer balls (round 3)</i>	287	3.14	0.90	1.00	3.00	6.00
<i>Potential jackpot (round 3)</i>	287	51,493.08	31,386.69	5,000.00	41,150.00	168,100.00
<i>Jackpot</i>	287	13,416.09	19,182.98	2.85	4,300.00	100,150.00
<i>Decision (split = 1)</i>	574	0.53	0.50	0.00	1.00	1.00
<i>Prize won</i>	574	4,850.55	11,821.06	0.00	38.75	100,150.00
<i>Prize won if nonzero</i>	303	9,188.82	15,004.52	1.83	2,175.00	100,150.00

Notes. This table shows selected characteristics for the British TV game show *Golden Balls*, extracted from our sample of 287 episodes. *Cash ball (overall)* is the monetary value of a cash ball drawn from the lottery machine in the first or second round of the game. *Cash ball (round 3)* is the monetary value of a cash ball that is in play at the start of the third round. *No. of killer balls (round 3)* describes the number of killer balls that are in play at the start of the third round. *Potential jackpot (round 3)* is the jackpot size that is attained during the third round if the best-case scenario would occur. *Jackpot* describes the actual size of the jackpot. *Decision* is a contestant's decision in the prisoner's dilemma at the end of the show, with a value of 1 for split and 0 for steal. *Prize won (if nonzero)* records the take home prize for a contestant who made it to the final (if she did not leave empty-handed). All monetary values are in UK pounds (£1.00 ≈ \$1.75).

For each episode we collected data on the relevant observables in the show, such as the hidden and visible ball values, statements made by contestants, the votes, the jackpot size, and the decision to split or steal at the end. Some variables were estimated based on contestants' physical appearance and on information provided in the introductory talk and other conversations during the show.

Table 1 displays some descriptive characteristics of the game. Cash balls drawn from the lottery machine during the first two rounds have a mean value of £5,654 and a median of £1,500. Clearly, the distribution is positively skewed. The mean value of the cash balls taken to round 3 is £6,775, which is statistically significantly greater than the average value of all cash balls in the show, implying that the contestants are successful in using their votes to keep high-value balls in play and eliminate small ones. The average number of killer balls in the game at the start of round 3 is 3.14, significantly less than the 3.67 we would statistically expect if voting was random. Contestants thus also seem successful in eliminating killer balls from the game. Unreported analyses of contestants' voting behavior clearly show that contestants indeed try to vote off the opponents that have the worst set of balls on their front row.

At the start of round 3, special attention is paid to the highest possible jackpot. Dependent on the cash balls and killer balls taken to this stage, this maximum varies between £5,000 and £168,100, with a mean of £51,493 and a median of £41,150. The actual jackpot for which contestants play the prisoner's dilemma

game is generally considerably smaller because of the skewed distribution of cash ball values and the effect of killer balls, but still has a mean size of £13,416 and a median of £4,300. These amounts are many times the amounts typically used in laboratory experiments and also large sums relative to the median gross weekly earnings of £397 in the United Kingdom in April 2009 (Office for National Statistics 2009). About half of the time, the jackpot in our show exceeds three months of median UK earnings, and 21% of the contestants decide over a jackpot that is even larger than a median annual salary (the third quartile in our sample is at £18,350). The stakes are also large compared to the two other game shows employed in earlier analyses of cooperative behavior: in *Friend or Foe*, the average is about \$3,500 (List 2004, 2006; Oberholzer-Gee et al. 2010); for the Dutch show, Belot et al. (2010) report a median of €1,683. The wide range of the jackpot in our sample is caused by its random construction, by the highly skewed distribution of cash ball values, and by the effect of killer balls. The largest jackpot was played for in an exhilarating episode from March 2008; trainee accountant Sarah stole the entire jackpot of £100,150 from collection agent Stephen.⁸

For the jackpot to be awarded, at least one player needs to cooperate. We find that 52.8% of the contestants decide to split. Although this might seem high, the rate is actually remarkably similar to earlier experimental evidence (see, for example, Sally 1995). In our sample, both players split the jackpot 31% of the time, one splits while the other one steals occurs in 44% of the shows, and in the remaining 25% of the shows both players steal. The efficiency rate in terms of the percentage of jackpots that is actually awarded

stole while they themselves chose to split) get a second chance. In four episodes, unlucky players who had been voted off in the first game round receive a second chance. We do not find that returning contestants behave differently, and, unless stated otherwise hereafter, excluding them from our analyses does not materially affect our results.

⁸ A video clip of this episode is widely available on the Internet, for example, through YouTube.

thus amounts to 75%. The efficiency rate obtained by dividing the sum of earnings across all episodes by the sum of all jackpots is slightly lower at 72%. (The difference in efficiency results from contestants' lower propensity to cooperate when the stakes are larger; we explore this effect in detail later.) These simple statistics are a first indication that contestants do not condition their behavior on that of their opponent. Given the average cooperation rate, we would expect to observe (split, split) in 28% of the cases and (steal, steal) 22% of the time if the individual decisions in our sample were randomly matched. Although the actual percentages are higher (31 and 25), the differences are relatively small considering that each pair of contestants operates under highly similar conditions (same jackpot, same potential jackpot, and many shared unobserved conditions).

On average, a finalist goes home with £4,851, but the median prize is only £39 because 47% of the contestants get nothing. The 303 contestants who end up with a nonzero prize take home £9,189 on average, with a median of £2,175. It is worth noting that would we have run this show as an experiment ourselves, the total costs in subject payoffs alone would have been £2.8 million.

2.3. Modeling the Decision to Split or Steal

In the following sections, we will analyze the decisions to split or steal the jackpot using a binary probit model. We assume that when people enter the final round they have a latent propensity to split y^* , where $y^* \in (-\infty, \infty)$. Furthermore, we assume that this latent propensity is a linear function of personal demographic characteristics x and context characteristics z , in the form $y^* = x'\beta + z'\gamma + u$, where β and γ are parameter vectors and u represents an unobserved stochastic component. We do not observe the latent propensity to split directly, but only the actual decision y , where $y = 1$ if a contestant chooses split and $y = 0$ if a contestant chooses steal. We impose the observation criterion $y = \mathbf{1}(y^* > 0)$, where $\mathbf{1}(\cdot)$ is the indicator function taking the value of 1 if $y^* > 0$ and 0 otherwise. Assuming that the stochastic component has a standard normal distribution, or $u \sim N(0, 1)$, leads to the binary probit model of the form $\Pr(y = 1 | x, z) = \Phi(x'\beta + z'\gamma)$, where $\Phi(\cdot)$ is the standard normal cumulative distribution function. Using this framework we estimate the parameter vectors β and γ using maximum likelihood estimation. We allow for the possibility that the decisions of contestants within the same episode are correlated by performing a clustering correction on the standard errors (see, for example, Wooldridge 2003).

Because coefficients in a probit model do not have an immediate intuitive economic meaning due to the inherent nonlinearities, we will follow the common

approach of reporting marginal effects instead. More specifically, the marginal effects we report apply to the medians of the explanatory variables, with two exceptions. To highlight some of the interaction effects that we find in our data, we set *Age* to be 20 and *Transmissions* (the number of times the show has aired at the time of recording) to be 0. The resulting "representative agent" is a 20-year-old white female without higher education, who lives in a relatively small town and plays the final of our game for a jackpot of £4,300, which potentially could have been £41,150. For dummy variables we consider the effect of a discrete change from 0 to 1. As noted by Ai and Norton (2003), the traditional way of calculating marginal effects and their standard errors is not valid for interaction terms, and we therefore apply the alternative method they propose. For the sake of consistency, we report significance levels that apply to the marginal effects, though these levels do not differ materially from the significance levels for the original regression coefficients. Original coefficients and their significance levels are available from the authors upon request.

3. Demographic Characteristics

First, we investigate how various demographic characteristics are related to the propensity to cooperate. Our later analyses include these demographic variables as control variables.

In previous studies examining the relations between demographic characteristics and cooperative behavior, most attention has been directed to gender. Psychologists have a long history when it comes to investigating the relation between gender and behavior, and, over the past decade, economists have become increasingly interested in gender effects as well. The standard finding is that women act more prosocially than males, but the reverse is also found.⁹ For contextual settings similar to ours, List (2004, 2006), Oberholzer-Gee et al. (2010), and Belot et al. (2010) report that women are more cooperative than men, although some results are only marginally significant.

For other demographic characteristics, the experimental findings are also mixed. Carpenter et al. (2004), for example, run public good experiments with symbolic but costly punishment in Bangkok and Ho Chi Minh City. They find that in Bangkok males and higher educated subjects contribute more, however there is no significant age effect. The same experiment

⁹One possible cause for the varying results is that males and females respond differently to specific contextual settings of the experiments (Croson and Gneezy 2009). For example, if women are more risk averse than men, this may lead to different social behavior in situations in which risk is involved (Eckel and Grossman 2008).

in Ho Chi Minh City, however, shows the opposite findings: males and higher educated subjects cooperate less and age increases cooperation. Gächter et al. (2004) find no influence of background characteristics in a one-shot public good experiment with Russian subjects.

To add to this literature we will explore the effect of various demographic characteristics on cooperative behavior. We employ the following set of variables:

- *Gender* is a dummy variable indicating whether a contestant is male (1) or female (0).

- *Age* is a continuous variable measuring the contestant's age in years. In many instances the contestant's age is not explicitly mentioned during the show. In these cases we estimate age on the basis of physical appearance and other helpful information such as the age of children.

- *Race* is a dummy variable indicating whether a contestant is white (1) or nonwhite (0). We apply such a broad distinction because the large majority of contestants are white.

- *City* and *London* are two dummy variables that are constructed to distinguish contestants that live in major urban areas from those that reside in more rural surroundings. Contestants' city or county of residence is always an integral part of the introductory talk. *City* indicates whether a contestant lives in a large urban area (1) or not (0). We define a large urban area as a conurbation with a population exceeding 250,000 inhabitants.¹⁰ For some contestants we only know their region and not their exact town or city; we then assume a small domicile. *London* indicates whether a contestant lives inside (1) or outside (0) the Greater London Urban Area.

- *Education* is a dummy variable for the level of education and differentiates between those with at least a bachelor degree (1) and those without (0). Players generally do not talk about their education during the show. We therefore estimate a contestant's level of education on the basis of her occupation, which is always explicitly mentioned when she is introduced, and on the basis of other information given in talks. Contestants who are currently enrolled in higher education and people whose job title suggests work experience equivalent to the bachelor level or higher are included in the higher education category. From the information that we have about each contestant, the proper binary values are generally clear.

- *Student* is a dummy variable indicating whether the contestant currently is a higher education (undergraduate or postgraduate) student (1) or not (0).

Estimates for *Age* and *Education* are based on the independent judgments of three research assistants, where each value is based on the assessments of two of them. When the estimates for *Education* were different, we decided on the most appropriate value ourselves. For *Age* we took the mean of the two judgments, and included our own assessment as a third input if the values of the coders diverged more than five years.

We have also attempted to collect data on contestants' marital status and the existence of children. These topics were, however, not systematically discussed in the program and values would therefore be unknown for the large majority of our contestants. Table 2 summarizes all the variables that are included in our analyses, including the demographic characteristics.

Model 1 in Table 3 shows the regression results for a model that includes demographic characteristics only. To be able to distinguish both general gender and age effects as well as a possible interaction effect, the interaction of gender and age is also included. The results show that, relative to our representative 20-year-old female agent, young males are 22 percentage points less likely to cooperate ($p = 0.002$). In line with past results by List (2004) and Carpenter et al. (2008), this difference disappears when age increases. The effect of age is significantly different for males and females ($p = 0.001$). Women do not become significantly more or less cooperative when age increases ($p = 0.422$). Men, on the other hand, do have a higher propensity to cooperate as they are older: their cooperation rate increases by more than one percentage point per year ($p = 0.000$; untabulated).¹¹ Contrary to the two previous studies, we find that the gender difference not only disappears as age increases, but actually reverses; males become significantly more likely to split from age 46 onward. Figure 1 displays observed cooperation rates at different age levels for both males, females, and aggregates, clearly depicting an age effect for men. Further analyses show that there is no evidence of a quadratic age effect, neither for men nor for women. We have also experimented with specifications where the (semi-) continuous age variable is replaced by a set of dummy variables that represent various age groups. The results are economically and statistically similar.

¹⁰ For England and Wales, the population data and the definitions of conurbations are taken from the UK Office for National Statistics (<http://www.statistics.gov.uk>). Similar information for Scotland, Northern Ireland, and Ireland is from the General Register Office for Scotland (<http://www.gro-scotland.gov.uk>), the Northern Ireland Statistics and Research Agency (<http://www.nisra.gov.uk>), and the Central Statistics Office Ireland (<http://www.cso.ie>), respectively.

¹¹ The effect of age for males could be related to increasing dependence on others (van Lange et al. 1997), or to hormonal or neurological changes as men grow older, but we are hesitant to draw conclusions in these directions for we cannot exclude that a generational effect (van Lange et al. 1997, List 2004) or a wealth effect is (partly) driving our finding.

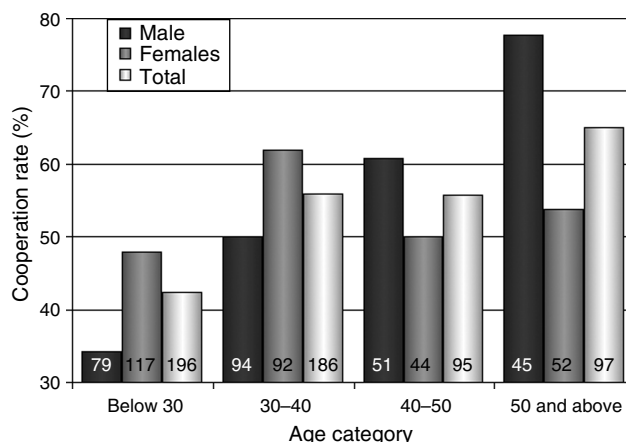
Table 2 Summary Statistics

	Mean	Std. dev.	Min	Median	Max
Demographic characteristics					
Age	36.78	11.76	18.00	34.40	73.00
Gender (male = 1)	0.47	0.50	0.00	0.00	1.00
Race (white = 1)	0.92	0.27	0.00	1.00	1.00
City (large = 1)	0.47	0.50	0.00	0.00	1.00
London (London = 1)	0.14	0.35	0.00	0.00	1.00
Education (high = 1)	0.34	0.47	0.00	0.00	1.00
Student (student = 1)	0.09	0.28	0.00	0.00	1.00
Stakes and context					
Actual stakes (log)	8.19	2.08	1.05	8.37	11.51
Potential stakes (log)	10.68	0.60	8.52	10.62	12.03
Transmissions	111.68	74.18	0.00	109.00	214.00
Reciprocal preferences					
Vote received from opp. (yes = 1)	0.05	0.22	0.00	0.00	1.00
Expectational conditional cooperation					
Promise (promise = 1)	0.53	0.50	0.00	1.00	1.00
Past deceitful behavior					
Lie round 1 (lie = 1)	0.41	0.49	0.00	0.00	1.00
Lie round 2 (lie = 1)	0.36	0.48	0.00	0.00	1.00
Lie cash ball round 1 (lie = 1)	0.24	0.42	0.00	0.00	1.00
Lie cash ball round 2 (lie = 1)	0.15	0.36	0.00	0.00	1.00
Lie killer ball round 1 (lie = 1)	0.21	0.41	0.00	0.00	1.00
Lie killer ball round 2 (lie = 1)	0.24	0.43	0.00	0.00	1.00

Notes. This table shows descriptive statistics for the explanatory variables in our analyses of cooperative behavior based on the decisions of 574 contestants to either split or steal the jackpot in the prisoner's dilemma at the end of the British TV game show *Golden Balls*. Age is the contestant's age measured in years. Gender, Race, City, London, Education, and Student are dummy variables taking the value of 1 if the contestant is male (Gender), is white (Race), lives in a conurbation with a population exceeding 250,000 inhabitants (City), is a resident of the Greater London Urban Area (London), has completed or is enrolled in higher education (bachelor degree or higher) or has equivalent working experience (Education), or is a student (Student), respectively. Actual stakes is the natural logarithm of the size of the jackpot in the prisoner's dilemma game. Potential stakes is the natural logarithm of the highest possible jackpot at the start of the third round. Transmissions expresses the number of episodes that was already aired when the current episode was recorded in the studio. Vote received from opp. is a dummy variable taking the value of 1 if the contestant's final opponent has tried to vote her off the program at an earlier stage of the game. Promise is a dummy variable taking the value of 1 if the contestant explicitly promised her opponent to split (or not to steal) the jackpot. Lie round 1 (round 2) is a dummy variable taking the value of 1 if the contestant has misrepresented her back row balls—either by overstating a cash ball or by hiding a killer ball—in the first (second) round. Lie cash (killer) ball round 1 (round 2) is a dummy variable taking the value of 1 if the contestant has overstated a cash ball (hidden a killer ball) in the first (second) round. Standard deviations for the two stakes variables and the transmissions variable are calculated across episodes ($N = 287$) to avoid the effect of clusters at the episode level. All monetary values are in UK pounds (£1.00 ≈ \$1.75).

When it comes to race we find weak evidence that whites are more likely (about 13 percentage points) to cooperate than nonwhites ($p = 0.101$; $p < 0.10$ in the models discussed hereafter). List (2004, 2006) and Oberholzer-Gee et al. (2010) report a similar pattern, yet in more conventional experiments the reverse is often found (see, for example, Cox et al. 1991). Because possible but unobservable wealth

Figure 1 Age and the Propensity to Cooperate for Males and Females



Notes. This figure displays the relative frequency of contestants who decide to split across various age intervals. Bars depict the percentage of cooperators within specific age brackets for males, females, and the aggregate, respectively. For each category, the number of contestants is displayed at the bottom of the bar.

effects could contribute to this result, it should be interpreted with caution.

Higher educated contestants are about nine percentage points more cooperative ($p = 0.062$), although this effect is only marginally significant in the current model and not consistently significant across the various regression models discussed hereafter ($0.041 < p < 0.070$). Similar to the effect of race, the effect of education could be spurious because of an unobservable wealth effect.

Students are frequently used as subjects in experiments, and the reliance on such a specific subject pool is often criticized. Sears (1986), for example, extensively describes how the use of student subjects might produce misleading or mistaken conclusions about social behavior. It is therefore interesting to investigate whether there is evidence that students behave differently from others, holding other observable characteristics constant. This turns out not to be the case. Controlling for demographics such as age and education, our regression results yield no indications of a different attitude toward cooperation among students ($p = 0.888$).¹²

None of the residence dummy variables have a significant effect. Possibly, relatively small social differences between urban and more rural areas in the United Kingdom explain this null result.¹³

¹² In a similar vein, van Lange et al. (1997) and Bellemare and Kröger (2007) do not detect a difference, whereas Carpenter et al. (2008) and Egas and Riedl (2008) do report a negative bias.

¹³ In experiments conducted in a region of Russia where there is a large gap, Gächter and Herrmann (2011) do find that rural residents are more cooperative than urban residents.

Table 3 Binary Probit Regression Results [1/2]

	Model 1	Model 2	Model 3	Model 4	Model 5
Demographic characteristics					
<i>Age</i>	0.002 (0.422)	0.002 (0.345)	0.003 (0.311)	0.002 (0.352)	0.002 (0.403)
<i>Gender</i> (male = 1)	−0.221 (0.002)	−0.241 (0.000)	−0.233 (0.001)	−0.236 (0.001)	−0.233 (0.001)
<i>Race</i> (white = 1)	0.134 (0.101)	0.143 (0.082)	0.139 (0.091)	0.142 (0.089)	0.149 (0.065)
<i>City</i> (large = 1)	−0.039 (0.396)	−0.043 (0.359)	−0.045 (0.335)	−0.039 (0.405)	−0.036 (0.439)
<i>London</i> (London = 1)	0.059 (0.402)	0.066 (0.348)	0.059 (0.400)	0.054 (0.444)	0.058 (0.406)
<i>Education</i> (high = 1)	0.088 (0.062)	0.093 (0.053)	0.094 (0.050)	0.094 (0.049)	0.091 (0.058)
<i>Student</i> (student = 1)	0.012 (0.888)	−0.002 (0.983)	−0.008 (0.923)	−0.013 (0.884)	−0.007 (0.933)
<i>Age</i> × <i>Gender</i>	0.011 (0.001)	0.010 (0.001)	0.010 (0.001)	0.011 (0.001)	0.010 (0.001)
Stakes and context					
<i>Actual stakes</i> (log)		−0.043 (0.000)	−0.048 (0.000)	−0.049 (0.000)	−0.048 (0.000)
<i>Potential stakes</i> (log)			0.057 (0.139)	0.174 (0.006)	
<i>Transmissions</i>				−0.000 (0.722)	
<i>Potential stakes</i> × <i>Transmissions</i>				−0.001 (0.037)	
<i>Second half</i> (second = 1)					0.030 (0.508)
<i>Potential stakes</i> × <i>First half</i>					0.139 (0.005)
<i>Potential stakes</i> × <i>Second half</i>					−0.035 (0.543)
Wald χ^2 (df)	34.87 (8)	51.61 (9)	52.57 (10)	57.87 (12)	62.35 (12)
Log pseudo-likelihood	−379.78	−371.55	−370.42	−368.29	−367.51
Pseudo R^2	0.043	0.064	0.067	0.072	0.074
<i>N</i>	574	574	574	574	574
Number of clusters	287	287	287	287	287

Notes. This table displays results from the probit regression analyses of contestants' decisions to split (1) or steal (0) the jackpot in the prisoner's dilemma at the end of the British TV game show *Golden Balls*. *First (Second) half* is a dummy variable taking the value of 1 if less (more) than 50% of the episodes in our sample were already aired when the current episode was recorded in the studio. Definitions of other variables are as in Table 2. For each explanatory variable, the marginal effect is shown for a representative agent who takes the median value on all variables, except for *Age* and *Transmissions*, which are set to 20 and 0, respectively. Standard errors are corrected for clustering at the episode level; *p*-values are in parentheses.

4. Stakes and Context

Economists typically argue that behavior will converge toward the prediction of rational self-interest if the stakes increase (e.g., Rabin 1993, Telser 1995, Levitt and List 2007). The evidence from lab and field experiments is, however, not generally supportive of this view. Except for the finding that people seem to become more willing to accept relatively low offers in ultimatum bargaining games when the stakes are high, empirical research generally finds no evidence that stake size affects behavior, even when the stakes are increased up to several months' wages.¹⁴

Given that the stakes in *Golden Balls* are widely ranging and, on average, considerably larger than in previous studies, the show provides an excellent opportunity to reexamine the relation between cooperation and stakes. In addition, compared to earlier game show studies on cooperation, an advantage of *Golden Balls* is that the stakes are mainly built up by a random process and not by contestants' answers

to trivia questions. The latter may lead to a spurious correlation because the ability to answer trivia questions may be related to unobserved background characteristics such as income, which in turn may well be related to the propensity to cooperate.

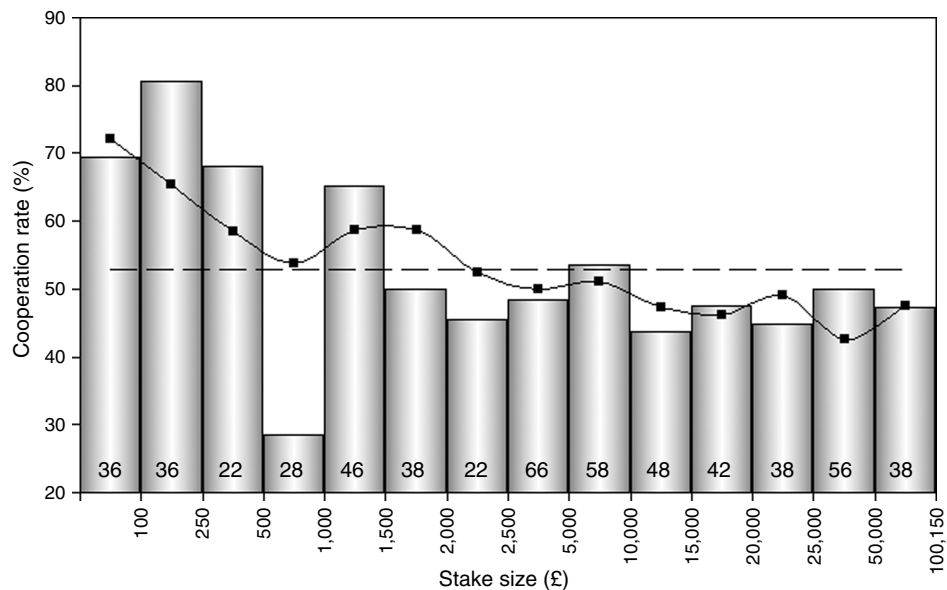
The variable that we use in our regressions is labeled *Actual stakes* and defined as the natural logarithm of the size of the jackpot.

Model 2 in Table 3 displays the regression results when the stakes are included. Clearly, cooperative behavior in our show is sensitive to the amount that is at stake. To illustrate this effect, Figure 2 depicts the actual and estimated cooperation rates for different stake levels. The fitted line based on our full regression model (Model 6 presented in Table 4) appears to capture the pattern rather well. Cooperation is high when the stakes are relatively small: For amounts up to £500, people on average cooperate 73.4% of the time. The rate drops to approximately 45% as the stakes increase and remains relatively stable for the largest amounts. An unreported test shows that we cannot reject that the relation becomes essentially flat for stakes larger than £1,500.

Although the absolute level of the stakes thus appears to have some influence on the propensity to cooperate, behavioral research suggests that people do not always evaluate prospects just in absolute

¹⁴ See, for example, Hoffman et al. (1996a), Slonim and Roth (1998), Cameron (1999), List and Cherry (2000, 2008), Fehr et al. (2002), Munier and Zaharia (2002), Carpenter et al. (2005), Johansson-Stenman et al. (2005), and Kocher et al. (2008). For TV game show data, List (2004, 2006) and Oberholzer-Gee et al. (2010) also find that cooperative behavior is practically invariant to the stakes. Belot et al. (2010) report some counterintuitive evidence that cooperation actually increases with stakes.

Figure 2 Stakes and the Propensity to Cooperate



Notes. This figure displays the relative frequency of contestants who decide to split across various stake intervals. Tick mark values represent the endpoints of the intervals. Each bar depicts the percentage of cooperators within a specific stake bracket. The dashed line reflects the average cooperation rate across our full sample, and the solid line connects the average estimate of the propensity to cooperate for each stake bracket. The estimates are computed using our “full model,” which is Model 6 in Table 4. For each interval, the number of contestants is displayed at the bottom of the bar.

terms, but rather they sometimes use relative comparisons to determine subjective values. This way, what comprises the context can strongly influence choices (Kahneman et al. 1999). In choice tasks, for example, one can increase the likelihood that a given option is chosen by adding an alternative to the choice set that is dominated by the given option but not by the other alternatives available (Huber et al. 1982). Also, as a consequence of the use of relative judgments, seemingly irrelevant anchors can influence how people value goods of various kinds (Green et al. 1998, Ariely et al. 2003, Simonson and Drolet 2004), even (risky) monetary prospects (Johnson and Schkade 1989).

For our purposes, the question of interest is whether the game show influences the contestants’ perceptions of what constitutes “serious money.” Suppose that some contestants decide that for serious money they are willing to bear the reputational costs, if any, of defecting on national TV, but if the stakes are small, so-called peanuts, then they will just cooperate to look good. In this scenario, we would observe the pattern of cooperation in our data: high cooperation rates for low stakes and lower cooperation for high stakes. The interesting point, however, is that the “small” stakes on this show, several hundred pounds, are quite large relative to most experiments. So even when the contestants are playing for what seems to be peanuts, these are big peanuts!

In *Deal or No Deal*, another game show that has even larger stakes than *Golden Balls*, Post et al. (2008) also find strong evidence of such a “big peanuts” phenomenon. Namely, when unlucky contestants faced

decisions near the end of the show that were “merely” for thousands of euros, they displayed little or no risk aversion. In fact, some of their contestants made risk-seeking choices in such situations. The authors provide further evidence for this behavior in classroom experiments designed to mimic the show at two levels of stakes: “low” and “medium.” In the low stakes treatment the average prize was €40 with a maximum of €500, and in their medium stakes treatment the average prize was €400 with a maximum of €5,000. Although risk aversion increased with stakes within each treatment, such an effect was not found across treatments: Despite the very different money amounts, risky choices were similar for the low and medium stakes session. Choices in both conditions were even remarkably similar to those made in the actual TV show, despite the huge stakes used there (average €400,000, maximum €5,000,000).

These results suggest that a context can convert a sum of money that would normally be considered consequential into perceived peanuts. In the *Golden Balls* scenario, earlier expectations about the jackpot size or a specific value from the game might operate as an anchor or reference value by which the actual size of the jackpot is evaluated.¹⁵ The most obvious benchmark contestants may use seems to be the maximum possible jackpot at the beginning of round 3.

¹⁵ We intentionally do not use the term “reference point” in order to avoid associations with prospect theory here, because we are hesitant to translate the elements of prospect theory to preferences in this game and to derive testable predictions.

Though the expected jackpot size might be an alternative candidate, it is neither salient nor easily calculated. In fact, even a rough assessment is rather complicated, particularly because of the influence of killer balls. The maximum potential jackpot, however, is always visually displayed and explicitly stressed by the game show host.

To test for such an effect, we include the variable *Potential stakes* in our analyses, defined as the highest possible jackpot at the start of round 3. As with the actual stakes, we take the natural logarithm. Not surprisingly, the actual and the potential stakes variable are significantly correlated, but because of the effect of killer balls and the skewed distribution of cash balls the degree is rather limited; the Pearson correlation coefficient is $\rho = 0.30$.

Model 3 in Table 3 shows the new results. The positive sign of the potential stakes coefficient is in line with what we would expect, but the effect is statistically insignificant for the entire sample ($p = 0.139$). Interestingly, the effect becomes marginally significant if we exclude the (returning) contestants who already appeared in a previous episode of the show ($p = 0.066$; untabulated). This gives rise to the idea that the anchoring effect of the potential jackpot may decrease over time as contestants become more familiar with the show by watching it on TV. Prior shows will give contestants an impression of expected payoffs, which may help them to evaluate whether the stakes they face themselves are high or low in the context of the game and reduce the role of an episode-specific reference value such as the maximum possible jackpot size.

We explore the effect of experience by testing whether the effect of the potential jackpot changes as contestants have watched more episodes on TV. As a (noisy) proxy for how many shows a contestant has watched, we define the variable *Transmissions* as the number of different episodes broadcast on TV prior to the studio recording of the current episode. Model 4 in Table 3 displays the results. There is no significant main effect of this variable ($p = 0.722$), indicating that there is no evidence of a trend in the cooperation rate over time. However, the interaction effect of the number of transmissions and the potential stakes is significantly negative ($p = 0.037$) and implies that the anchoring effect of the maximum possible jackpot decreases by a 0.10 percentage point for each previously aired episode. Controlling for this interaction effect, the effect of the maximum potential jackpot is highly significant in the early episodes ($p = 0.006$), where doubling the maximum potential jackpot increases cooperation by more than 12 percentage points.¹⁶

¹⁶ Because the potential stakes and the actual stakes are correlated, the interaction of *Potential stakes* and *Transmissions* might pick up

The pattern of a pronounced effect of the potential jackpot size in the earlier but not in the later shows also becomes apparent if we include dummy variables that subdivide our sample. Model 5 in Table 3, for example, uses a natural subdivision and employs a dummy variable for the first 149 episodes (0–112 transmissions prior to the recordings) and for the remaining 138 (149–214 prior transmissions). Clearly, the effect is significant across the first half our data ($p = 0.005$) and insignificant thereafter ($p = 0.543$).

5. Reciprocal Preferences

Reciprocity refers to a tendency to repay kindness with kindness and unkindness with unkindness. Reciprocal behavior in the field is generally embedded in long-term social interaction, and reputation concerns therefore form a plausible explanation for virtually all instances where people reciprocate (Sobel 2005). Reciprocal actions can sometimes also be explained by preferences over outcome distributions, most notably by a desire for equity or equality (Fehr and Schmidt 1999, Bolton and Ockenfels 2000). Interestingly, laboratory experiments in which such motivations are controlled for have convincingly demonstrated that people also have a real intrinsic preference for reciprocity.¹⁷

Golden Balls provides a neat opportunity to investigate the presence of reciprocal preferences outside of the laboratory and for substantial stakes. In particular, in the first two rounds contestants cast votes to determine who has to leave the show. Each vote carries a significant weight because of the small number of contestants, and voting against somebody can be viewed as unkind, particularly when the other votes were cast against other players. If people indeed have reciprocal preferences, we would expect that a contestant who makes it to the final in spite of her opponent's vote against her, has a lower propensity to cooperate.¹⁸

an effect of the interaction of *Actual stakes* and *Transmissions*. As a robustness check, we have therefore also added the latter to Model 4. The effect of this additional control variable is insignificant ($p = 0.573$; untabulated), confirming our interpretation.

¹⁷ See, for example, Kahneman et al. (1986), Blount (1995), Brandts and Solà (2001), Kagel and Wolfe (2001), McCabe et al. (2003), Charness and Rabin (2002), Cox (2002, 2004), Falk et al. (2003, 2008), Offerman (2002), Charness (2004), and de Quervain et al. (2004). For theoretical accounts, see, for example, Rabin (1993), Levine (1998), Dufwenberg and Kirchsteiger (2004), Falk and Fischbacher (2006), Cox et al. (2007), and Gul and Pesendorfer (2010).

¹⁸ Studying voting patterns in the internationally successful TV game show *The Weakest Link*, both Levitt (2004) and Antonovics et al. (2005) find evidence that people reciprocate against people who voted against them in past rounds. However, because players in this show have an incentive to vote off players that are more likely to vote against them, they cannot rule out that this strategic concern drives the effect.

Table 4 Binary Probit Regression Results [2/2]

	Model 6	Model 7	Model 8	Model 9	Model 10
<i>Age</i>	0.002 (0.387)	0.003 (0.283)	0.002 (0.457)	0.002 (0.372)	0.002 (0.355)
<i>Gender</i> (male = 1)	−0.249 (0.001)	−0.292 (0.000)	−0.258 (0.001)	−0.252 (0.001)	−0.247 (0.001)
<i>Race</i> (white = 1)	0.149 (0.079)	0.162 (0.077)	0.146 (0.078)	0.147 (0.089)	0.148 (0.083)
<i>City</i> (large = 1)	−0.034 (0.467)	−0.027 (0.552)	−0.035 (0.463)	−0.034 (0.463)	−0.037 (0.433)
<i>London</i> (London = 1)	0.041 (0.565)	0.039 (0.540)	0.050 (0.501)	0.042 (0.551)	0.037 (0.597)
<i>Education</i> (high = 1)	0.088 (0.068)	0.089 (0.041)	0.089 (0.067)	0.089 (0.065)	0.089 (0.066)
<i>Student</i> (student = 1)	0.001 (0.988)	−0.025 (0.768)	0.008 (0.924)	−0.002 (0.983)	0.001 (0.988)
<i>Age × Gender</i>	0.011 (0.000)	0.008 (0.010)	0.010 (0.009)	0.011 (0.001)	0.011 (0.001)
<i>Actual stakes</i> (log)	−0.050 (0.000)	−0.054 (0.000)	−0.052 (0.000)	−0.051 (0.000)	−0.052 (0.000)
<i>Potential stakes</i> (log)	0.183 (0.004)	0.174 (0.004)	0.170 (0.006)	0.180 (0.004)	0.186 (0.004)
<i>Transmissions</i>	−0.000 (0.660)	−0.000 (0.106)	−0.000 (0.608)	−0.000 (0.584)	−0.000 (0.497)
<i>Potential stakes × Transmissions</i>	−0.001 (0.026)	−0.001 (0.022)	−0.001 (0.030)	−0.001 (0.025)	−0.001 (0.029)
Reciprocal preferences					
<i>Vote received from opp.</i> (yes = 1)	−0.215 (0.019)	−0.237 (0.015)	−0.202 (0.020)	−0.214 (0.026)	−0.216 (0.025)
Exp. conditional cooperation					
<i>Promise</i> (promise = 1)		0.311 (0.000)			
<i>Promise opp.</i> (promise = 1)		−0.080 (0.053)			
<i>Age opp.</i>			0.001 (0.732)		
<i>Gender opp.</i> (male = 1)			−0.118 (0.140)		
<i>Race opp.</i> (white = 1)			−0.026 (0.775)		
<i>City opp.</i> (large = 1)			0.051 (0.288)		
<i>London opp.</i> (London = 1)			0.027 (0.713)		
<i>Education opp.</i> (high = 1)			0.116 (0.017)		
<i>Student opp.</i> (student = 1)			0.069 (0.442)		
<i>Age opp. × Gender opp.</i>			0.004 (0.279)		
Past deceitful behavior					
<i>Lie round 1 opp.</i> (lie = 1)				−0.013 (0.782)	
<i>Lie round 2 opp.</i> (lie = 1)				−0.027 (0.578)	
<i>Lie cash ball round 1 opp.</i> (lie = 1)					0.037 (0.491)
<i>Lie cash ball round 2 opp.</i> (lie = 1)					−0.064 (0.295)
<i>Lie killer ball round 1 opp.</i> (lie = 1)					−0.055 (0.330)
<i>Lie killer ball round 2 opp.</i> (lie = 1)					0.007 (0.894)
Wald χ^2 (df)	59.65 (13)	98.03 (15)	62.40 (21)	59.19 (15)	62.72 (17)
Log pseudo-likelihood	−365.86	−343.31	−359.26	−365.64	−364.52
Pseudo R^2	0.078	0.135	0.095	0.079	0.082
<i>N</i>	574	574	574	574	574
Number of clusters	287	287	287	287	287

Notes. This table displays results from the probit regression analyses of contestants' decisions to split (1) or steal (0) the jackpot in the prisoner's dilemma at the end of the British TV game show *Golden Balls*. The opponent variables measure the demographic characteristics of the contestant's opponent and are defined similar to the contestant's own demographic variables. Other definitions are as in previous tables.

Although the voting is anonymous, it is often straightforward to deduce who has voted against whom. If a contestant in round 1 (round 2) receives three (two) votes it is obvious that all others have voted against her and that she herself has voted against the contestant who received one vote. For the other possible distributions of votes, we can usually deduce the individual votes from the banter preceding the vote, or, in the case of a tie, from the discussion following the vote. In the banter leading up to a vote, contestants generally make abundantly clear whom they intend to vote against (possibly out of an attempt to coordinate voting with other contestants). In the case of a tie, contestants openly discuss who they want to leave the program; if it was not already clear from the banter whom they had originally voted against, this post-vote discussion generally makes it apparent. This procedure allows us to

determine a contestant's vote 95% of the time. For various reasons it is much more difficult to determine clear instances of someone going out of their way to be nice to another player, so we limit our analysis to negative rather than positive reciprocity.

Based on the voting information, we create a dummy variable entitled *Vote received from opponent*, taking the value of 1 if a contestant received a vote from her final opponent and 0 otherwise. If we could not establish whether a contestant received a vote from her final opponent, she is assigned the value of 0 as well (exclusion of these cases does not change our results). Because contestants who receive votes often do not make it to the next round, relatively few contestants qualify: as displayed in Table 2, 5% (28 subjects) of the final-round contestants received a vote from their opponent.

Model 6 in Table 4 includes the new dummy variable. In line with the idea that people have reciprocal preferences, the likelihood of a contestant to cooperate with her final opponent plummets by approximately 21 percentage points if this opponent has voted against her earlier in the game ($p = 0.019$). There are, however, three alternative explanations for such a behavioral pattern that are unrelated to a genuine preference for reciprocity. Although we cannot rule out that these alternative explanations explain part of the effect, they do not appear particularly strong.

First, the causality may not run from receiving a vote to cooperativeness but the other way around: players voting against contestants with a less cooperative disposition. This would imply that cooperation is also related to the number of votes received from other players, which appears not to be the case ($p = 0.231$; untabulated).

Second, a contestant may like to match her opponent's choice for reasons other than reciprocal concerns and interpret the earlier vote against her as a signal that her opponent dislikes her and will not cooperate. However, her interpretation would generally not be legitimate: players do not cooperate less with someone they voted against ($p = 0.403$; untabulated). Moreover, the next section finds little support for such expectational conditional cooperation.

Last, a contestant's lower propensity to cooperate with someone who voted against her may be out of reputation concerns ("I am not to be messed with") instead of an intrinsic preference for reciprocity. However, when asked to explain their choice after the final decisions, contestants never use this costless opportunity to strengthen their message and point to their reciprocal nature.

6. Expectational Conditional Cooperation

There is considerable evidence that many people have a preference for conditional cooperation, defined as the desire to match the cooperation of others. In laboratory and field experiments, about half of the subjects are more willing to cooperate if others do so as well (e.g., Fischbacher et al. 2001, Frey and Meier 2004). Conditional cooperation can arise from reciprocal preferences, but also for other reasons. Social norms or a desire for conformity might account for it, and, especially in the laboratory, egalitarian motives can often explain conditional cooperative behavior because equality in payoffs generally only arises if players coordinate on their level of cooperation, as is also the case in our show.

Experimental studies typically investigate conditional cooperation in settings where subjects have the

possibility to condition their behavior directly on the behavior of others. In everyday life, such clear-cut conditioning is usually not possible, especially in one-shot situations. Conditional cooperation then has to be based on expectations about the behavior of others, and the degree of coordination would depend on the predictive power of available information and on whether and how this information is interpreted. A natural question is whether conditional cooperation can be observed when the conditioning is only on an expectation of cooperation rather than on actual cooperation.

In *Golden Balls* it is not possible for a contestant to condition directly on her opponent's behavior because the two are playing a simultaneous move game. However, we can investigate whether contestants condition their behavior on factors that form reliable predictors of their opponent's behavior. That is, we can investigate the joint hypothesis that players make rational forecasts of their opponent's behavior *and* then condition their behavior on those expectations.

The first step in such an analysis is determining the factors that a contestant could use to form an expectation about their opponent's likelihood of cooperation. One such factor is whether an opponent made a promise to split. Although the literature on conditional cooperation is rather recent, literature investigating the role of communication, and especially promises in social dilemma situations, already pointed toward tendencies of conditional cooperation. In a meta-analysis of prisoner-dilemma experiments, Sally (1995), for example, finds that cooperation occurs more often when the other player makes an explicit though nonbinding promise that she will cooperate. The combination of a preference for conditional cooperation and a reluctance to lie (e.g., Charness and Dufwenberg 2005, 2006; Gneezy 2005) can explain why promises have such an effect: people like to cooperate if others do, and a promise is a reliable signal of others' behavior if they have a reluctance to lie.

In the conversation prior to the decision to either split or steal, many contestants explicitly promise to split or otherwise make a definitive statement of their intention to do so. Based on the statements made in this small talk, we create a dummy variable labeled *Promise*, indicating whether the contestant made an explicit, unambiguous promise or announcement that she will choose split (1) or not (0).¹⁹ As shown in Table 2, about half (53%) of the contestants make such

¹⁹ If a contestant responds affirmative to a question whether she will choose split or if she announces that she will not choose steal, *Promise* takes the value of 1 as well. The value is 0 in all other cases, including when people give the impression that they plan to split but do not explicitly express themselves as such, when they just refer to earlier intentions (for example, "I came here to split"),

a promise. We investigate both whether observing a promise is predictive of the cooperative behavior of the contestant making the promise and whether a contestant conditions her behavior on whether or not her opponent made a promise.

As shown in Model 7 in Table 4, a player's promise is a highly significant predictor of her propensity to cooperate ($p = 0.000$). Those who make a promise are about 31 percentage points more likely to cooperate. In fact, an explicit promise is the single most reliable predictor of whether someone will cooperate.²⁰

Although a promise is a strong signal of cooperation, contestants whose opponent made a promise do not have a higher propensity to choose split. In fact, as Model 7 also shows, if an opponent promises to be cooperative, the other player even displays a marginally significant *decrease* in the likelihood of choosing split. Belot et al. (2010) obtain a similar result.

An explicit promise is the strongest predictor of cooperation, but, as we have previously shown, there are also demographic factors that a contestant could use to forecast cooperation. For example, we have seen that young males cooperate less than young females. However, inferences from this sort of analysis have to be tentative because there could be an additional confound if opponents have a taste for cooperating with someone with particular demographic characteristics.

As Model 8 in Table 4 shows, we find little evidence that contestants condition their behavior on their opponents' background characteristics. The only weak evidence for conditional cooperation is that people cooperate significantly more frequently with higher educated opponents ($p = 0.017$), but education is not a very strong predictor of behavior. If we assess the joint significance of the various opponent background characteristics, we also find that they collectively do not have a significant effect on cooperation.²¹ Analyzing the show *Friend or Foe?*, Oberholzer-Gee et al. (2010) find no conditioning on opponent background characteristics in the first season, but they do find it in later seasons. They interpret this as conditional cooperation on the basis of learned expectations. We too have investigated whether conditioning

when they confine themselves to statements like "you can trust me" and "I will not let you down," and when they only express their preference for a coordinated outcome ("I want us to split" and "I do not want both of us to go home empty-handed").

²⁰ Of course, we do not interpret the promise as causing the cooperation. The direction of the causation could go the other way.

²¹ We also looked at more complex mechanisms related to the similarity of the contestant's own background characteristics and those of her opponent, such as whether people cooperate more with those who are more similar to them ("social-distance" effects). In our data, there is no evidence of such behavior.

arises as more episodes were transmitted, but we find no indication for such an effect.

As an alternative to the two models discussed above, we also examined a two-step approach. We first estimated each opponent's propensity to cooperate given her background characteristics and promise behavior, and then added the estimated propensity of opponents as an explanatory variable to our regression model. Again, we found no indication of conditionally cooperative behavior.

In summary, we find no evidence of expectational conditional cooperation. Apparently, either players cannot or do not forecast the behavior of their opponents, or they do not have conditionally cooperative preferences. Our evidence for reciprocal preferences in the previous section hints that it is the former rather than the latter interpretation that underlies this result.

Belot et al. (2012) also provide evidence that predicting one's opponent's behavior is difficult. They had subjects watch clips from the Dutch counterpart to *Golden Balls* and asked them to assess the likelihood of each contestant's cooperation. Although the estimated likelihood for cooperators was significantly higher than for defectors, the difference was only seven percentage points.

7. Past Deceitful Behavior

In this section, we investigate whether lies influence opponents' willingness to cooperate. In the early rounds of the show, contestants have numerous opportunities to lie about the values on their hidden balls, lies that are quickly revealed to everyone. These lies can be consequential. If someone hides low value and killer balls and in so doing manages to remain in the game, she will have reduced the potential payoff to the remaining contestants.

In the final, contestants might be less likely to cooperate with opponents who have lied, either out of reciprocal concerns (e.g., Brandts and Charness 2003) or because they interpret lying as evidence of a self-interested nature and a sign of an imminent steal decision (e.g., van Lange and Kuhlman 1994). Thus, past deceitful behavior is not a separate possible determinant of cooperation, but rather a special case of either reciprocity or conditional cooperation, or both.

We collected data on the statements made by contestants and the actual values of the balls that they possessed, allowing us to specify various measures for deceitful behavior. The analyses reported here are restricted to the use of dummy variables. We have also tried more complex, continuous variables for lying, but these approaches yielded similar results.

We apply separate variables for each game round. The general variables take the value of 1 if the contestant lied, irrespective of whether she overstated

the monetary value of a cash ball or failed to disclose a killer ball. To distinguish between these two types of lies, we also use specific variables for each type separately. It is not obvious which type of lie would be considered more objectionable. On the one hand, lying about killers is much more harmful to others than exaggerating the value of a cash ball, and, assuming a preference to reciprocate, doing so could then be expected to have a greater negative effect on an opponent's propensity to cooperate. On the other hand, lying about killer balls might also be more understandable, because killer balls have a much greater impact on a contestant's chances to be voted off the show. Players may realize that nearly everyone will fail to disclose a hidden killer ball, and thus not be inclined to punish such behavior. Lying about cash ball values might be more like gratuitous lying and be viewed more harshly and, consequently, have a greater effect on cooperation.

As shown in Table 2, lying is rather common on the show: 41% of the contestants who made it to the final lied about their back row balls in round 1, and 36% lied in round 2 (some did both). Furthermore, in the first round, 24% overstated the value of a cash ball, and 21% hid a killer ball (some did both). For the second round, these figures are 15% and 24%, respectively.²²

Table 4 displays the regression results when we add the general dummy variables (Model 9) and the dummy variables that distinguish between lying about killer balls and lying about cash balls (Model 10). We find that past lies of an opponent do not affect a contestant's propensity to cooperate: each of the six variables is insignificant ($0.295 < p < 0.894$). In addition to these simple tests, we also investigated whether lying is considered less fair and has more impact the more it is unexpected or "abnormal" given the circumstances, but again we found no significant effect.²³ Lying neither predicts a contestant's own cooperative behavior.²⁴ One plausible interpreta-

tion of these results is that lying is seen as an inherent part of *Golden Balls* and therefore unobjectionable behavior, much as bluffing is considered in the game of poker (Charness and Dufwenberg 2005).

8. Conclusions and Discussion

Golden Balls provides us with the possibility to examine cooperative behavior outside the conventional context of the laboratory with large sums of money at stake.

Our results provide support for the view that attitudes are strongly influenced by context. We find unusually high rates of cooperation when the luck of the game reduces the stakes to merely a few hundred pounds. Such amounts are tiny in the light of the thousands and even tens of thousands the game is often played for but would be considered very large in any laboratory setting. In the early days of the show, when the contestants have not had an opportunity to watch the show on TV and are still learning what kind of stakes are to be expected, cooperation rates appear to be influenced by the salient but normatively irrelevant value representing the maximum they could have been playing for with a lucky selection of balls. Over time, this effect vanishes, suggesting that expectations about stakes become well informed.

We label the tendency to be unusually cooperative for what would normally be considered high stakes a "big peanuts" result. Players seem to feel that when making a choice about a few hundred dollars when they might otherwise have been dividing tens of thousands, they are playing for peanuts, and cooperate, perhaps thinking that it is not worth stealing for what they perceive to be so little money. This finding reinforces a similar result for risk taking behavior in another game show, *Deal or No Deal*. In that context, where the stakes were even higher, amounts of money in the tens of thousands of dollars became perceived as peanuts, because hundreds of thousands of dollars had been on the line. These are very big peanuts indeed.

Using the interaction that occurs among contestants prior to the final, we also examined the effects that past opposition and lying have on cooperation. Using the votes we find evidence to support the view that people have reciprocal preferences. Contestants are less likely to cooperate if their opponent has tried to vote them off the show at an earlier stage of the game. Lying on the other hand has no significant effect. We investigated several measures, but none was significantly related to cooperation. Lying is evidently not frowned upon in *Golden Balls*, perhaps because it is expected. The different impact of opposition and lying might be related to their different nature in this

²² Conditional on having at least one killer ball on their back row, contestants hid a killer ball 50% (43%) of the time in round 1 (round 2).

²³ We used a two-stage procedure to express the abnormality of a lie. For each round, we estimated a regression model that explains a contestant's propensity to lie, given the ball values on her back row, the ball values on her front row, and the rank of her front row balls relative to those of the other players. For each final contestant, the "abnormality" of a lie we then measured as the difference between unity and this estimated lie propensity.

²⁴ Such a relation might be expected if the propensity to be honest and the propensity to cooperate are influenced by a similar preference for "prosocial," "kind," or "fair" behavior. It has, for example, been argued that the reluctance to lie is driven by guilt aversion (Charness and Dufwenberg 2005, 2006; Gneezy 2005), and empirical analysis suggests that guilt aversion is also a strong driving force behind cooperative behavior (Dufwenberg et al. 2012).

game. Voting is a directed and aggressive act toward one specific contestant. Lying, on the other hand, is an undirected and defensive act.

With *Golden Balls* we are also able to investigate an interesting, expectational form of conditional cooperation. Specifically, because explicit promises to cooperate are strong predictors of actual cooperation, we can see whether players are more likely to cooperate with someone who has made such a promise. We find no evidence of such behavior. More generally, we find that players do not appear to condition their choice of whether to cooperate on factors that predict the cooperation likelihood of their opponent. Players may lack the ability or ignore the possibility to reliably interpret information about the expected behavior of others, or they may not have a preference for matching the other's choice. Given our finding that people reciprocate votes against them, the former explanation seems more likely. For situations beyond the context of our game these results suggest that conditional cooperation is not a very important phenomenon, at the least when direct conditioning is not possible and people would need to form expectations about the behavior of others.

We conclude with some comments on the generalizability of our results. There are three primary concerns. First, selection procedures may have affected the average cooperation rate in our study. Subjects self-select into auditions, are then selected by the producer, and during the game they themselves have the opportunity to vote off opponents they would rather not play the final with. For some demographic variables, selection may perhaps also have affected the correlation with cooperation. Unfortunately, we cannot substantiate our intuition that such effects are negligible, nor could we have prevented them if they would exist. Note, however, that selection procedures are inevitable in any lab experiment or field setting. Moreover, the subjects in our sample vary widely in terms of their demographic characteristics and as a group they seem to resemble a (middle-class)²⁵ cross-section of the general population more closely than subjects in most conventional experiments.

Second, subjects' behavior in a game show might be influenced by what could be called "a drive to win the contest." However, an important but hard to answer question would then be what "winning" actually means in this context. After all, it seems like a matter of personal social preferences whether winning is equivalent to a successful stealing attempt or to a successful coordination attempt.

Last, our contestants are not strictly playing a one-shot game. In the setting we study decisions are made on national TV, under the scrutiny of a studio audience and millions of viewers. This undoubtedly influences the behavior we observe. However, we do not feel that these special circumstances render our findings less interesting or less predictive of behavior in other settings. The truth is that every setting is, in some way, special. Subjects in a laboratory experiment know that their behavior is being scrutinized to some extent as well. Field settings are also special; bargaining over the price of a car or a house is different from negotiating compensation with a new employer or the division of household chores with a spouse. Although it would be fascinating for researchers to be able to surreptitiously study the outcomes of these sorts of interactions from the "real world," the researchers would still only be able to speculate on how their results would generalize to different real-world settings. TV game shows offer a unique opportunity to study theoretically interesting behavior at stakes that are impossible to replicate in the lab. How the results compare with other contexts will be determined by future research.

In the absence of an ability to conduct such surreptitious field experiments in many domains, researchers are left with two alternatives: run experiments in the lab or the field, or study naturally occurring behavior in an interesting setting. This paper is an example of the latter strategy. Although a game show may seem like a strange environment, we think it may be closer to the situations that occur in the workplace than many other settings in which cooperation has been studied. Coworkers often must choose whether to cooperate, and their actions are often at least semipublic.

Finally, the big peanuts phenomenon, perhaps the most interesting finding in this paper, is one that does not appear to depend in any important way on the specific game show environment. As a U.S. senator once famously said, "a billion here, a billion there, pretty soon you're talking real money."²⁶

Acknowledgments

The authors are grateful to Guido Baltussen, Han Bleichrodt, Uri Gneezy, and the anonymous reviewers for their many constructive and valuable comments. The authors thank Endemol UK and in particular Tara Ali and Tom Blakeson for providing information and recordings of *Golden Balls*; and Shirley Kremer, Thomas Meijer, and Sanne van Bommel for their skillful research assistance. The paper benefited from discussions with seminar participants

²⁵ For whatever reason, whether it is the interest in applying or the preferences of the producers, contestants are rarely very rich or very poor.

²⁶ This quote is attributed to Everett McKinley Dirksen (1896–1969), an American politician of the Republican Party.

at the Erasmus University of Rotterdam and Tilburg University, and with participants of the Decision and Uncertainty 2010 Workshop at HEC Paris, the Behavioral Decision Research in Management (BDRM) 2010 Conference in Pittsburgh, the Society for the Advancement of Behavioral Economics (SABE) 2010 Conference in San Diego, and the Subjective Probability, Utility and Decision Making (SPUDM) 2011 Conference in Kingston upon Thames. The authors gratefully acknowledge support from the Erasmus Research Institute of Management, the Netherlands Organisation for Scientific Research (NWO), and the Tinbergen Institute.

References

- Ai, C., E. C. Norton. 2003. Interaction terms in logit and probit models. *Econom. Lett.* **80**(1) 123–129.
- Antonovics, K., P. Arcidiacono, R. Walsh. 2005. Games and discrimination: Lessons from *The Weakest Link*. *J. Human Resources* **40**(4) 918–947.
- Ariely, D., G. Loewenstein, D. Prelec. 2003. Coherent arbitrariness: Stable demand curves without stable preferences. *Quart. J. Econom.* **118**(1) 73–105.
- Beetsma, R. M. W. J., P. C. Schotman. 2001. Measuring risk attitudes in a natural experiment: Data from the television game show *Lingo*. *Econom. J.* **111**(474) 821–848.
- Bellemare, C., S. Kröger. 2007. On representative social capital. *Eur. Econom. Rev.* **51**(1) 183–202.
- Belot, M., V. Bhaskar, J. van de Ven. 2010. Promises and cooperation: Evidence from a TV game show. *J. Econom. Behav. Organ.* **73**(3) 396–405.
- Belot, M., V. Bhaskar, J. van de Ven. 2012. Can observers predict trustworthiness? *Rev. Econom. Statist.* Forthcoming.
- Bennett, R. W., K. A. Hickman. 1993. Rationality and the *Price Is Right*. *J. Econom. Behav. Organ.* **21**(1) 99–105.
- Berk, J. B., E. Hughson, K. Vandezande. 1996. The price is right, but are the bids? An investigation of rational decision theory. *Amer. Econom. Rev.* **86**(4) 954–970.
- Blount, S. 1995. When social outcomes aren't fair: The effect of causal attributions on preferences. *Organ. Behav. Human Decision Processes* **63**(2) 131–144.
- Bolton, G. E., A. Ockenfels. 2000. ERC: A theory of equity, reciprocity, and competition. *Amer. Econom. Rev.* **90**(1) 166–193.
- Brandts, J., G. Charness. 2003. Truth or consequences: An experiment. *Management Sci.* **49**(1) 116–130.
- Brandts, J., C. Solà. 2001. Reference points and negative reciprocity in simple sequential games. *Games Econom. Behav.* **36**(2) 138–157.
- Camerer, C. F. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press, Princeton, NJ.
- Cameron, L. A. 1999. Raising the stakes in the ultimatum game: Experimental evidence from Indonesia. *Econom. Inquiry* **37**(1) 47–59.
- Carpenter, J. P., C. Connolly, C. K. Myers. 2008. Altruistic behavior in a representative dictator experiment. *Experiment. Econom.* **11**(3) 282–298.
- Carpenter, J. P., A. G. Daniere, L. M. Takahashi. 2004. Cooperation, trust, and social capital in Southeast Asian urban slums. *J. Econom. Behav. Organ.* **55**(4) 533–551.
- Carpenter, J., E. Verhoogen, S. Burks. 2005. The effect of stakes in distribution experiments. *Econom. Lett.* **86**(3) 393–398.
- Charness, G. 2004. Attribution and reciprocity in an experimental labor market. *J. Labor Econom.* **22**(3) 665–688.
- Charness, G., M. Dufwenberg. 2005. Deception: The role of guilt. Working paper, University of Southern California, Santa Barbara.
- Charness, G., M. Dufwenberg. 2006. Promises and partnership. *Econometrica* **74**(6) 1579–1601.
- Charness, G., M. Rabin. 2002. Understanding social preferences with simple tests. *Quart. J. Econom.* **117**(3) 817–869.
- Cooper, D. J., J. H. Kagel. 2009. Other regarding preferences: A selective survey of experimental results. Working paper, Florida State University, Tallahassee.
- Cox, J. C. 2002. Trust, reciprocity, and other-regarding preferences: Groups vs. individuals and males vs. females. R. Zwick, A. Rapoport, eds. *Experimental Business Research*. Kluwer Academic Publishers, Norwell, MA, 331–350.
- Cox, J. C. 2004. How to identify trust and reciprocity. *Games Econom. Behav.* **46**(2) 260–281.
- Cox, J. C., D. Friedman, S. Gjerstad. 2007. A tractable model of reciprocity and fairness. *Games Econom. Behav.* **59**(1) 17–45.
- Cox, T. H., S. A. Lobel, P. L. McLeod. 1991. Effects of ethnic group cultural differences on cooperative and competitive behavior on a group task. *Acad. Management J.* **34**(4) 827–847.
- Croson, R., U. Gneezy. 2009. Gender differences in preferences. *J. Econom. Literature* **47**(2) 448–474.
- Dawes, R. M. 1980. Social dilemmas. *Ann. Rev. Psych.* **31** 169–193.
- Dawes, R. M., D. M. Messick. 2000. Social dilemmas. *Internat. J. Psych.* **35**(2) 111–116.
- Dawes, R. M., R. H. Thaler. 1988. Anomalies: Cooperation. *J. Econom. Perspect.* **2**(3) 187–197.
- de Quervain, D. J.-F., U. Fischbacher, V. Treyer, M. Schellhammer, U. Schnyder, A. Buck, E. Fehr. 2004. The neural basis of altruistic punishment. *Science* **305**(5688) 1254–1258.
- Dufwenberg, M., G. Kirchsteiger. 2004. A theory of sequential reciprocity. *Games Econom. Behav.* **47**(2) 268–298.
- Dufwenberg, M., A. Muren. 2006. Generosity, anonymity, gender. *J. Econom. Behav. Organ.* **61**(1) 42–49.
- Dufwenberg, M., S. Gächter, H. Hennig-Schmidt. 2012. The framing of games and the psychology of play. *Games Econom. Behav.* Forthcoming.
- Eckel, C. C., P. J. Grossman. 2008. Differences in the economic decisions of men and women: Experimental evidence. C. R. Plott, V. L. Smith, eds. *Handbook of Experimental Economics Results*, Vol. 1. Elsevier, Amsterdam, 509–519.
- Egas, M., A. Riedl. 2008. The economics of altruistic punishment and the maintenance of cooperation. *Proc. Roy. Soc. B: Biol. Sci.* **275**(1637) 871–878.
- Falk, A., U. Fischbacher. 2006. A theory of reciprocity. *Games Econom. Behav.* **54**(2) 293–315.
- Falk, A., E. Fehr, U. Fischbacher. 2003. On the nature of fair behavior. *Econom. Inquiry* **41**(1) 20–26.
- Falk, A., E. Fehr, U. Fischbacher. 2008. Testing theories of fairness-intentions matter. *Games Econom. Behav.* **62**(1) 287–303.
- Fehr, E., S. Gächter. 1998. Reciprocity and economics: The economic implications of Homo reciprocans. *Eur. Econom. Rev.* **42**(3–5) 845–859.
- Fehr, E., S. Gächter. 2000a. Cooperation and punishment in public goods experiments. *Amer. Econom. Rev.* **90**(4) 980–994.
- Fehr, E., S. Gächter. 2000b. Fairness and retaliation: The economics of reciprocity. *J. Econom. Perspect.* **14**(3) 159–181.
- Fehr, E., H. Gintis. 2007. Human motivation and social cooperation: Experimental and analytical foundations. *Ann. Rev. Sociol.* **33** 43–64.
- Fehr, E., K. M. Schmidt. 1999. A theory of fairness, competition, and cooperation. *Quart. J. Econom.* **114**(3) 817–868.

- Fehr, E., U. Fischbacher, E. Tougareva. 2002. Do high stakes and competition undermine fairness? Evidence from Russia. Working paper. Institute for Empirical Research in Economics, University of Zurich, Zurich.
- Fischbacher, U., S. Gächter. 2010. Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *Amer. Econom. Rev.* **100**(1) 541–556.
- Fischbacher, U., S. Gächter, E. Fehr. 2001. Are people conditionally cooperative? Evidence from a public goods experiment. *Econom. Lett.* **71**(3) 397–404.
- Frey, B. S., S. Meier. 2004. Social comparisons and pro-social behavior: Testing “conditional cooperation” in a field experiment. *Amer. Econom. Rev.* **94**(5) 1717–1722.
- Gächter, S., B. Herrmann. 2011. The limits of self-governance when cooperators get punished: Experimental evidence from urban and rural Russia. *Eur. Econom. Rev.* **55**(2) 193–210.
- Gächter, S., B. Herrmann, C. Thöni. 2004. Trust, voluntary cooperation, and socio-economic background: Survey and experimental evidence. *J. Econom. Behav. Organ.* **55**(4) 505–531.
- Gardner, A., S. A. West. 2004. Cooperation and punishment, especially in humans. *Amer. Naturalist* **164**(6) 753–764.
- Gertner, R. 1993. Game shows and economic behavior: Risk-taking on *Card Sharks*. *Quart. J. Econom.* **108**(2) 507–521.
- Gneezy, U. 2005. Deception: The role of consequences. *Amer. Econom. Rev.* **95**(1) 384–394.
- Green, D., K. E. Jacowitz, D. Kahneman, D. McFadden. 1998. Referendum contingent valuation, anchoring, and willingness to pay for public goods. *Resource Energy Econom.* **20**(2) 85–116.
- Gul, F., W. Pesendorfer. 2010. Interdependent preference models as a theory of intentions. Working paper, Princeton University, Princeton, NJ.
- Haley, K. J., D. M. T. Fessler. 2005. Nobody’s watching? Subtle cues affect generosity in an anonymous economic game. *Evolution Human Behav.* **26**(3) 245–256.
- Henrich, J., S. J. Heine, A. Norenzayan. 2010. The weirdest people in the world? *Behav. Brain Sci.* **33**(2–3) 61–83.
- Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis. 2004. *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. Oxford University Press, Oxford, UK.
- Henrich, J., R. Boyd, S. Bowles, C. F. Camerer, E. Fehr, H. Gintis, R. McElreath. 2001. Search of Homo economicus: Behavioral experiments in 15 small-scale societies. *Amer. Econom. Rev.* **91**(2) 73–78.
- Herrmann, B., C. Thöni, S. Gächter. 2008. Antisocial punishment across societies. *Science* **319**(5868) 1362–1367.
- Hoffman, E., K. A. McCabe, V. L. Smith. 1996a. On expectations and the monetary stakes in ultimatum games. *Internat. J. Game Theory* **25**(3) 289–301.
- Hoffman, E., K. A. McCabe, V. L. Smith. 1996b. Social distance and other-regarding behavior in dictator games. *Amer. Econom. Rev.* **86**(3) 653–660.
- Huber, J., J. W. Payne, C. Puto. 1982. Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *J. Consumer Res.* **9**(1) 90–98.
- Johansson-Stenman, O., M. Mahmud, P. Martinsson. 2005. Does stake size matter in trust games? *Econom. Lett.* **88**(3) 365–369.
- Johnson, E. J., D. A. Schkade. 1989. Bias in utility assessments: Further evidence and explanations. *Management Sci.* **35**(4) 406–424.
- Kagel, J. H., K. W. Wolfe. 2001. Tests of fairness models based on equity considerations in a three-person ultimatum game. *Experiment. Econom.* **4**(3) 203–219.
- Kahneman, D., J. L. Knetsch, R. H. Thaler. 1986. Fairness and the assumptions of economics. *J. Bus.* **59**(4) 285–300.
- Kahneman, D., I. Ritov, D. A. Schkade. 1999. Economic preferences or attitude expressions? An analysis of dollar responses to public issues. *J. Risk Uncertainty* **19**(1–3) 203–235.
- Kerr, N. L. 1999. Anonymity and social control in social dilemmas. M. Foddy, M. Smithson, S. Schneider, M. Hogg, eds. *Resolving Social Dilemmas*. Psychology Press, London, 103–119.
- Kocher, M. G., P. Martinsson, M. Visser. 2008. Does stake size matter for cooperation and punishment? *Econom. Lett.* **99**(3) 508–511.
- Ledyard, J. O. 1995. Public goods: A survey of experimental research. J. H. Kagel, A. E. Roth, eds. *Handbook of Experimental Economics*. Princeton University Press, Princeton, NJ, 111–194.
- Levine, D. K. 1998. Modeling altruism and spitefulness in experiments. *Rev. Econom. Dynam.* **1**(3) 593–622.
- Levitt, S. D. 2004. Testing theories of discrimination: Evidence from *Weakest Link*. *J. Law Econom.* **47**(2) 431–453.
- Levitt, S. D., J. A. List. 2007. What do laboratory experiments measuring social preferences reveal about the real world? *J. Econom. Perspect.* **21**(2) 153–174.
- Levitt, S. D., J. A. List. 2008. Homo economicus evolves. *Science* **319**(5865) 909–910.
- List, J. A. 2004. Young, selfish and male: Field evidence of social preferences. *Econom. J.* **114**(492) 121–149.
- List, J. A. 2006. Friend or foe? A natural experiment of the prisoner’s dilemma. *Rev. Econom. Statist.* **88**(3) 463–471.
- List, J. A., T. L. Cherry. 2000. Learning to accept in ultimatum games: Evidence from an experimental design that generates low offers. *Experiment. Econom.* **3**(1) 11–29.
- List, J. A., T. L. Cherry. 2008. Examining the role of fairness in high stakes allocation decisions. *J. Econom. Behav. Organ.* **65**(1) 1–8.
- Marwell, G., R. E. Ames. 1979. Experiments on the provision of public goods. I. Resources, interest, group size, and the free-rider problem. *Amer. J. Sociol.* **84**(6) 1335–1360.
- Marwell, G., R. E. Ames. 1980. Experiments on the provision of public goods. II. Provision points, stakes, experience, and the free-rider problem. *Amer. J. Sociol.* **85**(4) 926–937.
- McCabe, K. A., M. L. Rigdon, V. L. Smith. 2003. Positive reciprocity and intentions in trust games. *J. Econom. Behav. Organ.* **52**(2) 267–275.
- Metrick, A. 1995. A natural experiment in “Jeopardy!” *Amer. Econom. Rev.* **85**(1) 240–253.
- Munier, B., C. Zaharia. 2002. High stakes and acceptance behavior in ultimatum bargaining: A contribution from an international experiment. *Theory Decision* **53**(3) 187–207.
- Oberholzer-Gee, F., J. Waldfogel, M. W. White. 2010. Friend or foe? Cooperation and learning in high-stakes games. *Rev. Econom. Statist.* **92**(1) 179–187.
- Offerman, T. 2002. Hurting hurts more than helping helps. *Eur. Econom. Rev.* **46**(8) 1423–1437.
- Office for National Statistics. 2009. Statistical bulletin: 2009 Annual survey of hours and earnings. Office for National Statistics, Newport, UK. <http://www.ons.gov.uk/ons/rel/ashe/annual-survey-of-hours-and-earnings/2009-revised/stb-ashe-2009.pdf>.
- Ostrom, E., J. Walker, R. Gardner. 1992. Covenants with and without a sword: Self-governance is possible. *Amer. Political Sci. Rev.* **86**(2) 404–417.
- Post, G. T., M. J. van den Assem, G. Baltussen, R. H. Thaler. 2008. Deal or no deal? Decision making under risk in a large-payoff game show. *Amer. Econom. Rev.* **98**(1) 38–71.
- Rabin, M. 1993. Incorporating fairness into game theory and economics. *Amer. Econom. Rev.* **83**(5) 1281–1302.
- Rapoport, A. 1988. Experiments with *n*-person social traps I: Prisoner’s dilemma, weak prisoner’s dilemma, volunteer’s dilemma, and largest number. *J. Conflict Resolution* **32**(3) 457–472.
- Raub, W., C. Snijders. 1997. Gains, losses, and cooperation in social dilemmas and collective action: The effects of risk preferences. *J. Math. Sociol.* **22**(3) 263–302.

- Rege, M., K. Telle. 2004. The impact of social approval and framing on cooperation in public good situations. *J. Public Econom.* **88**(7–8) 1625–1644.
- Sears, D. O. 1986. College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *J. Personality Soc. Psych.* **51**(3) 515–530.
- Sally, D. 1995. Conversation and cooperation in social dilemmas. A meta-analysis of experiments from 1958 to 1992. *Rationality Soc.* **7**(1) 58–92.
- Simonson, I., A. Drolet. 2004. Anchoring effects on consumers' willingness-to-pay and willingness-to-accept. *J. Consumer Res.* **31**(3) 681–690.
- Slonim, R., A. E. Roth. 1998. Learning in high stakes ultimatum games: An experiment in the Slovak Republic. *Econometrica* **66**(3) 569–596.
- Sobel, J. 2005. Interdependent preferences and reciprocity. *J. Econom. Literature* **43**(2) 392–436.
- Telser, L. G. 1995. The ultimatum game and the law of demand. *Econom. J.* **105**(433) 1519–1523.
- Tenorio, R., T. N. Cason. 2002. To spin or not to spin? Natural and laboratory experiments from *The Price Is Right*. *Econom. J.* **112**(476) 170–195.
- van Lange, P. A. M., D. M. Kuhlman. 1994. Social value orientations and impressions of partner's honesty and intelligence: A test of the might versus morality effect. *J. Personality Soc. Psych.* **67**(1) 126–141.
- van Lange, P. A. M., W. Otten, E. M. N. de Bruin, J. A. Joireman. 1997. Development of prosocial, individualistic, and competitive orientations: Theory and preliminary evidence. *J. Personality Soc. Psych.* **73**(4) 733–746.
- West, S. A., A. S. Griffin, A. Gardner. 2007. Evolutionary explanations for cooperation. *Current Biol.* **17**(16) R661–R672.
- Wooldridge, J. M. 2003. Cluster-sample methods in applied econometrics. *Amer. Econom. Rev.* **93**(2) 133–138.