



## Experimental Tests of a Sequential Equilibrium Reputation Model

Colin Camerer; Keith Weigelt

*Econometrica*, Vol. 56, No. 1. (Jan., 1988), pp. 1-36.

Stable URL:

<http://links.jstor.org/sici?sici=0012-9682%28198801%2956%3A1%3C1%3AETOASE%3E2.0.CO%3B2-M>

*Econometrica* is currently published by The Econometric Society.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/econosoc.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

## EXPERIMENTAL TESTS OF A SEQUENTIAL EQUILIBRIUM REPUTATION MODEL

BY COLIN CAMERER AND KEITH WEIGELT<sup>1</sup>

We test whether a model of reputation formation in an incomplete information game, using sequential equilibrium, predicts behavior of players in an experiment. Subjects play an abstracted lending game: a *B* player lends or does not lend; then if *B* lends, an *E* player can pay back or renege. The game is played 8 times, and there is a small controlled probability that the *E* player's induced preferences make him prefer to pay back (but usually he prefers to renege). In sequential equilibrium, even *E* players who prefer to renege should pay back in early periods of the game, and renege with increasing frequency in later periods, to establish reputations for preferring to pay back. After many repetitions of the 8-period game, actual play is roughly like the sequential equilibrium, except that *E* players pay back later in the game and more often than they should. This behavior is rational if *B* players have a "homemade" prior probability of .17 (in addition to the controlled probability) that *E* players will prefer to pay back. We conclude that sequential equilibrium with homemade incomplete information describes actual behavior well enough that it is plausible to apply it to theoretical settings where individuals make choices (e.g., product markets, labor markets, bargaining).

KEYWORDS: Experiments, sequential equilibrium, reputation, incomplete information.

### 1. INTRODUCTION

WE TEST WHETHER a sequential equilibrium model of reputation formation in an incomplete information repeated game predicts behavior of players in an experiment. After subjects have experience their play is roughly like the sequential equilibrium, except they seem to have a "homemade" belief about the prior probability of the other players' payoff type, in addition to the prior probability we created in the experiments.

The mathematical analysis of reputations in repeated games with incomplete information is one of the most fertile areas of current research in mathematical social science (see Wilson's 1985 review). These games assume incomplete information about players' "types" (privately known characteristics). The common belief about a player's type is the player's reputation. In such games, modellers usually search for "sequential equilibria" (SE; Kreps and Wilson, 1982a)—equilibria in which play is rational in every subgame of play from any point to the end (as in "subgame perfect" equilibria; Selten, 1975). In SE, beliefs about players' types are updated based on play before each subgame, using Bayes' rule when possible.

In many games, reputations seem to form even though complete-information theories (assuming perfection) predict otherwise—players cooperate in the finitely-repeated prisoners' dilemma, for instance (Kreps, et al., 1982; cf. Axelrod, 1982), and firms deter entry (Kreps and Wilson, 1982b; Milgrom and Roberts,

<sup>1</sup> Thanks to James Friedman, Charles Holt, David Kreps, Robert Wilson, two anonymous referees, and participants at the 1986 Public Choice Society Meetings, the 1986 Summer Econometric Society Meetings, the University of Iowa, New York University, and the Wharton Decision Processes bag lunch seminar, for comments. This research was funded by the New York University Center for Entrepreneurial Studies.

1982). Empirical reputation-building can be rationalized by assuming that people play SE with certain kinds of incomplete information. However, equilibria which rationalize reputation-building are often so complicated that it is reasonable to ask whether people actually play SE in naturally-occurring games. This question is important given the increasing popularity of theories in which individuals are assumed to use SE (or similarly complicated equilibria) to make decisions under incomplete information—in product markets (e.g., Allen, 1984), in labor markets (e.g., Shapiro and Stiglitz, 1984), and in bargaining settings (e.g., Fudenberg, Levine, and Tirole, 1985; Rubinstein, 1985). It is not unreasonable to think that firms might play elaborate games rationally (perhaps by hiring consultants or pooling knowledge), but it seems less plausible that individuals can calculate SE. Of course, SE is not necessarily meant to describe how people play games, but in many applications individual agents are assumed to play SE. If SE does not describe individual play in simple games, perhaps we should be skeptical about the application of SE to more complex settings like product markets, labor markets, and bargaining.

The assumptions underlying sequential equilibrium theories of reputation building—such as common knowledge about the game, including the incomplete information—are hard to verify in natural settings. Therefore, we test SE in laboratory experiments, where we can create an information structure which satisfies the assumptions of SE.

Many others have experimented on reputation formation in economic settings. Miller and Plott (1985) found that sellers in experimental product markets developed reputations for selling high-quality products, and commanded premium prices for doing so (as predicted by Klein and Leffler (1981), Shapiro (1983), et al.). DeJong, et al. (1985) found similar results in an agency setting, and Daughety and Forsythe (1987a, 1987b) found reputation-building in experimental duopolies. Roth and Schoumaker (1983) found that past histories of players, artificially inseminated into a bargaining experiment, affected bargaining outcomes. Isaac and Smith (1985) found that subjects *qua* firms did *not* build entry-detering reputations.

Experiments on games have been enormously popular among psychologists and sociologists, especially in the 1960's and 1970's (Wrightsmann, O'Connor, and Baker, 1972, cite more than a thousand papers). In these experiments, a simple noncooperative, non-zero-sum game (often the two-person prisoners' dilemma) typically played repeatedly, with some variation in incentives, instructions, social characteristics of the players. These experiments might be good tests for measuring attitudes and behavior in a social setting; but they are not especially good tests of theories of equilibrium play, because so much information is left uncontrolled.

In prisoners' dilemmas with many repetitions, for instance, subjects typically cooperate initially, then defect toward the end of the game (e.g., Rapoport and Chammah, 1965), with some variation in results across studies (see Colman, 1983). Initial cooperation may be equilibrium play in the game the subjects perceive, but we can't be sure in many of these experiments (though see Stoeckli, 1983) because the game they perceive includes lots of uncontrolled incomplete

information—who the other player is (often a confederate), the number of repetitions, whether payoffs are real, etc. (Colman, 1983, 117–118).<sup>2</sup>

In contrast with these experiments, we limit players' incomplete information by making everything common knowledge, except for one kind of incomplete information we control—we randomly assign players to payoff structures which only they know (see also Hogatt, et al., 1978). By assuming that their knowledge of payoffs is the only incomplete information in the game, we can derive a SE path and test whether players follow it.

Of course, we cannot completely limit all sources of information or beliefs of players. If players do not follow the SE path, they may be playing the SE of a game based on some "homemade" incomplete information that we did *not* create. All we can do is measure their homemade beliefs from data, and make predictions about new experiments based on our measurement. If those predictions are accurate, we can conclude that subjects are playing *some* SE (though not the one we initially predicted), using their homemade incomplete information along with the controlled incomplete information.

## 2. EXPERIMENTAL DESIGN

Our experiments test SE in a game where one player with uncertain payoffs plays a finitely-repeated game against a series of different opponents with certain payoffs. (Thus, the reputation is "one-sided." "Two-sided" reputation models may be tested in further work.)

A mnemonic story that makes the discussion easier to follow is that an entrepreneur (*E*) borrows from a different banker (*B*) each period for several periods (cf. Sobel, 1985, pp. 568–570). In each period, a *B* decides whether to lend to *E*, at a fixed interest rate. If *E* gets a loan he must decide whether to pay it back. An honest-type *E* (called a "Y-type" in our experiments) prefers to pay back loans. A dishonest-type *E* (an "X-type") prefers to renege (not pay back) on loans. *E* knows his own type, but *B* only knows the chances that *E* is of either type. (In our experiments, we used blander labels than these, to avoid creating nonmonetary preferences for strategies.<sup>3</sup>)

The experiment game tree in Figure 1 shows the sequential structure of choices and payoffs, for one period of the game. (Payoffs are denominated in "francs," an artificial currency equal to \$.01 for *B* subjects, and \$.0015 for *E* subjects.) Note how *E*'s payoffs depended upon his or her type: If *B* chose to make a loan (*L*), a Y-type *E* got a higher payoff from paying back (*P*) than from reneging

<sup>2</sup> In this regard, note Ledyard's (1986) demonstration that in a reasonably general class of games, with appropriately transformed utility functions any equilibrium could be a Bayesian equilibrium.

<sup>3</sup> *E*'s were called "*B* players," and bankers (*B*'s) were called "*A* players." Their choices were *A*1 (no loan, or *N*) and *A*2 (loan, *L*); and given an *A*2, the "*B*'s" could choose either *B*1 (pay back, *P*) or *B*2 (renege, *R*). We chose bland labels—"B2," instead of "renege," or "sell crack"—because the SE might change dramatically if subjects think there is even a possibility that the connotation of words creates a nonmonetary preference for certain strategies. (Natural labels might produce detectable changes in the size of the "homemade" priors measured below.) The price of this control is that subjects may be bored by a nonsalient task, but this didn't seem to happen. In future work we intend to see whether natural labels might cue subjects to appropriate behavior early in the experiment, eliminating the error-ridden learning that we observed with bland labels.

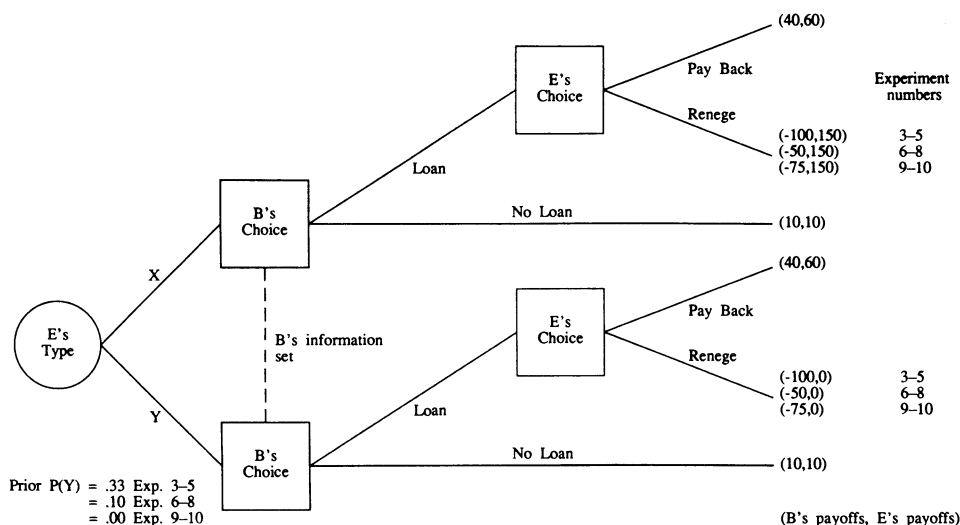


FIGURE 1.—The experiment game tree.

(*R*). An *X*-type *E* got a higher payoff from renegeing than from paying back. (Thus, the monetary payments subjects earned in our experiment induced non-monetary preferences like those associated with honesty (*Y*-type) and dishonesty (*X*-type) in the natural world.)

A single *E* subject played a sequence of 8 periods ("rounds," to them) against a series of different *B* subjects. *E*'s type was chosen from a bingo cage before each sequence of eight periods. The type was announced to *E* (but not to the *B*'s), and was fixed for an entire 8-period sequence. (After the sequence, the type was then announced to the *B*'s.) Both sides knew the prior probability that an *X*- or *Y*-type would result (the prior was .33 in experiments 3-5, .10 in experiments 6-8, and .00 in experiments 9-10). By reading the instructions aloud to all the subjects simultaneously, *everything* about the experiment's structure was made common knowledge (except for *E*'s type).

Subjects were MBA students at NYU or the University of Pennsylvania.<sup>4</sup> Each subject participated in only one experiment. Subjects were paid the sum of the payoffs from all their decisions, in cash, at the end of the experiment. (Average payoffs were about \$18 for the 2-1/2 hour experiment.)

Each experiment had 75-100 repetitions of 8-period sequences, to give the subjects lots of experience. Thus, the subjects actually observed 600-800 periods of play. If a single *E* subject played throughout an entire experiment, he might consider the experiment a long supergame with several hundred periods. But the theory models each 8-period sequence as if it stands alone—players are expected to regard period 8 as the last play of the game, though in the experiment they

<sup>4</sup> NYU MBA's were used in experiments 3, 5-6, 9-10; and University of Pennsylvania MBA's were used in experiments 7-8. Experiment 4 contained a mixture of University of Pennsylvania MBA's and undergraduates. The data suggest no striking differences between these subject pools.

know more 8-period sequences lie ahead. We would like to have the players see enough repetitions of the entire 8-period game to be able to learn a sophisticated equilibrium concept, while regarding each 8-period game as separate. To avoid this problem, we used *pools* of  $E$  and  $B$  subjects. Three  $E$  subjects sat in a room together, with one experimenter, and the eight  $B$  subjects sat in another room with a second experimenter. Before each sequence, we used pre-determined random numbers to choose *which* of the three  $E$  subjects would play that sequence, and this number was not announced to the  $B$  subjects. Similarly, in each sequence a random pre-generated order of the same  $B$  subjects was used, and that order was not known by the  $E$  subjects. In theory, this randomness made it difficult for either side to develop any stable reputation other than that we deliberately seeded through the  $X$ - and  $Y$ -type payoffs (as we show in Section 4.5.3 below).

After the experimenters read the instructions (see Appendix B) aloud to both  $E$  and  $B$  subjects, the  $B$  subjects stayed together in one room, and the  $E$  subjects went to an adjoining room. The experimenters communicated with each other and with the subjects in each room by walkie-talkie, to restrict communication so that subjects could not identify their partners by voice.<sup>5</sup>

### 3. COMPETING HYPOTHESES

#### 3.1. *The Sequential Equilibrium Prediction*

We will sketch the derivation of SE for the parameters in experiments 3–5, starting with the last (8th) period first.

SE play must be an equilibrium in every possible subgame. In subgames along the equilibrium path, players are assumed to use Bayes' rule to update their information about others based on observed play. In subgames off the equilibrium path, an *SE* must specify how players will update their beliefs (since Bayes' rule will usually not apply), and in these cases the *SE* is somewhat arbitrary. We solve for an *SE* by beginning in the last period, and calculate optimal play (including updating of beliefs) in the last period as a function of beliefs entering that period. Then we roll back to the second-to-last period, and calculate optimal play as a function of beliefs, taking into account the effect of second-to-last period actions on last-period beliefs, and hence on last-period play. We proceed this way back to the first period.

In the last period, indexed  $T$ , the  $B$  player knows that an  $X$ -type  $E$  player will certainly renege. Therefore, if  $B$  thinks the probability that a player is  $Y$ -type is  $P_T$ ,  $B$ 's expected value from choosing  $L$  is  $40P_T - 100(1 - P_T)$  (for the parameters of experiments 3–5). (We assume risk-neutrality, additively separable utility,

<sup>5</sup> One advantage of using this method is that when  $E$ 's were  $Y$ -types, they quickly learned that paying back was a dominant strategy. They chose to pay back very quickly to signal their type, since  $X$ -types had to think about whether to renege and  $Y$ -types did not, so a slow response from  $E$  could tip off  $B$ 's that  $E$  was an  $X$ -type. (This even happened in the experiments with  $P(Y) = 0$ , even though there were no  $Y$ -types.) To prevent speed from signaling  $E$ 's type, the experimenters "held" quick responses for a second or two before communicating them.

and no time discounting in making these calculations; but we generalize for other risk tastes below.) The gain from lending exceeds the sure gain from not lending, 10, if and only if  $P_T$  is greater than the "threshold" of  $110/140$ . Note that this  $P_T$  is  $E$ 's reputation, a measure of what type of player  $E$  is thought to be.

In period  $T-1$ , an  $X$ -type  $E$  could renege and get 150, and another 10 in period  $T$  (since  $B$  would not lend to  $E$  in period  $T$ ). Or  $E$  can play a mixed-strategy, choosing to pay back with probability  $S_{T-1}$  and reneging with probability  $1 - S_{T-1}$ .  $E$  will want to choose  $S_{T-1}$  so that when  $B$  observes  $E$ 's pay-back in period  $T-1$  and updates her beliefs about  $E$ , her updated posterior probability  $P_T$  is above the threshold of  $110/140$ . Then  $E$ 's total expected earnings from periods  $T-1$  and  $T$  are

$$(1) \quad S_{T-1}(60 + 150) + (1 - S_{T-1})(150 + 10) = 150 + 10 + S_{T-1}(60 - 10).$$

Intuitively, since these expected earnings are increasing in  $S_{T-1}$ ,  $E$  wants to choose  $S_{T-1}$  as large as possible, provided  $S_{T-1}$  makes the posterior probability  $P_T$  above  $B$ 's period  $T$  threshold of  $110/140$ . (More precisely, if the probability  $P_T$  is exactly at its threshold,  $B$  will be indifferent between lending and not lending.  $B$  will choose a mixed-strategy probability which makes  $E$  indifferent between reneging and paying back, and  $E$  will choose a mixed-strategy probability which makes the posterior probability  $P_T$  equal to its threshold.)

If  $B$  uses Bayes' rule to update probabilities, the posterior probability  $P_T$  is given by

$$(2) \quad P_T = 1 * P_{T-1} / [1 * P_{T-1} + S_{T-1}(1 - P_{T-1})].$$

For this posterior  $P_T$  to exceed the lending threshold  $110/140$  ( $= .786$ ) requires

$$(3) \quad S_{T-1} < P_{T-1}(1 - .786) / (1 - P_{T-1}) * .786.$$

A rational  $E$  will choose  $S_{T-1}$  to make (3) hold as an equality.

Now in period  $T-1$ ,  $B$  will lend if and only if

$$(4) \quad 40[P_{T-1} + S_{T-1}(1 - P_{T-1})] - 100(1 - S_{T-1})(1 - P_{T-1}) > 10.$$

Since  $E$ 's will choose  $S_{T-1}$  to satisfy (3), we can plug  $S_{T-1}$  from (3) into (4) and derive a threshold for  $P_{T-1}$ , which is simply

$$(5) \quad P_{T-1} > (.786)^2.$$

In period  $T-k$ ,  $E$ 's mixed-strategy probability should satisfy

$$(6) \quad S_{T-k} = P_{T-k}(1 - .786^k) / (1 - P_{T-k})(.786^k)$$

and  $B$ 's lending threshold in period  $T-k$  is

$$(7) \quad P_{T-k} > .786^{k+1}.$$

(That is, since even  $X$ -type  $E$ 's are likely to pay back in early periods,  $B$ 's require less and less assurance that  $E$  is a  $Y$ -type to convince them to lend, so the threshold is lower.)

If the game begins with some commonly-known prior probability  $h$  that  $E$  is a  $Y$ -type, the sequential equilibrium is for  $B$  to lend and for  $E$  to pay back, as long as the prior  $h$  is above the threshold in (7). But as  $k$  gets smaller (the end draws

near), the right-hand side of (7) grows, so  $h$  will eventually be *less* than  $.786^{k+1}$ . At that point, the  $E$  player must then do something to enhance his reputation—that is, to increase  $B$ 's posterior probability that  $E$  is a  $Y$ -type—otherwise  $B$  will refuse to lend in the remaining periods. In the period just before the inequality  $h > .786^{k+1}$  is violated,  $E$  begins playing mixed-strategies with probabilities of paying back given by (6).

Once mixed-strategy play begins,  $E$ 's choice of  $S_T$  in equilibrium makes  $B$  indifferent between strategies, and vice versa.  $B$  chooses to lend with probability  $M$ , where  $M$  makes  $E$  indifferent between values of  $S_{T-1}$  in the expression

$$(8) \quad S_{T-1}[60 + 150M + 10(1 - M)] + (1 - S_{T-1})[150 + 10] \\ = 150 + 10 + S_{T-1}[60 - 150 + M(150 - 10)].$$

For the parameters in all our experiments,  $M = (150 - 60)/(150 - 10)$ , or .643. This lending probability is optimal in every period where the posterior probability  $P_{T-k}$  exactly equals  $B$ 's threshold.

Sequential equilibria are usually not unique because theorists have some freedom in choosing what beliefs agents will hold after out-of-equilibrium moves. (Bayes' rule doesn't apply to out-of-equilibrium events, which have zero probability.) In the SE we describe, the only out-of-equilibrium moves to which Bayes' rule doesn't apply are renegees in the early periods (periods 1–3 in experiments 3–5, and periods 1–2 in experiments 6–8). Thus, to complete our SE we only need to specify what  $B$ 's will think after they observe an early-period (out-of-equilibrium) renege. In deriving the SE above, we assumed that if  $B$ 's observe an early-period renege they will believe  $E$  is an  $X$ -type, since renegeing is a dominated strategy for  $Y$ -types.<sup>6</sup>

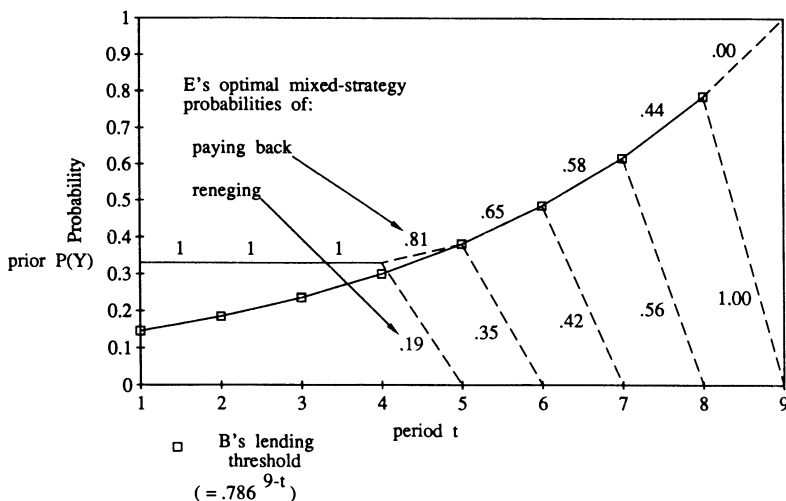
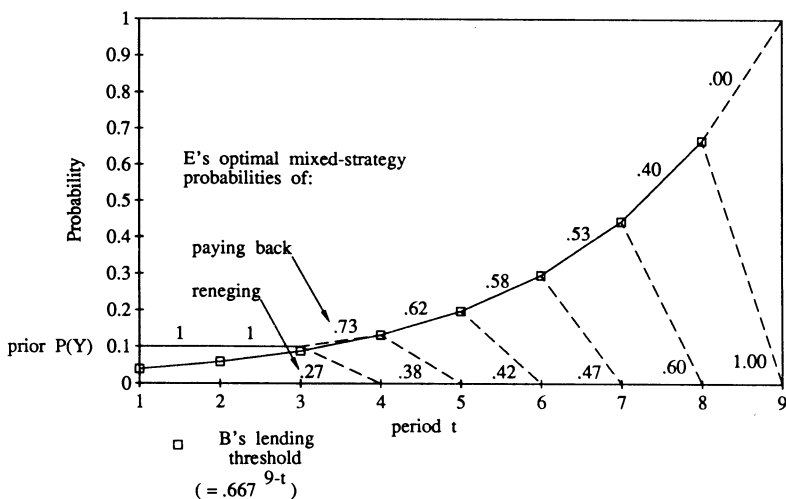
The parameters in our experiments were chosen so that the equilibrium makes testable implications about three kinds of probabilities: renegeing, “following,” and lending.

### 3.1.1. *Reneging Probabilities*

The graph in Figure 2a illustrates the equilibrium path of  $P(Y)$ , for an  $X$ -type  $E$  player, for the parameters in experiments 3–5. (Figure 2b is an analogous illustration for experiments 6–8, and these predictions are summarized in Table I.) In periods 1–3, the prior probability of  $1/3$  is above  $B$ 's threshold; so  $B$  should always lend and an  $X$ -type  $E$  should always pay back. Period 4 is a crucial period, because  $X$ -type  $E$ 's must begin playing mixed strategies to enhance their reputations, to make the probabilities  $P_{T-k}$  meet larger and larger thresholds. For instance, in experiments 3–5  $E$  should pay back with probability

<sup>6</sup> We have not explored whether this belief satisfies the many criteria proposed for refining sequential equilibria (e.g., Grossman and Perry, 1986; and see Cho and Kreps, 1987, for a criterion and a review). However, we shall consider the data for some guidance about what  $B$ 's believe after an early-period renege. We can also rule out many implausible SE, like the one in which all  $E$  types renege in the early periods,  $B$  does not lend, and if  $B$  lends and  $E$  pays back,  $B$  thinks  $E$  is an  $X$ -type.  $B$ 's latter belief is “inconsistent” in Grossman and Perry's (1986) sense (and probably can be ruled out by weaker criteria), since  $Y$  is one of the  $E$  types who prefers to pay back and  $B$ 's belief should reflect that.



FIGURE 2a.—The equilibrium path of  $P(Y)$  (for an  $X$ -type  $E$ ), experiments 3–5.FIGURE 2b.—The equilibrium path of  $P(Y)$  (for an  $X$ -type  $E$ ), experiments 6–8.

.81 in period 4, to boost the prior of  $1/3$  up to the next period's threshold of  $.786^4$ . (In Figure 2a, this is shown by a .81 probability on the branch leading to a higher  $P_5$ , and a .19 probability on the branch leading to  $P_5 = 0$ .) Note that the mixed-strategy probabilities of payback get smaller and smaller, since the prior must be boosted more and more to reach each higher threshold. In the 8th and last period—if  $E$  makes it that far without ruining his reputation by renegeing and sending  $P(Y)$  to 0—an  $X$ -type  $E$  will certainly renege; so in period 9 (i.e., after the sequence is over)  $P(Y) = 0$ .

TABLE I  
SUMMARY OF THEORETICAL PREDICTIONS

Experiments 3-5			Period	Experiments 6-8		
<i>E</i> 's Reneging Probability	<i>B</i> 's Lending Probabilities	"Following" Probability		<i>E</i> 's Reneging Probability	<i>B</i> 's Lending Probabilities	"Following" Probability
1.00	.643	1.00	8	1.00	.643	1.00
.560	.643	1.00	7	.597	.643	1.00
.415	.643	1.00	6	.473	.643	1.00
.347	.643	1.00	5	.416	.643	1.00
.192	1.00	.000	4	.380	.643	1.00
.000	1.00	.000	3	.269	1.00	.000
.000	1.00	.000	2	.000	1.00	.000
.000	1.00	—	1	.000	1.00	—

### 3.1.2. "Following" Probabilities

SE makes a nonobvious prediction about what will happen following a period with no loan. In a period where there is no loan the prior  $P_{T-k}$  does not get revised at all (since *E* has no chance to either maintain or ruin his reputation). In periods 1-3, even if there is no loan the initial prior of 1/3 is still large enough, even unrevised, to exceed the threshold in the following period. Therefore, no loan in periods 1-3 should be followed by a loan in the next period.

However, if there is no loan in period 4 the prior of 1/3 does not get revised, but it is below the period 5 threshold. The missed opportunity to build reputation in period 4 means that *B*'s should be unwilling to lend in period 5. No loan in period 4 (or in any subsequent period) should be followed by a string of no loans.

Thus, sequential equilibrium predicts a change in "following" probabilities (that is, the probability of no loan following no loan), from 0 to 1, in period 5 of experiments 3-5 and period 4 of experiments 6-8, as shown in Table I.

### 3.1.3. *B*'s Lending Probabilities

SE has another nonobvious property: *The payoffs for an X-type E do not affect E's equilibrium play at all* (assuming that the reneging payoff is less than the payoff from paying back). *E*'s payoffs only affect the equilibrium through *B*'s choice of lending probability in later periods.

As shown above, *B*'s optimal lending probability is .643 in all periods where the probability  $P_{T-k}$  is at the lending threshold. (Mixing strategies in this way is rational in the usual weak sense: *B* does not benefit from playing a mixed-strategy, but no pure strategy is better.)

Changes in *E*'s payoffs produce counter-intuitive changes in *B*'s lending probabilities. For instance, as *E*'s reneging payoff of 150 increases, one might think that indifferent *B*'s would be *less* likely to lend. The opposite is true: As 150 gets large compared to *E*'s payoffs from paying back and from getting no loan, the fraction *M* approaches 1—if they are indifferent, *B*'s are almost sure to lend to make *E*'s indifferent between reneging now and reneging later. In this

paper we do not test whether changing  $E$ 's payoffs actually changes  $B$ 's mixed-strategy play in the predicted way, but we can do so in further research.

One caveat: All these predictions about players' behavior depend on the assumption that others are acting in accord with SE, and predictions in later periods depend upon the assumption that an SE has been played in earlier periods. Thus, we must judge the accuracy of predictions about later-period play conditional on observed early-period play. If play is not consistent with SE in early periods, for instance, then later-period play might be inconsistent with the overall SE described above, but consistent with the SE prediction beginning off the equilibrium path.

### 3.2. *Competing Hypotheses*

In experimental economics we try to test hypotheses against serious alternative hypotheses, rather than against toothless null hypotheses, to make "strong inferences" (Platt, 1964). There are no alternative hypotheses as precise as SE, but we shall try to specify a few plausible reasons why results may depart from SE.

**COMPETING THEORY #1: *Unraveling*.** If players ignore the possibility that  $E$ 's may be honest, then "unraveling" will result:  $B$ 's will think that  $E$  will certainly renege in the 8th period, so  $E$  will then renege in the 7th period, and so forth. No loans will result. Persistent reputation-building (paying back by  $X$ -type  $E$ 's) will falsify this theory.

**COMPETING THEORY #2: *Social Norms*.** The cognitive or social force of traditions, or social norms, might freeze players into an equilibrium other than SE. (In economic terms, players may substitute the costs of thinking, or a taste for conformity to implicit group standards of behavior, for money gains.) As typically stated (e.g., Bettenhausen and Murnighan, 1985), this social norms prediction is hard to falsify, but we suppose that a social norm implies consistency of behavior. For instance, the norm "never renege until period 4" might develop among  $E$  players. This theory can be falsified if we shift parameters so that SE predicts renegeing in period 3 (as in experiments 6–8), and we then observe renegeing in period 3. Within an experiment, we can judge the strength of norms by whether each series of decisions varies markedly from sequence to sequence, and by whether individual players exhibit the same patterns of behavior.

SE makes certain predictions which contradict squarely the notion that deterministic norms govern behavior. For instance, SE predicts that no loans will be followed by no loans only in later periods, but not in early periods. A conformity theory which predicts that  $B$ 's do what others do, always following no loans with no loans, can be falsified if the SE prediction proves correct.

**COMPETING THEORY #3: *Risk Tastes*.** In our experiments we deliberately do not control or measure the risk-aversion of subjects (cf. Roth, 1983; Berg, et al.,

1986; Harrison and McKee, 1985). However, it is easy to show<sup>7</sup> that if  $B$  players are risk-averse (-seeking), the lending threshold is higher (lower). This means that under risk-aversion (-seeking),  $E$ 's should begin to renege earlier (later), but they will renege less often (more often) than predicted by SE under risk-neutrality.<sup>8</sup> If  $E$ 's are risk-averse,  $B$ 's will lend less often (when they do play the mixed-strategy); but if  $E$ 's are risk-seeking,  $B$ 's lend more often.

**COMPETING THEORY #4: *Altruism, or Envy.*** Call players "altruistic" (or "guilty") if they are happier when other players earn more (or, if they are unhappier when others earn less); and call them "envious" if they are happier when other players earn less. We can model such attitudes by assuming a  $B$  player earns a fraction  $g_B$  of the  $E$  player's payoff (and similarly for  $E$ 's). Then the lending threshold becomes  $(110 - 140g_B)/(140 - 90g_B)$ , and  $B$ 's mixed-strategy probability becomes  $(90 - 140g_E)/(140 - 110g_E)$ . Thus, if  $B$ 's are altruistic ( $g_B > 0$ ) the lending threshold is lower, so renegeing begins later but  $E$ 's renege more often (see footnote 7) than predicted by  $g_B = 0$ . If  $B$ 's are envious ( $g_B < 0$ ), renegeing begins earlier but  $E$ 's renege less often. If  $E$ 's are altruistic ( $g_E > 0$ ),  $B$ 's lend less often than predicted; and if  $E$ 's are envious ( $g_E < 0$ ),  $B$ 's lend more often.

**COMPETING THEORY #5: *Homemade Priors.*** The previous experimental work reviewed in the introductory section strongly suggests that subjects may have homemade priors about the information or tastes of others, along with the prior probabilities that we deliberately introduce into the experiment. For instance, even if we set the prior  $P(Y)$  equal to zero, players may act as if some fraction of the  $E$ 's behave honestly, though we have not induced them to do so. (For instance, the homemade prior might be the percentage of  $E$  players who are so altruistic that they will always pay back rather than renege.) If subjects do use a homemade prior that increases the effective  $P(Y)$ , then renegeing will begin later in the game than we predict; but once renegeing has started, the amount of renegeing will be the same as predicted.

Some kinds of homemade priors should not affect the game. For instance, suppose the subjects think some  $B$ 's will be so altruistic (or irrational) that they always choose to lend (even after a renege reveals that  $E$  is an  $X$ -type). Unless this percentage is quite high (e.g., for experiments 3–5, around 36 per cent), it

<sup>7</sup> For this simple game, risk attitudes can be easily summarized by defining the utility of the 10-franc  $N$  payoff as  $x$ . If  $x = 10$ , players are risk-neutral;  $x < 10$  means players are risk-seeking; and  $x > 10$  means players are risk-averse. The critical threshold then becomes  $(100 + x)/140$ , so if players are risk-averse and  $x > 10$ , this threshold is larger. If players are risk-seeking, the threshold is smaller. For  $E$  players' risk-tastes, let the parameter  $60 = x$ , and the ratio  $M$  that determines  $B$ 's mixed-strategy lending is then  $(150 - x)/(150 - 10)$ . This ratio is lower if  $E$ 's are risk-averse and higher if  $E$ 's are risk-seeking.

<sup>8</sup> Suppose play is on the equilibrium path, and call the threshold  $t$ . Then in period  $T - k - 1$ ,  $E$  chooses a mixed-strategy  $S_{T-k-1}$  to satisfy  $S_{T-k-1} = t^{T-k-1}(1 - t^{T-k})/t^{T-k}(1 - t^{T-k-1})$ . Differentiation shows that this expression is increasing in  $t$ , so if the thresholds are higher (lower) by risk-aversion (-seeking), then  $E$ 's will pay back more (less) than predicted under risk-neutrality.

will not affect  $E$ 's play at all.<sup>9</sup> The reason is that as long as  $E$ 's payoffs in any period are increasing in the mixed-strategy probability of paying back,  $E$  will choose that probability to be as large as possible (though constrained by the requirement that the posterior probability of honesty is above  $B$ 's lending threshold). Unless the chance of a  $B$  always lending is high,  $E$ 's payoffs will be increasing in the mixed-strategy probability, and the game is unaffected. (Keep in mind that  $B$ 's "type" does not matter much because in each period a different  $B$  plays, so  $B$ 's cannot build up reputation effects of their own.)

Note that competing theories #3–5 are only competing theories in a specialized sense. In each case, we have asked what happens if some perturbation of subjects' payoffs that we cannot entirely control—risk-tastes, interdependence in utilities, or homemade priors about others' utilities—actually comes into play, and we have worked out the SE for each perturbation. These competing SE's are not competing solution concepts, they are simply attempts to anticipate possible deviations between predicted SE and actual behavior, based on uncontrolled payoffs. If SE with homemade priors, or SE with risk-aversion, turns out to describe behavior well, then we can still take the data as supportive of SE as a descriptive solution concept, provided that the subjects' homemade beliefs or risk tastes are empirically systematic.

#### 4. RESULTS

We conducted three experiments (numbered 3–5) with a prior  $P(Y)$  of .33, and three experiments (6–8) with a prior of .10. All raw data from these experiments are shown in Appendix A, in Tables A.1–A.6 (except for the orders in which the  $B$  players actually played in each sequence, which are available from the authors). There is evidence of reputation-building in *all* experiments. The majority of  $E$  players did *not* renege at the first opportunity in the first sequence they played.

The weakest requirement of SE in this game is that  $E$ 's should not renege in early periods, and  $B$ 's should not lend after observing a previous renege in that sequence. In the first 30 sequences or so, there were some violations of these minimal conditions, but the violations disappeared (with a few exceptions) as subjects gained experience. Since most authors discuss games with more than 8 periods when explaining reputation-building,<sup>10</sup> it is significant that we observe reputation-building in games with only 8 periods. (We also saw reputation-building in pilot experiments with 6 periods.)

Table I summarizes the predictions of SE about renegeing, lending, and following probabilities, in the two sets of experiments. We test these predictions with

<sup>9</sup> Take the period  $T-1$  play as illustrative. Suppose the chance of a perfectly altruistic  $B$  is  $K$ . Then  $E$  chooses  $S_{T-1}$  to maximize  $S_{T-1}(60 + 150) + (1 - S_{T-1})(150 + 150K + 10(1 - K))$ . This reduces to  $160 + 140K + S_{T-1}(50 - 140K)$ . The expression in parentheses is only negative, thereby changing  $E$ 's optimal choice, if  $K > 50/140$  (around .36). This threshold for  $K$  goes down as the game rolls back to the beginning, but for the parameters we have worked through it never gets so low that we expect  $E$ 's to renege earlier than predicted by SE.

<sup>10</sup> E.g., Selten (1978, p. 153) hints that games with 2–4 periods will have no reputation-building, but 20-period games will.

cross-sectional data from each experiment, and with data pooled across experiments.

For example, to estimate the reneging probability  $P(R_{T-k}|\text{no previous } R)$ , we counted the fraction of times in each round  $T-k$  that  $X$ -type  $E$  subjects who had not reneged earlier in that sequence were given a loan and then reneged.<sup>11</sup> (We excluded periods in which  $E$ 's had reneged earlier in the sequence because late-period lending in these sequences could be disequilibrium errors.)

In each of the experiments we divided the sample of 8-period sequences into thirds, and analyzed each third separately. (In experiment 3, for instance, we analyzed sequences 1–30, 31–60, and 61–90 separately.) This simple way of dividing the data helps distinguish the period of learning and disequilibrium, which mostly occurs in the first third of the experiment, from equilibrium behavior exhibited in the last two thirds of the experiment. We often pooled data from the second and third thirds of the sequences in each experiment, since these thirds were generally quite similar. (Indeed, that similarity is our assurance that we are observing equilibrium behavior.)

We shall discuss in turn each of the three kinds of probabilistic predictions: Reneging, lending, and following.

#### 4.1. *Reneging Probabilities*

Sequential equilibrium predicts that the mixed-strategy probability of an  $X$ -type  $E$  reneging,  $P(R_T|\text{no previous } R)$  rises monotonically, but erratically, across rounds (recall Table I, or Figures 2a–b for the predictions).

Because data from replications of the same experiment were very similar, we pooled data to get more statistical power in estimating the reneging probabilities. Tables II and III show reneging probability estimates from experiments 3–5 pooled (prior  $P(Y) = .33$ ) and experiments 6–8 pooled ( $P(Y) = .10$ ). Figures 3 and 4 show the predicted reneging probabilities and the reneging estimates from the last two-thirds of the sequences, pooled across experiments (with 90 per cent confidence intervals around the estimated means).

If we take the mixed-strategy probability predicted by SE to be a null hypothesis about a binomial proportion, we can test that hypothesis with the standard normal approximation to the binomial distribution.<sup>12</sup> These test statis-

<sup>11</sup> Note that the number of informative observations could be quite low in later periods, since it was not often that a dishonest  $E$  made it to a late period without having reneged, and then got a loan. This is the primary reason why we ran many repetitions of a short game, rather than a longer game with few repetitions.

<sup>12</sup> That is, we calculate standard errors as  $(p(1-p)/n)^{1/2}$ , where  $p$  is the probability predicted by SE and  $n$  is the sample size for a particular round. Then the test statistic,  $p$  minus the estimated  $p$  divided by the standard error, is normally distributed around zero with variance of one if the null hypothesis is true. If the hypothesized proportion is zero or one, we can estimate the standard error using the observed proportion. We also checked that observations were independent, because tests of binomial proportions rely on the assumption of independence. The choices in a given round were remarkably independent across sequences—for instance, choices in the 6th round when a loan was made to an  $X$ -type  $E$  were approximately uncorrelated from sequence to sequence, perhaps because a different  $B$  subject was randomly chosen to play in the 6th round in each sequence.

TABLE II  
PREDICTED VS. ACTUAL RENEGING BY *E* PLAYERS, EXPERIMENTS 3–5  
 $P(R \text{ in } t | \text{no previous } R)$ ;  $P(Y) = .33$

Period	First Third Actual	Second Third Actual	Last Third Actual	Last Two Thirds Pooled Actual	Pre- dicted	z-score
8	.909 (10/11)	.909 (10/11)	.818 (9/11)	.864 (19/22)	1.000	—
7	.600 (15/25)	.625 (5/8)	.445 (4/9)	.529 (9/17)	.560	.26
6	.167 (6/36)	.320 (8/25)	.240 (6/25)	.280 (14/50)	.415	1.93
5	.053 (2/38)	.156 (7/45)	.119 (5/42)	.138 (12/87)	.347	4.10
4	.075 (3/40)	.082 (5/61)	.018 (1/56)	.051 (6/117)	.192	3.92
3	.089 (4/45)	.000 (0/59)	.033 (2/61)	.017 (2/120)	.000	—
2	.087 (4/46)	.048 (3/62)	.016 (1/62)	.032 (4/124)	.000	—
1	.222 (10/45)	.081 (5/62)	.015 (1/65)	.047 (6/127)	.000	—

TABLE III  
PREDICTED VS. ACTUAL RENEGING BY *E* PLAYERS, EXPERIMENTS 6–8  
 $P(R \text{ in } t | \text{no previous } R)$ ;  $P(Y) = .10$

Period	First Third Actual	Second Third Actual	Last Third Actual	Last Two Thirds Pooled Actual	Pre- dicted	z-score
8	.920 (11/12)	1.00 (7/7)	1.00 (6/6)	1.00 (13/13)	1.000	—
7	.500 (11/22)	.467 (7/15)	.556 (10/18)	.515 (17/33)	.597	.96
6	.184 (7/38)	.429 (15/35)	.120 (3/25)	.300 (18/60)	.473	2.70
5	.146 (7/48)	.157 (8/51)	.150 (6/40)	.154 (14/91)	.416	5.07
4	.020 (2/49)	.067 (4/60)	.088 (5/57)	.077 (9/117)	.380	6.75
3	.111 (6/54)	.077 (5/65)	.000 (0/65)	.031 (4/130)	.269	5.94
2	.036 (2/55)	.047 (3/64)	.046 (3/65)	.047 (6/129)	.000	—
1	.093 (5/54)	.016 (1/65)	.046 (3/66)	.031 (4/131)	.000	—

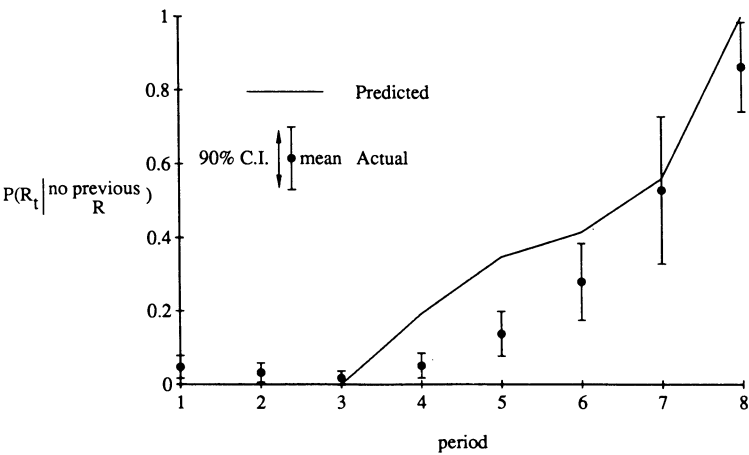


FIGURE 3.—Predicted vs. actual probabilities of reneging by *E* players (with 90 per cent confidence intervals), experiments 3–5 pooled.

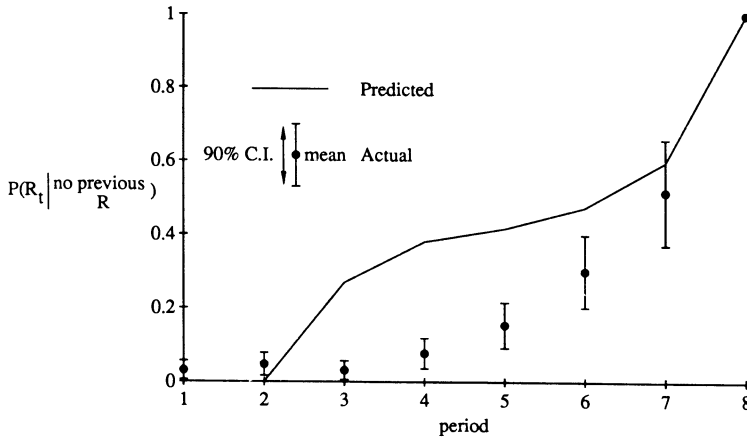


FIGURE 4.—Predicted vs. actual probabilities of reneging by *E* players (with 90 per cent confidence intervals), experiments 6–8 pooled.

tics are reported in the right-hand column (“*z*-scores”) in Tables II and III, and shown graphically in Figures 3 and 4 by the 90 percent confidence intervals (with confidence interval standard errors calculated from the estimated proportions, rather than from predicted proportions).

The actual proportions of reneging were generally lower than the SE predictions (so the *z*-scores are almost always positive), significantly so in early rounds 3–5. (The SE predictions in rounds 6 and 7 are more accurate, but the small samples in those later rounds don’t give us much power to distinguish between a true null hypothesis and alternative hypotheses close to SE.)

The hypothesis tests suggest that the deviations between SE and the data are much too large to be due to chance, except in round 7 from both sets of experiments. However, the estimated proportions of reneging *do* rise almost monotonically: Note from Figures 3 and 4 that the upper bound of each round’s confidence interval generally is below (or almost below) the lower bound of the next round’s confidence interval.

#### 4.2. Lending Probabilities

The estimates of *B*’s lending probabilities,  $P(L_T | \text{no } N_{T-1}, \text{no previous } R)$  are shown in Tables IV–V, for data pooled across experiments 3–5 and experiments 6–8. Generally, the probabilities are very close to one (*B*’s always lend) in early rounds, and less than one in later rounds. In experiments 3–5 (Table IV) we see a drop in the actual proportion of lending, from .913 to .721, between periods 4 and 5 where SE predicts a drop from 1.00 to .643.<sup>13</sup> (Note that this drop is absent in the first third of the sequences—the change between rounds 4 and 5 is from .894 to .897—suggesting that it takes subjects a little while to learn to begin using mixed strategies in round 5.) Testing the hypothesis that  $p = .643$  for each of

<sup>13</sup> A large-sample normally-approximated *z*-test for the difference in proportions between rounds 4 and 5 yields  $z = 5.26, 1.16$ , and  $1.80$  for experiments 3–5 separately and  $z = 6.00$  for the last two thirds of the sequences pooled across experiments.



TABLE IV  
 PREDICTED VS. ACTUAL LENDING BY *B* PLAYERS, EXPERIMENTS 3–5  
 $P(L \text{ in } t | \text{No } N \text{ in } t-1, \text{ no previous } R); P(Y) = .33$

Period	First Third Actual	Second Third Actual	Last Third Actual	Last Two Thirds Actual	Predicted	z-score
(1–4)	.819 (244/298)	.930 (304/327)	.959 (328/342)	.945 (632/669)	1.00	—
(5–8)	.834 (151/181)	.626 (87/139)	.608 (101/166)	.616 (188/305)	.643	1.00
8	.870 (20/23)	1.00 (5/5)	.538 (7/13)	.667 (12/18)	.643	–.21
7	.667 (30/45)	.318 (7/22)	.419 (13/31)	.377 (20/53)	.643	4.03
6	.891 (49/55)	.548 (23/42)	.615 (32/52)	.585 (55/94)	.643	1.18
5	.897 (52/58)	.743 (52/70)	.700 (49/70)	.721 (101/140)	.643	–1.95
4	.894 (59/66)	.937 (74/79)	.890 (73/82)	.913 (147/161)	1.00	—
3	.900 (63/70)	.949 (75/79)	.965 (82/85)	.957 (157/164)	1.00	—
2	.829 (63/76)	.951 (78/82)	.977 (85/87)	.964 (163/169)	1.00	—
1	.686 (59/86)	.885 (77/87)	1.00 (88/88)	.943 (165/175)	1.00	—

TABLE V  
 PREDICTED VS. ACTUAL LENDING BY *B* PLAYERS, EXPERIMENTS 6–8  
 $P(L \text{ in } t | \text{No } N \text{ in } t-1, \text{ no previous } R); P(Y) = .10$

Period	First Third Actual	Second Third Actual	Last Third Actual	Last Two Thirds Actual	Predicted	z-score
(1–3)	.889 (177/199)	.986 (210/213)	.980 (201/205)	.983 (411/418)	1.00	—
(4–8)	.862 (175/203)	.825 (175/212)	.784 (149/190)	.806 (324/402)	.643	–6.79
8	.625 (10/16)	.500 (5/10)	.429 (3/7)	.471 (8/17)	.643	1.48
7	.735 (25/34)	.500 (13/26)	.682 (15/22)	.583 (28/48)	.643	.87
6	.848 (39/46)	.745 (38/51)	.683 (28/41)	.717 (66/92)	.643	–1.48
5	.961 (50/52)	.917 (55/60)	.778 (42/54)	.850 (97/114)	.643	–4.60
4	.928 (51/55)	.985 (64/65)	.924 (61/66)	.954 (125/131)	.643	–7.40
3	.952 (60/63)	1.00 (69/69)	1.00 (65/65)	1.00 (134/134)	1.00	—
2	.923 (60/65)	1.00 (72/72)	.985 (67/68)	.993 (139/140)	1.00	—
1	.803 (57/71)	.958 (69/72)	.958 (69/72)	.958 (138/144)	1.00	—

rounds 5–8 separately, we see z-scores (Table IV) barely consistent with the SE prediction (and far off in round 7). Testing for the data from rounds 5–8 combined (in the second line of the table), we see very close convergence, with an estimated proportion of .616 (and with a sample large enough to convince us that the accuracy of the .643 prediction is probably not due to chance).

The data from experiments 6–8 in Table V are less supportive of SE. We do not see the sharp drop from rounds 3 to 4 as predicted (though the drop in last two-thirds data, pooled, is statistically significant<sup>14</sup>), but there is a monotonic drop in *B*'s propensity to lend in every round from rounds 4 through 8. (And note that z-scores for the hypothesis  $p = .643$  tested on each round's data

<sup>14</sup> That is, a large-sample z-test for the difference between the proportions 1.00 and .954 yields  $z = 2.56$ .

TABLE VI  
PREDICTED VS. ACTUAL FOLLOWING BY *B* PLAYERS, EXPERIMENTS 3–5  
 $P(N \text{ in } t | N \text{ in } t-1, \text{ no previous } R); P(Y) = .33$

Period	Experiment 3 (Seqs. 31–90) Actual	Experiment 4 (Seqs. 31–90) Actual	Experiment 5 (Seqs. 27–82) Actual	Experiments 3–5 Pooled Actual	Predicted
(2–4)	.333 (1/3)	.100 (1/10)	.200 (2/10)	.174 (4/23)	.000
(5–8)	.692 (36/52)	.833 (40/48)	.694 (25/36)	.743 (101/136)	1.00
8	.500 (5/10)	1.00 (11/11)	.813 (13/16)	.784 (29/37)	1.00
7	1.00 (16/16)	.929 (13/14)	.692 (9/13)	.884 (38/43)	1.00
6	.681 (15/22)	.769 (10/13)	.600 (3/5)	.700 (28/40)	1.00
5	.000 (0/4)	.600 (6/10)	.000 (0/2)	.375 (6/16)	1.00
4	(0/0)	.125 (1/8)	.000 (0/1)	.111 (1/9)	.000
3	.000 (0/1)	(0/0)	.000 (0/5)	.000 (0/6)	.000
2	.500 (1/2)	.000 (0/2)	.500 (2/4)	.375 (3/8)	.000

TABLE VII  
PREDICTED VS. ACTUAL FOLLOWING BY *B* PLAYERS, EXPERIMENTS 6–8  
 $P(N \text{ in } t | N \text{ in } t-1, \text{ no previous } R); P(Y) = .10$

Period	Experiment 6 (Seqs. 24–70) Actual	Experiment 7 (Seqs. 26–77) Actual	Experiment 8 (Seqs. 24–69) Actual	Experiments 6–8 Pooled Actual	Predicted
(2–3)	(0/0)	.000 (0/7)	(0/0)	.000 (0/7)	.000
(4–8)	.737 (14/19)	.934 (57/61)	.878 (43/49)	.884 (114/129)	1.00
8	.867 (13/15)	1.00 (18/18)	.900 (18/20)	.925 (49/53)	1.00
7	(0/0)	1.00 (19/19)	.842 (16/19)	.921 (35/38)	1.00
6	(0/0)	.867 (13/15)	.889 (8/9)	.875 (21/24)	1.00
5	(0/0)	.750 (6/8)	1.00 (1/1)	.778 (7/9)	1.00
4	(0/0)	1.00 (1/1)	(0/0)	1.00 (1/1)	1.00
3	(0/0)	.000 (0/1)	(0/0)	.000 (0/1)	.000
2	(0/0)	.000 (0/6)	(0/0)	.000 (0/6)	.000

separately are not far from zero in rounds 6–8, but are very far off in rounds 4–5, and for the pooled data.)

#### 4.3. *B*'s Following Probabilities

SE predicts the probability of no loan following a no loan in the previous period will change from 0 to 1 between rounds 4 and 5 in experiments 3–5, and between rounds 3 and 4 in experiments 6–8. (In general, the last round with zero following probability is the round in which *E* begins playing mixed-strategies.)

Tables VI–VII show the data from experiments 3–5 and 6–8 (with data from the last two thirds of each experiment pooled because samples were very small). We do see changes in the following probability as predicted, but the samples are too small to permit powerful tests. Pooling rounds with the same predicted following probability (2–4 and 5–8 in Table VI; 2–3 and 4–8 in Table VII) and testing for differences in proportions yields  $z = 5.42$  for experiments 3–5. The

small samples in experiments 6–8 make the  $z$ -test inappropriate, but the observed probabilities—0/7 and 114/129—are consistent with the following-probability predictions of SE.

#### 4.4. *Competing Theories*

The data are generally consistent with the qualitative predictions of SE about the proportions of reneging, lending, and following. The data cannot distinguish subtle differences between perfect or sequential equilibria<sup>15</sup> and theories of heuristic or boundedly rational play which approximate perfect equilibrium, but they can rule out many simple theories. We think the data are best taken as an antidote to the argument that SE couldn't possibly describe behavior because it is too complicated. Since the precise numerical predictions of SE are often rejected, especially in the consistent under-reneging of  $E$  players, we now consider whether competing theories can explain the systematic deviations between SE and actual behavior.

Clearly, competing theory #1 (unraveling) is falsified. (Below, we report experiments with  $P(Y) = 0$  in which unraveling is the theoretical prediction of perfect and sequential equilibrium, but it is falsified there also.)

Competing theory #2 (social norms) is harder to falsify. If social norm theories imply consistency between players, they are falsified by the systematic variations in players' strategies (see Tables XIII–XIV for data from experiment 7). Players do *not* seem to obey group norms and act similarly. If social norm theories imply consistency across sequences, they are contradicted by the chronic variation across sequences, which is easily seen from the raw data in Appendix A. Any norm theory which suggests different subjects will all do the same thing in a particular period cannot stand up to the data. Theories like “lend until period 8,” “never renege until period 6,” etc., are clearly false.

Theories which predict the same norms will arise in different experiments, perhaps because of focal points in the game structure, are ruled out by the responsiveness of the data to parameter shifts. For instance, the theory that people in experiments 3–5 obey the norm “don't renege until halfway through the game (period 4)” is falsified by reneging in period 3 of experiments 6–8 (where SE predicts reneging). Any norm theories which are invariant to the proportion of ( $Y$ ) players are similarly falsified by comparing experiments 3–5 and 6–8. The theory that  $B$  players conform whenever other players choose no loan before them is falsified by the change in following probabilities between periods 4–5 (experiments 3–5) and periods 3–4 (experiments 6–8)—in early periods,  $B$ 's do not conform.

In fairness, we note that these experiments are not very conducive to the development of norms because players cannot communicate, deviance cannot be

<sup>15</sup> A more demanding test would involve changing  $E$ 's payoffs, or any other parameters which should produce large changes according to competing theories and small changes according to SE. We shall do this in future work, and also gather probability judgments from players during the game, to test whether strategies *and* beliefs are in equilibrium as SE predicts.

punished, and the experiments do not last very long. However, norms or conformity pressures might explain some of the disequilibrium behavior in early sequences of an experiment, like the tendency for *B* players to lend even after the *E* player reneged on a loan earlier in the same sequence.

The basic predictions of SE seem correct—*E*'s renege increasingly toward the end, *B*'s lend sporadically in late periods—but the observed proportions of renegeing and lending do not match the proportions predicted by SE. Since competing theories #3–5—risk-tastes, altruism or envy, and homemade priors—predict *some* SE, but with probabilities different than those we have specified, these competing theories might help explain the deviations from SE.

The homemade prior theory (#5) predicts that renegeing (and mixed-strategy lending) will begin *later* than we predict, but once renegeing begins, the proportions of renegeing and lending should be as predicted. By contrast, the risk-tastes (#4) and altruism or envy (#5) theories predict that renegeing will start later, and *E*'s will over-renege (for altruism or risk-seeking); or renegeing will start earlier, and *E*'s will under-renege (for envy or risk-aversion). The homemade prior theory has no effect on *B*'s probability of lending, but the other theories predict more lending (risk-seeking or envy) or less lending (risk-aversion or altruism).

The stylized facts seem to be that *E*'s start to renege later, and then under-renege, while *B*'s lend about the correct amount of the time (perhaps over-lending in experiments 6–8). None of the competing theories explains these stylized facts entirely well, but after observing experiments 3–8 we realized that the homemade prior theory could explain many of the facts.<sup>16</sup> This theory has the distinct advantage of predicting the same amount of *B* lending as SE predicts. Since the lending predictions seemed accurate, we fit the homemade prior theory to the experiment 3–8 data, and ran more experiments (described below) to see how well the fitted theory could predict. We note also that the homemade prior theory is supported by the earlier empirical work which finds reputation formation where no incomplete information has been explicitly introduced, as in frequent cooperation in the repeated prisoners' dilemma.

#### 4.5. *The Homemade Prior Theory, Revisited and Jeopardized*

We focus on a particular homemade prior theory: subjects think some proportion of the *X*-type *E*'s will always pay back (like *Y*-types). We first estimated this proportion from the data on experiments 3–5 and 6–8, then we ran further experiments to test whether the theory that subjects play an SE with the same honesty proportion we measured in experiments 6–8.

The total prior  $P(Y)$  will consist of the prior we created (either .33 or .10), along with some fraction  $q$  of the proportion  $1 - P(Y)$  of *X*-type *E*'s, who behave like *Y*-types despite our efforts to induce *X*-type preferences. We can try to estimate this total prior (then calculate  $q$  from the estimated total) by seeing in which period *E*'s begin renegeing with any frequency.

<sup>16</sup> We thank the referees and David Kreps for pointing us in this direction.

#### 4.5.1. *Estimating the Homemade Prior in Experiments 3–8*

According to SE,  $E$ 's should begin reneging in the period just before the threshold exceeds the prior. For instance, in experiments 3–5, there is little reneging (i.e., less than 5 per cent, with no apparent upward trend) until period 5, when  $X$ -type  $E$ 's reneged 13.8 per cent of the time (Table II). The thresholds in periods 5 and 6 are .381 and .485, respectively. Thus, we estimate the total prior to be between .381 and .485, but we can be even more precise. If the  $X$ -type  $E$ 's are playing optimally, then their reneging probability in period 5 is chosen to make the posterior  $P(Y)$  in period 6 equal to the threshold of .485. (However, note that the observed fraction of reneging, .138, is an average of zero reneging by the  $q$  proportion of  $X$ -types who behave like  $Y$ 's, and a proportion of reneging .138/(1 –  $q$ ) by the 1 –  $q$  proportion of  $X$ -types.) Thus, using Bayes' rule and assuming optimality, we can infer that the total prior  $P(Y)$  satisfies

$$(10) \quad .485 = \frac{1 \cdot P(Y)}{1 \cdot P(Y) + \left(1 - \frac{.138}{1 - q}\right)(1 - P(Y))}.$$

Since  $P(Y) = 1/3 + 2q/3$  by assumption, solving (10) for  $q$  yields  $q = .161$ .

The calculation for experiments 6–8 works the same way. Since  $E$ 's reneged only 7.7 per cent of the time in period 4 (Table III), and 15.4 per cent of the time in period 5, we take period 5 to be the first reneging period. The threshold in period 6 is .296, so the analogous condition to (10) for the total prior  $P(Y)$  is

$$(11) \quad .296 = \frac{1 \cdot P(Y)}{1 \cdot P(Y) + \left(1 - \frac{.154}{1 - q}\right)(1 - P(Y))}.$$

Substituting  $P(Y) = .1 + .9q$  and solving for  $q$  yields  $q = .172$ . This is quite close to the estimate of  $q = .161$  from experiments 3–5, which used different subjects and different parameters.

#### 4.5.2. *Testing the Homemade Prior Theory*

Since the homemade prior appears to have some empirical regularity, we ran two experiments with an initial prior  $P(Y) = 0$ , to test the theory that subjects act as if  $P(Y)$  is around .17. (Zero-prior experiments can also falsify the unraveling theory, and create some continuity between our work and the large number of earlier experiments with no controlled incomplete information.)

Experiments 9–10 were run exactly like experiments 3–8, except that we used no bingo cage to randomly determine whether  $E$  was an  $X$ -type or a  $Y$ -type, and we changed one parameter as a test of robustness ( $B$ 's reneging payoff was –75, not –50 or –100 as in earlier experiments). The raw data are shown in Appendix A, Tables A.7–A.8, and summary statistics are provided in Tables VIII–XII. Using a (completely homemade) prior  $P(Y)$  of .17, we predict that

TABLE VIII  
PREDICTED VS. ACTUAL RENEGING BY *E* PLAYERS, EXPERIMENT 9  
 $P(R \text{ in } t | \text{no previous } R); P(Y) = 0$

Period	First Third Actual	Second Third Actual	Last Third Actual	Last Two Thirds Pooled Actual	Pre- dicted
8	(0/0)	.667 (2/3)	.750 (3/4)	.714 (5/7)	1.000
7	1.00 (1/1)	.400 (2/5)	.588 (10/17)	.545 (12/22)	.575
6	(0/0)	.333 (2/6)	.100 (2/20)	.154 (4/26)	.436
5	.667 (2/3)	.000 (0/4)	.048 (1/21)	.040 (1/25)	.374
4	.600 (3/5)	.111 (1/9)	.091 (2/22)	.097 (3/31)	.332
3	.600 (6/10)	.375 (3/8)	.103 (3/29)	.162 (6/37)	.274
2	.400 (8/20)	.286 (4/14)	.000 (0/29)	.093 (4/43)	.000
1	.148 (8/54)	.153 (9/56)	.034 (2/59)	.093 (11/118)	.000

TABLE IX  
PREDICTED VS. ACTUAL RENEGING BY *E* PLAYERS, EXPERIMENT 10  
 $P(R \text{ in } t | \text{no previous } R); P(Y) = 0$

Period	First Third Actual	Second Third Actual	Last Third Actual	Last Two Thirds Pooled Actual	Pre- dicted
8	.667 (2/3)	(0/0)	(0/0)	(0/0)	1.000
7	.667 (2/3)	.333 (1/3)	(0/0)	.333 (1/3)	.575
6	.364 (4/11)	1.00 (2/2)	(0/0)	1.00 (2/2)	.436
5	.143 (2/14)	.300 (3/10)	.667 (2/3)	.385 (5/13)	.374
4	.043 (1/23)	.444 (8/18)	.250 (1/4)	.409 (9/22)	.332
3	.032 (1/31)	.156 (5/32)	.353 (6/17)	.224 (11/49)	.274
2	.000 (0/29)	.030 (1/33)	.226 (7/31)	.125 (8/64)	.000
1	.000 (0/29)	.000 (0/33)	.030 (1/33)	.015 (1/66)	.000

TABLE X  
PREDICTED VS. ACTUAL RENEGING BY *E* PLAYERS, EXPERIMENTS 9–10  
 $P(R \text{ in } t | \text{no previous } R); P(Y) = 0$

Period	First Third Actual	Second Third Actual	Last Two Last Third Actual	Thirds Pooled Actual	Pre- dicted	z-score
8	.667 (2/3)	.667 (2/3)	.750 (3/4)	.714 (5/7)	1.000	—
7	.750 (3/4)	.375 (3/8)	.588 (10/17)	.520 (13/25)	.575	.56
6	.363 (4/11)	.500 (4/8)	.100 (2/20)	.214 (6/28)	.436	2.36
5	.235 (4/17)	.214 (3/14)	.125 (3/24)	.158 (6/38)	.374	2.75
4	.143 (4/28)	.333 (9/27)	.115 (3/26)	.226 (12/53)	.332	1.63
3	.171 (7/41)	.200 (8/40)	.196 (9/46)	.198 (17/86)	.274	1.58
2	.163 (8/49)	.106 (5/47)	.117 (7/60)	.112 (12/107)	.000	—
1	.148 (8/54)	.153 (9/56)	.034 (2/59)	.093 (11/118)	.000	—

TABLE XI  
PREDICTED VS. ACTUAL LENDING BY *B* PLAYERS, EXPERIMENTS 9-10  
 $P(L \text{ in } t | \text{No } N \text{ in } t-1, \text{ no previous } R); P(Y) = 0$

Period	First Third Actual	Second Third Actual	Last Third Actual	Last Two Thirds Actual	Predicted	z-score
(1-3)	.873 (144/165)	.880 (146/166)	.912 (165/181)	.896 (311/347)	1.00	—
(4-8)	.506 (40/79)	.571 (40/70)	.736 (78/106)	.670 (118/176)	.643	-.47
8	1.00 (1/1)	.200 (1/5)	.429 (3/7)	.333 (4/12)	.643	1.48
7	.143 (1/7)	.500 (2/4)	.833 (15/18)	.773 (17/22)	.643	-.80
6	.231 (3/13)	.455 (5/11)	.857 (18/21)	.719 (23/32)	.643	-.56
5	.458 (11/24)	.500 (9/18)	.739 (17/23)	.634 (26/41)	.643	.08
4	.706 (24/34)	.719 (23/32)	.677 (25/37)	.696 (48/69)	.643	-.58
3	.873 (41/47)	.816 (40/49)	.836 (46/55)	.827 (86/104)	1.00	—
2	.891 (49/55)	.870 (47/54)	.968 (60/62)	.922 (107/116)	1.00	—
1	.857 (54/63)	.937 (59/63)	.922 (59/64)	.929 (118/127)	1.00	—

TABLE XII  
PREDICTED VS. ACTUAL FOLLOWING BY *B* PLAYERS, EXPERIMENTS 9-10  
 $P(N \text{ in } t | N \text{ in } t-1, \text{ no previous } R); P(Y) = 0$

Period	Experiment 9 (Seqs. 31-90) Actual	Experiment 10 (Seqs. 34-100) Actual	Experiments 9-10 Pooled Actual	Predicted
(2-3)	.333 (5/15)	.667 (2/3)	.389 (7/18)	.000
(4-8)	.600 (30/50)	.890 (105/118)	.803 (135/168)	1.00
8	.700 (7/10)	1.00 (28/28)	.921 (35/38)	1.00
7	.500 (5/10)	.903 (28/31)	.805 (33/41)	1.00
6	.667 (8/12)	.960 (24/25)	.865 (32/37)	1.00
5	.556 (5/9)	.680 (17/25)	.647 (22/34)	1.00
4	.556 (5/9)	.889 (8/9)	.722 (13/18)	1.00
3	.571 (4/7)	.500 (1/2)	.556 (5/9)	.000
2	.125 (1/8)	1.00 (1/1)	.222 (2/9)	.000

reneging should start in period 3. Reneging predictions for all periods are given in Tables VIII-X.

The two experiments were very different. In experiment 9, subjects were very slow to learn. There were many sequences, until very late in the experiment, in which *E* reneged in an early period, then *B* made a loan and *E* reneged again (in one cases, five times in a single sequence). In the final third of the sequences, behavior came slightly close to SE with the homemade prior, except that *E*'s did not renege nearly as much as predicted until periods 7 and 8 (see Table VIII). Lending probabilities did not drop much from periods 3 to 4 (as SE predicts; see Table XI), and following probabilities were not especially accurate either (Table XII).

Experiment 10 was entirely the opposite: Subjects learned to build reputation almost immediately, and the frequencies of reneging (Table IX), were close to predicted in almost every period. Lending, when pooled with experiment 9, did

drop between period 3 and period 4 (Table XI), and aggregate late-period lending was .670, close to the prediction of .643. Following jumped somewhat from period 3 to period 4, but the samples are small (Table XII).

Experiment 9 could probably be excluded from the analysis, because two *B* subjects said (in written comments) they thought the *E* choices were made randomly or by the experimenters, but we shall pool it with experiment 10, and analyze the pooled results.<sup>17</sup> The pooled results from the last two-thirds of sequences (Table X) indicate an increase in renegeing from periods 2 to 3, but *E*'s under-renegeed in period 3 (.198 actual, .274 predicted). *E*'s also under-renegeed in periods 4–6 (a result of strong under-renegeing in experiment 9).

If we use period 3 renegeing to calculate the value of  $q$  implicit in behavior (as we did for experiments 3–5,  $q = .161$ , and 6–8,  $q = .172$ ), we estimate  $q = .184$  and  $q = .171$  for experiments 9 and 10 separately, and  $q = .176$  for the experiments pooled. Since there were parameter changes (and different subjects) in all three sets of experiments, the degree of concordance of these estimates of the homemade prior  $q$  is a pleasant surprise.

Furthermore, the subjects' homemade prior is roughly consistent with the behavior they observed. For instance, in experiment 9 *E*'s renegeed only 5 times out of 7 in period 8 (excluding sequences containing earlier renegees), so a subject who thought 17 per cent of the *E*'s were honest would have that expectation roughly fulfilled by the data—*E*'s *did* behave honestly, 29 per cent of the time. In experiments 3–5 there were 19 of 22 renegees in period 8, which is easily consistent with  $q = .161$ ; but there were 13 of 13 renegees in period 8 of experiments 6–8 (which should occur only 8.6 per cent of the time if  $q = .172$  is right).

When we look back at experiments 3–5 and 6–8 with the belief that subjects used a homemade prior of 17 per cent, along with the controlled prior, some of the deviations between SE and observed behavior can be explained. With a homemade prior of 17 per cent renegeing should only start in period 5 for both sets of experiments, so the only renegeing probabilities that we expect to be correct are those in periods 6–8. In these periods, the predictions are generally quite accurate (although period 6 renegeing is a little low, and the samples are small). This reanalysis, along with the accuracy of the prediction that *B*'s lend with probability .643, leads us to conclude that SE, with a homemade prior of 17 per cent, describes subjects' aggregate behavior reasonably well.

#### 4.5.3. *Overarching Reputation-Building Is Suboptimal*

One potential problem with our design is that the *E* player in one sequence has a 1/3 chance of playing the next sequence, and the *B* players know this. This gives *E* players some incentive to develop a reputation which arches over

<sup>17</sup> Plott (1986) gives a rationale for rejecting such data, and some examples. Experiments are joint tests of whether controls worked ("internal validity"), and whether theories are true. If controls did not work—in our case, some experiment 9 subjects thought they were playing against experimenters, or random devices—then we can legitimately exclude the data.



different sequences, but we now show that this incentive is too small to make reputation-building across sequences optimal, so it cannot explain the homemade prior we observe.

Denote the equilibrium expected value to  $E$  of a game which begins with prior  $h$  by  $V(h, X)$  if  $E$  plays like an  $X$ , and  $V(h, Y)$  if  $E$  plays like a  $Y$ . The crucial question is whether it pays for an  $X$ -type  $E$  to pay back in the 8th period of the current game and continue to play honestly (if he gets a chance) in sequences which follow immediately afterward, rather than renege and play like an  $X$ -type. We consider only the sequences which follow immediately afterward because we assume the other  $E$ 's, if they are  $X$ 's, are playing like  $X$ -types.<sup>18</sup> In experiments 6–8, by reneging  $E$  earns 150 and his expected earnings from immediately following sequences,  $V(.1, X)/3 + V(.1, X)/9 + V(.1, X)/27 + \dots$ . Assuming the sequence of terms continues infinitely (a good approximation), reneging thus yields  $150 + V(.1, X)/2$ . Since  $V(.1, X) = 276.2$ , reneging in the 8th period thus yields 288.1. By behaving honestly in the 8th period, an  $X$ -type  $E$  earns 60 and has a  $1/3$  chance of playing the next sequence, with a prior of  $(1/3)(1) + (2/3)(.1)$ , or .4. (This assumes that after paying back in the 8th period, the  $B$ 's take  $E$  to be a perfectly honest type in all future sequences, except they know there is only a  $1/3$  chance that  $E$  will play the next sequence.) An  $X$ -type who behaves honestly earns  $60 + V(.4, Y)/2$ , or 244.2. (The analogous figures for experiments 3–5 are 334.4 for reneging, and 290.0 for honesty. These calculations omit the homemade prior. Adding it makes the advantage of reneging over honesty even larger.)

Thus, with 3  $E$  players it is not optimal to pay back in the 8th period, because there is not enough probability of playing the subsequent sequence(s), and even if the same  $E$  plays, the  $B$ 's do not know that. (It is also not optimal with 2  $E$  players, as in experiment 4, but the difference in payoffs is smaller.)

These calculations assume that an  $E$  player who pays back in the 8th period does so alone. We do not know how to model the development of collusive group reputations, and they rarely occurred (perhaps because we restricted verbal communication between players). In future experiments, we shall dilute the incentive to form overarching reputations even further, by not allowing  $E$ 's to play two sequences in a row. Then, only group reputations could form.

#### 4.6. Individual Behavior and Self-Insight

The formal statistical analyses, and the reanalysis (and new experiments) to detect any homemade prior, suggest that SE predicts well; but there is much behavior that is not easily captured by summary statistics.

The behavior of subjects in our experiments provides some support for the usual defense of complicated theories in economics—people act “as if” they

<sup>18</sup> This is another approximation, since another  $Y$ -type  $E$  might play between two sequences where the same honest-behaving  $E$  plays, thus maintaining the  $B$ 's beliefs that the honest-behaving  $E$  is playing. However, this will happen rarely, especially in experiments 6–8, so our approximation is close.

TABLE XIII  
PROPORTION OF RENEGING BY INDIVIDUAL *E* PLAYERS  
Sequences 26–77, Experiment 7

Period	Player <i>E</i> 1	Player <i>E</i> 2	Player <i>E</i> 3
8	4/4	1/1	1/1
7	1/4	1/1	2/2
6	0/4	1/1	1/3
5	1/8	2/6	0/4
4	0/14	4/14	3/8
3	0/16	2/17	1/12
2	0/16	1/17	3/15
1	0/13	0/17	1/13

maximize, though they make no calculations. Our subjects knew no formal game theory, made almost no calculations, and were often puzzled or skeptical when we described the sequential equilibrium to them after the experiment. Yet their aggregate behavior is not badly described by SE. How?

The “testimony” of subjects, the answers they gave to simple questions asked after the experiment, provides some clues. Most subjects described decision rules that were consistent with sequential equilibrium, but not nearly so elaborate. Said *E*3: “I was trying to disguise the fact that I was an *X* type, and waiting before I made a [renege] decision.” Many other subjects talked about reputation in roundabout terms—as in “not killing the golden goose,” or, “developing trust.” (See Table XIII for actual renegeing probabilities in experiment 7 for each *E* player.)

Their self-reported decision rules were usually less sophisticated than their apparent behavior. Said *B*2: “The deciding factor, somehow seemed to be the round number... The farther along the rounds, the more likely I would choose [no loan].” The word “somehow” is telling: Like many other subjects, *B*2 seemed to use a near-optimal strategy without knowing why it was near-optimal.

One reason aggregate play could be close to SE, even if each individual subject was not exactly following SE, is that the choice of *E* player and the sequence of *B* players were randomly varied from sequence to sequence. Even if all players used pure strategies (like *B*1: “If my round number was 6 or greater, I definitely picked [no loan]”), the random choice among such players using *different* pure strategies could look exactly like mixed-strategy play. (This argument is used to explain how animal populations can achieve intricate mixed-strategy equilibria, even though individual animals presumably play pure strategies. See, e.g., Dawkins, 1976, or Maynard-Smith, 1982.) Indeed, about half the players seemed to use simple cutoff strategies—pay back until period 4, lend until period 6, and so on—but their cutoffs varied (see the data on individual lending strategies in experiment 7, Table XIV). For instance, *B* subjects 4, 5, and 7 *always* made loans in periods 5–8; *B*1 and *B*8 *never* did. But since the choice of which *B* subjects played in periods 5–8 varied, the result was that periods 5–8 loans were made 59 per cent of the time (cf. the SE prediction of .643).

TABLE XIV  
PROPORTION OF LENDING BY INDIVIDUAL *B* PLAYERS  
Sequences 26–77, Experiment 7

Period	Player <i>B</i> 1	Player <i>B</i> 2	Player <i>B</i> 3	Player <i>B</i> 4	Player <i>B</i> 5	Player <i>B</i> 6	Player <i>B</i> 7	Player <i>B</i> 8
8	0/0	1/1	0/0	0/0	1/1	0/0	1/1	0/0
7	0/0	0/0	1/2	1/1	0/0	0/0	1/1	0/1
6	0/4	1/1	1/3	1/1	1/1	0/3	2/2	0/1
5	0/2	2/4	2/3	4/4	4/4	3/4	3/3	0/6
4	6/6	9/9	3/4	7/7	5/5	5/5	1/4	1/3
3	4/4	4/4	5/5	11/11	4/4	6/6	5/5	7/7
2	6/6	3/3	9/9	3/3	7/7	7/7	6/7	8/8
1	8/8	9/9	6/7	3/6	5/7	4/4	5/6	5/5

When subjects *did* use mixed strategies, they rationalized them as the selection of pure strategies based on an internal thought process (which is unpredictable to an opponent). Said *B*1: “If my round number was 5, it was a toss up—I would operate on a hunch.” *B*6: “If [renege] was picked once, and then [no loan], I went with my ‘gut feeling.’” By basing their randomization on privately-known hunches, subjects are “purifying” their mixed strategies, or thinking of them as equivalent to choices of pure strategies based on random realizations of private information (e.g., Harsanyi, 1973; Aumann, et al., 1983).

##### 5. CONCLUSION, METHODOLOGICAL LESSONS, AND FUTURE RESEARCH

We conclude that SE predicts reasonably well, given its complexity. However, formal statistical tests reject SE strongly for some periods of the game. Subjects failed to renege as early in the game, or as often, as predicted.

Because other predictions of SE about lending and following were fairly accurate, we sought a competing theory that could explain the observed under-reneging without otherwise changing the predicted SE. One plausible competing theory is that subjects bring into the experiment a “homemade” prior belief that some subjects will not renege, even when we tried to induce a preference for reneging. This homemade prior theory is consistent with the large body of evidence that subjects in repeated games cooperate more often than predicted, including recent economics experiments showing reputation formation.

We estimated the homemade prior,  $q$ , from experiments 3–5 ( $q = .161$ ) and experiments 6–8 ( $q = .172$ ). We then ran two experiments (9–10) in which we did not induce preferences that would lead any subjects to not renege, and we made predictions assuming 17 per cent of the subjects would not renege. In these experiments the homemade prior is about .176. Thus, some of the descriptive failures of SE are diminished when we include the homemade incomplete information about preferences which subjects create themselves. The existence of this homemade prior is a failure of our experimental control, but we were able to

estimate that homemade prior reliably from different experiments (using different subjects, and parameter changes).

The data suggest that people are reasonably good intuitive game players, though other evidence suggests people are poor intuitive statisticians, and are poor intuitive scientists—e.g., Kahneman, Slovic, and Tversky (1982). Perhaps everyday life does not prepare people for the surprises of Bayes' rule, or to search for disconfirming data, but life does provide a lot of practice interacting with other people of unknown character. Still, SE requires Bayesian updating along with optimal strategy choice. It is interesting that subjects are apparently able to use Bayes' rule to approximate a sophisticated SE, while erring in simpler applications of Bayes' rule (e.g., Camerer, 1987).

We draw one methodological lesson from our work. It takes time for subjects to learn in this experiment. In the first 30 sequences or so, which took about an hour, subjects invariably made mistakes. An experiment like ours which lasted only one hour (30 sequences) would yield very different conclusions about the descriptive accuracy of SE. Also, subjects need to observe several actual plays of the end of the game, before they can play optimally in early periods. (This tendency to learn the last period first is called the "swingback hypothesis" in research on asset markets, e.g., Forsythe, Palfrey, and Plott, 1982). Experiments in which subjects play a long repeated game only once (as in most earlier game experiments) may misleadingly suggest that subjects play sub-optimally.

There are several avenues for further research. First, we suspect that giving subjects a history of previous play (e.g., on a computer screen) will aid in convergence to SE. Second, we can test SE predictions about the effects of parameter changes which are especially counter-intuitive, like varying  $E$ 's payoffs and testing whether the renegeing proportions vary (they shouldn't, unless  $E$ 's payoffs affect the homemade prior). Third, experimental data may provide a way to specify the off-equilibrium path behavior of subjects. The importance of updating beliefs (out of equilibrium) is the chief characteristic of SE, but out-of-equilibrium beliefs are often arbitrary or difficult to refine. By asking subjects their beliefs about  $E$ 's type or strategy each period, we can directly measure how beliefs are updated after off-path moves.<sup>19</sup> These data could complement the theoretical work on refinements of SE which seek to rule out irrational out-of-equilibrium beliefs (e.g., Cho and Kreps, 1987; Grossman and Perry, 1986). Fourth, the fact that SE predicts behavior reasonably well in this simple setting suggests that subjects could behave coherently in much more elaborate settings—with two-sided reputation, "third-party" reputation between consumers and producers (such as product reviewers; see Yao and Faulhaber, 1985), and networks or grapevines of personal reputation. Finally, the long period of disequilibrium behavior early in these experiments raises the important

<sup>19</sup> Of course, we don't expect to see many out-of-equilibrium moves by experienced subjects. However, we can gather data from each of several subjects after each such move, and we can possibly provoke off-path moves by changing parameters frequently.

question of how people learn to play complicated games. The data could be fit to statistical learning models (e.g., Selten and Stoecker, 1986), though new experiments or new models might be needed to explain learning adequately. Indeed, even equilibrium behavior could conceivably be better explained by a heuristic model in which people adapt or learn to approximate SE, than by SE itself.

*Department of Decision Sciences, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104, U.S.A.*

*and*

*Department of Management, Graduate School of Business Administration, New York University, 100 Trinity Place, New York, NY 10006, U.S.A.*

*Manuscript received November, 1985; final revision received December, 1986.*

## APPENDIX A: RAW DATA

TABLE A1  
CHOICES OF SUBJECTS  
Experiment 3 [Pr of  $Y = .331$ ]

1	-2,1	2,3,3,3,3,3,3,3	26	-3,1	1,1,1,1,1,1,2,3	51	-1,2	1,1,1,1,3,3,3,3	76	-2,1	1,1,1,1,3,1,1,3
2	-1,1	2,3,3,2,3,3,3,3	27	-2,1	1,1,1,1,3,1,3,3	52	-3,1	1,1,1,1,2,3,3,3	77	-2,1	1,1,1,1,1,1,3,3
3	-2,1	2,3,3,3,3,2,3,3	28	-3,1	1,1,1,1,1,3,2,3	53	-3,1	1,1,1,1,1,3,3,3	78	-3,1	1,1,1,1,3,2,3,3
4	-3,1	3,3,2,3,3,3,3,3	29	-2,1	1,1,1,1,1,3,3,2	54	-1,1	2,3,3,3,3,3,3,3	79	-3,1	1,1,1,1,3,3,3,3
5	-2,1	2,3,3,3,3,3,3,3	30	-3,1	1,1,1,1,1,1,3,3	55	-3,1	1,1,1,1,3,3,3,3	80	-2,1	1,1,1,1,2,3,3,3
6	-3,2	3,3,1,1,1,1,1,1	31	-1,1	2,3,3,3,3,3,3,3	56	-3,1	1,1,1,1,3,3,3,3	81	-3,1	1,1,1,1,3,3,3,3
7	-3,1	3,3,2,2,3,3,3,3	32	-1,1	3,2,3,3,3,3,3,3	57	-3,1	1,1,1,1,3,3,3,3	82	-1,2	1,1,1,1,1,1,3,1
8	-1,2	3,3,3,1,1,1,1,1	33	-2,2	1,1,1,1,1,1,3,3	58	-2,1	1,1,1,2,3,3,3,3	83	-2,1	1,1,1,1,3,1,3,2
9	-2,2	1,1,1,1,1,1,1,1	34	-3,1	1,1,1,1,1,3,3,3	59	-1,2	1,1,1,1,1,1,3,1	84	-3,1	1,1,1,1,1,3,3,3
10	-3,1	3,3,2,3,3,3,3,3	35	-2,1	1,1,1,1,1,3,3,3	60	-2,1	1,1,1,1,3,3,3,3	85	-2,1	1,1,1,1,1,2,3,3
11	-1,1	2,3,3,3,3,3,3,3	36	-2,1	1,1,1,1,2,3,3,3	61	-2,1	1,1,1,1,3,3,3,3	86	-3,1	1,1,1,1,1,3,3,3
12	-3,1	2,3,3,3,3,3,3,3	37	-3,1	1,1,1,1,3,1,3,3	62	-1,2	1,1,1,1,1,3,3,3	87	-1,2	1,1,1,1,1,1,1,3
13	-2,1	1,1,1,1,1,1,1,3	38	-1,2	1,1,1,1,3,1,1,1	63	-3,2	1,1,1,1,3,1,3,3	88	-3,1	1,1,1,3,3,3,3,3
14	-3,1	3,3,3,3,1,1,2,3	39	-3,1	1,1,1,1,1,3,3,3	64	-2,1	1,1,1,1,2,3,3,3	89	-2,1	1,1,1,1,1,3,3,2
15	-1,2	1,1,1,1,1,1,3,1	40	-2,1	3,3,1,1,1,3,3,2	65	-3,1	1,1,1,1,3,3,3,3	90	-3,1	1,1,1,1,3,3,2,3
16	-2,1	1,1,1,1,1,1,2,3	41	-2,1	1,1,1,1,1,3,3,3	66	-2,2	1,1,3,1,3,1,1,3			
17	-1,2	1,1,1,1,1,3,1,3	42	-3,2	1,1,1,1,1,3,3,3	67	-3,1	1,1,1,1,1,1,3,3			
18	-2,1	3,1,1,1,1,1,3,3	43	-1,2	1,1,1,1,3,3,3,3	68	-3,1	1,1,1,1,1,2,3,3			
19	-3,1	1,1,1,1,1,1,3,3	44	-3,1	1,1,1,1,1,3,3,3	69	-1,2	1,1,1,1,1,3,3,1			
20	-3,1	1,2,3,3,3,3,3,3	45	-2,1	1,1,1,3,1,1,3,2	70	-3,1	1,1,1,1,3,3,3,3			
21	-2,1	3,1,1,1,1,2,3,3	46	-2,1	1,1,1,1,2,3,3,3	71	1,2	1,1,1,1,3,3,3,3			
22	-1,2	1,1,1,1,3,3,1,1	47	-2,1	1,1,1,1,3,3,3,3	72	-3,2	1,1,1,1,3,1,1,1			
23	-1,2	1,1,1,1,1,1,3,3	48	-3,2	1,1,1,1,1,3,3,3	73	-2,1	1,1,1,1,1,3,3,2			
24	-1,1	2,3,3,3,3,3,3,3	49	-2,1	3,1,1,1,3,2,3,3	74	-3,1	1,1,1,1,1,1,3,2			
25	-3,1	3,1,1,1,1,1,1,2	50	-1,2	1,1,1,1,1,3,3,3	75	-1,2	1,1,1,3,1,1,1,1			

### LEGEND

1st number = round  
2nd number =  $E$ 's number  
3rd number =  $E$ 's type ( $1 = X$ )  
4-11 numbers = round choices:  
1 = loan, payback;  
2 = loan, renege; 3 = no loan.

TABLE A2  
CHOICES OF SUBJECTS  
Experiment 4 [Pr of  $Y = .33$ ]

1	-3,1,	2,3,3,3,3,3,2,2	26	-2,1,	1,1,1,1,1,1,3,2	51	-3,2,	1,1,1,1,3,3,3,3	76	-2,2,	1,1,1,3,1,1,3,3
2	-2,2,	3,1,1,1,1,1,1,1	27	-2,1,	1,1,1,3,1,1,3,3	52	-2,1,	1,1,1,1,3,3,2,3	77	-3,1,	1,1,2,3,3,3,3,3
3	-3,1,	3,3,3,3,3,3,2,3	28	-3,1,	1,1,1,1,1,3,2,3	53	-3,1,	1,1,1,1,1,1,3,3	78	-2,1,	1,1,1,1,1,3,3,3
4	-2,1,	3,1,1,1,2,2,3,3	29	-2,2,	1,1,1,1,1,1,3,3	54	-3,1,	1,1,1,2,3,3,3,3	79	-3,1,	1,1,1,1,1,1,3,3
5	-2,1,	1,1,3,1,3,2,3,2	30	-3,1,	1,1,1,1,3,1,2,3	55	-3,1,	1,1,1,1,1,2,3,3	80	-3,2,	1,1,1,1,1,3,1,1
6	-3,1,	1,2,2,3,3,3,2,3	31	-2,1,	1,1,1,1,3,3,3,3	56	-2,1,	1,1,1,1,1,3,3,3	81	-3,1,	2,3,3,3,3,3,3,3
7	-3,1,	1,2,3,3,3,3,3,3	32	-2,1,	1,1,3,1,1,2,3,3	57	-2,1,	1,1,1,1,2,3,3,3	82	-2,1,	1,1,1,1,1,3,3,3
8	-3,1,	1,1,2,2,3,3,3,3	33	-2,1,	1,1,1,1,1,3,3,3	58	-2,2,	1,1,1,3,3,1,1,3	83	-3,2,	1,1,1,3,3,3,3,3
9	-2,1,	3,1,1,3,1,1,2,3	34	-3,1,	1,2,3,1,3,3,3,3	59	-2,1,	1,1,1,1,2,3,3,3	84	-3,2,	1,1,1,1,3,1,3,3
10	-3,2,	3,1,1,3,1,1,1,1	35	-2,1,	3,1,1,1,1,2,3,3	60	-3,2,	1,1,1,1,3,3,3,3	85	-2,1,	1,1,3,1,1,3,3,3
11	-2,1,	1,1,1,1,1,1,1,2	36	-3,1,	2,3,1,3,2,3,3,3	61	-3,1,	1,1,1,1,2,3,3,3	86	-2,2,	1,1,1,1,1,1,1,1
12	-3,2,	1,1,1,3,1,1,1,1	37	-3,2,	1,1,1,1,3,3,3,3	62	-2,1,	1,1,1,3,3,3,3,3	87	-3,1,	1,1,1,1,3,3,3,3
13	-3,1,	1,1,1,1,1,2,2,2	38	-3,1,	3,1,3,2,3,3,3,2	63	-2,1,	1,1,2,3,3,3,3,3	88	-3,1,	1,1,3,3,1,1,1,1
14	-2,1,	1,3,1,1,1,1,1,3	39	-2,1,	1,1,1,1,1,3,3,2	64	-3,2,	1,1,1,1,3,3,3,3	89	-3,1,	1,1,1,1,3,3,3,2
15	-2,1,	1,1,1,1,1,1,1,3	40	-3,2,	1,1,1,3,3,1,3,3	65	-3,2,	1,1,1,3,3,3,1,1	90	-2,2,	1,1,1,1,1,3,3,3
16	-3,2,	1,1,1,1,1,1,1,1	41	-2,1,	1,1,1,1,3,3,3,3	66	-2,1,	1,1,1,3,1,3,3,3			
17	-2,1,	1,1,1,1,1,1,1,2	42	-2,1,	1,1,1,2,3,3,3,3	67	-2,1,	1,1,1,1,1,1,3,3			
18	-3,1,	1,3,1,1,1,2,3,2	43	-3,1,	1,2,1,3,3,3,3,2	68	-3,2,	1,1,1,3,1,3,3,1			
19	-3,1,	1,1,1,1,1,2,3,2	44	-3,2,	1,1,1,3,3,3,3,3	69	-2,1,	1,1,1,1,1,1,1,3			
20	-2,1,	1,1,1,1,1,1,1,2	45	-2,2,	1,1,1,1,3,1,3,3	70	-2,1,	1,1,1,3,1,1,3,3			
21	-2,1,	1,1,1,1,1,1,2,3	46	-3,1,	1,1,3,2,3,3,3,3	71	3,1,	1,1,1,1,3,3,2,3			
22	-3,1,	1,1,1,1,1,2,2,3	47	-2,1,	2,3,2,3,3,3,3,3	72	-2,1,	1,1,1,1,1,1,3,3			
23	-2,1,	1,1,1,1,1,3,1,2	48	-2,1,	1,1,1,1,1,3,3,2	73	-3,1,	1,1,1,1,3,1,2,3			
24	-3,1,	3,1,1,2,3,3,3,3	49	-3,1,	2,3,3,3,3,3,3,3	74	-2,1,	1,1,1,1,1,3,3,2			
25	-3,2,	1,1,1,1,1,1,3,1	50	-2,1,	1,1,1,3,3,3,3,3	75	-3,1,	1,1,1,1,3,3,3,3			

TABLE A3  
CHOICES OF SUBJECTS  
Experiment 5 [Pr of  $Y = .33$ ]

1	-1,2,	1,3,1,2,1,3,1,3	26	-3,1,	3,1,1,1,1,1,1,1	51	-2,1,	1,1,1,1,1,2,3,3	76	-3,2,	1,1,1,3,1,1,1,3
2	-3,1,	1,1,1,3,2,1,3,2	27	-1,1,	1,1,1,1,1,1,1,2	52	-3,2,	1,1,1,1,1,3,3,3	77	-1,1,	1,1,1,1,1,1,3,2
3	-1,2,	1,2,1,1,3,1,1,1	28	-2,1,	3,1,1,1,1,1,3,2	53	-2,1,	1,1,1,1,3,1,3,3	78	-	was skipped
4	-1,2,	3,3,3,1,3,2,3,2	29	-1,1,	1,1,1,1,1,3,3,2	54	-1,1,	1,1,1,1,1,3,2,3	79	-1,1,	1,1,1,1,1,1,3,3
5	-3,1,	1,1,1,1,1,1,1,2	30	-2,1,	1,1,1,1,1,1,3,3	55	-2,1,	1,1,1,1,1,3,3,3	80	-2,1,	1,1,1,1,1,3,3,3
6	-2,2,	2,3,3,3,1,1,1,1	31	-1,1,	3,3,1,1,2,3,3,3	56	-3,1,	1,1,1,1,1,1,3,3	81	-3,2,	1,1,1,1,1,1,1,3
7	-1,1,	3,2,2,3,3,2,3,3	32	-2,1,	1,1,1,1,1,2,3,2	57	-1,1,	1,1,1,1,3,3,3,3	82	-2,1,	1,1,1,1,1,2,3,3
8	-3,1,	1,1,1,1,1,1,3,3	33	-3,1,	1,1,1,1,1,1,2,3	58	-3,2,	1,1,1,1,1,1,3,3			
9	-1,1,	3,1,3,2,3,3,2,3	34	-3,1,	1,1,1,1,1,3,1,1	59	-2,1,	1,1,1,1,1,3,3,2			
10	-3,2,	3,1,1,1,1,3,1,1	35	-1,1,	1,1,3,1,2,3,3,3	60	-3,1,	1,1,1,1,1,3,3,3			
11	-1,1,	3,1,1,2,2,3,3,3	36	-2,1,	1,1,1,1,1,2,3,3	61	-1,1,	1,3,1,2,3,3,3,3			
12	-2,2,	1,1,3,1,1,1,1,1	37	-2,1,	1,1,1,1,1,1,2,3	62	-1,1,	1,2,3,3,1,3,3,3			
13	-1,1,	3,1,1,1,1,1,2,2	38	-3,2,	1,1,1,1,1,3,3,3	63	-2,1,	1,1,1,1,2,2,3,3			
14	-1,2,	3,3,1,1,1,1,3,3	39	-1,1,	3,3,1,1,3,3,3,3	64	-2,1,	1,1,1,1,2,3,3,3			
15	-3,1,	1,1,1,1,1,1,2,1	40	-2,1,	1,1,1,1,1,1,3,3	65	-3,2,	1,1,1,1,1,1,3,3			
16	-1,1,	3,2,3,3,1,1,1,1	41	-3,1,	1,3,1,1,1,1,3,2	66	-1,1,	1,1,1,1,3,3,3,2			
17	-2,1,	3,3,1,1,1,1,3,3	42	-3,2,	1,1,1,1,1,1,3,3	67	-1,1,	1,1,1,1,1,1,1,3			
18	-2,2,	1,1,1,1,1,1,3,3	43	-2,1,	1,1,1,1,1,1,3,2	68	-3,1,	1,1,1,1,1,1,1,1			
19	-1,2,	3,1,1,1,1,1,1,1	44	-1,1,	1,1,1,1,1,1,1,2	69	-1,1,	1,3,1,1,1,1,3,3			
20	-2,1,	1,1,1,3,1,1,2,3	45	-2,1,	1,1,1,1,1,1,3,3	70	-2,1,	1,1,1,1,1,2,2,3			
21	-2,1,	1,1,1,1,1,1,3,3	46	-3,1,	1,1,1,1,1,2,3,3	71	3,2,	1,1,1,1,1,3,1,3			
22	-3,1,	1,3,1,1,1,1,2,3	47	-3,1,	1,1,1,1,1,1,3,3	72	-1,1,	1,1,1,1,1,1,3,3			
23	-3,1,	3,3,1,1,1,1,3,3	48	-2,1,	3,1,1,1,3,1,3,3	73	-1,1,	1,1,1,1,1,3,3,3			
24	-2,2,	1,1,1,1,1,1,3,3	49	-2,1,	1,1,1,1,1,3,2,3	74	-2,1,	1,1,1,3,1,2,3,3			
25	-1,1,	1,1,1,1,3,1,1,2	50	-1,1,	1,1,1,1,1,1,2,3	75	-1,1,	1,1,1,1,1,3,3,3			

LEGEND

1st number = round  
2nd number =  $E$ 's number  
3rd number =  $E$ 's type ( $1 = X$ )  
4-11 numbers = round choices:  
1 = loan, payback;  
2 = loan, renege; 3 = no loan.

TABLE A4  
CHOICES OF SUBJECTS  
Experiment 6 [Pr of  $Y = .10$ ]

1	-2,1,	2,3,3,3,3,3,3,3	26	-3,1,	1,1,1,1,1,2,3,3	51	-2,1,	1,1,1,1,1,1,1,3
2	-3,1,	1,3,1,1,2,3,2,3	27	-2,2,	1,1,1,1,1,1,3,3	52	-3,1,	1,1,1,1,1,1,3,3
3	-3,1,	1,1,3,1,1,2,3,3	28	-1,1,	1,1,1,1,1,1,3,3	53	-1,1,	1,1,1,1,1,1,1,3
4	-1,1,	1,1,1,1,1,1,1,2	29	-3,1,	1,1,1,1,1,2,3,3	54	-2,1,	1,1,1,1,1,1,2,3
5	-1,1,	1,1,1,1,1,1,1,2	30	-1,1,	1,1,1,1,1,1,1,2	55	-3,1,	1,1,1,1,1,2,3,3
6	-2,2,	3,1,1,3,1,1,1,1	31	-3,1,	1,1,1,1,1,1,1,3	56	-2,1,	1,1,1,1,1,2,3,2
7	-3,1,	3,1,1,1,1,1,1,2	32	-2,1,	1,1,1,1,1,1,3,3	57	-1,1,	1,1,1,1,1,1,1,3
8	-1,1,	1,1,1,1,1,1,1,2	33	-3,1,	1,1,1,1,1,1,2,3	58	-3,1,	1,1,1,1,1,1,2,3
9	-1,1,	1,1,1,1,1,1,1,3	34	-1,1,	1,1,1,1,1,3,2,1	59	-2,1,	1,1,1,1,1,1,3,3
10	-2,1,	1,1,1,1,1,1,1,3	35	-1,1,	1,1,1,1,1,1,3,3	60	-2,1,	1,1,1,1,1,3,2,3
11	-2,1,	3,1,1,1,1,1,2,3	36	-2,1,	1,1,1,1,1,2,3,3	61	-1,1,	1,1,1,1,1,1,1,3
12	-3,1,	1,1,1,1,1,1,3,3	37	-3,1,	1,1,1,1,1,2,3,3	62	-3,1,	1,1,1,1,1,1,2,3
13	-3,1,	1,1,1,1,1,2,2,2	38	-2,1,	1,1,1,1,2,2,3,3	63	-3,1,	1,1,1,1,1,1,2,3
14	-2,1,	1,1,1,1,2,2,3,3	39	-2,1,	1,1,1,1,1,1,2,3	64	-1,1,	1,1,1,1,1,1,3,3
15	-1,1,	1,1,1,1,1,1,3,3	40	-3,1,	1,1,1,1,1,2,3,3	65	-2,2,	1,1,1,1,1,1,3,1
16	-2,1,	1,1,1,1,1,2,2,3	41	-1,2,	1,1,1,1,1,1,1,3	66	-2,1,	1,1,1,1,1,1,3,3
17	-2,1,	1,1,1,1,1,1,2,3	42	-1,1,	1,1,1,1,1,1,3,3	67	-3,1,	1,1,1,1,1,1,2,3
18	-3,1,	1,1,1,1,1,3,3,2	43	-2,2,	1,1,1,1,1,1,1,3	68	-3,1,	1,1,1,1,1,1,2,3
19	-1,1,	1,1,1,1,1,1,1,3	44	-3,1,	1,1,1,1,1,1,1,3	69	-1,1,	1,1,1,1,1,1,2,2
20	-2,1,	1,1,1,3,1,1,3,3	45	-1,1,	1,1,1,1,1,1,3,3	70	-2,1,	1,1,1,1,1,1,3,2
21	-3,1,	1,1,1,1,1,1,3,3	46	-2,1,	1,1,1,1,2,3,3,3			
22	-2,1,	1,1,1,1,1,2,3,2	47	-3,1,	1,1,1,1,1,2,3,3			
23	-1,1,	1,1,1,1,1,3,1,3	48	-1,1,	1,1,1,1,1,1,1,2			
24	-1,1,	1,1,1,1,1,1,1,3	49	-3,1,	1,1,1,1,1,1,2,3			
25	-3,1,	1,1,1,1,1,2,3,3	50	-2,1,	1,1,1,1,1,1,2,3			

TABLE A5  
CHOICES OF SUBJECTS  
Experiment 7 [Pr of  $Y = .10$ ]

1	-2,2,	3,1,1,1,1,1,1,1	26	-2,2,	1,1,1,1,1,1,1,1	51	-3,1,	1,1,1,3,1,3,3,2	76	-1,1,	1,1,1,1,1,1,1,2
2	-2,1,	3,3,3,3,3,1,1,2	27	-1,1,	3,1,1,1,1,1,3,3	52	-3,1,	1,2,2,3,3,3,3,2	77	-3,1,	1,1,1,1,1,3,3,3
3	-3,1,	3,3,2,2,3,3,3,3	28	-3,1,	3,1,1,1,1,3,2,2	53	-1,1,	1,1,1,1,1,3,3,3			
4	-1,1,	1,1,1,1,1,1,2,1	29	-1,1,	3,1,1,1,1,1,1,2	54	-3,1,	2,1,1,1,3,3,3,3			
5	-3,1,	3,2,1,2,3,3,3,3	30	-2,1,	1,1,1,1,1,2,3,3	55	-2,1,	1,3,1,2,3,3,3,3			
6	-2,1,	3,3,3,1,2,3,2,3	31	-1,1,	1,1,1,1,1,3,3,3	56	-1,1,	1,1,1,1,3,3,3,2			
7	-1,1,	3,1,1,1,1,1,2,2	32	-2,1,	1,1,1,1,2,2,3,3	57	-1,1,	1,1,1,3,3,3,2,3			
8	-1,1,	1,1,1,1,1,1,1,3	33	-1,1,	1,1,1,1,1,3,3,3	58	-2,1,	1,1,1,1,1,3,3,2			
9	-1,1,	1,1,1,1,1,1,3,2	34	-3,1,	1,1,1,2,3,3,3,2	59	-2,1,	1,1,1,2,3,3,3,3			
10	-3,1,	1,1,2,3,2,3,3,2	35	-1,1,	3,1,1,1,3,3,3,3	60	-3,1,	3,2,3,3,3,3,2,3			
11	-2,1,	2,3,3,3,3,3,3,3	36	-2,1,	1,1,1,2,3,2,3,2	61	-3,2,	2,3,3,3,3,3,3,3			
12	-3,1,	3,1,2,3,3,3,2,3	37	-3,1,	1,1,1,2,3,3,3,3	62	-1,1,	1,1,1,1,3,3,3,3			
13	-3,1,	1,2,3,3,3,2,3,3	38	-3,1,	1,1,1,1,1,1,2,3	63	-2,1,	1,1,1,3,3,3,3,3			
14	-3,2,	1,3,1,2,3,1,3,1	39	-2,1,	1,1,1,1,2,3,3,3	64	-2,1,	3,1,1,1,1,3,3,3			
15	-1,2,	3,1,1,1,1,1,1,1	40	-3,1,	1,1,1,1,3,2,3,3	65	-1,1,	1,1,1,1,1,3,3,3			
16	-2,1,	3,1,1,1,1,1,2,3	41	-2,1,	1,1,1,1,3,3,2,3	66	-1,1,	1,1,1,3,3,3,3,3			
17	-2,1,	3,1,3,1,1,3,2,3	42	-2,1,	1,1,1,1,1,3,3,3	67	-3,1,	3,1,1,3,3,3,3,2			
18	-1,1,	1,1,1,1,3,1,1,2	43	-3,1,	1,1,1,1,3,1,3,3	68	-1,1,	1,1,1,1,2,3,3,3			
19	-2,1,	3,1,1,1,2,3,3,2	44	-1,1,	1,1,1,1,3,3,3,3	69	-2,1,	1,1,1,1,3,3,3,3			
20	-1,1,	3,1,1,1,1,1,3,2	45	-3,1,	1,1,1,2,3,3,3,2	70	-2,1,	1,1,1,1,3,3,3,3			
21	-1,1,	3,1,1,1,1,1,2,2	46	-2,1,	1,2,3,3,3,3,3,3	71	3,1,	1,1,1,3,3,3,3,3			
22	-3,1,	1,1,1,1,2,3,3,2	47	-3,1,	1,1,2,3,2,3,3,3	72	-1,1,	1,1,1,1,3,3,3,3			
23	-2,1,	1,1,1,1,2,3,3,3	48	-1,1,	1,1,1,1,3,1,1,2	73	-3,1,	1,2,3,3,2,3,3,3			
24	-3,1,	1,1,2,3,3,2,3,3	49	-2,1,	1,1,2,1,2,3,2,3	74	-2,1,	1,1,1,1,3,3,3,3			
25	-3,1,	1,1,1,1,1,2,2,3	50	-2,1,	1,1,2,3,1,2,3,3	75	-2,1,	1,1,1,2,3,3,3,3			

## LEGEND

1st number = round  
2nd number =  $E$ 's number  
3rd number =  $E$ 's type ( $1 = X$ )  
4-11 numbers = round choices:  
1 = loan, payback;  
2 = loan, renege; 3 = no loan.

TABLE A6  
CHOICES OF SUBJECTS  
Experiment 8 [Pr of  $Y = .10$ ]

1	-1,1,	1,1,1,1,1,1,2,2	26	-1,1,	1,1,1,1,1,3,2,3	51	-1,1,	1,1,1,1,2,3,3,3
2	-2,1,	2,3,3,1,1,1,3,2	27	-3,1,	1,1,1,1,1,2,2,3	52	-3,1,	1,1,1,1,1,3,3,3
3	-1,1,	1,1,1,1,1,1,2,2	28	-2,1,	1,1,1,1,1,2,3,3	53	-3,1,	1,1,1,1,1,3,3,3
4	-3,1,	1,1,1,2,3,1,2,3	29	-2,1,	1,1,1,1,1,2,3,3	54	-2,1,	1,1,1,1,1,3,3,3
5	-3,1,	1,1,2,1,3,3,2,3	30	-1,1,	1,1,1,1,1,2,3,3	55	-1,1,	1,1,1,1,2,3,3,3
6	-2,1,	1,1,1,1,1,1,2,2	31	-3,1,	1,1,1,1,1,3,3,3	56	-3,1,	1,1,1,1,2,3,3,3
7	-1,1,	1,1,1,1,1,1,1,3	32	-3,1,	1,1,1,1,2,3,2,3	57	-2,1,	1,1,1,2,3,3,3,3
8	-3,1,	1,1,1,3,2,3,3,3	33	-2,1,	1,1,1,1,1,1,1,2	58	-2,1,	1,1,1,1,3,3,3,3
9	-1,1,	1,1,1,1,1,1,2,3	34	-1,1,	1,1,1,1,1,3,3,3	59	-1,1,	1,1,1,1,2,3,3,3
10	-2,1,	1,1,1,1,1,1,1,3	35	-1,1,	1,1,1,1,2,3,3,3	60	-3,1,	1,1,1,1,2,3,3,3
11	-1,1,	1,1,1,1,1,1,3,2	36	-1,1,	2,3,3,3,3,3,3,3	61	-1,1,	1,1,1,2,3,3,2,3
12	-3,1,	1,1,1,1,1,1,1,3	37	-3,1,	1,1,1,1,1,2,3,3	62	-2,1,	1,1,1,1,3,3,3,3
13	-3,1,	1,1,1,1,1,1,3,2	38	-2,1,	1,1,1,1,1,1,3,2	63	-2,1,	1,1,1,1,3,3,3,3
14	-2,2,	1,1,1,1,1,1,1,3	39	-2,1,	1,1,1,1,1,3,3,3	64	-3,1,	1,1,1,1,3,3,3,3
15	-2,1,	1,1,1,2,1,3,3,3	40	-1,1,	1,1,1,1,2,3,3,3	65	-1,1,	1,1,1,3,3,3,3,3
16	-1,1,	1,1,1,1,1,2,3,3	41	-3,1,	1,1,1,1,1,3,3,2	66	-3,1,	2,2,3,3,3,3,3,3
17	-3,1,	2,1,2,1,3,3,3,3	42	-3,1,	1,1,1,1,2,3,3,3	67	-1,1,	1,1,1,1,3,3,3,3
18	-2,1,	2,3,1,3,2,3,2,3	43	-2,1,	1,1,2,3,3,3,3,3	68	-2,1,	1,1,1,1,1,1,3,3
19	-2,1,	1,1,1,1,1,1,1,2	44	-2,2,	1,1,1,1,1,1,3,3	69	-1,1,	1,1,1,1,3,3,3,3
20	-3,1,	2,2,3,3,3,3,3,3	45	-1,1,	1,1,1,1,1,1,3,3			
21	-2,1,	1,1,1,1,1,1,1,3	46	-3,1,	1,1,1,1,1,2,3,3			
22	-1,1,	1,1,1,1,1,2,2,3	47	-1,1,	1,1,1,1,1,2,3,3			
23	-3,2,	1,1,1,1,1,3,3,3	48	-2,1,	1,1,1,1,1,3,2,3			
24	-2,1,	2,2,3,3,3,3,3,3	49	-2,1,	2,3,3,3,3,3,2,2			
25	-3,1,	1,2,3,3,3,3,3,3	50	-3,2,	1,1,1,1,1,3,3,3			

TABLE A7  
CHOICES OF SUBJECTS  
Experiment 9 [Pr of  $Y = .00$ ]

1	-1,1,	1,1,1,2,3,2,3,2	26	-3,1,	1,2,3,1,3,3,2,3	51	-3,1,	2,3,3,3,3,3,3,3	76	-3,1,	1,1,1,1,1,1,1,2
2	-2,1,	3,2,1,3,3,2,2,2	27	-2,1,	1,3,3,2,3,2,3,3	52	-2,1,	1,3,1,1,1,1,1,3	77	-1,1,	3,1,1,3,1,2,3,2
3	-2,1,	1,1,2,2,3,3,3,3	28	-1,1,	3,1,3,1,3,3,3,3	53	-2,1,	3,2,3,3,3,3,2,2	78	-2,1,	1,1,1,1,3,1,2,3
4	-3,1,	2,3,3,2,3,2,2,3	29	-3,1,	3,1,3,3,2,2,3,2	54	-3,1,	1,1,1,1,1,1,1,3	79	-3,1,	1,1,1,1,1,1,1,3
5	-1,1,	3,1,1,2,2,3,3,2	30	-2,1,	1,2,2,2,3,3,3,2	55	-2,1,	2,3,3,3,3,3,3,3	80	-1,1,	1,1,2,3,1,2,3,3
6	-2,1,	2,3,2,1,3,3,3,3	31	-2,1,	2,3,3,3,2,3,3,3	56	-1,1,	1,1,3,2,3,3,3,3	81	-2,1,	1,1,1,3,1,1,2,3
7	-3,1,	2,2,3,3,3,2,3,3	32	-1,1,	1,3,3,3,3,2,3,2	57	-2,1,	1,3,1,1,1,1,3,3	82	-3,1,	1,1,1,1,1,1,3,2
8	-2,1,	2,3,3,3,1,1,3,2	33	-1,1,	1,1,2,3,3,3,3,3	58	-3,1,	3,1,3,1,3,3,2,3	83	-1,1,	1,1,1,2,3,3,3,3
9	-3,1,	1,2,3,1,3,2,2,3	34	-2,1,	1,1,3,1,3,3,3,3	59	-2,1,	1,1,1,1,1,1,3,2	84	-3,1,	1,1,1,1,1,3,2,3
10	-1,1,	1,1,2,3,3,3,3,3	35	-2,1,	1,3,3,3,3,3,3,2	60	-1,1,	1,1,2,3,3,3,3,1	85	-2,1,	1,1,1,1,1,1,2,3
11	-1,1,	1,1,2,2,1,3,3,3	36	-3,1,	3,1,3,1,3,3,3,3	61	-1,1,	1,1,2,3,1,3,2,3	86	-1,1,	3,1,1,1,1,1,3,3
12	-2,1,	1,1,1,3,3,3,2,2	37	-1,1,	1,2,3,3,3,3,2,2	62	-2,1,	3,1,1,1,1,1,1,3	87	-2,1,	1,1,1,1,1,1,2,3
13	-3,1,	2,3,3,1,1,2,2,3	38	-1,1,	1,1,3,3,3,3,3,3	63	-3,1,	1,1,1,1,1,1,1,1	88	-3,1,	1,1,1,3,1,2,3,3
14	-2,1,	1,1,1,3,1,3,3,3	39	-3,1,	1,3,2,3,3,3,3,2	64	-2,1,	1,1,1,1,1,1,1,2	89	-2,1,	1,1,1,1,1,1,3,3
15	-3,1,	2,3,3,3,3,3,3,1	40	-3,1,	1,2,3,3,3,3,3,3	65	-2,1,	1,1,1,1,1,1,2,1	90	-1,1,	1,1,1,1,1,1,2,3
16	-1,1,	1,1,2,3,3,2,2,3	41	-2,1,	2,2,3,2,2,3,2,3	66	-1,1,	1,1,1,3,2,2,3,3			
17	-3,1,	3,1,2,1,2,3,1,1	42	-1,1,	1,3,3,3,3,3,2,3	67	-3,1,	3,1,1,1,3,3,2,3			
18	-1,1,	1,3,3,1,2,3,2,2	43	-1,1,	2,2,1,3,3,2,3,3	68	-2,1,	1,1,1,1,1,1,1,3			
19	-2,1,	1,2,1,2,3,3,3,2	44	-3,1,	2,3,2,3,3,3,2,3	69	-1,1,	2,3,3,3,3,3,2,3			
20	-2,1,	1,2,1,3,3,3,3,3	45	-2,1,	2,3,3,3,3,3,3,3	70	-3,1,	1,1,1,1,1,1,1,3			
21	-1,1,	1,1,2,3,3,3,2,3	46	-3,1,	3,3,3,3,3,2,3,2	71	3,1,	1,1,1,1,1,1,2,3			
22	-3,1,	2,3,1,3,3,3,2,3	47	-3,1,	2,3,3,3,3,3,2,3	72	-1,1,	1,1,2,3,3,1,2,3			
23	-1,1,	1,2,1,3,3,2,2,3	48	-2,1,	2,2,2,3,3,3,3,3	73	-1,1,	1,1,1,2,1,3,3,3			
24	-2,1,	1,2,3,2,2,3,3,3	49	-3,1,	2,1,2,3,3,3,3,3	74	-2,1,	1,1,1,1,3,1,2,3			
25	-3,1,	2,3,3,1,3,3,3,2	50	-1,1,	1,1,1,1,3,3,1,1	75	-3,1,	1,1,1,1,1,1,3,3			

LEGEND

1st number = round  
2nd number =  $E$ 's number  
3rd number =  $E$ 's type ( $1 = X$ )  
4-11 numbers = round choices:  
1 = loan, payback;  
2 = loan, renege; 3 = no loan.



TABLE A8  
CHOICES OF SUBJECTS  
Experiment 10 [Pr of  $Y = .00$ ]

1	-3,1,	3,3,3,3,3,1,3,3	26	-3,1,	1,1,1,1,3,3,3,3	51	-3,1,	1,1,1,2,3,3,3,3	76	-2,1,	1,2,3,3,3,3,3,3
2	-3,1,	1,1,1,1,3,1,2,3	27	-1,1,	1,1,1,1,3,3,1,1	52	-3,1,	1,1,1,1,1,3,3,3	77	-2,1,	1,1,1,3,2,3,3,3
3	-1,1,	3,1,1,1,2,1,2,2	28	-2,1,	1,1,1,1,1,1,3,3	53	-1,1,	1,1,1,3,1,3,3,3	78	-1,1,	1,1,1,3,3,3,3,3
4	-3,1,	1,1,3,1,1,3,3,2	29	-2,1,	1,1,1,1,1,3,3,3	54	-2,1,	1,1,2,3,3,3,3,3	79	-3,1,	1,1,3,3,3,3,3,3
5	-2,1,	1,3,1,3,3,2,3,3	30	-1,1,	1,1,1,3,1,3,3,3	55	-1,1,	1,1,1,2,3,3,1,3	80	-1,1,	1,1,2,3,3,3,3,3
6	-1,1,	1,1,1,1,3,3,3,3	31	-3,1,	1,1,1,1,3,3,2,2	56	-3,1,	1,1,1,1,3,3,3,3	81	-3,1,	1,1,3,3,1,3,3,3
7	-1,1,	1,1,1,3,3,3,3,3	32	-2,1,	1,1,1,3,3,2,2,3	57	-2,1,	1,1,2,3,3,3,3,3	82	-3,1,	1,1,2,3,3,3,3,3
8	-2,1,	1,1,1,1,1,3,3,3	33	-1,1,	1,1,1,1,1,1,3,3	58	-2,1,	1,1,1,3,3,3,2,3	83	-1,1,	1,1,1,3,3,3,3,3
9	-2,1,	1,1,1,2,3,2,3,3	34	-2,1,	1,1,1,2,3,3,3,2	59	-1,1,	1,1,1,1,3,3,3,3	84	-2,1,	1,1,2,3,3,3,3,3
10	-1,1,	1,1,1,1,3,3,3,3	35	-3,1,	1,1,1,3,2,3,3,3	60	-3,1,	1,1,1,2,3,1,3,3	85	-1,1,	1,1,1,3,3,3,3,3
11	-2,1,	1,1,2,3,3,3,3,3	36	-3,1,	1,1,1,1,3,3,3,3	61	-1,1,	1,1,1,2,3,3,3,3	86	-3,1,	1,2,3,3,3,3,3,3
12	-3,1,	1,1,1,3,1,3,2,3	37	-1,1,	1,1,1,1,1,3,3,3	62	-2,1,	1,1,2,3,3,3,3,3	87	-2,1,	1,1,3,3,3,3,3,3
13	-3,1,	1,1,1,1,3,1,3,3	38	-2,1,	1,2,1,3,3,3,3,3	63	-3,1,	1,1,1,1,1,3,3,3	88	-1,1,	1,1,2,3,3,3,3,3
14	-2,1,	1,1,1,3,2,3,3,2	39	-2,1,	1,1,2,3,3,3,3,3	64	-2,1,	1,1,1,3,3,2,3,3	89	-1,1,	1,1,3,3,3,3,3,3
15	-1,1,	1,1,1,3,3,1,3,3	40	-2,1,	1,1,2,3,3,3,3,3	65	-3,1,	1,1,1,2,2,3,3,3	90	-3,1,	1,2,3,3,3,3,3,3
16	-2,1,	3,1,1,1,3,2,3,3	41	-1,1,	1,1,1,1,1,2,3,3	66	-1,1,	1,1,1,1,3,3,3,3	91	-1,1,	1,3,3,3,3,3,3,3
17	-3,1,	1,3,1,1,3,3,3,3	42	-1,1,	2,3,3,3,3,3,1,3	67	-1,1,	1,1,1,3,2,3,3,3	92	-3,1,	2,3,3,3,3,3,3,3
18	-3,1,	1,1,1,3,1,3,3,3	43	-2,1,	1,1,1,2,3,2,3,3	68	-2,1,	1,1,2,3,3,3,3,3	93	-1,1,	1,2,3,3,3,3,3,3
19	-1,1,	3,1,1,1,1,3,3,3	44	-1,1,	1,1,1,3,3,3,1,3	69	-3,1,	1,1,1,3,3,3,3,3	94	-2,1,	1,2,3,3,3,3,3,3
20	-2,1,	1,3,1,1,1,3,3,3	45	-3,1,	1,1,1,1,2,3,3,3	70	-2,1,	1,2,3,3,3,3,3,3	95	-1,1,	1,1,3,3,3,3,3,3
21	-1,1,	1,1,1,1,1,3,3,3	46	-2,1,	1,1,2,2,3,3,3,3	71	-1,1,	1,1,1,1,3,3,3,3	96	-1,1,	1,1,3,1,3,3,3,3
22	-2,1,	1,1,1,1,3,3,3,3	47	-3,1,	1,1,1,3,1,3,1,3	72	-2,1,	1,1,1,3,3,3,3,3	97	-3,1,	1,1,3,3,3,3,3,3
23	-1,1,	1,1,1,1,1,1,3,3	48	-1,1,	1,1,1,3,1,3,3,3	73	-3,1,	1,1,1,3,3,3,3,3	98	-2,1,	1,2,3,3,3,3,3,3
24	-3,1,	1,1,1,1,3,2,3,3	49	-2,1,	1,1,1,1,3,3,3,3	74	-2,1,	1,1,2,3,3,3,3,3	99	-3,1,	3,3,1,1,3,3,3,3
25	-3,1,	1,1,1,1,3,3,3,3	50	-1,1,	1,1,1,3,3,3,3,3	75	-1,1,	1,1,1,2,3,3,3,3	00	-1,1,	1,1,3,3,2,3,3,3
									01	-1,1,	1,1,1,3,3,3,3,3

## APPENDIX B: EXPERIMENTAL PROCEDURE AND INSTRUCTIONS

### *Reading Instructions*

After the subjects have checked the contents of their folders, the instructions were read aloud. This ensured that the instructions, and stated parameters were “common knowledge” to all subjects, which is especially important in this experiment because a little spark of uncertainty (about what others know) can start reputation-building, over which you have no experimental control. Reading written instructions also allows the experimenter to use a standardized format, thus preventing biases and demand effects that can be created in an informal, ad libbed, explanation of procedure. The following is the set of instructions we used. Comments in brackets [ ] explain physical actions of the experimenters. All the diagrams and Record and Profit Sheets that are referred to in the instructions are available from the authors.

### *Instructions*

This is an experiment in decision making. Various research foundations have provided funds for this research. The instructions are simple and if you follow them carefully and make good decisions, you can earn a considerable amount of money which will be paid to you in cash. [At this point, the experimenter pulls out a wad of money from an envelope, and hold it up so all the subjects could see the cash, to reinforce the reality of the monetary rewards.]

### *General Instructions*

Each decision maker has been randomly assigned to be a member of the *A* group, or a member of the *B* group. Members of the *A* group have an “*A*” marked on their folders, along with a number (for example, *A*2). Group *A* players will remain in this room (the “*A* room”) during the experiment. Group *B* members have a “*B*” marked on their folders, and a number (for example, *B*3). Group *B* players will sit in a different room (the “*B* room”) across the hall. [The experimenter then announces the number of the “*B* room”—e.g. Room 307.] One experimenter will be in each room, and the experimenters will communicate with each other by speaking and listening on walkie-talkies.

In this experiment, there will be several sequences of decision rounds. Each sequence consists of eight decision rounds. Each round is a choice by one *A* player, followed by a choice by a *B* player.

*How Players are Selected*

A single *B* player will be chosen before each sequence of rounds, to make choices in all eight rounds in that sequence. Which *B* player will be chosen in a specific sequence has been predetermined by a random selection process. For instance, if the number 1 is assigned to a sequence, then player *B1* is the player who will make all the choices in the eight rounds in that sequence.

In each sequence of rounds, each of the eight *A* players will make a choice in one of the decision rounds. Which player makes a choice in which round has been predetermined by a random number selection process. [In our pilot experiments we used another bingo cage to determine the ordering of *A* players. This proved to be very time consuming, so we predetermined the ordering before the experiment. After several sequences, subjects could see that the ordering appeared random, and they did not seem concerned that we had pre-ordered.]

*Choices in Each Decision Round*

In each decision round, the *A* player will make the first choice. The *A* player can choose either *A1* or *A2*. The *A* player in each round must record his or her choice on the "*A* Player Information and Record Sheet" [the experimenter holds up a sample copy of this sheet] and that choice will be recorded on the blackboards in both the *A* room and the *B* room.

After *A*'s choice is announced to the *B* players in the *B* room by walkie-talkie, the *B* player who has been chosen to make the choices in all eight rounds of that sequence must make his or her choice. The *B* player can choose either *B1* or *B2*. The *B* player must record that choice on the "*B* Player Information and Record Sheet" [the experimenter holds up a sample copy of this sheet], and that choice will be recorded on the blackboards in both the *A* room and the *B* room. After the *B* player's choice is made, the decision round is concluded, and players will record their earnings on the appropriate Information and Record Sheets. Then we will proceed to the next decision round. Do not discuss your choice with other players between rounds, or at any time during the experiment. You should not speak to any other player during the experiment. After a sequence of eight decision rounds are concluded, we proceed to the next sequence. [The experimenter then asks if there are any questions regarding this section.]

*Earnings From Choices*

In a particular decision round, the earnings to each participating player depend on the choices of both the *A* player and the *B* player. You only receive earnings if you make a choice in that decision round.

In this experiment, all earnings are in terms of "francs." Each franc is worth \$0.XX [a specific number is given here, but not read aloud—instead, we say "some number"] to you. This number is your own private information and you are not to reveal it to anyone. At the end of the experiment, your francs will be converted to dollars and paid to you in cash. [The conversion rate can be considered as a balancing act. While you want to provide subjects with a large enough monetary incentive so they consider the experiment "worthwhile," you want to run as many experiments as possible, and thus don't want to waste your limited resources. One result of this is that it is always better to err on the low side. It is much easier to give subjects "bonuses" at the end of the experiment (e.g., for filling out the questionnaires) than to try explaining to subjects why you want to only give them part of their earnings. Generally, the best method for determining a conversion rate is to calculate the subjects' expected value. For example, the expected value of the *A* players in experiments 3, 4, 5 was \$15.48. This was derived by calculating the *A* players' expected value when the *B* player's type was *X* (and multiplying this by .667), when the *B*'s player type was *Y* (and multiplying this by .333), adding these totals together, then multiplying by the expected number of sequences, and dividing by 8. It should be noted that this expected value calculation is subject to the law of large numbers, and it is quite possible that some *A* players may receive bad random draws (e.g., have many of their scheduled periods late in the sequences). The *B* players' expected earnings were calculated in a parallel fashion. In this experiment it was important that both *A* and *B* players received similar earnings, because their exchange rates were necessarily different, to balance expected earnings. After the experiment, players of both types tend to ask each other how much they earned, and they get upset (harming goodwill and making future recruiting difficult) if their figures are systematically different. Our mean payoff for the experiments was \$207.00, with a high payoff of \$228.00, and a low payoff of \$188.00.]

The possible earnings, which depend on the choices of the *A* and *B* players, are best shown in a diagram written on the blackboard [the experimenter would point to a diagram like Figure 1 in the text]. We shall explain them as well.

[The experimenter now walks to where the diagram is drawn on the blackboard.] The *A* player choosing first, may consider *A1* or *A2*. [He points to *A*'s decision node.] If the player chooses *A1*, then the *B* player has no choice, and both players earn 10 francs. [The experimenter traces the *A1* decision branch.] If the *A* player chooses *A2*, then the *B* player must make a choice. [The experimenter traces the *A2* decision branch to *B*'s decision node.] If the *B* player chooses *B1*, the *A* player earns 40 francs. If the *B* player chooses *B2*, the *A* player loses 100 francs. [Both *B2* decision branches are traced.]

The earnings of the *B* player are slightly more complicated. If the *A* player chooses *A1*, then the *B* player has no choice, and the *B* player earns 10 francs. [The *A1* decision branch is traced again.] If the *A* player chooses *A2*, then the *B* player has a choice—either *B1* or *B2* [the experimenter traces along the *A2* branch from *A*'s decision node to *B*'s decision node.] If the *B* player chooses *B1*, then the *B* player receives 60 francs. [The *B1* node is traced.] But if the *B* player chooses *B2*, the *B* player's earnings depends upon what "type" of earnings the *B* player has.

Remember that a single *B* player is chosen by random drawing to make all eight choices in a specific sequence. After that player is announced, a bingo cage is used to determine what type of earnings that *B* player will receive. [The experimenter goes to a second blackboard diagram showing the bingo cage structure.] The bingo cage contains 3 balls [for experiments 3–5, 10 balls for experiments 6–8, and there is no bingo cage for experiments 9–10] numbered 1 through 3. If a ball numbered 1 or 2 is drawn, the *B* player has "type *X*" earnings, and whichever *B* player has been selected to choose in that sequence should record an "*X*" next to "Type" at the top of their "*B* Player Information and Record Sheet." [The experimenter holds up the "*B* Player Information and Record Sheet," showing where the "Type" space is.] If a ball numbered 3 is drawn, the *B* player has "type *Y*" earnings, and the *B* player should record a "*Y*" next to "Type" at the top of their "*B* Player Information and Record Sheet." [The experimenter again holds up the "*B* Player Information and Record Sheet."] The outcome of this drawing to determine what type of earnings the *B* player has will be announced in the *B* room, but not in the *A* room. (However, after a sequence of eight decision rounds is concluded, then we will announce which type of earnings the *B* player had, to everyone in the *A* room.)

[The experimenter walks back to the decision tree.] Assume the *A* player has chosen *A2* and the *B* player has chosen *B2* [these branches are followed to their terminal nodes]. Then the *B* player earns 150 francs if he or she has type *X* earnings, and the *B* player earns 0 francs if he or she has type *Y* earnings. Notice that for all other patterns of choices—for instance, if the *A* player chooses an *A1*, or if the *A* player chooses *A2* and the *B* player chooses *B1*—type *X* and type *Y* earnings are identical for *B* players. [While reading this last sentence, the experimenter traces these branches to their terminal nodes.]

The pattern of potential earnings, depending upon the choices which the *A* player and the *B* player make, is shown in the diagram on the blackboard. [The experimenter points to the diagram again.]

#### *Total Earnings From the Experiment*

At the end of the experiment, total your earnings from all the sequences, on your "Profit Sheet." [The experimenter holds up an *A* player and a *B* player profit sheet.] Group *A* players will have made a choice in one round in each sequence. Group *B* players will have made eight choices, in each round of a sequence, in several different sequences.

Record the total profit from all your choices, in francs, in row *T1* of your profit sheet. [The experimenter holds up a profit sheet and points to row *T1*.] Multiply this total by the number of dollars per franc in row *T2* of the Profit Sheet [he points to row *T2*], and record the total dollars profit in row *T3*. [He points to row *T3*.] This amount will be paid to you in cash at the end of the experiment. You are free to earn as much as you wish in the experiment.

#### REFERENCES

- ALLEN, FRANKLIN (1984): "Reputation and Product Quality," *Rand Journal of Economics*, 15, 311–327.
- AXELROD, ROBERT (1982): *The Evolution of Cooperation*. New York: Basic Books.
- AUMANN, ROBERT J., Y. KATZNELSON, ROY RADNER, ROBERT W. ROSENTHAL, AND B. WEISS (1983): "Approximate Purification of Mixed Strategies," *Mathematics of Operations Research*, 8, 327–341.

- BERG, JOYCE E., LANE A. DALEY, JOHN W. DICKHAUT, AND JOHN R. O'BRIEN (1986): "Controlling Preference for Lotteries on Units of Experimental Exchange," *Quarterly Journal of Economics*, 101, 281-306.
- BETTENHAUSEN, KENNETH, AND KEITH J. MURNIGHAN (1985): "The Emergence of Norms in Competitive Decision-making Groups," *Administrative Science Quarterly*, 30, 350-372.
- CAMERER, COLIN F. (1987): "Do Biases in Probability Judgment Matter in Markets? Experimental Evidence," *American Economic Review*, 77, forthcoming.
- CHO, IN-KOO, AND DAVID KREPS (1987): "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, 102, 179-221.
- COLMAN, ANDREW (1983): *Game Theory and Experimental Work*. London: Pergamon Press.
- DAUGHETY, ANDREW F., AND ROBERT FORSYTHE (1987a): "Regulatory-Induced Industrial Organization: A Laboratory Investigation," *Journal of Law, Economics, and Organization*, forthcoming.
- (1987b): "Regulation and the Formation of Reputations: A Laboratory Analysis," in *Public Regulation: New Perspectives on Institutions and Policies*, ed. by E. Bailey. Cambridge, Mass.: MIT Press.
- DAWKINS, RICHARD (1976): *The Selfish Gene*. New York: Oxford University Press.
- DEJONG, DOUGLAS V., ROBERT FORSYTHE, AND RUSSELL J. LUNDHOLM (1985): "Ripoffs, Lemons and Reputation Formation in Agency Relationships: A Laboratory Market Study," *Journal of Finance*, 40, 809-820.
- FORSYTHE, ROBERT, THOMAS PALFREY, AND CHARLES PLOTT (1984): "Futures Market and Informational Efficiency: A Laboratory Examination," *Journal of Finance*, 39, 955-981.
- FUDENBERG, DREW, DAVID LEVINE, AND JEAN TIROLE (1985): "Infinite-Horizon Models of Bargaining With One-sided Incomplete Information," in *Game-theoretic Models of Bargaining*, ed. by A. E. Roth. Cambridge: Cambridge University Press.
- GROSSMAN, SANFORD J., AND MOTTY PERRY (1986): "Perfect Sequential Equilibrium," *Journal of Economic Theory*, 39, 97-119.
- HARRISON, GLENN W., AND MICHAEL MCKEE (1985): "Monopoly Behavior, Decentralized Regulation, and Contestable Markets: An Experimental Evaluation," *Rand Journal of Economics*, 16, 51-69.
- HARSANYI, JOHN C. (1973): "Games With Randomly Disturbed Payoffs: A New Rationale for Mixed-Strategy Equilibrium Points," *International Journal of Game Theory*, 2, 1-23.
- HOGGATT, AUSTIN, REINHARD SELTEN, D. CROCKETT, S. GILL, AND J. MOORE (1978): "Bargaining Experiments With Incomplete Information," in *Bargaining Behavior*, ed. by H. Sauermann. Tübingen: J. Mohr.
- ISAAC, R. MARK, AND VERNON L. SMITH (1985): "In Search of Predatory Pricing," *Journal of Political Economy*, 93, 320-345.
- KAHNEMAN, DANIEL, PAUL SLOVIC, AND AMOS TVERSKY (1982): *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.
- KLEIN, BENJAMIN, AND KEITH B. LEFFLER (1981): "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy*, 89, 615-641.
- KREPS, DAVID M., PAUL MILGROM, JOHN ROBERTS, AND ROBERT WILSON (1982): "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma," *Journal of Economic Theory*, 27, 245-252.
- KREPS, DAVID M., AND ROBERT WILSON (1982a): "Sequential Equilibria," *Econometrica*, 50, 863-894.
- (1982b): "Reputation and Imperfect Information," *Journal of Economic Theory*, 27, 253-279.
- LEDYARD, JOHN (1986): "The Scope of the Hypothesis of Bayesian Equilibrium," *Journal of Economic Theory*, 39, 59-82.
- MAYNARD-SMITH, JOHN (1982): *Evolution and the Theory of Games*. Cambridge: Cambridge University Press.
- MILGROM, PAUL, AND JOHN ROBERTS (1982): "Predation, Reputation, and Entry Deterrence," *Journal of Economic Theory*, 27, 280-312.
- MILLER, ROSS, AND CHARLES R. PLOTT (1985): "Rent, Quality and Reputation in Experimental Markets," paper presented at the 1985 Summer Econometric Society Meetings.
- PLATT, JOHN R. (1964): "Strong Inference," *Science*, 146, 347-353.
- PLOTT, CHARLES R. (1986): "Rational Choice in Experimental Markets," *Journal of Business*, 59, S301-S327.
- RAPOPORT, ANATOL, AND ALBERT CHAMMAH (1965): *Prisoner's Dilemma*. Ann Arbor: University of Michigan Press.
- ROTH, ALVIN E. (1983): "Toward a Theory of Bargaining: An Experimental Study in Economics," *Science*, 220, 687-691.

- ROTH, ALVIN E., AND FRANCOISE SCHOUMAKER (1983): "Expectations and Reputations in Bargaining: An Experimental Study," *American Economic Review*, 73, 362-372.
- RUBINSTEIN, ARIEL (1985): "A Bargaining Model with Incomplete Information About Time Preferences," *Econometrica*, 53, 1151-1172.
- SELTEN, REINHARD (1975): "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," *International Journal of Game Theory*, 4, 25-55.
- (1978): "The Chain-Store Paradox," *Theory and Decision*, 9, 127-159.
- SELTEN, REINHARD, AND ROLF STOECKER (1986): "End Behavior in Sequences of Finite Prisoner's Dilemma Supergames," *Journal of Economic Behavior and Organization*, 7, 47-70.
- SHAPIRO, CARL (1983): "Premiums for High Quality Products as Returns to Reputations," *Quarterly Journal of Economics*, 98, 659-679.
- SHAPIRO, CARL, AND JOSEPH STIGLITZ (1984): "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review*, 74, 433-444.
- SOBEL, JOEL (1985): "A Theory of Credibility," *Review of Economic Studies*, 52, 557-573.
- STOECKER, ROLF (1983): "Das erlernte Schlussverhalten-eine experimentelle Untersuchung," *Zeitschrift für die gesamte Staatswissenschaft* (transl. *Journal of Institutional and Theoretical Economics*), 100-121.
- WILSON, ROBERT (1985): "Reputations in Games and Markets," in *Game-theoretic Models of Bargaining*, ed. by A. E. Roth. Cambridge: Cambridge University Press.
- WRIGHTSMAN, LAWRENCE S., JR., JOHN O'CONNOR, AND NORMA BAKER (1972): *Cooperation and Competition: Readings on Mixed-Motive Games*. Belmont, CA: Wadsworth.
- YAO, DENNIS A., AND GERALD R. FAULHABER (1985): "Reputation, 'Fly-by-Night' Firms, and the Market for Product Review," The University of Pennsylvania, Fishman-Davidson Center for the Study of the Service Center Discussion Paper #12.