Journal of Economic Theory **104**, 48–88 (2002)
doi:10.1006/jeth.2001.2910, available online at http://www.idealibrary.com on IDEAL®

A Backward Induction Experiment[1]

Ken Binmore

Department of Economics, University College London, London WC1E 6BT, United Kingdom
Uctpa97@ucl.ac.uk

John McCarthy

ELSE Experimental Laboratory, University College London, London WC1E 6BT, United Kingdom

Giovanni Ponti

Department of Economics, University of Alicante, 03071 Alicante, Spain
giuba@merlin.fae.ua.es

Larry Samuelson[2]

Department of Economics, University of Wisconsin, Madison, Wisconsin 53706-1320
LarrySam@ssc.wisc.edu

and

Avner Shaked

Department of Economics, University of Bonn, Adenaurallee 24-26, Bonn, Germany
shaked@glider.econ3.uni-bonn.de

Received July 31, 2001

[1] We thank Menesh Patel and John Straub for research assistance and thank Vince Crawford for helpful comments. The instructions and data for the experiments reported in this paper are posted at http://www.nyu.edu/jet/supplementary.html. Financial support from the ESRC Centre for Economic Learning and Social Evolution at University College London, the National Science Foundation, and the Deutsche Forschungsgemeinschaft, SFB 303 at the University of Bonn, is gratefully acknowledged.

[2] To whom correspondence should be addressed.

0022-0531/02 $35.00

This paper reports experiments with one-stage and two-stage alternating-offers bargaining games. Payoff-interdependent preferences have been suggested as an explanation for experimental results that are commonly inconsistent with players' maximizing their monetary payoffs and performing backward induction calculations. We examine whether, given payoff-interdependent preferences, players respect backward induction. To do this, we break backward induction into its components, subgame consistency and truncation consistency. We examine each by comparing the outcomes of two-stage bargaining games with one-stage games with varying rejection payoffs. We find and characterize systematic violations of both subgame and truncation consistency. *Journal of Economic Literature* Classification Numbers: C70, C78.

Key Words: bargaining; experiments; backward induction; subgame-perfect equilibrium; interdependent preferences.

1. INTRODUCTION

Experimental subjects frequently fail to play subgame-perfect equilibria in one-stage and two-stage alternating-offers bargaining games. A common response is to question the implicit assumption that players' monetary payoffs and utilities are synonymous. A variety of alternative utility functions have been suggested, typically allowing for "interdependence," or the possibility that a player's utility depends upon his opponent's as well as his own monetary payoff.

These alternative utility functions allow some reconciliation of the theory and experimental results, but leave open the original question: Does play respect backward induction? And if not, how can the departures from backward induction be characterized? This paper reports an experiment which investigates these questions.[3]

Once we abandon the equivalence of monetary payoffs and utility, we are left without a precise idea of what determines utility. Then how can we examine backward induction? Section 2 makes this question more precise and sets the stage for our analysis by splitting backward induction into two components, subgame consistency and truncation consistency. Section 3 describes the experimental procedure used to examine subgame and truncation consistency. Section 4 presents and discusses the results. We find systematic violations of backward induction that cannot be explained by payoff-interdependent preferences. For example, proposers are less aggressive in the second stage of a two-stage bargaining game than in an equivalent one-stage game (violating subgame consistency). Players are less responsive to variations in the expected value of playing a subgame than to

[3] The instructions used in the experiment and the data are posted at http://www.nyu.edu/jet/supplementary.html.

equivalent variations in terminal payoffs (violating truncation consistency). Section 5 concludes.

2. BACKGROUND

Bargaining Games. Figure 1 presents one-stage and two-stage alternating-offers bargaining games. We take the total surplus to be 100 and measure divisions of the surplus in terms of the percentage allocated to player 1, speaking of 1's actions as "demands" and 2's actions as "offers." The one-stage game is commonly called the Ultimatum Game.

The subgame-perfect equilibrium prediction is that player 1 receives all of the surplus in the Ultimatum Game (or at least all but a penny, if divisions must be made in multiples of pennies), and receives $100(1-D)$ of the surplus (with $100D$ going to player 2) in the two stage game, where D is the common discount factor. However, in the original study of the Ultimatum Game, Güth *et al.* [22] found that player 1's modal demand claimed only half of the surplus, and significantly more aggressive demands were often rejected. Binmore, Shaked and Sutton [7] found similar results, as have a multitude of subsequent studies, surveyed in Bolton and Zwick [13], Davis and Holt [17], Güth and Tietz [23], Roth [34], and Thaler [40].

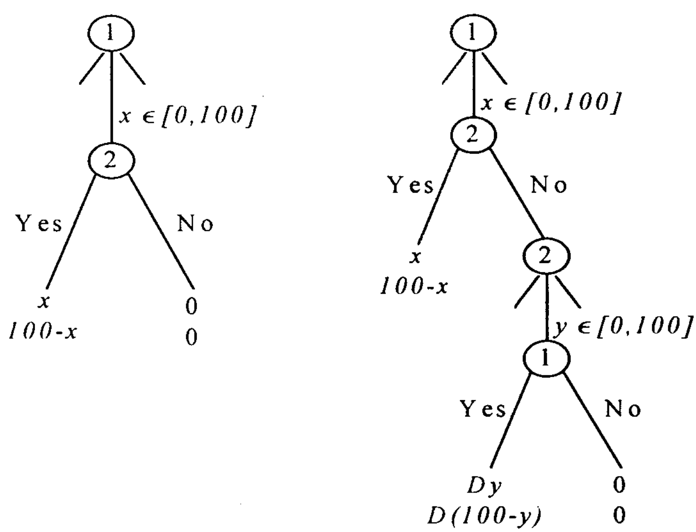

FIG. 1. One-stage and two-stage alternating-offers bargaining games.

Experimental outcomes in the two-stage game similarly tend to be less extreme than the subgame perfect equilibrium.[4]

Payoff-interdependent preferences. The experimental results are commonly interpreted as indicating that players have *interdependent preferences*, meaning that preferences depend upon more than simply one's own monetary payoff.[5] Bolton [10], for example, suggests that utility is increasing in one's own payoff and decreasing in the ratio of one's opponent's to one's own payoff, as do Ochs and Roth [31].

We concentrate on *payoff-interdependent* preference theories, in which preferences depend *only* upon the payoffs received by the players. For example, applying Bolton and Ockenfels' ERC (Equity, Reciprocity and Competition) [12] theory to two-player bargaining games, we can write player i's utility function as

$$u_i(\pi_i, \pi_j) = v_i\left(\pi_i, \frac{\pi_i}{\pi_i + \pi_j}\right), \tag{1}$$

where v_i is assumed to be increasing and concave in its first argument, and to be strictly concave in its second argument, with a zero derivative in the second argument when the latter equals $1/2$.[6] Player i thus prefers higher payoffs but dislikes inequality, and hence may prefer to reject quite asymmetric payoff allocations.

Alternatively, Fehr and Schmidt [21] work with a utility function (in two-player games)

$$u_i(\pi_i, \pi_j) = \pi_i - \alpha_i \max\{\pi_j - \pi_i, 0\} - \beta_i \max\{\pi_i - \pi_j, 0\}, \tag{2}$$

where $0 < \beta_i < \alpha_i$, so that player i dislikes inequality, and especially dislikes inequality in which i has the smaller payoff. Costa-Gomes and Zauner [16] examine a utility function whose deterministic part (supplemented by an error designed to facilitate empirical application) is given by

$$u_i(\pi_i, \pi_j) = \pi_i + \alpha_i \pi_j, \tag{3}$$

[4] Figure 5.6 of Davis and Holt [17, p. 272] provides a convenient summary of experiments with two-stage games. The results of Camerer *et al.* [14] and Johnson *et al.* [27], who examine the information-gathering patterns of experimental subjects, raise further questions concerning backward induction.

[5] A variety of experiments have investigated the fairness considerations which are often invoked to motivate interdependent preferences. Examples include Abbink *et al.* [1], Andreoni *et al.* [2], Andreoni and Miller [3], Bolton *et al.* [11], Bolton and Zwick [13], Dufwenberg and Gneezy [18], Kagel *et al.* [28], Ruffle [36], Slembeck [38], Straub and Murnighan [39], Winter and Zamir [41], Zwick and Chen [42], and Zwick and Weg [43].

[6] The function v is continuous and the utility when $\pi_i = \pi_j = 0$ is defined to equal $v_i(0, 1/2)$ (so that $u_i(\pi_i, \pi_j)$ is not continuous).

where α_i may be positive or negative, reflecting a positive or negative concern for the opponent's payoff.

We shall use (1)–(3) as illustrations, but our results apply to any *payoff-interdependent* utility function $u_i(\pi_i, \pi_j)$ that is strictly quasi-concave on sets of the form $\{(\pi_i, \pi_j): \pi_i + \pi_j = C\}$ (for some constant C).[7]

A more general interdependent utility specification would allow preferences to be based not only on realized payoffs, but also upon other characteristics of one's opponent or the structure of the game, including the alternative payoffs offered by unreached outcomes. In Levine [29], player i's utility may be increasing in j's payoff if j himself is relatively altruistic, while i's utility may be decreasing in j's payoff if j is similarly spiteful. Building on the theory of psychological games, Dufwenberg and Kirchsteiger [19], Falk and Firshbacher [20], and Rabin [33] offer alternatives in which the structure of the game, coupled with beliefs about opponents' intentions, plays an important role. This allows player i to prefer to be kind to kind opponents, but allows i's assessment of whether j has been kind to depend upon i's beliefs about what j believed about the consequences of j's actions.

We can be assured of the ability to construct interdependent preferences capable of reconciling experimental data and backward induction, as long as we allow sufficiently flexible preferences and examine a sufficiently narrow class of games.[8] For the interdependent-preferences approach to backward induction to be useful, we require a relatively parsimonious specification of preferences that is readily applicable across a relatively broad class of games. We say that such preferences are relatively "portable." Payoff-interdependent preferences are attractive because their simplicity makes them eminently portable. Coupled with the observation that such preferences are consistent with many experimental results, including violations of backward induction (Bolton and Ockenfels [12], Costa-Gomes and Zauner [16], Fehr and Schmidt [21]), this makes payoff-interdependent preferences particularly interesting.

Backward induction. In the Ultimatum Game, backward induction requires player 1 to choose 1's most preferred allocation, from the set of allocations that player 2 at least weakly prefers to disagreement. But when preferences exhibit payoff-interdependence, we do not have a precise idea of the latter set. Then how can we examine backward induction?

[7] Strict quasiconcavity ensures that backward induction solutions are unambiguous.

[8] For example, Dufwenberg and Kirchsteiger [19] argue that their theory creates sufficiently flexible self-referential links across the stages of the game as to render the concept of backward induction vacuous.

Our experimental approach begins by separating backward induction into its three components:[9]

- **Rationality**: Given a choice between two (vectors of) payoffs, a player chooses the most preferred.
- **Subgame consistency**: Play in a subgame is independent of the subgame's position in a larger game.
- **Truncation consistency**: Replacing a subgame with its equilibrium payoffs does not affect play elsewhere in the game.

In generic, finite games of perfect information, these three requirements are equivalent to backward induction, as captured by the equilibrium notion of subgame perfection.[10] In the case of the Ultimatum Game with ordinary preferences, rationality ensures that a player will always choose a positive amount of money rather than zero. Subgame consistency ensures that a player will accept when this same decision appears as the result of an opponent's offer. Next, truncation consistency allows us to replace this accept/reject decision with its equilibrium payoffs, and then rationality is once again invoked to examine the proposal that opens the game.

Interdependent-preference theories are designed to preserve the maintained assumption of rationality. In order to assess backward induction, our analysis accordingly presumes rationality and examines issues of subgame consistency and truncation consistency.

3. THE EXPERIMENTS

The games. Figure 2 presents the games involved in the experiments. Game III is the two-stage game of Fig. 1. We refer to game I as the Ultimatum Game, though the presence of the rejection payoffs (V_1, V_2) causes the game to differ from a standard Ultimatum Game. The rejection payoffs (Z_1, Z_2) in game IV are subject-specific, and are calculated on the basis of the subjects' realized payoffs in game II (details below).

[9] The concepts of subgame consistency and truncation consistency are taken from Harsanyi and Selten [24]. An alternative approach, used by Holt [25] to examine coordination games, would estimate subjects' utility functions, use these estimates to calculate the backward-induction solution, and then compare the calculated solution with the outcomes of further experiments.

[10] In nongeneric, finite games of perfect information, subgame perfection may also require appropriate tie-breaking rules (Harsanyi and Selten [24, pp. 106–109]). Strict quasiconcavity ensures that the relevant genericity condition is satisfied in our games. In games of imperfect information, the possibility of subgames with multiple Nash equilibria obviously allows subgame-perfect equilibria to violate subgame consistency.

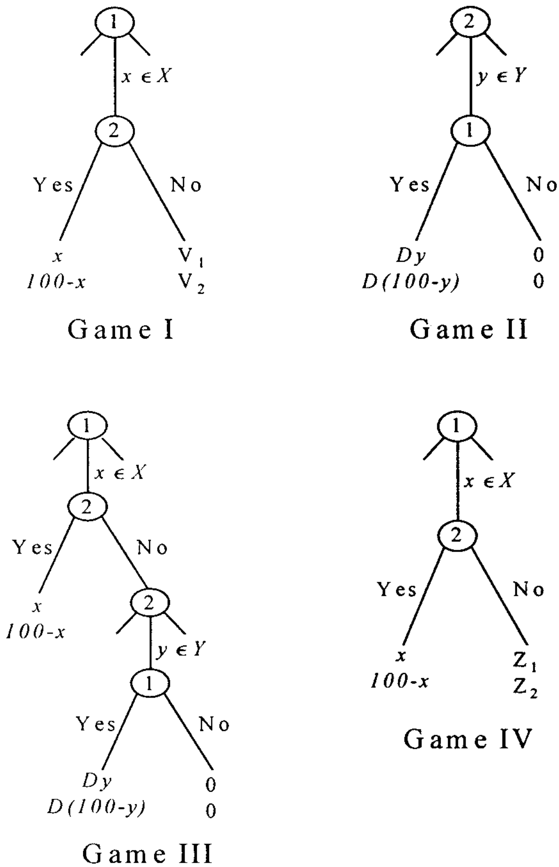

FIG. 2. Experimental games.

Twenty four treatments were run, with all four of games I-IV played in each treatment, and with one treatment for each of the twenty four elements of the set

$$\{(10, 10), (70, 10), (10, 60)\} \times \{.2, .3, .4, .5, .6, .7, .8, .9\},$$

where the first element identifies a rejection-payoff pair (V_1, V_2) that appeared in game I (only) and was common to all of the subjects in the treatment, and the second element identifies the discount factor (again common to all subjects within a treatment) that appeared in games II and III of the treatment.

The Ultimatum Game. The first of our four games, the Ultimatum Game, serves three purposes.

First, the Ultimatum Game with rejection payoffs (10, 10) provides a control. We regard results similar to those obtained in other Ultimatum Game experiments as an essential indication that there is nothing in our procedures that prevents replication of standard experimental results, and would reexamine our procedures in the absence of such results. We choose the Ultimatum Game with the rejection payoffs (10, 10) as a control, rather than the standard Ultimatum Game, to check that the mere introduction of (relatively small) rejection payoffs does not significantly change subjects' behavior.

Second, the Ultimatum Game provides a check on our intuition as to how players respond to varying rejection payoffs. We expect play in the (10, 60), (10, 10), and (70, 10) rejection-payoff games to differ, with player 1 becoming increasing aggressive across these three games, and would again reexamine our procedures if this were not the case.

Finally, the (10, 60) rejection-payoff game provides a setting in which a common form of payoff-interdependence makes a particularly sharp prediction. Payoff-interdependent models typically assume that utility is increasing in one's own payoff and (possibly weakly) decreasing in inequality, as do Bolton and Ockenfels [12], Fehr and Schmidt [21], and as does the model of Costa-Gomes and Zauner [16] when $-1 < \alpha_i < 0$ (in which case it is illuminating to let $u_i(\pi_i, \pi_j) = (1+\alpha)\,\pi_i + \alpha(\pi_j - \pi_i)$). Hence, player 2 should prefer to accept any allocation in which 2 receives at least sixty percent of the surplus. Player 1 will then demand at least 40, and player 2 will accept 1's demand.

Subgame consistency. The second stage of the two-stage bargaining game is itself an Ultimatum Game, with player 2 making the initial proposal and with the total surplus given by $100D$ rather than 100. Game II duplicates this second stage as a separate game. We shall refer to game II as the *continuation game*. We investigate subgame consistency by comparing play in the continuation game with play in the second stage of the two-stage game:[11]

- Subgame consistency indicates that play in the second stage of the two-stage game, for those cases in which it is reached, should duplicate play in the continuation game.

[11] Violations of subgame consistency are readily found in games with imperfect information and hence multiple backward-induction equilibria, in which the case for subgame consistency is less obvious. See, for example, the Cooper *et al.* [15] coordination-game experiments. Much less is known about games with unique backward induction outcomes. In an experiment involving centipede games of varying length, McKelvey and Palfrey [30] find encouraging results concerning subgame consistency.

Isn't the mere fact that the second stage is reached evidence that backward induction fails? If players' preferences are commonly known, the answer is yes. However, different subjects may have different (interdependent) preferences. Player 1 may then be uncertain about the interdependent preferences of the (anonymous) opposing player 2, and hence may optimally make a first-stage demand that some player 2*s* reject, leading to the second stage.

In the presence of such heterogeneity, differing player-2 offers in the second stage of the two-stage game and in the continuation game may reflect not a failure of backward induction, but rather that player 2 has inferred something about player 1, and hence about 2's optimal second-stage offer, from the demand that 2 rejected to reach the second stage of the two-stage game.[12] Depending upon how prior beliefs are specified and how beliefs are updated in response to zero-probability demands, we can construct subgame-perfect equilibria that will account for virtually any outcome. But can this be done with beliefs that are sufficiently straightforward as to yield a useful theoretical model?[13] The evidence suggests that the updating of beliefs helps very little in explaining player 2's observed play in the second stage of the two-stage game. Player 2s reject a wide variety of demands in the first stage of game III. If player 2 can draw inferences about player 1 from the latter's period-1 demand, then we would expect player 2's period-2 offer to vary significantly with the identity of the demand rejected by player 2 to reach the second stage. We find no evidence of such a relationship.

Heterogeneous preferences also raise the possibility that the player 2s who reach the second stage of the two-stage game are a biased sample of the complete set of player 2s who participate in the continuation game. We can eliminate this potential selection bias by restricting attention to the continuation-game play of those player 2s who reach the second stage when playing the two-stage game. Doing so only exacerbates (slightly) the observed behavioral differences between the two games.

Truncation consistency. We next turn to truncation consistency. A rejected demand in game IV leads to the rejection payoffs (Z_1, Z_2). A pair of values (Z_1, Z_2) is assigned to each experimental subject (according to a

[12] In the presence of incomplete information, we must now work with perfect Bayesian or sequential rather than subgame-perfect equilibrium, as well as appropriate generalizations of subgame consistency and truncation consistency in terms of Markov perfection. Because we find that incomplete information does not play an important explanatory role, we forego a formal development, following the common practice of retaining the terms subgame perfection, subgame consistency, and truncation consistency.

[13] This consideration is reminiscent of the observation that *some* preferences must exist which support backward induction, while our interest centers on preferences that are sufficiently portable, such as payoff-interdependent preferences.

method that will be important in the next subsection but is not relevant here). When two subjects are matched to play game IV, the value of Z_1 for that interaction is the corresponding value assigned to the subject who plays as player 1, while Z_2 is the corresponding value assigned to the subject who plays as player 2 in game IV. The values Z_1 and Z_2 thus vary across instances of game IV, with subjects always completely informed as to the relevant values.

We can estimate a function describing the relationship between player-1 demands in game IV and the rejection payoffs (Z_1, Z_2). Similarly, we can examine a function describing the relationship between player-1 demands in game III and the anticipated values (Z_1^{III}, Z_2^{III}) of play in the second stage of game III, where we estimate the latter values on the basis of the observed second-stage play in game III. If truncation consistency holds, then a change in a game-IV rejection payoff should have the same effect on player-1 demands as an equivalent change in the anticipated value of the game-III second stage:

- Truncation consistency indicates that play in game IV should bear the same relationship to the rejection values (Z_1, Z_2) as does play in the first stage of game III to the anticipated payoffs (Z_1^{III}, Z_2^{III}) of the second stage of the two-stage game.

The primary difficulty here involves identifying and estimating the appropriate anticipated value (Z_1^{III}, Z_2^{III}) of playing the second stage of game III. We find that our results are insensitive to a variety of alternative measures of (Z_1^{III}, Z_2^{III}).

Subgame and truncation consistency. Games III and IV differ in that a rejection of a first-period demand in game III leads to a copy of the continuation game, while a rejection in game IV leads to the fixed rejection payoffs (Z_1, Z_2). The latter payoffs are calculated on the basis of observed play in the *continuation* game. A pair of values (Z_1, Z_2) is calculated for each experimental subject, one describing the experience of that subject as player 1 in the continuation game, and one describing the subject's experience as player 2 in the continuation game. When two subjects are matched to play game IV, Z_1 is the estimated continuation-game value for the subject who plays as player 1 in game IV, and Z_2 the estimated continuation-game value for the subject who plays as player 2 in game IV.

If subgame consistency holds, then the continuation-game payoffs (Z_1, Z_2) provide an estimate of the value of entering the second stage of the two-stage game. If truncation consistency holds, then it should not matter whether a first-stage rejection leads to the second-stage game or to the payoff pair (Z_1, Z_2). Hence:

- Subgame and truncation consistency indicate that experimental play in game IV should duplicate that of the first stage of game III.

The primary difficulty here involves ensuring that Z_1 and Z_2 are good estimates of the value of playing the continuation game. Notice that the problem now involves not (Z_1^{III}, Z_2^{III}), which are estimates that appear only in our *analysis* of truncation consistency and whose properties we can examine and adjust in the course of our empirical investigation, but values (Z_1, Z_2), which appear in the specification of game IV and hence whose calculation must be fixed as part of the experimental design.[14]

Procedures. The experiments were conducted at University College London in the fall of 1998 with undergraduate subjects. Each of the twenty four (one for each possible combination of three rejection payoff pairs ((10, 10), (70, 10), and (10, 60)) and eight discount factors (.2, .3, .4, .5, .6, .7, .8, and .9)) treatments involved ten subjects, for a total of 240 subjects. Each treatment consisted of eighty rounds, with the ten subjects matched into five pairs for each round, with each pair playing one game. Game I was played in the first twenty rounds, game II in the next twenty, game III in the penultimate twenty rounds, and game IV in the final twenty rounds. We thus have a total of 400 games in each treatment of the experiment and an overall total of 9600 games, 2400 each of games I, II, III, and IV. In each of the four games, each subject played about half of the time as player 1 and half of the time as player 2, with the "about" reflecting the fact that roles were assigned randomly. Each of the ten subjects in a treatment could be matched with each of the nine opponents.

All subjects play the four games in the same order. A more complete experimental design would add another dimension to the definition of a treatment, corresponding to different orders in which the four games are played and allowing us to test for the possibility that the results are sensitive to the order of play. Our theoretical design places some constraints on this order, in that game II must be played before game IV so that game-II outcomes can be used in defining the game-IV rejection payoffs (Z_1, Z_2). Even after incorporating this constraint, investigating all possible orderings would require twelve times as many treatments. We discuss possible evidence of order effects as we proceed.

Instructions were provided via a self-paced, interactive computer program that introduced and described the experiment, and provided practice in how to make choices in each of the four games. The surplus was pictured as a wedge-shaped slice of "cake," as shown in Fig. 3. In games I

[14] Fortunately, this problem does not arise in the test of truncation consistency described in the preceding subsection, where (Z_1, Z_2) need not bear any relationship to the value of the continuation game or second stage of game III.

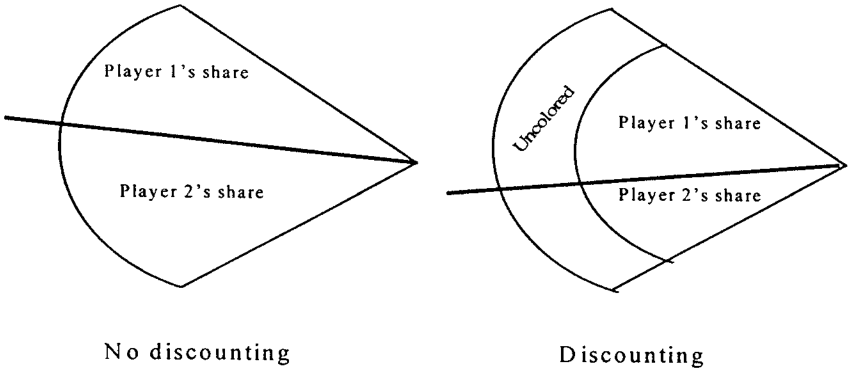

FIG. 3. Representation of the bargaining games. The wedges were outlined in white against a black background. The interior of the entire wedge was colored light brown in the no-discounting case, while only the area within the inner boundary was colored in the discounting case.

and IV, the left wedge of Fig. 3 appeared, along with two smaller wedges with areas corresponding to the appropriate rejection payoffs of players 1 and 2.

In game II, the right wedge of Fig. 3 appeared, capturing the fact that payoffs in the event of an agreement were discounted. In game III, both wedges appeared, one corresponding to each stage of the game, with the second stage being somewhat fainter while the first stage was being played. Discounting was captured by coloring only an inner segment of the second wedge whose area corresponded to the discounted value of the cake.

Demands and offers were made by using the arrow keys to move a line that rotated about the point of the wedge, with player 1's share lying above the line and player 2's below. To avoid suggesting focal points, there were initially no numbers on the screen. Once a tentative division was proposed, the percentage of the cake going to each player was indicated, as was the equivalent number of "tickets" going to each player. The percentages always added to 100. The number of tickets added to 100 in the absence of discounting, and in a discounted stage (game II or the second stage of game III) was given by $100D$, where D was the discount factor. Players then had a chance to confirm or revise their choice.

After each twenty rounds, and hence after each of games I, II, III, and IV, an electronic roulette wheel was spun whose surface was divided into "win" and "lose" areas, with the former being proportional to the number of tickets won in the previous twenty rounds of play.[15] A win paid six

[15] Our purpose was not so much to control for risk aversion, as expected-utility maximizers are likely to be risk neutral over the relatively small amounts of money involved in the experiment, but to provide an interlude when switching from one game to the next.

Rounds	(V_1, V_2)	Observations	Mean demand	Median	5th %tile	95th %tile
1–10	(10, 10)	400	64.9	65	50	80
10–20	(10, 10)	400	66.8	68	55.5	76.5
1–10	(70, 10)	400	82.8	85	64	90
10–20	(70, 10)	400	83.4	84	79	88
1–10	(10, 60)	400	39.8	38.5	28	56.5
10–20	(10, 60)	400	36.0	36	26	49

		All demands		Demands in [30, 40]		Demands in [70, 80]	
		Obs.	R %	Obs.	R %	Obs.	R %
1–10	(10, 10)	400	24			111	48
10–20	(10, 10)	400	19			146	34
1–10	(70, 10)	400	30			72	1
10–20	(70, 10)	400	22			61	0
1–10	(10, 60)	400	36	266	23		
10–20	(10, 60)	400	30	313	24		

FIG. 4. Player-1 demands (measured in terms of the percentage of the surplus demanded by player 1) and player-2 rejection rates (R%) in the Ultimatum Game. There were five games per round in each of eight treatments for each (V_1, V_2) specification, for a total of 800 observations (or "Obs.") over twenty rounds. No demands from [30, 40] were rejected in the (10, 10) and (70, 10) cases, and none from [70, 80] were accepted in the (10, 60) case.

pounds, which was then worth slightly less than ten dollars. Together with a six-pound initial fee, subjects' earnings were then drawn from the set $\{6, 12, 18, 24, 30\}$ pounds, with these amounts being won by 4, 28, 79, 106, and 23 subjects, respectively, for an experiment that took about two hours.

4. RESULTS

4.1. *Game I: The Ultimatum Game*

We begin with game I. Figure 4 reports player-1 demands and provides information on player 2's response to those demands.

First, the results for rejection payoff (10, 10) are much like those of conventional Ultimatum-Game experiments. The mean and median demands are both near two-thirds of the surplus. A significant number of demands are rejected, with higher rejection rates for higher demands. Demands are slightly higher in the final ten rounds than in the first ten rounds of play, and the distribution of demands is somewhat tighter in the final ten rounds (cf. the 5th and 95th percentiles, and notice that this tightening contributes to the reduction in rejection rates), but the changes are small relative to the variation in Ultimatum-Game results reported in the literature.[16] Our experiment replicates familiar Ultimatum Game results.

Second, player-1 demands increase, as expected, as the rejection payoffs change from (10, 60) to (10, 10) to (70, 10). These differences are significant: over the final ten rounds of play, the 90-percentile intervals for the observed demands made under the three specifications, given by

$$[26, 49], \qquad [55, 77], \qquad [79, 88],$$

are disjoint.

When rejection payoffs are (70, 10), player 1 is ensured a payoff (70) larger than player 1 conventionally receives in the ordinary Ultimatum Game. Player 1's mean demand in this case is approximately 83, with a 90-percentile interval (over the last ten rounds) of [79, 88], leaving player 2 with little more than the rejection payoff of 10. This willingness of player 1s to make such aggressive demands reflects the sensitivity of rejection rates to rejection payoffs. Figure 4 reports that when rejection payoffs are (10, 10), 40 percent (103 of 257, over all 20 rounds) of demands in the interval [70, 80] are rejected. Only .75 percent (1 of 133) are rejected when the rejection payoffs are (70, 10). Figure 5 provides an additional summary of the behavior of player 2s. Many player 2s are thus willing to settle for 20 or 25 percent of the surplus if 1 has a rejection payoff of 70, but not if 1's rejection payoff is a mere 10, revealing interdependent preferences for these player 2s of the form:[17]

$$(10, 10) \succ (75, 25) \succ (70, 10). \tag{4}$$

The (70, 10) outcome is consistent with the behavior that would appear if players mentally "assigned" the rejection payoffs 70 to player 1 and 10 to

[16] We frequently compare results for the first and last ten rounds, and often focus on the last ten rounds, in order to isolate any initial subject confusion. The differences are quite small compared to those involved in the hypotheses of interest.

[17] Would player 2 exhibit the preferences $(10, 10) \succ (75, 25)$ if the choice were given exogenously, rather than arising as a result of player 1's choice? If not, we must question the portability of payoff-interdependent preferences. The generalized dictator games of Andreoni and Miller [3], in which some dictators give away all of the surplus when faced with exogenously-imposed tradeoffs that make it efficient to do so, suggest a negative answer. The experiments of Slembeck [38], in which rejection rates were higher when subjects were presented with exogenously-determined choices, and of Blount [9], in which subjects did not have significantly smaller rejection thresholds when facing a "disinterested" proposer (who did not receive part of the surplus) as when facing an ordinary proposer, are less clear.

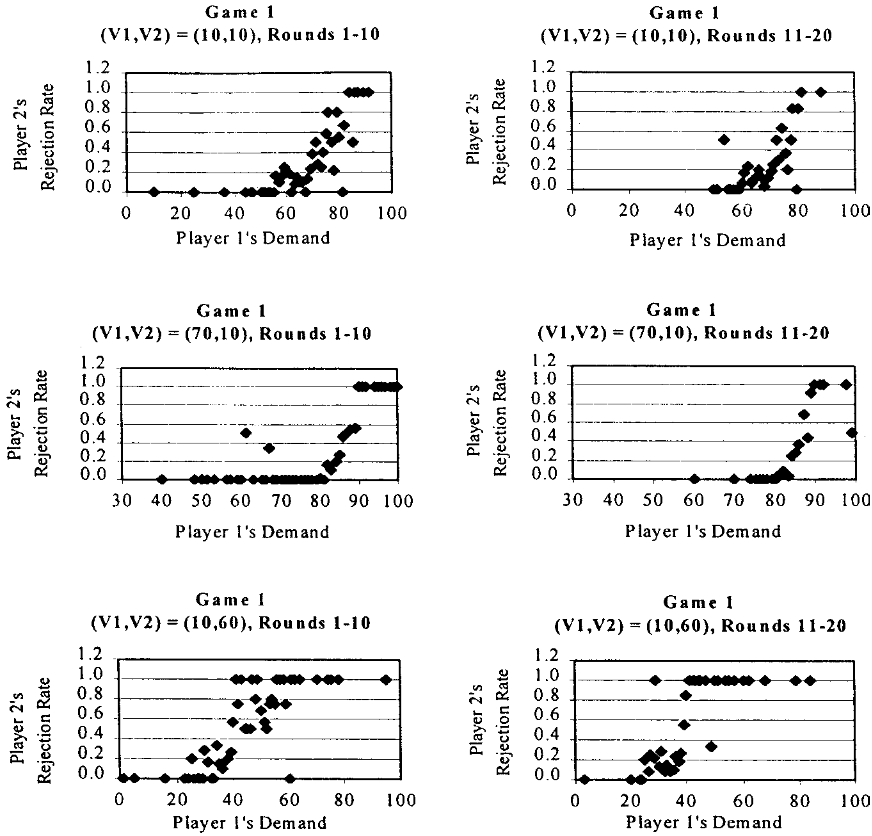

FIG. 5. Rejection rates in the Ultimatum Game.

player 2, and then bargained over the remaining surplus as if they were in an ordinary Ultimatum Game with a total surplus of size twenty. For example, player 1's mean demand allocates slightly less than two-thirds of this remaining surplus to player 1. This reaction to rejection payoffs contrasts with the findings of Binmore *et al.* [5, 6, 8], where players appear to ignore outside options that pose no constraint on the agreement that would be reached in the absence of such an option and make demands close to constraining outside options. In the (70, 10) case, such behavior would produce the agreement that comes closest to that of the ordinary Ultimatum Game while still respecting player 1's rejection payoff, giving player 1 a payoff of (perhaps just over) 70. More importantly, approaching the surplus remaining after rejection payoffs have been covered as an ordinary Ultimatum Game yields results that contrast sharply with those of the (10, 60) rejection-payoff case, described below.

Finally, the rejection payoffs (10, 60) allow an examination of the most common form of payoff-interdependent theories. If subjects value their own monetary payoff but dislike inequality, as in the models of Bolton and Ockenfels [12] and Fehr and Schmidt [21], then subgame perfection calls for player 1 to demand at least forty percent of the surplus, and for such a demand to be accepted.

Figure 4 shows that player 1s initially demand about forty percent of the surplus (the mean player-1 demand over the first ten rounds is 39.8 and the median 38.5), coming very close to the subgame-perfect equilibrium, and with their demands drifting downward (a mean demand of 36 over the last ten rounds). However, rejection rates are the highest in this treatment, with almost a quarter of the cases in which player 1 demands between 30 and 40 percent of the surplus ending in rejection.[18] A payoff of sixty or slightly higher is not enough to ensure acceptance from player 2.

In summary, we find that (1) our game-I results include replications of standard results for the Ultimatum Game; (2) subjects respond to rejection payoffs, with proposers increasing their demand in response to a high rejection payoff and decreasing their demand when the opponent's rejection payoff is high; and (3) findings for the (10, 60) rejection-payoff specification suggest that something in addition to payoff-interdependent preferences, at least in inequality-aversion forms such as (1)–(2) and (3) (with $\alpha_i \in (-1, 0)$), lies behind the results.

4.2. *Game II: The Continuation Game*

Figure 6 provides information on offers and rejection rates in game II. The results in Figure 6 are reported in terms of the share of the surplus *offered* to player 1, noting that player 1 was the responder in this game.

Game II is an Ultimatum Game, with rejection payoffs of zero. Once we make allowance for the reversal of roles, we again obtain results consistent with previous Ultimatum-Game experiments. On average, the proposer demands between sixty and seventy percent of the surplus. Higher demands on the part of the proposer are likely to be rejected, with about twenty percent of plays ending in rejection. Figure 7 shows the mean and median percentage of the surplus offered to player 1 by round, showing some initial adjustment followed by relatively unchanging behavior.

[18] Behind the reduction in mean player-1 demand is a more pronounced tightening of the distribution of demands. Over the course of the twenty rounds, there are 146 cases (out of 800) in which player 1 demands more than forty percent, and hence offers player 2 less than her rejection payoff. However, half (72 of 146) of these demands come in the first four rounds of play. As Fig. 5 shows, some demands that leave player 2 with less than the rejection payoff of 60 are accepted in the early rounds of play, but this behavior has virtually disappeared in the final ten rounds.

Rounds	(V_1, V_2)	Observations	Mean offer	Median	5th %tile	95th %tile
1-10	All	1200	33.4	33	15	50
11-20	All	1200	32.2	31	20	48.5
1-10	(10, 10)	400	32.8	33	19	48
11-20	(10, 10)	400	30.5	30	20	41.5
1-10	(70, 10)	400	28.3	27.5	12	45
11-20	(70, 10)	400	28.0	28	18.5	40
1-10	(10, 60)	400	39.2	41	16.5	55
11-20	(10, 60)	400	38.0	38	24	50

Discount factor	Rejection rates (%)		
	(10, 60)	(70, 10)	(10, 10)
0.2	27	15	22
0.3	22	23	23
0.4	22	23	18
0.5	27	18	16
0.6	25	18	19
0.7	17	22	18
0.8	20	27	12
0.9	17	23	6

FIG. 6. Player-2 offers (the percentage of the surplus offered to player 1) and player-1 rejection rates in game II. There were five games per round in each of eight treatments per rejection payoffs, for a total of 800 offers in each specification. Rejection rates are given for each combination of rejection payoff and discount factor, over all twenty rounds of play.

Subjects who faced different rejection payoffs in game I face precisely the same game II. However, Figs. 6 and 7 suggest that offers in game II vary systematically with the rejection payoffs that prevailed in game I. Subjects who faced rejection payoffs (10, 60) in game I offer more of the surplus to the responder than subjects who experienced rejection payoffs (10, 10), who in turn offer slightly more than those who experienced rejection payoffs (70, 10). Hence, rejection payoffs that induced more asymmetric divisions of the surplus in game I correspond to game II outcomes with more asymmetric divisions of the surplus. We report tests indicating that these differences are statistically significant in the next section, in the course of comparing games II and III.

It thus appears as if the different game-I specifications condition players to coordinate on different outcomes in game II.[19] On the one hand, this finding provides evidence that subgame consistency fails. At the same time, these results suggest that there are spillovers between games, and hence that

[19] This is reminiscent of the experimental findings of Roth *et al.* [35], who find different conventions in Ultimatum-Game experiments in different countries.

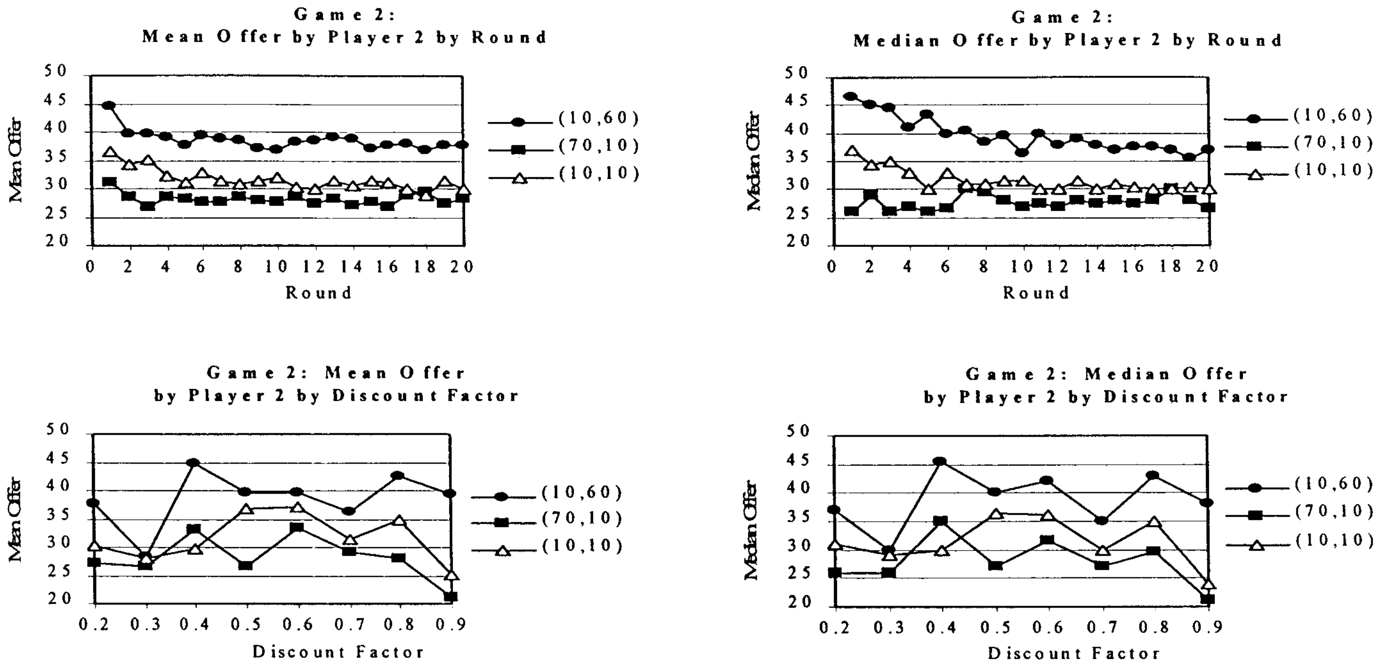

FIG. 7. Mean and median offers made by player 2 to player 1 in game II, by round and by discount factor, measured as a percentage of the surplus.

the order in which the games are played can matter. The next section shows that play in game III does not vary significantly with the game-I rejection payoff, perhaps because games I and III are less similar than I and II. Could it be that subgame consistency appears to fail, in the form of differing behavior in game II and the second stage of game III, simply because game II (and not game III) was affected by previous experience in game I? The differences in game II and the second stage of game III persist even when each of the game-I rejection-payoff cases is examined separately. In addition, these differences take the same direction in each case, even though the three game-I rejection payoffs involve quite different allocations between the two agents and hence would be expected to push game-II behavior in quite different directions. Finally, the effect of game-I rejection payoffs on game-II outcomes are small in comparison to the differences between game II and game III, suggesting that order effects do not lie behind the results.

In contrast to rejection payoffs, differing discount factors, which determine the size of the surplus to be divided, have little effect on the outcome. Figure 7 shows mean and median offers by discount factor, revealing no systematic relationship.

In summary, our game-II results are again consistent with standard Ultimatum-Game findings.

4.3. *Games II and III: Subgame Consistency*

We now investigate subgame consistency by comparing behavior in the continuation game with that of the second stage of the two-stage game. Of the 2400 initial demands made in the two-stage game, 501 (20.875%) were rejected. Figure 8 summarizes behavior in the second stage of the two-stage game. As in game II, game-I rejection payoffs are irrelevant in game III, but we will typically report the results for different game-I rejection payoff cases separately.

The mean and median offers for the three rejection-payoff specifications are much closer to one another than they were in game II. Their ranking has also shifted, with the most generous offer now attached to rejection payoff (10, 10) rather than (10, 60). Finally, there is also little pattern to the relationship between discount factors and offers, and there is relatively little difference between the first and last ten rounds of play.

Figures 6 and 8 show that play in the continuation game and the second stage of game III differ. Proposers are more generous in the second stage of game III, offering a mean percentage of 43.5 (median 44) of the surplus to player 1, as opposed to only 32.8 (median of 32) in the continuation game. This pattern of more generous offers in the second stage of game III holds for every rejection payoff and every discount rate.

Specification	Rounds	Observations	Mean offer	Median	5th %tile	95th %tile
All	1-10	276	43.3	44	23	60
All	11-20	225	43.8	45	25	60
All	1-20	501	43.5	44	25	60
(10, 60)	1-20	136	43.3	44	24	52
(70, 10)	1-20	197	42.0	44	20	59
(10, 10)	1-20	168	45.4	45	32	62
.2	1-20	37	48.6	45	40	62
.3	1-20	47	46.0	47	26	60
.4	1-20	39	43.7	48	15	52
.5	1-20	56	40.3	44	10	75
.6	1-20	44	38.9	37.5	16	53
.7	1-20	70	44.6	45	26	59
.8	1-20	90	41.7	40.5	25	52
.9	1-20	118	44.7	45	33	55

FIG. 8. Player-2 offers in the second stage of game III, the two-stage game, measured as the percentage of the surplus offered to player 1. "Observations" reports the number of games in which player 2 caused the second stage to be reached by rejecting player 1's first-stage demand. The first three lines report all cases, with a total of 2400 possible observations over the course of twenty rounds. There are 800 possible observations for each game-I rejection-payoff specification reported in the next three lines, and 300 possible observations for each discount-factor specification.

Are these differences statistically significant? To address this question, we require an analysis that respects the panel nature of the data. Figure 9 reports the results of a random effects regression, with a transformation of player 2's offer as the dependent variable and with the independent variables including an intercept capturing the base case of game II and rejection payoff (10, 10), and five dummy variables that identify the five remaining combinations of a game (II or III) and one of the three possible rejection payoffs.[20]

The "Game II, (70,10)" and "Game II, (10,60)" coefficients reported in Fig. 9 identify departures of the game II, (70,10) and game II, (10,60)

[20] The transformation of player 2's offer y to $\log y/(100-y)$, taking $[0, 100]$ into $[-\infty, \infty]$, allows us to capture the restriction that offers must lie in the interval $[0, 100]$. The random effects estimator allows us to capture the fact that the multiple offers of a single player are likely to be correlated. Offers may depend upon the history of opponents' actions observed by the offerer. It is appropriate to omit this history from the regression as long as it is not correlated with the offerer-specific error term. Such a correlation could appear, as player i's play could affect the subsequent behavior of opponent j and hence the subsequent history of opponent actions observed by player i, and our implicit assumption is that the resulting correlation is not too large.

Independent variable	Estimated coefficient	Standard error	p-value
Intercept (Game II, (10,10))	-.80	.038	.000
Game II, (70,10)	-.19	.054	.001
Game II, (10,60)	.30	.054	.000
Game III, (10,10)	.59	.033	.000
Game III, (70,10)	.56	.059	.000
Game III, (10,60)	.49	.062	.000

FIG. 9. Random-effects regression results. The dependent variable is $\log y/(100-y)$, where y is player 2's offer in either the continuation game or the second stage of game III. There are 2901 observations, 2400 from the continuation game and 501 from the second stage of game III. Independent variables include an intercept capturing the base case of game II and rejection payoff (10, 10), and dummies capturing departures from the base case for the five remaining combinations of games (II or III) and game-I rejection payoff ((10, 10),(70, 10), or (10, 60)). "p-value" is the probability that, given a parameter value of zero, a test statistic appears with absolute value (i.e., a two-tailed test) at least that of the calculated statistic.

rejection payoff cases from the game II, (10,10) base case captured by the intercept (shown in bold). These coefficients show that game-II offers are significantly related to game-I rejection payoffs, being highest in the (10, 60) case and lowest in the (70, 10) case. In contrast, a test of the game-III coefficients reveals that offers in game III do not vary significantly in the game-I rejection payoff.

More importantly, the estimated "Game III, (10,10)" coefficient indicates that for the case in which game-I rejection payoffs were (10, 10), game-III offers are higher than game-II offers at any conventional significance level (i.e., a two-tailed test p-value of .000), in contrast to the prediction of subgame consistency. It is straightforward to calculate that game-III offers are also higher, at similar significance levels, for the (70,10) and (10,60) rejection-payoff cases.

Why are proposers more generous in the second stage of the two-stage game? Figure 10 compares rejection rates in the continuation game and the second stage of the two-stage game.

The rates are not too dissimilar, being 29 percent in the second stage of the two-stage game and 20 percent in the continuation game. However, these aggregate rates hide the fact that offers are significantly higher in the second stage of the two-stage game. There are much larger differences in rejection rates conditional on offers. The range of offers [30–40] lies below a typical offer in the second stage of the two-stage game (a mean of 43.5 and median 44) while containing near its bottom end a typical offer in the continuation game (mean of 32.8 and median of 32). Figure 10 shows that the rejection rate in the second stage of game III is just over three times that of the continuation game, over this range of offers. The range [35–45] lies above a typical offer in the continuation game, and contains near its upper end a typical offer in the second stage of the two-stage game. Here,

Game	Observations	Offers: All	30–40	35–45
II	2400	20%		
III	501	29%		
II	1064		15%	
III	131		47%	
II	733			9%
III	190			36%

FIG. 10. Rejection rates for the continuation game (game II) and the second stage of the two-stage game (game III), in percentages. There were 2400 plays of each game. In the two-stage game, 501 of these plays reached the second stage.

we find rejection rates four times higher in the second stage of the two-stage game. Proposers have good reason to be more generous in the second stage of the two-stage game, because responders are much more likely to reject less generous offers.

Are the differences in behavior between the second stage of the two-stage game and the continuation game economically important? Let y_m^{II} be the median offer made in the continuation game, and let y_m^{III} be the median offer in the second stage of the two-stage game. How much would a proposer sacrifice by making offer y_m^{III} in game II? How much by making offer y_m^{II} in game III? Figure 11 reports the results. The first eight lines report, for each discount factor, the expected payoff that one would achieve in game II by making the game-II median offer y_m^{II}, and by making the game-III second-stage median offer y_m^{III}. The second eight lines report the payoffs that these offers would receive in the second stage of game III.[21]

In game II, for every discount factor, one is better off making the game-II median offer than the (higher) game-III median, with the latter sacrificing between 3 (D = .6) and 27 (D = .9) percent of the former's expected payoff. These relatively small differences reflect the fact that higher offers reduce the surplus from each agreement, but sacrifice no agreements, eliciting a somewhat higher acceptance probability. Results are more dramatic in the second stage of game III. The game-II median offer is sufficiently low as to garner no acceptances in six of the eight cases, hence sacrificing

[21] To calculate the expected payoff of an offer y, we must estimate the expected acceptance rate attached to the offer. We first calculated the observed acceptance rate of each offer that appears in the data, given by the proportion of the times the offer was accepted. The "expected" acceptance rate of offer y is then chosen to minimize the sum of the number of higher offers with lower observed acceptance rates and the number of lower offers with higher observed acceptance rates. The minimizer was unique in 24 cases and was an interval in the remaining 8, in which case we chose the midpoint of the interval.

Game	D	II-median, y_m^{II}	Payoff	III-median, y_m^{III}	Payoff
II	.2	31	13	45	11
II	.3	28	19	47	16
II	.4	35	23	48	21
II	.5	33.5	33	44	28
II	.6	35	34	37.5	33
II	.7	31	41	45	39
II	.8	35	52	40.5	48
II	.9	25	68	45	50
III	.2	31	0	45	7
III	.3	28	0	47	13
III	.4	35	0	48	9
III	.5	33.5	0	44	18
III	.6	35	32	37.5	38
III	.7	31	0	45	39
III	.8	35	15	40.5	27
III	.9	25	0	45	41

FIG. 11. Comparison of expected monetary payoffs in the continuation game (Game II) and the second stage of the two-stage game (Game III), by discount factor. The y_m^{II} and y_m^{III} columns report the median offers made in games II and III for the relevant discount factor. In the first eight lines, each offer is followed by the expected payoff the offer would receive if made in Game II. In the second eight lines, each offer is followed by the expected it would receive if made in the second stage of Game III.

100 percent of the expected payoff. This reflects the rejection dangers associated with more aggressive offers. The implication is that the differences in game-II and game-III behavior have important payoff consequences, with it being disastrous to treat the second stage of game III as if it were game II.

The differing outcomes of the continuation game and the second stage of game III would be convincing evidence of a failure of subgame consistency if preferences were known and identical across players. However, if preferences are payoff-interdependent, then players may be incompletely informed about their opponents' possibly heterogeneous preferences. This raises two considerations.

First, could proposers in the second stage of the two-stage game be simply reacting to information gleaned about their opponents' preferences from the offer made by their opponents in the first stage? It is not *a priori* clear which direction this information updating should take. A relatively aggressive demand on the part of player 1 may reveal that 1 is intent on a large share, and hence that player 2 should make a relatively generous offer to player 1 in the next round. Alternatively, an aggressive demand may indicate that 1 is relatively unconcerned with relative-payoff considerations, and hence that 2 can safely make a quite niggardly offer.

	Accepted demands			Rejected demands		
D	5th %tile	median	95th %tile	5th %tile	median	95th %tile
.2	55	68	77	63	75	81
.3	52	66	77	61	68	80
.4	50	64	70	58	65	80
.5	46	58	66	53	65	78
.6	50	58	66	55	62	70
.7	50	60	65	60	63	79
.8	40	52	60	50	57.5	73
.9	45	51	56	50	55	76

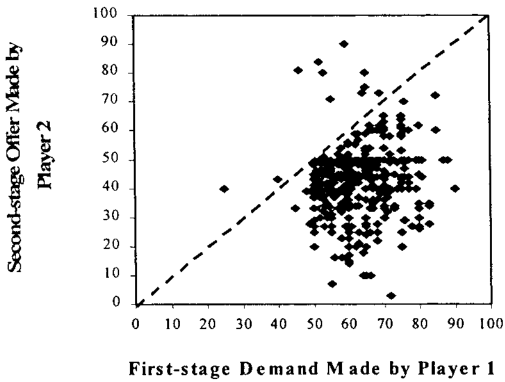

FIG. 12. The tables show ranges of accepted and rejected demands in the first stage of game III. The figure shows player 2's offer to player 1 in the second stage of the two-stage game, as a function of player 1's demand in the first stage, for those demands that were rejected. Both axes measure the percentage of the surplus accruing to player 1.

In the experiment, a wide variety of player-1 demands are rejected. Figure 12 shows that there is considerable overlap between the set of accepted and rejected demands. If the value of player 1's demand reveals significant information about player 1, then player-2 offers in the second stage of game III should be systematically related to the value of the player-1 demand that was rejected in order to reach the second stage.

Figure 12 plots player 2's offer to player 1 in the second stage of the two-stage game, as a function of player 1's demand in the first stage, for those 501 demands that were rejected.[22] As expected, the observations cluster

[22] In 217 of the 501 cases in which player 2 rejected, the subsequent offer was "disadvantageous," in the sense that it provided player 2 (if accepted) with a discounted *monetary* payoff lower than the monetary payoff 2 would have secured by accepting player 1's first-stage demand. Ochs and Roth [31] draw attention to disadvantageous counteroffers in two-stage bargaining experiments, citing them as evidence that subjects must be concerned with more than simply their own monetary payoffs.

below the diagonal: player 2 is generally less generous to player 1 than is player 1. More importantly, there appears to be little relationship between the rejected demand and the subsequent offer. However, Fig. 12 aggregates over all the discount factor specifications. We expect first-stage demands to vary systematically with the discount factor (as Fig. 16 below confirms), rendering such aggregation suspect. Figure 13 reports the results of a random-effects regression of (a transformation of) player 2's second-stage offer on player 1's first-stage demand, once again finding little relationship. In the base case of $D = 0.5$, player 2's second-stage offer does not vary significantly in player 1's first-stage demand. Only the cases $D = 0.7$ and $D = 0.4$ show significant departures from this base case, with the latter somewhat weaker than the former. Player 2s do not appear to be drawing useful inferences from the magnitude of player 1's rejected demand.

An examination of player 1's behavior in the second stage of game III suggests that there is little for player 2 to learn from observing player 1's first-period demand. Figure 14 shows player 1's rejection rate of player 2's offers in the second stage of game III, as a function of the first-period demand that player 1 had rejected in order to reach the second stage. There is scant evidence of a systematic relationship. Decomposing these data by discount factor and controlling for player 2's second-stage offer, though

Variable	Estimated coefficient	Standard error	*p*-value
Intercept, $D = .2$	-.29	1.2	.82
Intercept, $D = .3$	-1.2	.72	.098
Intercept, $D = .4$	-2.4	1.2	.045
Intercept, $D = .5$	-.17	.54	.75
Intercept, $D = .6$	.084	.98	.93
Intercept, $D = .7$	-1.8	.74	.015
Intercept, $D = .8$	.34	.65	.96
Intercept, $D = .9$	-.31	.61	.61
Player-1 demand, $D = .2$	.0073	.017	.67
Player-1 demand, $D = .3$	.015	.010	.152
Player-1 demand, $D = .4$	.036	.018	.044
Player-1 demand, $D = .5$	-.0019	.0082	.82
Player-1 demand, $D = .6$	-.00078	.015	.96
Player-1 demand, $D = .7$	.030	.011	.008
Player-1 demand, $D = .8$	.0016	.010	.88
Player-1 demand, $D = .9$	.0084	.0095	.37
$(V_1, V_2) = (70, 10)$	-.16	.10	.11
$(V_1, V_2) = (10, 60)$	-.064	.10	.54

FIG. 13. Random effects regression results. The dependent variable is $\log y/(100 - y)$, where y is player 2's offer in the second stage of game III. "Player-1 demand" is the demand rejected by player 2 to reach the second stage. Dummy variables are used to estimate the deviation of the intercept and slope term (on player-1 demand), for each discount factor, from the base-case relationship (shown in bold) of $D = 0.5$.

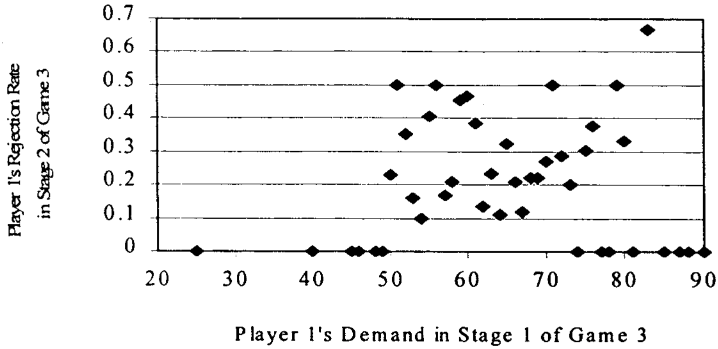

FIG. 14. Player 1's rejection rate in the second stage of game III, as a function of the first-period demand that player 1 had rejected in order to reach the second stage.

hampered by very small sample sizes, yields similar results. An appeal to incomplete information cannot readily reconcile the observed behavior with payoff-interdependent preferences.

Second, the second stage of game III can only be reached if player 2 rejects player 1's initial demand. Could play in the continuation game differ from that of the second stage of game III because the latter game is not played by a random sample of player 2s? Notice that one's initial intuition here works the wrong way. We would expect player 2s who reject to be more aggressive than those who do not, leading to lower rather than higher offers in the second stage of game III. Figure 15 reports results for game II, analogous to those reported in Fig. 6, but restricts attention to experimental subjects who rejected at least one demand when playing as player 2 in game III. A comparison with Fig. 6 shows that the differences are slight and the directions are mixed. The game-II average offer of those with

Rounds	(V_1, V_2)	Observations	Mean offer	Median	5th %tile	95th %tile
1-10	All	777	33.5	33	16	50
11-20	All	780	32.4	31	20	49
1-10	(10, 10)	248	32.9	32	20	48
11-20	(10, 10)	242	30.3	30	21	41
1-10	(70, 10)	285	27.6	26	10	45
11-20	(70, 10)	283	27.5	27	17	40
1-10	(10, 60)	244	40.8	43	24	52
11-20	(10, 60)	255	39.9	40	28	50

FIG. 15. Player-2 offers in game II, measuring the percentage of the surplus offered to player 1, as in Fig. 6, but for those subjects who reject a demand in game III.

(10, 60) rejection-payoffs who rejected at least one offer in game III is larger than the overall average, with the opposite relation holding for (70, 10) rejection payoffs. There is little evidence that the second stage of game III is played by a sufficiently atypical group of player 2s as to reconcile game-II and game-III second-stage behavior.[23]

In summary, we find that subgame consistency fails, and does so systematically. Players make less aggressive offers in the second stage of a two-stage game than in an equivalent, stand-alone game. There is evidence that information is incomplete, in the form of rejected first-stage offers in the two-stage game, but the failure of second-stage behavior to depend upon the magnitude of the rejected first-stage demand suggests that this does not provide a useful explanation for the differences between game II and the second stage of game III. Similarly, the evidence is that self-selection bias in determining which player 2's participate in the second stage of the two-stage game does not provide a useful explanation.

4.4. *Games III and IV: Truncation Consistency*

We next consider truncation consistency, which we examine by comparing the initial demands in games III and IV. Figure 16 shows the mean and median demands made by player 1 in the first stage of the two-stage game, as a function of the discount factor and the game-I rejection payoffs. (Again, the game-I rejection payoffs are relevant only in game I.) As the discount factor increases from 0.2 to 0.9, player 1's mean demand falls from about seventy to about fifty percept of the surplus. Subgame perfection predicts that player 1's demand should decrease in the discount factor, but at a more precipitous rate, falling from eighty percent to ten percent as D increases from 0.2 to 0.9. The data shown in Fig. 16 are quite similar to data displayed in Fig. 5.6 of Davis and Holt [17, p. 272], which summarizes a variety of experiments with two-stage games.

Each subject i in game IV was characterized by a pair $(Z_1(i), Z_2(i))$, which varied across subjects. Each time two subjects i and j were matched to play the game (in roles 1 and 2), the rejection payoffs were given by $Z_1(i)$ (for player i in role 1) and $Z_2(j)$ (player j in role 2). The following

[23] We can pursue this possibility further by examining "chronic rejecters," namely subjects who frequently rejected demands while playing as player 2 in game III. Consider subjects who rejected five or more demands in the first stage of game III. The average game-II offers made by these subjects (all rounds) were 27.5 (all rejection payoffs), 30.2 (the 10,10 case), 22.1 (the 70,10 case) and 35.7 (the 10,60 case). Hence, chronic rejecters made even smaller offers when playing game II, rendering it all the less likely that the relatively large offers encountered in the second stage of game III can be attributed to nonrandomness in the selection of player 2s.

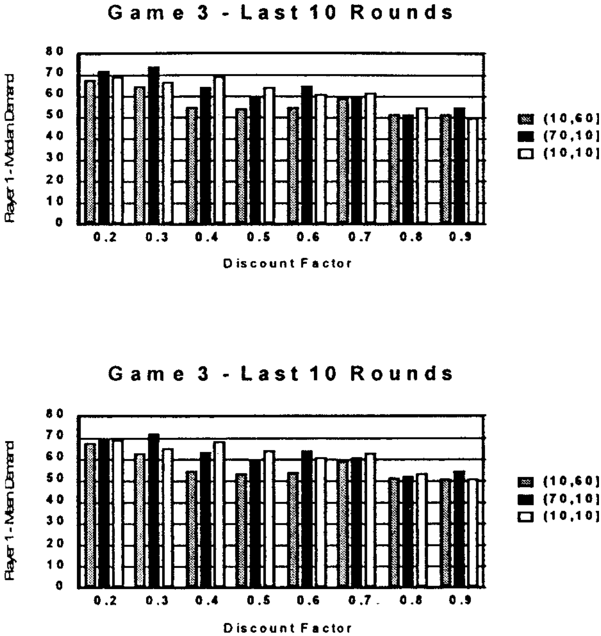

FIG. 16. Mean and median demands made by player 1 in the first stage of the two-stage game, measured as a percentage of the surplus.

section discusses how the values $(Z_1(i), Z_2(i))$ were determined. For the purposes of this section, it matters only that these values varied across subjects, and were commonly known in each game, so that first stage demands can be expected to vary systematically as do the values (Z_1, Z_2).

We can similarly think of first-stage demands in game III as depending upon the player-1 and player-2 values of proceeding to the second stage, which we denote by (Z_1^{III}, Z_2^{III}). If truncation consistency holds, and if play reflects rational behavior given payoff-interdependent preferences, then initial demands in game IV should bear the same relationship to $(Z_1(i), Z_2(j))$ as initial demands in game III to (Z_1^{III}, Z_2^{III}). Our investigation of truncation consistency thus estimates (Z_1^{III}, Z_2^{III}), and then examines the relationship between initial demands in games III and IV and the values (Z_1^{III}, Z_2^{III}) and (Z_1, Z_2), respectively.

Because we are interested in the explanatory power of rational behavior, given payoff-interdependency, we use observed play in the second stage of game III to estimate the expected payoffs (Z_1^{III}, Z_2^{III}). However, some difficulties are raised by the potential heterogeneity in preferences among anonymously matched opponents. In the presence of such heterogeneity, first-stage demands in principle depend upon player 1's expectation of

(Z_1^{III}, Z_2^{III}), as well as 1's expectation of 2's expectation of (Z_1^{III}, Z_2^{III}), and 1's expectation of 2's expectation of 1's expectation of (Z_1^{III}, Z_2^{III}), and so on. We cannot estimate this entire infinite hierarchy, and must focus on what we expect to be the most salient variables. The greater is the amount of variation in second-stage payoffs explained by variations in the discount factor, and the less important are player idiosyncrasies, the more likely is this hierarchy to be captured by a single pair of values. This observation motivates the wide range of discount factors incorporated in our experimental design.

Since we are interested in the determinants of player 1's initial demand, we first consider player 1's expectation of (Z_1^{III}, Z_2^{III}), which we denote by $(Z_1^{III1}(i), Z_2^{III1}(i))$, where i is the subject acting as player 1 in game III. We take $Z_1^{III1}(i)$ to be the average of the player-1 payoffs in those second-stage games in which player i participated, and take $Z_2^{III1}(i)$ to be the average of the player-2 payoffs in those games.[24]

Figure 17 presents the results of random-effects regression with data drawn from games III and IV.[25] The dependent variable is a transformation of the demand made by player 1 in the first period of the game, where this is a demand in game III for some observations and a demand in game IV for others. The independent variables include an intercept term and two intercept dummy variables to identify the (70, 10) and (10, 60) game-I rejection payoff cases (the base case is (10, 10)). The variables "Player-1 payoff" and "Player-2 payoff" identify the expected payoffs following a rejection in game III or IV. These variables are given by $(Z_1^{III1}(i), Z_2^{III1}(i))$ for observations taken from game III and by $(Z_1(i), Z_2(j))$ for observations taken from game IV. In addition, we include two slope dummy variables, ("Player-1 payoff, Game III" and "Player-2 payoff, Game III)", to capture differences, across games III and IV, in the relationship between initial demands and the payoffs that follow a rejection. If truncation consistency holds, these latter dummy variables should be zero.

The coefficient on "Player-2 payoff" in Fig. 17 is negative. Hence, the larger is the rejection payoff for player 2, the more moderate is 1's initial demand. The coefficient on the "Player-2 payoff, game III" dummy identifies how the dependence of game-III initial demands on player-2 rejection payoffs differs from that of game IV. This coefficient is (significantly) positive, and smaller in absolute value than the "Player-2 payoff" coefficient. Hence, initial demands in game III are again decreasing in player 2's

[24] The more dispersed are realized payoffs, the greater is the extent to which an approach based on such averages requires utility functions that are not too nonlinear, as in (2) or (3).

[25] In game III, the sample is restricted to subjects who participated in the second stage of game III at least once, and hence for whom we can estimate $(Z_1^{III1}(i), Z_2^{III1}(i))$. These subjects played game III as player 1 a total of 2229 times which, together with 2400 instances of game IV, gives us 4629 observations.

Independent variable	Estimated coefficient	Standard error	p-value
Intercept	.63	.027	.000
Player-1 payoff	.030	.0011	.000
Player-2 payoff	-.027	.00058	.000
Player-1 payoff, Game III	-.035	.0017	.000
Player-2 payoff, Game III	.025	.0012	.000
(70,10)	.024	.031	.44
(10,60)	-.21	.032	.000

FIG. 17. Random effects estimates of player 1's first-stage demand in Games III and IV. The dependent variable is $\log x/(100-x)$, where x is player 1's demand in the first stage of game III or game IV. The player-1 and player-2 payoffs are (Z_1, Z_2) for game-IV observations and (Z_1^{III1}, Z_2^{III1}) for game-III observations. "Player-1 payoff, Game III" and "Player-2 payoff, Game III" are dummy variables, indicating how the coefficients on player-1 and player-2 payoffs for game III differ from the game-IV base case. The game-I rejection payoff (10, 10) is the intercept base case, with dummy variables for the (70, 10) and (10, 60) cases.

rejection payoff, but are much less sensitive to the latter.[26] The coefficient on "Player-1 payoff" is positive, indicating that player 1 tends to be more aggressive in game IV when 1 has a larger rejection payoff. But the "Player-1, game III" dummy is negative, again indicating that this sensitivity is attenuated in game III.[27]

Our basic result is then that player-1 initial demands are significantly less sensitive to rejection payoffs in game III, which appear as the result of play in a continuation game, than to rejection payoffs in game IV, which are part of the specification of the game. In contrast to the prediction of truncation consistency, players react more sharply to variations in fixed terminal payoffs than they do to equivalent variations in the expected value of a continuation game.

Are the differences shown in Fig. 17 economically relevant? From Fig. 17, which gives $\log x/(100-\log x)$ as a function of V_2, we can calculate that, in game IV, $dx/dV_2 \approx -.8$, so that eighty percent of an increase in player 2's rejection payoff V_2 is translated into a decrease in player 1's initial demand x. This is a smaller response than the derivative of -1 that would characterize subgame perfection given monetary payoff maximization, but a much larger response than that described by the corresponding derivative for game III, where $dx/dV_2 \approx -.2$. The latter calculation, which is consistent with the results shown in Fig. 16, shows that players react

[26] That is, the sum of the "Player-2 payoff" and "Player-2 payoff, Game III" coefficients is negative, but smaller in absolute value than the "Player-2 payoff" coefficient.

[27] In this case, adding the dummy to the base coefficient gives a negative value, indicating that 1's demands in game III are *inversely* related to 1's rejection payoffs. This reflects the fact that Z_1^{III1} and Z_2^{III1} tend to be positively correlated, as both vary positively in the discount factor, with the dominant effect on player-1 demands being the inverse relationship with Z_2^{III1}.

quite sluggishly to changes in rejection payoffs generated by continuation games.

Are these results robust to our estimation of game III second-stage payoff expectations? We can explore alternatives. First, the average $Z_1^{III1}(i)$ may involve second-stage player-1 payoffs from cases in which subject i occupied the role of player 2. These in turn may involve accept/reject decisions that subject i would have made differently, and which hence may present a misleading estimate of subject i's expected payoff from playing the second stage as player 1. To examine this possibility, we restrict the calculation of $Z_1^{III1}(i)$ to those cases in which i plays the second stage as player 1, calling the estimate $\hat{Z}_1^{III1}$. In addition, player 1's expectation of Z_2^{III} may be less important than 1's estimate of 2's estimate of Z_2^{III}, since the latter is likely to play the major role in shaping 2's accept/reject decision in stage 1. We accordingly replace Z_2^{III1} with $\hat{Z}_2^{III21}$, 1's expectation of 2's expectation of 2's payoff in the second stage, calculated as the average payoff realized by player 2 in those second-stage games in which agent i fills the role of player 1.

Figure 18 duplicates the analysis of Fig. 17, using the alternative measures $(\hat{Z}_1^{III1}(i), \hat{Z}_2^{III21}(i))$. The results are familiar. The coefficient on "Player-2 payoff" is (significantly) negative, so that player 1's game-IV initial demand is more moderate for larger player-2 rejection payoffs. The "Player-2 payoff, game III" dummy is (significantly) positive, and smaller in absolute value. Hence, initial demands in game III are again decreasing in player 2's rejection payoff, but are much less sensitive to the latter. In this case, the calculated derivatives are $dx/dV_2 \approx -.7$ in game IV, and $dx/dV_2 \approx -.1$ in game III. Once again, players are much more sensitive to changes in terminal payoffs than to equivalent changes in the expected value of a continuation game.

Next, we would like to investigate the effect of simply using 2's expectation of 2's payoff, rather than 1's expectation of 2's expectation, which suggests replacing $\hat{Z}_2^{III21}(i)$ with $\hat{Z}_2^{III2}(j)$ (when subject i plays j), where the

Independent variable	Estimated coefficient	Standard error	p-value
Intercept	.64	.026	.000
Player-1 payoff	.029	.0011	.000
Player-2 payoff	-.027	.00058	.000
Player-1 payoff, Game III	-.030	.0015	.000
Player-2 payoff, Game III	.022	.0010	.000
(70,10)	.028	.032	.38
(10,60)	-.19	.033	.000

FIG. 18. Random effects estimates of player 1's first-stage demand in Games III and IV, as in Fig. 17, but with $(Z_1^{III1}(i), Z_2^{III1}(i))$ replaced by $(\hat{Z}_1^{III1}(i), \hat{Z}_2^{III21}(i))$.

Independent variable	Estimated coefficient	Standard error	*p*-value
Intercept	1.1	.030	.000
Player-1 payoff	.018	.0013	.000
Player-2 payoff	-.032	.00066	.000
Player-1 payoff, Game III	-.027	.0015	.000
Player-2 payoff, Game III	-.022	.00098	.000
(70,10)	.015	.026	.57
(10,60)	-.16	.027	.000

FIG. 19. Random effects estimates of player 1's first-stage demand in Games III and IV, as in Fig. 17, but with $(Z_1^{III1}(i), Z_2^{III1}(i))$ replaced by $(\tilde{Z}_1^{III1}(i), \tilde{Z}_2^{III2}(i))$.

latter measures the average payoff earned in those second-stage games in which subject j acted as player 2.[28] In addition, we note that when calculating $\hat{Z}_1^{III1}$, those cases in which subject i, in the role of player 1, rejects a second-stage offer add a zero payoff to $Z_1^{III1}(i)$ while having no effect on $\hat{Z}_2^{III21}(i)$ (or $\hat{Z}_2^{III2}(j)$). This is likely to underestimate i's payoff, since i has revealed that i's realized utility, from payoffs $(0, 0)$, is higher than the utility of accepting 2's offer, which in turn is likely to exceed the utility of the effectively recorded outcome $(0, Z_2^{III2}(j))$. The best available correction is to calculate 1's payoff as the average of the *offers* made to subject i when playing the second stage as player 1 (though this still potentially underestimates i's utility in those cases in which i rejects), denoted by $\tilde{Z}_1^{III1}(i)$. Similarly, we are likely to underestimate 2's utility in those cases in which 2 makes a second-stage offer that is rejected. In this case, we do not have any attractive alternative estimates of 2's utility available, since (unlike the situation of player 1) we cannot conclude that 2 preferred that the offer be rejected. We accordingly restrict our calculation of 2's payoff to those cases in which 2's offer is accepted, denoted by $\tilde{Z}_1^{III2}(j)$.

Figure 19 reports the corresponding estimates, again with familiar results. The coefficient on "Player-2 payoff" is (significantly) negative, while the "Player-2 payoff, Game III" dummy is (significantly) negative but smaller in absolute value. Hence, initial demands in games III and IV are both decreasing in player 2's rejection payoff, but are much less sensitive to the latter in game III. In this case, the estimated derivatives are $dx/dV_2 \approx -.8$ in game IV and $dx/dV_2 \approx -.2$ in game III.

In summary, truncation consistency does not hold. It makes a difference whether a rejected offer is followed by a pair of fixed payoffs, or by a continuation game whose expected outcome matches those fixed payoffs.

[28] If this change makes little difference, then we have evidence that our results are not sensitive to which expectation involving player 2's payoff we choose from the infinite hierarchy of possibilities. More generally, there are numerous alternatives for examining the robustness of the results. We found none that made a significant difference.

Initial demands are much more sensitive to changes in terminal payoffs than to equivalent changes in the expected value of a continuation game.[29] These results are consistent across a variety of methods for estimating the expected payoffs following a game-III first-stage rejection.

4.5. *Games II, III and IV: Subgame and Truncation Consistency*

This section provides a joint test of subgame and truncation consistency, based on comparing first-period demands in game III with demands in game IV.

Each experimental subject i in game IV was characterized by an idiosyncratic pair of rejection payoffs $(Z_1(i), Z_2(i))$, one when playing as player 1 and one when playing as player 2. In each play of game IV, rejection payoffs were commonly known, and given by $(Z_1(i), Z_2(j))$, where player 1 was subject i and player 2 was j. Our intention was that the rejection payoffs (Z_1, Z_2) would equal the subjects' expected payoffs from playing game II, the continuation game. If subgame consistency holds, then these payoffs would also equal the expected payoffs of the second stage of the two-stage game. If truncation consistency also holds, play in the first stage of game III should be identical to play in game IV.

The previous subsection described a variety of alternatives estimating the expected payoff of playing the second stage of game III or, equivalently, playing the continuation game. Our experimental design required one of these estimates to be built into the experiment in the calculation of Z_1 and Z_2. In making this choice, we were anxious to provide the most favorable environment for payoff-interdependent preferences, and hence were anxious not to underestimate *utility* when offers are rejected. We accordingly employed the final alternative investigated in the previous subsection, taking $Z_1(i)$ to be the average offer received by subject i when playing as player 1 in the continuation game, and taking $Z_2(j)$ to be the average payoff realized by subject j player in those periods in which j played as player 2 in the continuation game and made an offer that was accepted.

[29] Beard and Beil [4] suggest a similar conclusion. They examine a game in which player 1 can either choose L, ending the game with a known pair of monetary payoffs, or choose R, in which case player 2 chooses between l or r, each ending the game with known payoffs. The payoffs are chosen so that R, r is the unique subgame-perfect equilibrium (if utility depends only upon one's own earnings), but so that R, l is worse for player 1 than L. Their experimental finding is that player 1s quite often choose the "safe" outcome of L rather than risk a suboptimal choice of l on the part of player 2, with the incidence of such choices depending in expected ways upon payoff magnitudes. They suggest that player 1s appear to be more responsive to the payoff following L than to the expected payoff of the subgame following R, attributing this to a preference for certain payoffs which players can ensure over uncertain ones which players cannot ensure.

Discount	Obs.	Z_1	Z_2	$Z_1 + Z_2$	Surplus	$Z_1\%$	$Z_2\%$
.2	30	6.3	13.4	19.7	20	32	67
.3	30	8.4	21.5	29.9	30	28	72
.4	30	14.4	25.2	39.6	40	36	63
.5	30	17.2	32.4	49.6	50	34	65
.6	30	22.0	37.4	59.4	60	37	62
.7	30	22.6	46.7	69.3	70	32	67
.8	30	28.1	51.1	79.2	80	35	64
.9	30	25.8	63.7	89.5	90	28	71

FIG. 20. Mean rejection payoffs for game IV by discount factor. There were three treatments for each discount factor (one for each game-I rejection payoff (V_1, V_2)), with ten subjects in each treatment, for a total of 30 rejection payoffs for each discount factor and player role. The mean of these 30 payoffs is reported in each case. $Z_1\%$ and $Z_2\%$ are player 1 and 2's average rejection payoff as a percentage of the total surplus.

Figure 20 reports the resulting mean rejection payoffs for game IV. As expected, rejection payoffs are larger for larger discount factors. The rejection payoffs allocate about two-thirds of the surplus to player 2 and one-third to player 1. The latter percentage varies with the discount factor, but again in no systematic way. The mean rejection payoffs virtually exhaust the surplus in each case, consistent with a rejection-payoff calculation designed to capture expected utilities, where player 1 prefers disagreement to the offers 1 rejects.

Figure 21 compares player-1 demands in games III and IV. We concentrate on the final ten rounds of play in this section, though expanding to all twenty rounds makes virtually no difference. (Once again, the rejection payoffs (V_1, V_2), being (10, 10), (10, 60), or (70, 10), are irrelevant for games III and IV). If subgame and truncation consistency hold, then player-1 demands in games III and IV should be identical. Figure 21 indicates that for low discount factors, mean and median demands are similar. However, as the discount factor increases, the mean and median demands fall much more rapidly in game IV than in game III. As a result of this sluggish game-III response, proposers are more aggressive in game III than in game IV, for high discount factors.

Figure 22 provides evidence that the differing behavior in games III and IV is important, comparing the mean amount of surplus offered to player 2 in the first stage of games III and IV with player 2's mean rejection payoff in game IV, for large discount factors. In every case, the game-IV mean player-1 demand yields a higher payoff to player 2 than does the mean rejection payoff. If subgame and truncation consistency hold, we would expect the same of the game-III mean demand. However, in every case, the game-III mean player-1 demand is sufficiently aggressive as to leave player 2 with a lower payoff than the game-IV mean rejection payoff.

Discount	Obs.	III mean	III median	IV mean	IV median
		Rejection payoffs (10, 10)			
.2	50	69.3	69	70.9	71.5
.3	50	65.5	66.5	65.1	65
.4	50	68.6	69	61.5	61
.5	50	63.9	64	61.1	61
.6	50	61.0	61	55.0	55
.7	50	62.8	62	47.0	47
.8	50	53.4	55	42.3	41
.9	50	50.9	50	29.6	30
		Rejection payoffs (10, 60)			
.2	50	68.0	68	70.5	70
.3	50	62.8	65	67.6	68
.4	50	55.1	55	54.8	55.5
.5	50	53.7	54	54.4	54
.6	50	54.5	55	56.5	55
.7	50	59.0	59	52.4	53
.8	50	51.8	52	50.4	50
.9	50	51.4	51.5	44.5	46
		Rejection payoffs (70, 10)			
.2	50	69.0	72	70.1	70
.3	50	72.5	74	66.4	66
.4	50	63.8	64	57.2	58
.5	50	59.3	59	53.1	53.5
.6	50	64.1	65	51.6	50
.7	50	61.0	60	43.0	43.5
.8	50	52.3	52	36.4	36
.9	50	54.7	55	26.7	26

FIG. 21. Player-1 demands in the first stage of game III and game IV. Data are taken from the last ten rounds in each case. For each discount-factor and rejection-payoff combination, there were ten rounds of five games each, for 50 observations.

It is intuitive that there should be little difference between games III and IV when discount factors are small. In this case, the rejection payoffs in game IV are small, and the second stage in game III is relatively unimportant. As the discount factor grows, rejection payoffs become larger in game IV and the second stage becomes more important in game III, magnifying behavioral differences.

We can illustrate the difference between games III and IV. For each of the 240 subjects, we can calculate the subject's mean demand as player 1 in games III and IV. Figure 23 shows the demands. (Analogous results obtain for median demands.) Low discount factors give rise to relatively

Discount	(V_1, V_2)	Obs.	100-(III mean)	mean Z_2	100-(IV mean)
.7	(10, 10)	50	37.2	47.3	53.0
.8	(10, 10)	50	46.6	52.0	57.7
.9	(10, 10)	50	49.1	67.2	70.4
.7	(10, 60)	50	41.0	44.2	47.6
.8	(10, 60)	50	48.2	44.5	49.6
.9	(10, 60)	50	48.6	53.7	55.5
.7	(70, 10)	50	39.0	48.3	57.0
.8	(70, 10)	50	47.7	56.7	63.6
.9	(70, 10)	50	45.3	70.1	73.3

FIG. 22. Amount of the surplus that mean player-1 demands allocate to player 2, in the first stage of game III (100-(III mean)) and in game IV (100-(IV mean)).

large demands, in which case game-III and game-IV demands are similar. However, higher discount factors give rise to lower demands, in which case player 1s demand significantly more in game III than in game IV, reflecting the relatively sluggish response of game-III demands to discount factors.

To examine the significance of these differences, Fig. 24 reports estimations of subjects' mean and median initial demands in game III as a function of their initial median demands in game IV. Subgame and truncation consistency combine to predict a zero intercept and unitary slope, indicating that there is no systematic difference between the two games. Instead, the intercept is greater than zero and the slope is less than one (both at a p-value of .000), as we would expect if player 1 consistently demands more in game III than in game IV when the discount factor is high.

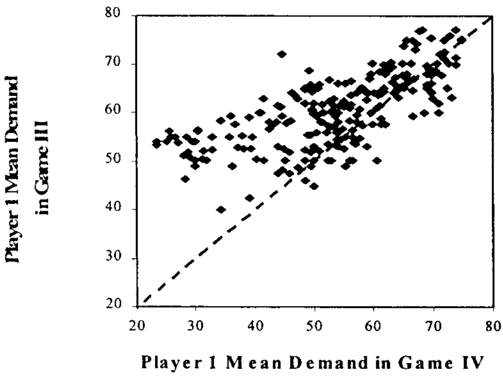

FIG. 23. Plot of player-1 mean demand in game IV (horizontal axis) and the first stage of game III (vertical axis), taken from the last ten rounds of play. There are 240 observations, one for each of the 240 experimental subjects.

Dependent variable	Observations	Intercept	Game-IV demand
Game-III mean demand	240	.37 (.015, (.34, .40))	.43 (0.028, (.37,.48))
Game-III median demand	240	.37 (.016, (.34, .40))	.42 (0.029, (.37,.48))

FIG. 24. Linear regressions of transformations of player 1's mean and median demand in the first stage of game III on player 1's median and mean demand in game IV confidence intervals.

In summary, our comparison of games III and IV suggests a failure of at least one of subgame and truncation consistency, leading to systematic differences in play in the first stage of game III and play in game IV. The differences appear primarily for high discount factors, when rejection payoffs in game IV are high and the second stage of game III is relatively important. In such cases, opening demands are more aggressive in game III than in game IV.

These results are consistent with our separate tests of subgame and truncation consistency. When the discount factor is small, games III and IV are both quite similar to an Ultimatum Game, and yield similar play. As the discount factor rises, so does player 2's payoff in the continuation game, and hence 2's rejection payoff in game IV, leading to lower player-1 demands. A similar force appears in game III as the second stage becomes more valuable. As Section 4.3 shows, however, player 2's value in the second stage of game III is less than that of the continuation game, reflecting 2's less aggressive play (and 1's more aggressive play) in game III's second stage. In addition, Section 4.4 shows that initial play is less responsive to changes in the expected value of a second stage than to changes in a corresponding terminal payoff. These failures of subgame and truncation consistency reinforce one another. A rising discount factor causes a smaller increase in player 2's rejection value in game III than in game IV, and player 1 is less sensitive to changes in the rejection value in game III than in game IV. Together, the result is that player 1's demands show less variation in game III than in game IV, leading to the result shown in Fig. 23.

5. CONCLUSION

Our experimental results provide several indications that payoff-interdependent preferences and backward induction, in the form of subgame and truncation consistency, are inconsistent. The second stage of the two-stage game features more generous player-2 offers than does the (identical) continuation game. This is a failure of subgame consistency: players regard the second stage of the two-stage game and a seemingly stand-alone equivalent as different strategic situations. Making an offer to someone whose demand you have just rejected, in the second stage of the two-stage game, is not viewed as equivalent to opening the seemingly identical continuation game, with no history of interaction. Truncation consistency also fails. Players are more responsive to variations in future prospects when a rejection leads to a fixed pair of rejection payoffs, as opposed to the case in which a rejection leads to a game involving another offer and response.

Attention now turns either to alternative formulations of preferences or to models of behavior that do not depend upon backward induction. Because the self-references or additional arguments built into more complicated preference formulations can deprive backward induction of its content, it is not clear that these are distinct alternatives.

Our findings reinforce those of Andreoni *et al.* [2], who show that payoff interdependence alone cannot account for behavior in public-good provision experiments.[30] Instead, changes in the extensive form of the game prompt changes in behavior that are inconsistent with preferences that depend upon only payoffs. Our results are similar in spirit, suggesting that preferences in seemingly identical games depend upon the larger context in which the games are played. Andreoni *et al.* [2] suggest incorporating the specification of the game into the utility function, allowing players to have different preferences over identical monetary payoff vectors in different games. Given the mounting experimental evidence, such an approach seems inevitable if the results are to be explained in terms of more elaborate utility functions. However, the results will be useful only if some portability of the preferences can be recovered, in the form of some systematic view of the relationship between the specification of the game and preferences.

We suspect the key to such portability lies in a more systematic investigation of how people think about games. Psychologists direct attention to the use of analogy when reasoning about novel problems (e.g., Holyoak and Thagard [26]). We envision players as analyzing unfamiliar games or subgames by drawing analogies to more familiar contexts. Subgame consistency will then obtain if the considerations that shape these analogies are precisely those captured by the extensive-form specification of a game. As a result, subgame consistency and backward induction would be compelling in the classical view of game theory, in which games are complete, literal representations of strategic interactions. But game theory is typically used not as a literal description but as a *model* of a more complicated strategic interaction, and there is no reason to believe that the extensive form constructed by an analyst exactly captures the considerations used by players to analyze the interaction. If not, subgame and truncation consistency can

[30] Prasnikar and Roth [32] explore similar games and issues.

be expected to fail. Anticipating this failure, however, makes many seemingly anomalous experimental findings less puzzling. Framing effects are now expected, for example, as differing details of the experimental environment trigger varying analogies. Nor is it a surprise that rejecting an offer might bring a new analogy into play, or that fixed rejection payoffs and continuation games trigger different analogies. Our hope is that a theory of reasoning-by-analogy might lead to a more useful model of behavior in games. Samuelson [37] begins the construction of such a theory.

REFERENCES

1. K. Abbink, G. E. Bolton, A. Sadrieh, and F.-F. Tang, Adaptive learning versus punishment in ultimatum bargaining, mimeo, University of Bonn and Penn State University, 1996.
2. J. Andreoni, P. M. Brown, and L. Vesterlund, What produces fairness? Some experimental results, *Games Econ. Behav.*, forthcoming.
3. J. Andreoni and J. H. Miller, Giving according to GARP: An experimental test of the rationality of altruism, *Econometrica* **70** (2002), 737–754.
4. T. R. Beard and R. Beil, Do people rely on the self-interested maximization of others? An experimental test, *Manage. Sci.* **40** (1994), 252–262.
5. K. Binmore, P. Morgan, A. Shaked, and J. Sutton, Do people exploit their bargaining power? An experimental study, *Games Econ. Behav.* **3** (1991), 295–322.
6. K. Binmore, C. Proulx, L. Samuelson, and J. Swierzbinski, Hard bargains and lost opportunities, *Econ. J.* **108** (1998), 1279–1298.
7. K. Binmore, A. Shaked, and J. Sutton, Testing noncooperative bargaining theory: A preliminary study, *Amer. Econ. Rev.* **75** (1985), 1178–1180.
8. K. Binmore, A. Shaked, and J. Sutton, An outside option experiment, *Quart. J. Econ.* **104** (1989), 753–770.
9. S. Blount, When social outcomes aren't fair: The effect of causal attributions on preferences, *Organ. Behav. Human Decision Processes* **63** (1995), 131–144.
10. G. E. Bolton, A comparative model of bargaining: Theory and evidence, *Amer. Econ. Rev.* **81** (1991), 1096–1136.
11. G. E. Bolton, J. Brandts, and A. Ockenfels, Measuring motivations for the reciprocal responses observed in a simple dilemma game, *Exper. Econ.* **1** (1998), 207–219.
12. G. E. Bolton and A. Ockenfels, ERC: A theory of equity, reciprocity and competition, *Amer. Econ. Rev.* **90** (2000), 166–193.
13. G. E. Bolton and R. Zwick, Anonymity versus punishment in ultimatum bargaining, *Games Econ. Behav.* **10** (1995), 95–121.
14. C. F. Camerer, E. J. Johnson, T. Rymon, and S. Sen, Cognition and framing in sequential bargaining for gains and losses, *in* "Frontiers of Game Theory" (K. Binmore, A. Kirman, and P. Tani, Eds.), pp 27–48, MIT Press, Cambridge, MA, 1993.
15. R. Cooper, D. V. DeJong, R. Forsythe, and T. W. Ross, Alternative institutions for resolving coordination problems: Experimental evidence on forward induction and preplay communication, *in* "Problems of Coordination in Economic Activity" (J. W. Friedman, Ed.), pp. 129–146, Kluwer Academic, Boston, 1994.

16. M. Costa-Gomez and K. G. Zauner, Ultimatum bargaining behavior in Israel, Japan, Slovenia, and the United States: A social utility analysis, *Games Econ. Behav.* **34** (2001), 238–270.
17. D. D. Davis and C. A. Holt, "Experimental Economics," Princeton Univ. Press, Princeton, NJ, 1993.
18. M. Dufwenberg and U. Gneezy, Efficiency, reciprocity, and expectations in an experimental game, Discussion paper 9679, CentER for Economic Research, Tilburg University, 1996.
19. M. Dufwenberg and G. Kirchsteiger, A theory of sequential reciprocity, Discussion paper 9837, CentER for Economic Research, Tilburg University, 1998.
20. A. Falk and U. Fishbacher, A theory of reciprocity, Mimeo, University of Zurich, 1999.
21. E. Fehr and K. M. Schmidt, A theory of fairness, competition and cooperation, *Quart. J. Econ.* **114** (1999), 817–868.
22. W. Güth, R. Schmittberger, and B. Schwarze, An experimental analysis of ultimatum bargaining, *J. Econ. Behav. Organ.* **3** (1982), 367–388.
23. W. Güth and R. Tietz, Ultimatum bargaining behavior: A survey and comparison of experimental results, *J. Econ. Psych.* **11** (1990), 417–449.
24. J. C. Harsanyi and R. Selten, "A General Theory of Equilibrium Selection in Games," MIT Press, Cambridge, MA, 1988.
25. D. J. Holt, An empirical model of strategic choice with an application to coordination games, *Games Econ. Behav.* **27** (1999), 86–105.
26. K. J. Holyoak and P. Thagard, "Mental Leaps," MIT Press, Cambridge, MA, 1996.
27. E. J. Johnson, C. Camerer, S. Sen, and T. Rymon, Detecting failures of backward induction: Monitoring information search in sequential bargaining, *J. Econ. Theory* **104** (2002), 16–47.
28. J. H. Kagel, C. Kim, and D. Moser, Fairness in ultimatum games with asymmetric information and asymmetric payoffs, *Games Econ. Behav.* **13** (1996), 100–110.
29. D. K. Levine, Modeling altruism and spitefulness in experiments, *Rev. Econ. Dynam.* **1** (1998), 593–622.
30. R. D. McKelvey and T. R. Palfrey, An experimental study of the centipede game, *Econometrica* **60** (1992), 803–836.
31. J. Ochs and A. E. Roth, An experimental study of sequential bargaining, *Amer. Econ. Rev.* **79** (1989), 355–384.
32. V. Prasnikar and A. E. Roth, Considerations of fairness and strategy: Experimental data from sequential games, *Quart. J. Econ.* **106** (1992), 865–888.
33. M. Rabin, Incorporating fairness into game theory and economics, *Amer. Econ. Rev.* **83** (1993), 1281–1302.
34. A. E. Roth, Bargaining experiments, *in* "Handbook of Experimental Economics" (J. Kagel and A. E. Roth, Eds.), pp 253–348, Princeton Univ. Press, Princeton, NJ, 1995.
35. A. E. Roth, V. Prasnikar, M. Okuno-Fujiwara, and S. Zamir, Bargaining and market power in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An experimental study, *Amer. Econ. Rev.* **81** (1991), 1068–1095.
36. B. J. Ruffle, More is better, but fair is fair: Tipping in dictator and ultimatum games, *Games Econ. Behav.* **23** (1998), 247–265.
37. L. Samuelson, Analogies, adaptation, and anomalies, *J. Econ. Theory* **97** (2001), 320–367.
38. T. Slembeck, As if playing fair—Experimental evidence on the role of information in ultimatum bargaining, mimeo, University College London, 1998.
39. P. G. Straub and J. Keith Murnighan, An experimental investigation of ultimatum games: Information, fairness, expectations, and lowest acceptable offers, *J. Econ. Behav. Organ.* **27** (1995), 345–364.
40. R. H. Thaler, Anomalies: The ultimatum game, *J. Econ. Perspect.* **2** (1988). 195–206.

41. E. Winter and S. Zamir, An experiment with ultimatum bargaining in a changing environment, Discussion Paper 159, Hebrew University of Jerusalem Center for Rationality and Interactive Decision Making, 1997.
42. R. Zwick and X. P. Chen, What price for fairness? A bargaining study, mimeo, Hong Kong University of Science and Technology, 1997.
43. R. Zwick and E. Weg, An experimental study of buyer-seller negotiation: Self-interest versus other-regarding behavior, mimeo, Hong Kong University of Science and Technology and Purdue University, 1996.