

Nash Bargaining

- Today we're going to discuss bargaining experiments. This will give us an opportunity to talk about how details of experimental designs grow out of the questions they are supposed to answer, and the theories they are intended to explore.
- The first part of the story will have to do with how experimental results helped change the kind of theories of bargaining that economists concentrate on.
- That being the case, I'll have to start the story by reminding you about some of the salient features of John Nash's (1950) model of bargaining. It was perhaps the principal model of bargaining in the economic literature for a long time--certainly through the beginning of the 80s--and experimental results played a large role in exposing its shortcomings as a descriptive theory. (These experiments in turn led to the exploration of a number of robust empirical regularities, which we'll explore later, and to some new theory, in several different directions.)

Nash's model of bargaining

- 2 bargainers are faced with a set A of alternatives. The rules are that, if they both agree on some alternative a in A , then a will be the outcome. Otherwise (i.e. if they fail to agree on an outcome) there is a fixed disagreement outcome d which will be the result.
- Under these rules, each player may veto any outcome other than d .
- Let u_1 and u_2 be expected utility functions representing the preferences of players' 1 and 2.
- Let S be the set of feasible utility payoffs from an agreement, i.e.
 - $$S = \{x = (u_1(a), u_2(a)) \mid a \text{ is in } A\}$$
- And let d be the utility payoffs to the players from a disagreement, i.e.
 - $$d = (u_1(d), u_2(d))$$
- The “complete information hypothesis” is that, when the set (S,d) is known (*) to both bargainers, it provides all the information needed to “solve” the bargaining problem. Formally, Nash proposed to study solutions to the bargaining problem embodied in functions which would determine utility payoffs to both bargainers from the data (S,d) .
 - $$f: \{(S,d)\} \rightarrow R$$

- But how to make the utilities, (S,d) known? I.e. how should the experiment be designed so that we know what a particular *theory* $f:\{(S,d)\} \rightarrow R$ *predicts*?
- Early experiments assumed linear utility in money: $u_i: R \rightarrow R : u_i(x)=x$.
- So if $A = \{(x_1, x_2) \mid x_1 + x_2 \leq k\}$, these experiments were interpreted as if $(S,d) = (A,0)$
- This kind of design turned out to be unpersuasive to game theorists, who felt that an experiment that *assumes* everyone's utility function is the same cannot do justice to a theory that models *all* individual differences as differences in utilities.

So Roth and Malouf (1979)

Introduced binary lottery games:

$$A = \{\text{All lotteries over } (\$L1, \$s2), (\$s1, \$L2)\} \cup \{(\$s1, \$s2)\}$$

Where L is the large prize, s is the small prize, and people bargain over probabilities of getting the large prize L.

Theorem: If u_i is an arbitrary utility function for money, and q is a lottery that gives player i probability q of receiving $\$L1$, then we can find the unit and the zero point such that

$$u_i(q \$L1 + (1-q)\$s1) = q.$$

(i.e. we set $u_i(\$L1)=1$ and $u_i(\$s1) = 0$)

So (S,d) is (up to affine transformations of each axis) the convex hull of $(0,0)$, $(1,0)$, and $(0,1)$.

Nash characterized a particular solution f via four axioms (1,2,3,5)

1. Independence of equivalent utility transformations:

$$f(T(S,d)) = T(f(S,d)) \text{ for } T(x_i) = a_i x_i + b_i \quad a_i > 0$$

$$\text{(i.e. } f_i(T(S,d)) = a_i f_i(f(S,d)) + b_i)$$

[This is *intended* and mostly interpreted to be merely a statement about *meaningfulness* (*)]

2. Independence of Irrelevant Alternatives: if S' contains S , and $f(S',d)$ is in S , then $f(S,d) = f(S',d)$

(this was the controversial axiom)

3. Pareto optimality

$$f(S,d) = x \text{ such that for no } y \text{ in } S \text{ is } y \geq x$$

4. Individual rationality $f(S,d) \geq v$ where v is the disagreement payoff

5. Symmetry: If S is symmetric then $f_1(S,d) = f_2(S,d)$

[S symmetric means (x_1, x_2) is in S iff (x_2, x_1) is also in S .]

The Nash Bargaining Solution:

Requirements:

S is convex, compact, and there exists an x such that $x > d$ for both players where d is the threat point payoff.

Players have complete information over S, d .

The negotiated outcome maximizes $(x_1 - d_1)(x_2 - d_2)$ where x_i is player i 's negotiated payoff and d_i is the threat point payoff

The binary lottery game design controls for the predictions of the theory, i.e. it allows us to know (S,d) , and therefore $f(S,d)$. [It isn't a control for the behavior of the bargainers...]

The complete information hypothesis (and the auxiliary assumption that players are utility maximizers with utility functions $u_i(x_i)$ that have as arguments only their own payoff) now predicts that

information about the prizes won't effect the outcome of bargaining. (!)

- That is, up to choice of origins and scales, (S,d) isn't sensitive to the prizes $\$L_i$ and $\$s_i$, and under the assumption that each player's preferences concern his own payoffs only, it isn't sensitive to whether each player knows the other's prizes.

The experimental environment:

- Each participant sat at a visually isolated terminal of a networked computer laboratory.
- Participants could send each other text messages (which passed through a monitor's terminal) containing anything other than information about personal identity (e.g. "I am sitting in station 24 of the foreign language building, wearing a blue windbreaker" was not allowed).
- They could also send numerical proposals.
- After a message or a proposal was entered, it appeared on the screen with a prompt asking whether you wanted to edit it or transmit it to your bargaining partner.
- In order to accept a numerical proposal (i.e. shares of lottery tickets to win the high prize), you had to transmit the identical proposal back. (e.g. my share = 67%, your share = 33%)
- There was a fixed time period, and a clock on the screen counted off the time.
- If agreement on a numerical proposal had not been reached by the end of the time period, the game ended with disagreement.

Early experiments

The strong prediction of the complete information hypothesis is easily refuted by experiments in which e.g.

$$s_1 = s_2 = 0, \text{ and } L_1 > L_2.$$

- When neither player knows the other's prizes, outcomes cluster closely around 50-50 agreements at which each player gets half the lottery tickets.
- But when both players know L_1 and L_2 , player 2 suddenly wants (and often gets) more than half of the lottery tickets.

- So experiments show us an unpredicted effect of information about prize values.
- One of the nice things about experimentation is that, when we have a robust phenomenon, we can explore it in a detail unimaginable in field data.
- In particular, we see that when the players know each others' prizes, the outcomes are different than when they do not. From a variety of theory, we know that common knowledge is different from simple shared information.
- In the lab we can explore these relatively subtle but potentially important issues.

digression on common knowledge

Something is called “common knowledge” if not only do

- *I know it* and
- *you know it*, but
- *I know that you know it*, and *you know that I know it*, and
- *I know that you know that I know it*, and *you know that I know that you know it*, and

Every sentence of that form of any length is true.

- (Also, “*I know that you know that he knows that she knows...*”)

Loosely speaking, we think of events becoming common knowledge when they happen in public, so I see it, you see it, I see you see it, etc.

Common knowledge is a convenient “overkill” assumption, when we don’t want to assume that something is well known.

It also has subtle, non-obvious consequences, that can make the assumption that a bit of information is common knowledge far more powerful than simply assuming that they all know it.

Consider the famous story of the red hats...

"The red hat problem"

- Consider a group of people, say 100, sitting in a circle, each wearing a hat that is either red or blue. Each person can see every hat but his own. As it happens, all 100 hats are red, but no one knows the color of his own hat. No one is allowed to speak. (It is common knowledge that everyone in the room is "perfectly rational.")
- A master of ceremonies stands in the middle of the circle, and makes the following announcement. "Every sixty seconds I will ring a bell. If, when I ring the bell, you know that you are wearing a red hat, you should get up and leave the room."
- It should be reasonably clear that no one ever leaves the room: no information is conveyed by the ringing of the bell.
- Suppose instead, the master of ceremonies prefaces his announcement with the following observation:
 -
 - "There is at least one red hat in the room."
- Notice that, since everyone in the room can see that there are at least 99 red hats, this announcement will not surprise anyone. Nevertheless, this public announcement fundamentally changes the situation: now, the bell rings 99 times and no one leaves, but on the 100th ring of the bell, everyone gets up and leaves the room.

- To see why, first consider the case of 2 people, instead of 100. You and I each see each other's red hat, but don't know the color of our own. So we each know that there is at least one red hat. But when this is publicly announced, it becomes common knowledge, so now we each not only know that there is a red hat, we each know that the other knows this. So, after the first bell, when you see that I do not get up and leave (as I would if your hat were blue) you know that your hat is red, and so at the second bell you get up and leave, as do I.

Before going to the general case of n people, consider the case $n=3$.

- You and I both see that the third person's hat is red, and we see each other seeing it. Furthermore, we both know that the third person knows that there is at least one red hat (I know it because I know he sees your hat, you know it because you know he sees my hat). But until the announcement, you don't know that I know that he knows there is one red hat. Now for the bells: if your hat were blue, he and I would both ignore you and treat the problem as in the case of 2 people, and leave at the second bell. So when you see that we do not leave at the second bell, you know your hat is red, and so you (and we) leave at the third bell.
- You can now prove the proposition for a group of any size by induction on n ... 😊

- In experiments, when we want to make something (approximately) common knowledge, we do so by making it public information.
- Consider the following experimental design, involving bargaining in a binary lottery game between two players, one of whom has a \$20 prize and one of whom has a \$5 prize. A player in any cell of the experimental design always knows his own prize. (Roth and Murnighan, *Econometrica*, 1982)

**DESIGN OF THE SHARED-INFORMATION
COMMON-KNOWLEDGE EXPERIMENT**

Information	Common Knowledge	Non-Common Knowledge
Neither player knows both prizes	1	5
Only the \$20 player knows both prizes	2	6
Only the \$5 player knows both prizes	3	7
Both players know both prizes	4	8

- In the common knowledge conditions, the common instructions informed both players about binary lottery games without specifying the prizes, and then told them that, after the common instructions were read, each player would receive private information of the sort indicated in the design. (The private info would then say, e.g. something like “your prize is \$5, you are not being told the other player’s prize, but he is being told the value of your prize.”)
- In the not common knowledge conditions, they were told that in the private information may or may not include information about the other players’ prize. (The private info would then say something like “your prize is \$5, you are not being told the other player’s prize, and he may or may not know the value of your prize.”)

Frequency of Agreements in Terms of the Percentage of Lottery Tickets Obtained by the \$20 Player

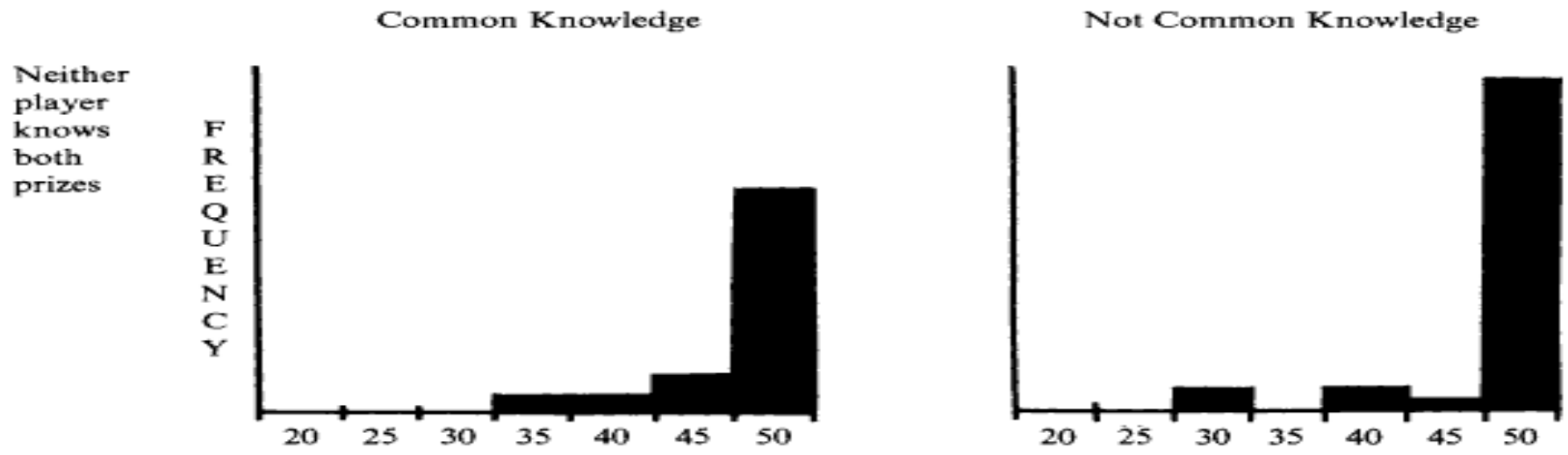


FIGURE 1.

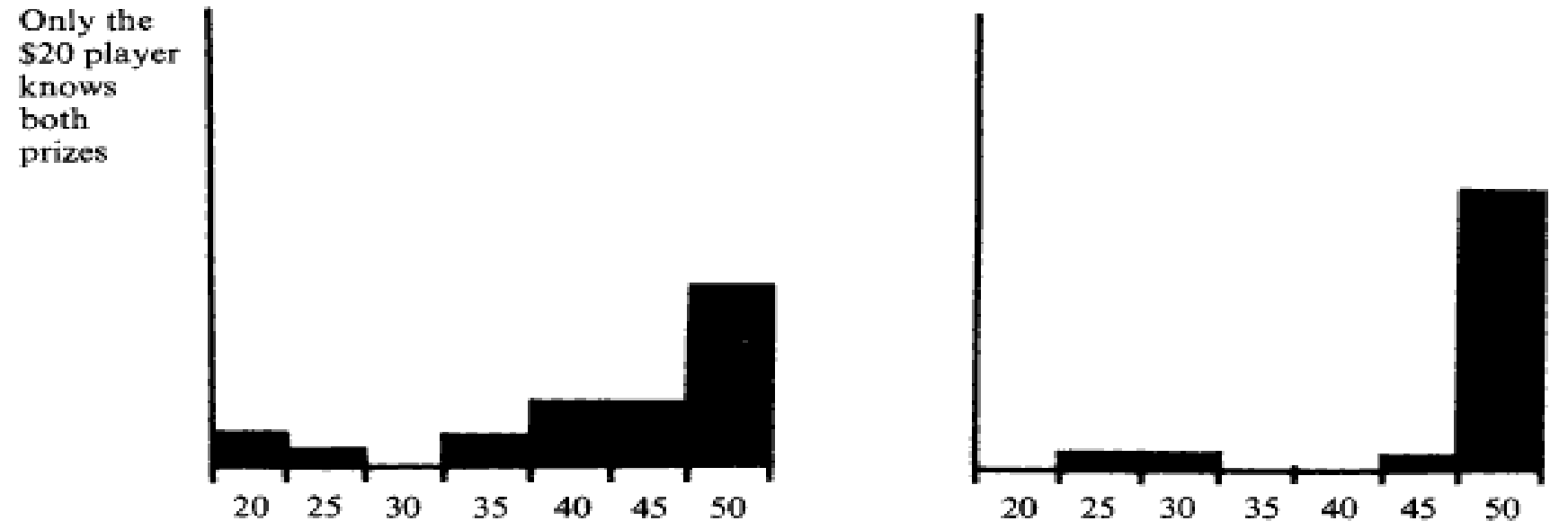


FIGURE 2.

Common Knowledge

Not Common Knowledge

Only the \$5 player knows both prizes

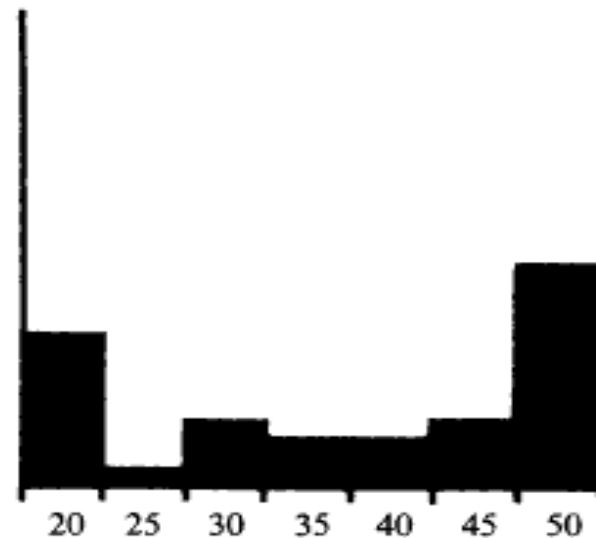
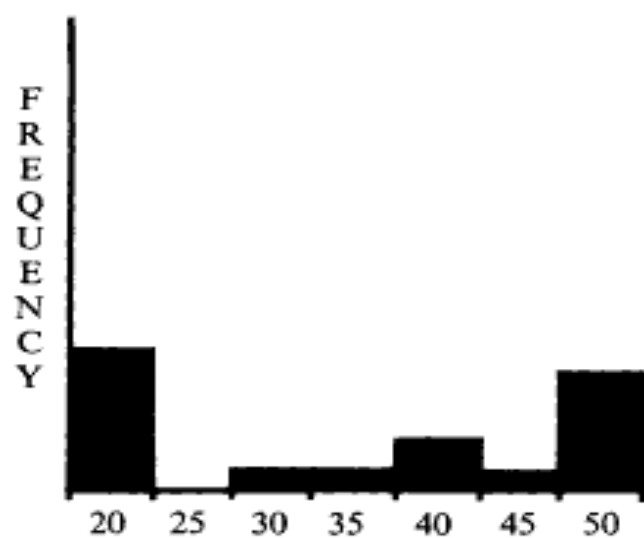


FIGURE 3.

Both players know both prizes

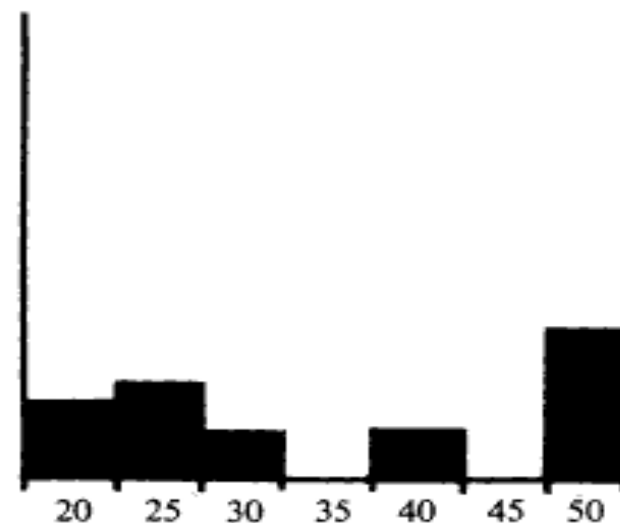
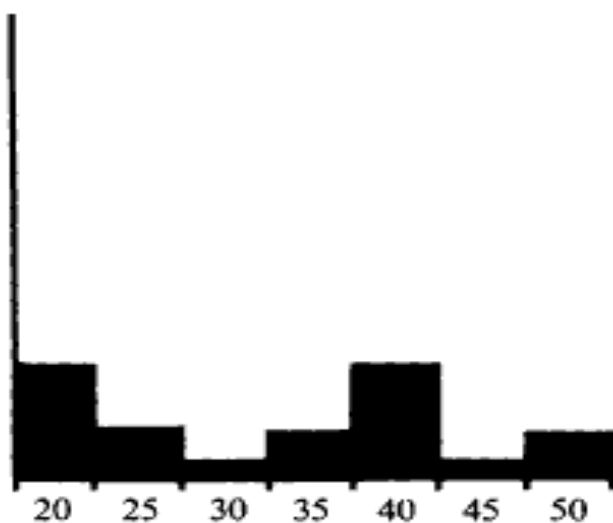


FIGURE 4.

Disagreements:

Disagreements (%)	Common Knowledge	Non-Common Knowledge
Neither player knows	14%	8%
Only \$20 player knows	20%	17%
Only \$5 player knows	19%	33%
Both Players know	17%	26%

MEAN OUTCOMES TO THE \$20 AND \$5 PLAYERS IN EACH
INFORMATION/COMMON KNOWLEDGE CONDITION OVER ALL INTERACTIONS
(DISAGREEMENTS INCLUDED AS ZERO OUTCOMES)

Information	Common Knowledge		Not Common Knowledge	
	\$20 Player	\$5 Player	\$20 Player	\$5 Player
Neither player knows both prizes	41.6 _{ab}	43.3 _c	43.5 _a	48.2
Only the \$20 player knows both prizes	34.9 _{bc}	45.1 _{bc}	40.9 _a	42.4
Only the \$5 player knows both prizes	27.2 _c	53.6 _{ab}	25.0 _b	42.0
Both players know both prizes	27.2 _c	56.4 _a	25.5 _b	48.8

NOTE: Within a column, means with common subscripts are *not* significantly different from one another using the Mann-Whitney *U* test ($\alpha = .01$); none were significantly different in the Not-Common-Knowledge conditions for the \$5 player.

conclusions

- whether players know each other's prizes makes a difference in the outcome.
- distribution of agreements primarily reflects whether \$5 player knows both prizes
- common knowledge influences frequency of disagreements
- in the not-common knowledge conditions, the relationships among outcomes in observationally equivalent cells is consistent with equilibrium behavior

Serendipity

- After many of these experiments, Roth et al analyzed the distribution of agreements over time, and found some striking regularities.
- (Roth, Murnighan and Schoumaker, "The Deadline Effect in Bargaining: Some Experimental Evidence," AER, 1988.)
- The distributions over time all look like this...

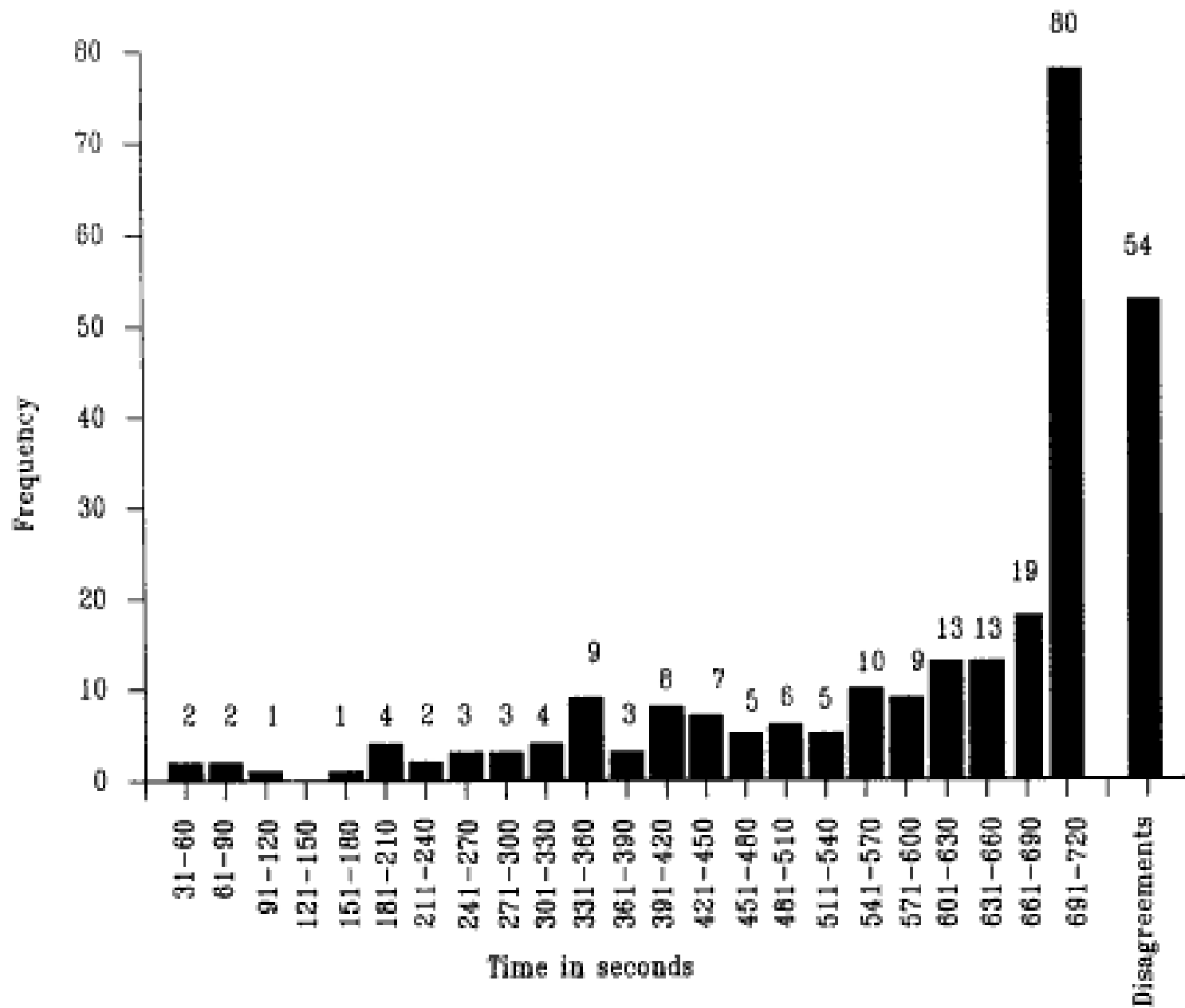


FIGURE 3A. THE FREQUENCY OF AGREEMENTS AND DISAGREEMENTS FROM ROTH AND MURNIGHAN (1982)

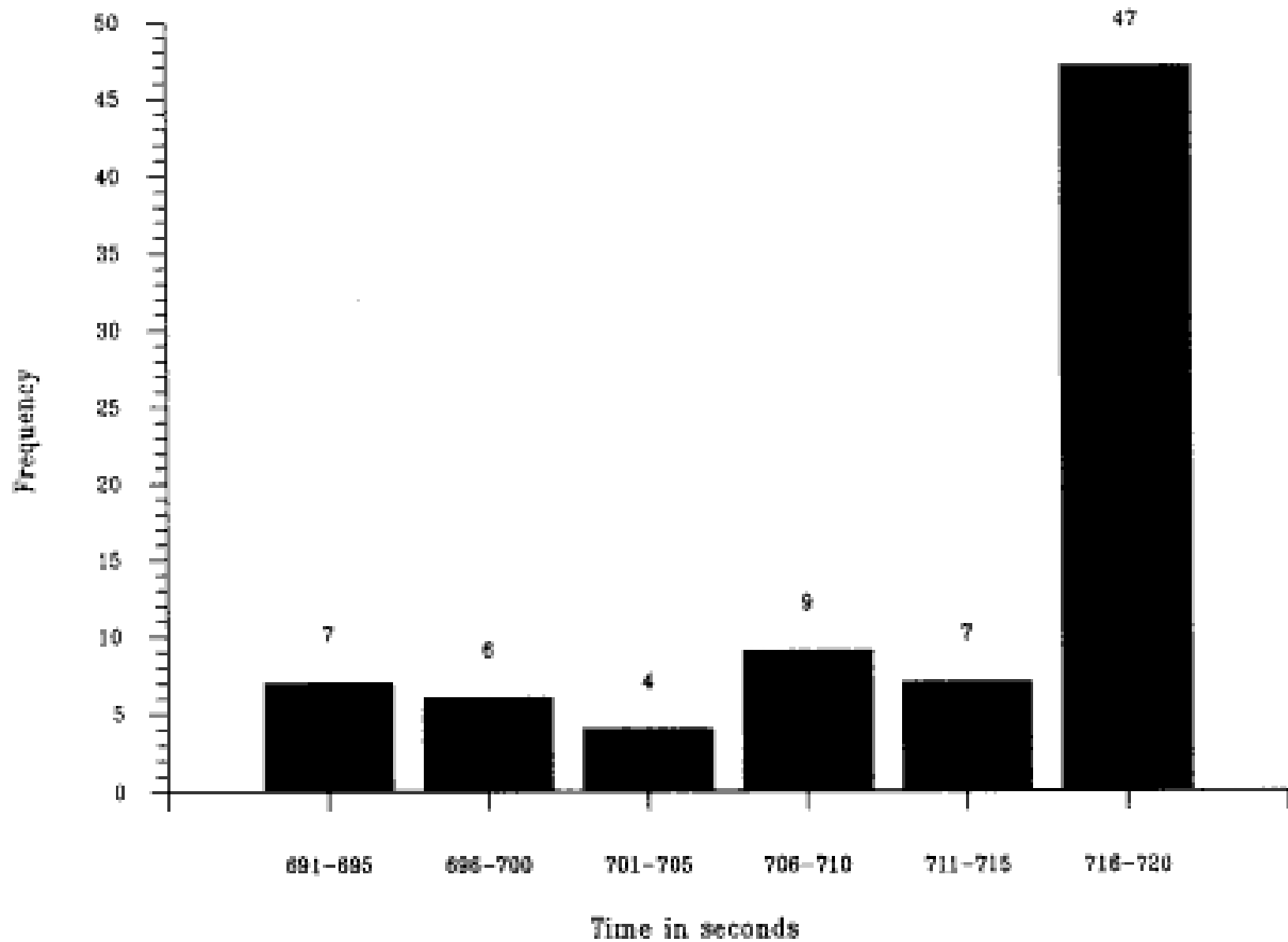


FIGURE 3B. THE FREQUENCY OF AGREEMENTS REACHED IN THE LAST 30 SECONDS OF BARGAINING IN ROTH AND MURNIGHAN (1982)

Expectations as an independent variable (Roth and Schoumaker, AER, 1983)

- So far, we've talked about experiments which show that there is a big unpredicted effect of information about prizes. The experimental designs attempt to control for expected utilities, and to some extent for strategies, so one place left in a rational actor model is expectations. Perhaps information about prizes changes bargainers' expectations about what agreements will prove acceptable.
- But many game-theoretic models of bargaining treat expectations as something to be derived from the data of the game. The next experiment was designed to see if expectations were manipulable independently of the game data, in a two-stage bargaining environment studied by Harsanyi.
- All games are binary lottery games, in which one player has a \$10 prize and one has a \$40 prize.
- Stage 1: each player states a demand p_i . If $p_1 + p_2 \leq 1$, the game ends, and player i receives probability p_i of winning.
- If $p_1 + p_2 > 1$, the game proceeds to stage 2.
- Stage 2: Each bargainer i can either state $q_i = p_i$ or
- $q_i = (1 - p_j)$. If $q_1 + q_2 \leq 1$, each bargainer receives probability q_i of winning. Otherwise each bargainer receives probability 0.

The experimental conditions were motivated by the following thought experiment.

- Suppose a randomly selected individual plays a large number of these games, and although he doesn't know it, his opponents are all confederates of the experimenter, and they never demand more than say, 20% for themselves.
- After a while, he gets used to getting 80%. He's written up in the newspapers...
- Now it is *your* turn to bargain with him, on your own behalf, not as a confederate. What should you do?

Note in passing that the above thought experiment involves deception. Furthermore, if the deception were simply removed, the results would likely be different, so the deception plays an important role in this design. (But a different, cleverer design might be able to explore the same issues without deception.)

Something to think about: design such an experiment.

Experimental design:

Each player played 25 2-stage games.

“Although players were told that they bargained with another individual in each game, each individual in fact played against a programmed opponent (the computer) in the first 15 games”

[N.B. This was in the early 1980's. A.R. himself would think hard before using a design involving deception today.]

Both players knew both prizes (\$40 and \$10), and each player had the same prize throughout all 25 games.

Subjects were divided into 3 experimental conditions:

- “20-80”: the agent whose prize was \$40 bargained with a computer programmed to randomly select a first demand between 75 and 80%, and to repeat its demand in the second stage. The programmed opponent of the \$10 player randomly selected a demand between 20 and 25%, and in stage two it accepted any offer giving it at least 20%.
- “50-50”: the programmed opponent of the \$40 player randomly selected a first demand between 70 and 75%, and in stage two accepted any offer giving it at least 50%. The programmed opponent of the \$10 player randomly selected a demand between 45 and 50%, and in stage two always repeated its demand.
- “control”: subjects never bargained against a computer, always against other subjects.

After trial 15, new instructions appeared on the screen, and explained that in the remaining 10 trials, each player would see his counterpart’s history for trials 11 through 15. (i.e. they now have a “reputation”)

In trials 16-25, *subjects in each group bargained with other members of that group.* Each game was played with a different, anonymous opponent. Bargainers also were told their opponent's first demand, whether he repeated it or accepted his opponent's offer, and which agreement, if any, was reached in each of trials 11 to 15 (i.e. the final 5 games against a programmed opponent, when each player's behavior reflected his experience.)

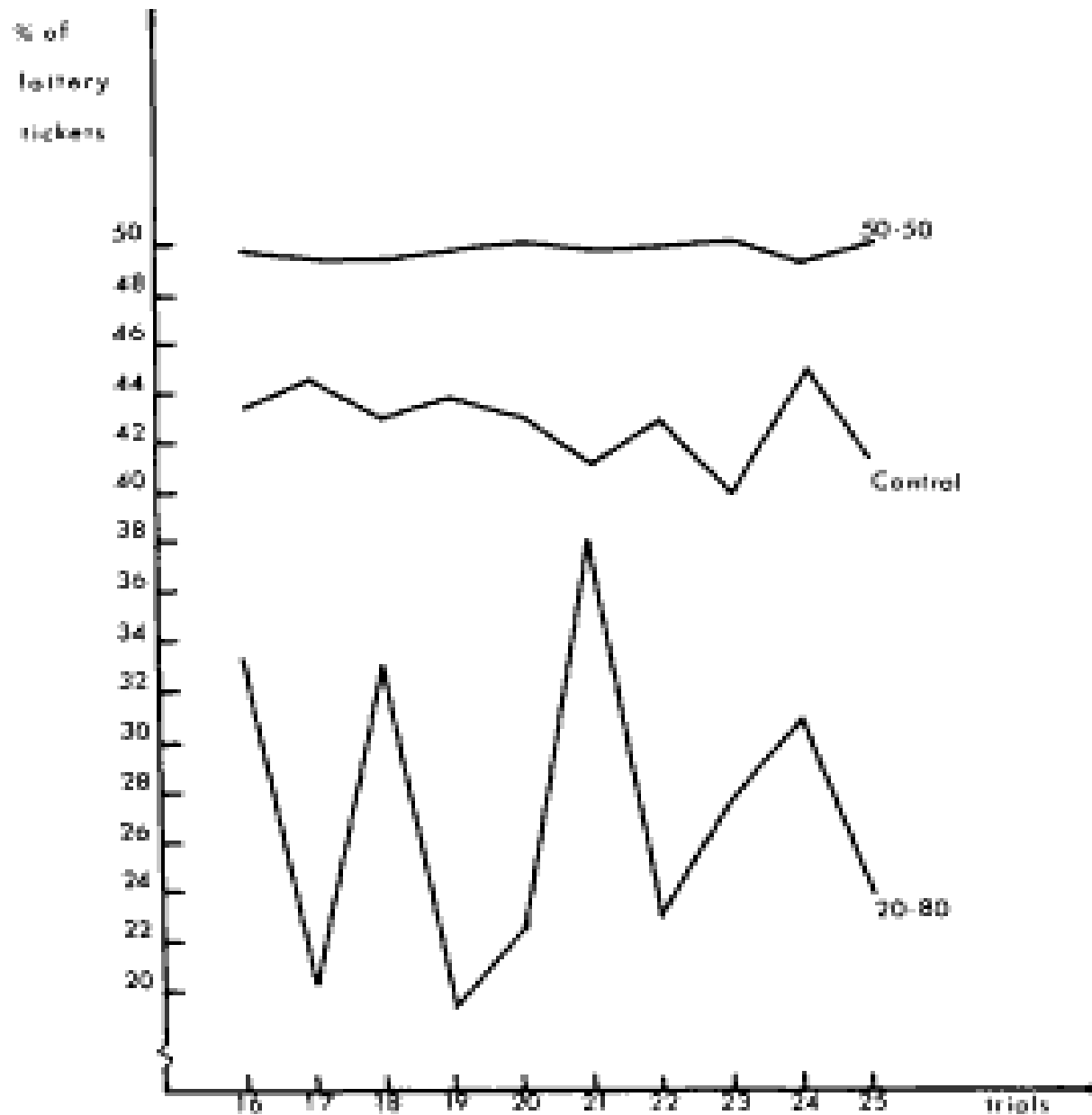


FIGURE 1. AVERAGE PERCENTAGE OF LOTTERY TICKETS OBTAINED BY \$40 PLAYER WHEN AGREEMENT WAS REACHED IN TRIALS 16-25

- So..... a long series of experiments showed a robust and unpredicted effect of information.
- It identified a clear deadline effect, in an environment that ruled out some of the traditional explanations for deadline effects.
- And it also produced evidence confirming subtle aspects of the theory's predictions about the qualitative effects of risk aversion.
- But many of the hypotheses raised about the causes of these effects depend on individual behavior that is hard to observe in the collective outcomes. To isolate individual behavior, it is convenient to look at games in which the players move *sequentially*.
- We'll discuss such games next, with particular attention to the remarkably simple ultimatum game.