A simple test of expected utility theory using professional traders

John A. List[a,b,c,d] and Michael S. Haigh[b,e]

Departments of [a]Economics and [b]Agricultural and Resource Economics, University of Maryland, College Park, MD 20742; [c]National Bureau of Economic Research, Cambridge, MA 02138; and [e]Office of the Chief Economist, U.S. Commodity Futures Trading Commission, Washington, DC 20581

Communicated by Marc Nerlove, University of Maryland, College Park, MD, November 1, 2004 (received for review May 17, 2004)

We compare behavior across students and professional traders from the Chicago Board of Trade in a classic Allais paradox experiment. Our experiment tests whether independence, a necessary condition in expected utility theory, is systematically violated. We find that both students and professionals exhibit some behavior consistent with the Allais paradox, but the data pattern does suggest that the trader population falls prey to the Allais paradox less frequently than the student population.

Allais paradox | experiments | futures traders

Expected utility (EU) theory remains the dominant approach for modeling risky decision-making and has been considered the major paradigm in decision making since World War II, being used predictively in economics and finance, prescriptively in management science, and descriptively in psychology (1). Furthermore, EU is the common economic approach for addressing public policy decision-making: a comparison of the expected costs and benefits of a proposed public policy implicitly assumes that economic agents maximize EU.

The mathematical form of EU stems back to the 17th century during the development of modern probability theory. Mathematicians Blaise Pascal and Pierre de Fermat assumed that the attractiveness of a lottery offering payoff $(x_1, \ldots, x_n)$ with probabilities $(p_1, \ldots, p_n)$ was simply given by its expected value $\bar{x} = \Sigma x_i p_i$. However, the fact that individuals evaluate more than expected value was dramatically illustrated by Nicholas Bernoulli in 1728 by using an example known today as the St. Petersburg paradox.[f]

Daniel Bernoulli (2), who was Nicholas's cousin, offered a solution to this paradox. Specifically, he outlined intuitive reasons why individuals would pay only a small amount for a game with an infinite mathematical expectation. He argued that a \$100 gain was not necessarily worth more than twice as much as a gain of \$50, suggesting that individuals maximize EU rather than expected monetary value. The utility function $U(x)$ he proposed was logarithmic, which exhibited diminishing increases in utility for equal increments in wealth and did, in fact, show that the EU is indeed finite in such cases. [Bernoulli proposed the following function: $U(x) = b\ln[(\alpha + x)/\alpha]$. $dU(x)/dx = b/(\alpha + x)$, which is inversely proportional to wealth. Moreover, $d^2U(x)/dx^2 < 0$.]

The work of John von Neumann and Oskar Morgenstern (3) proved that several basic axioms guarantee that there exists a utility index such that the ordering of lotteries based on their expected utilities fully coincides with the person's actual preferences.[g] Although EU represents a convenient and tractable approach to measuring utility, it continues to be the focus of much theoretical, empirical, and experimental research simply because EU theory is often shown to be systematically violated under certain conditions. Indeed, in experimental investigations, the weight of evidence suggests that student subjects make decisions inconsistent with EU theory (see refs. 4–9). Despite finding violations of EU, the theory has been shown to be a "pretty good" approximation to individuals' true preferences. Indeed, according to some researchers, learning and familiarization with the decision tasks are required before true preferences settle on the genuine underlying form (10).

Understanding whether experienced agents, like professional traders, make similar decisions is important in light of some recent studies (e.g. refs. 11 and 12) that find that market anomalies in the realm of riskless decision-making are attenuated among economic players who have intense market experience.[h] To provide a stringent test of EU theory, we follow Starmer and Sugden (7) and use choice problems that correspond very closely to one of the classic forms of the Allais paradox (18), the "common consequence effect," a test of the independence assumption in EU theory.[i]

In our experimental investigation of EU theory (specifically the test of the independence assumption), we find that both students and professional traders from the Chicago Board of Trade (CBOT) exhibit behavior consistent with the Allais paradox (thus violating EU theory), but the data pattern does suggest that the trader population falls prey to the Allais paradox slightly less frequently than the student population. Furthermore, we find that professional traders behave in accordance with the reduction principle (reducing compound lotteries to simple ones via the calculus of probabilities (see ref. 20), whereas students did not exhibit this tendency.

ECONOMIC SCIENCES

Abbreviations: EU, expected utility; CBOT, Chicago Board of Trade.

[d]To whom correspondence should be addressed. E-mail: jlist@arec.umd.edu.

[f]The St. Petersburg paradox is as follows: Imagine someone offers to toss a coin repeatedly until it comes up tails, and to reward you \$1 if this happens on the very first toss, \$2 if it takes two tosses, \$4 if it takes three tosses, \$8 for four tosses, etc. What is the largest sure gain you would be willing to forego in order to undertake a single play of the game? Because the gamble offers 1/2 chance of winning \$1, 1/4 of winning \$2, etc., the expected value is $(1/2) \times \$1 + (1/4) \times \$2 + (1/8) \times \$4 = 1/2 + 1/2 + 1/2 + \cdots + 1/2 = \$\infty$, so it should be preferred to any finite gain. However, it is clear that very few individuals would forego more than a moderate amount for a one-shot play.

[g]Specifically, if a person's preferences satisfy three axioms (ordering, continuity, and independence), it is appropriate to model his behavior as if he were maximizing EU. For a more complete discussion of the origin of EU theory and a discussion of the axioms, see Schoemaker (1).

[h]Christensen-Szalanski and Beach (13) make an even stronger argument, noting that the experimental literature using students is biased (see also ref. 14). Their main contention is that experimental findings of major anomalies have largely used student subjects, whereas those few studies that have employed professionals have usually reported performance more in line with mainstream theory. This point is strengthened given that there are numerous reasons to suspect that professional behavior may differ from nonprofessional behavior due to training, regulation, etc. (15, 16). Also, see List (17) for the use of experienced players in field experiments testing market theory.

[i]Because the only difference between pair wise sets of safe and risky options that are presented to agents (the usual set up for testing the "common consequence effect") are common to each respective lottery, EU theory suggests that either all safe lotteries, or all risky lotteries, are picked in all lottery choices (or indifference is shown). A common consequence effect is manifest if choices switch systematically between the riskier and safer options. Readers are directed to Machina's (6, 19) pictorial representation of this scenario. Indeed, Machina's use of probability triangles popularized the concept to the extent that the representation is sometimes referred to as the "Machina Triangle."

Table 1. Experimental design

Subject type	Treatment 1	Treatment 2	Treatment 3
Students	32	32	30
Traders	27	27	54

Data represent sample sizes. Treatment 1 (2) had subjects choosing between R_1 and S_1 (R_2 and S_2). Treatment 3 had subjects choosing between R_1 and S_1 and then R_2 and S_2, where it was understood that one of the choices would be carried out for real earnings.

The Allais Paradox

One important violation of EU's independence assumption is the Allais paradox.[j] Indeed, a survey conducted by Allais in 1952 showed that the majority of real decision makers order risky prospects in a way that is inconsistent with the postulate that choices are independent of irrelevant alternatives, thus casting doubt on the validity of EU theory. However, it is precisely this postulate that permits economists to represent preferences over risky prospects as a linear function of the utilities of the basic outcomes, and hence have the ability to measure utility. A classic paper by Marschak (25) discusses the rationality of this approach by using examples from both statistics and business.

As presented by Allais (18), a person chooses between lotteries a_1 and a_2, and between lotteries a_3 and a_4, where a_1, a_2, a_3, and a_4 are: a_1, 1.00 chance of \$1,000,000; a_2, .10 chance of \$5,000,000, 0.89 chance of \$1,000,000, and 0.01 chance of \$0; a_3, 0.10 chance of \$5,000,000, 0.90 chance of \$0; and a_4, 0.11 chance of \$1,000,000; 0.89 chance of \$0.

In this example, if an agent prefers a_1 to a_2 and a_4 to a_3, (or a_2 to a_1 and a_3 to a_4), the individual does not violate the independence axiom and the evidence is consistent with the agent having indifference curves that are parallel straight lines (as surmised by EU theory) (this is illustrated pictorially in refs. 6 and 19). Laboratory experimentation, however, has shown that agents commonly choose a_1 and a_3, violating the independence assumption. To understand the nature of the paradox, note that the expected value of a_1 is \$1 million and the expected value of a_2 is \$1.39 million. By preferring a_1 to a_2, an agent is presumably maximizing EU, not expected value. If $a_1 > a_2$, then $u(1) > 0.1u(5) + 0.89u(1) + 0.01u(0)$, implying that $0.11u(1) > 0.1u(5) + 0.1u(0)$, which in turn [adding $0.89u(0)$ to each side] implies $0.11u(1) + 0.89u(0) > 0.1u(5) + 0.90u(0)$. This suggests that an EU-maximizing agent must prefer a_4 to a_3. Of course, the expected value of a_4 is \$110,000, whereas the expected value of a_3 is \$500,000, so if the agent is maximizing expected value, he should in fact prefer a_4 to a_3. However, his choice in the first stage is inconsistent with his choice in the second stage, and thus the paradox emerges.

Experimental Design

Our experimental design is a slightly different example of the Allais paradox most commonly referred to as the "common consequence effect." [Other examples include "the certainty effect" and the "Bergen paradox." See Machina (6, 19) or Thaler (9) for a more in-depth treatment of these cases and the literature involving independence violations.] In our example, we ask subjects to make simple lottery choices. We make use of a straightforward 2×3 experimental design (see Table 1). Because one important purpose of our research is to explore whether agents who are professional traders exhibit behavior in line with EU, we use undergraduate students as our experimental control group.[k] Using a between-person experimental design, we include both undergraduate students and professional traders in two distinct treatments: treatments 1 and 2. In a third treatment, we had a different student group participate, but had the same 54 traders who participated in treatments 1 and 2 complete treatment 3. (See Fig. 1 for our experimental instructions for treatment 3. Instructions for the other two treatments are similar, with the necessary adjustments.) The traders first took part in treatment 1 or 2 (determined randomly) and then proceeded through some

This experiment consists of a simple lottery choice. Below we provide two lotteries:

1. A. You win \$7 with certainty.
   B. You win \$7 with 75% probability (a draw of numbers 1-75); \$10 with 20% probability (a draw of numbers 76-95); and \$0 with 5% probability (a draw of numbers 96-100).

2. A. You win \$7 with 25% probability (a draw of numbers 1-25); \$0 with 75% probability (a draw of numbers 26-100).
   B. You win \$10 with 20% probability (a draw of numbers 1-20); \$0 with 80% probability (a draw of numbers 21-100).

Your task is to choose between A and B for *each* of the two lotteries above. BUT ONLY ONE OF THESE 2 LOTTERIES WILL ACTUALLY INVOLVE CASH PAYMENTS. In other words, after you have made your choices for both of the 2 lotteries, we will choose one of them that we're really going to use – we'll call this the "Binding Lottery." The binding lottery will be determined by a flip of a fair coin. A flip of heads and the binding lottery is #1; a flip of tails and the binding lottery is #2.

We'll then play the lottery selected. I have an empty envelope here with 100 slips of paper. Each slip of paper has a number, 1, 2, 3…up to 100. I will show the empty bag and the 100 slips to one of your colleagues; if you would also like to see them, just come right on up to the front of the room and inspect them. Once all of us have made our choices, I will first flip this fair coin to decide which lottery is binding and then I'll ask your colleague to reach into the bag and pull out a slip of paper. Whatever number is on that piece of paper will define the outcome of the lottery.

Any questions? Please begin.

Fig. 1. Experimental instructions for treatment 3 given to traders.

[j]Alternative theories to EU that can accommodate the Allais paradox include: Generalized Expected Utility Theory (19); Rank Dependent Theory (21, 22), and Prospect Theory (23). However, despite their ability to accommodate the paradox, all theories produce divergent predictions in certain cases (24).

[k]We refer to all subjects recruited from the CBOT as traders. However, the 54 traders recruited consisted of locals, brokers, clerks, and exchange employees (e.g., floor managers or market reporters) who worked in the open outcry environment. We found no statistical difference between floor participant types; therefore, we pool participants and collectively call them "traders." Note that this finding is intuitive because the average non-local/broker had accumulated ≈9 years of floor experience, and many reported to have had several years of experience as either a local or broker. Finally, the average trader (including non-locals/brokers) was involved with ≈537 traded contracts daily.

unrelated experiments for ≈1 h, and then participated in treatment 3 at the end of the session.[l]

In treatment 1, subjects were confronted with a simple one-shot choice question (following Starmer and Sugden (7), S (R) represents "safer" ("riskier") choice): R_1, 0.05 chance of zero winnings, 0.20 chance of winning \$10, 0.75 chance of winning \$7.00; S_1, \$7 with certainty. In treatment 2, the choice was over a different pair: R_2, 0.80 chance of zero winnings, 0.20 chance of winning \$10; S_2, 0.75 chance of zero winnings, 0.25 chance of winning \$7.

In this case, EU theory predicts that if R_1 (S_1) is chosen instead of S_1 (R_1), then R_2 (S_2) should be chosen instead of S_2 (R_2). Alternatively, the Allais paradox and several other decision-making theories predict that subjects will "switch" by choosing S_1 and R_2 (see ref. 7). This paradoxical choice has been termed the "common consequence effect," and is a classic form of the Allais paradox.

In treatment 3, subjects were confronted with the two identical choice questions and informed that they should make a choice for both questions, but that only one of the questions would be binding, i.e., carried out for real money. This treatment, which is commonly referred to as the "random payment" or "random lottery" procedure in experimental economics, was used to provide information on whether subjects use the "reduction principle," reducing compound lotteries to simple ones via the calculus of probabilities (see ref. 20). Under our design the lotteries reduce as follows: (R_1R_2) = (\$10, 0.20; \$7, 0.375; 0, 0.425); (R_1S_2) = (\$10, 0.10; \$7, 0.50; 0, 0.40); (S_1R_2) = (\$10, 0.10; \$7, 0.50; 0, 0.40); (S_1S_2) = (\$7, 0.625; 0, 0.375).

For example, the choices of R_1 and R_2 in treatment 3 provide \$10 with 20% chance, \$7 with 37.5% chance, and \$0 with 42.5% chance. In this case, (R_1S_2) and (S_1R_2) should be chosen with equal probability if the reduction principle holds because these pairs are isomorphic in reduced form. It seems intuitive to us that, if two options are identical, they will be chosen with equal probability. We should stress that a test of the reduction principle relies on this proposition.

We recruited 94 subjects for our student treatments from the undergraduate student body at the University of Maryland (College Park). Each treatment was run in a large classroom on the College Park campus. To ensure that decisions remained anonymous, the subjects were seated far apart from each other. The trader subject pool included 54 professional traders from the CBOT. Each of the trader treatments was run in a large room on site at the CBOT. As in the case of the students, communication between the subjects was prohibited and the traders were seated such that no subject could observe another individual's decision (and payoffs).

Before moving to a discussion of the experimental results, a few noteworthy aspects of our experimental design merit further consideration. First, all treatments were run with pencil and paper. After subjects made their decisions and the lottery results were realized, experimenters circulated to ensure that individual payoffs were calculated correctly. Second, in treatment 3, we varied the order in which the questions were given with no apparent ordering effect in the results, so we pooled the data. Third, in the student treatments, the exchange rate was 1:1 (1 cent for each unit), and in the trader treatments the exchange rate was 4:1 (4 cents for each unit). Our decision to quadruple the stakes for traders was based on a detailed discussion with CBOT officials about trader earnings. However, we should highlight that all conclusions should be viewed in light of this technical qualification.

[l]For comparability purposes with respect to treatment 3 data, we also ran an experiment with 46 students that mirrored the trader treatment: these students first took part in treatment 1 or 2 (determined randomly), proceeded through some unrelated experiments for ≈1 h, and then participated in treatment 3 at the end of the session. This procedure produced similar data patterns as those gained from the treatments without learning possibilities, so we suppress further discussion.

Table 2. Experimental results summary

Subject type	Treatment 1	Treatment 2	Treatment 3
Students	$17R_1$ $15S_1$	$24R_2$ $8S_2$	$10R_1R_2$ $3R_1S_2$ $13S_1R_2$ $4S_1S_2$
Traders	$19R_1$ $8S_1$	$24R_2$ $3S_2$	$30R_1R_2$ $9R_1S_2$ $7S_1R_2$ $8S_1S_2$

Data represent the number of subjects making each choice. For example, in the student treatment 1, 17 subjects chose R_1 and 15 chose S_1. In treatment 3, data represent the patterns of choice. For example, $10R_1R_2$ shows that 10 students chose R_1 in the first question and R_2 in the second.

Results

Table 2 provides a summary of our data. Data in the treatment 1 and treatment 2 columns of Table 2 represent the number of subjects making each choice. For example, in the student treatment 1 data, 17 subjects chose R_1 and 15 subjects chose S_1. In treatment 3, data represent the patterns of choice. For example, $10R_1R_2$ shows that 10 students chose R_1 in the first question and R_2 in the second question.

The major issue at hand is whether subjects' choices are in line with the Allais paradox. We examine this question with a null hypothesis of $\Pr(R_1) = \Pr(R_2)$ and with a one-sided "common consequence effect" alternative: $\Pr(R_2) > \Pr(R_1)$. Among students, 17 of 32 (53%) chose R_1, whereas 24 of 32 (75%) chose R_2. Traders show a preference in a similar direction: 19 of 27 (70%) chose R_1, whereas 24 of 27 (89%) chose R_2. Although these differences are both statistically significant at the $P < 0.05$ level when a Fischer's exact test is used, suggesting that both traders and students' behavior is line with the Allais paradox, the data pattern does suggest that the trader population falls prey to the Allais paradox less frequently than the student population.

Our next exploration concerns whether the reduction principle holds. In this case, the null hypothesis is that $\Pr(R_1S_2) = \Pr(S_1R_2)$, and for a common consequence effect to be evident it must be the case that $\Pr(S_1R_2) > \Pr(R_1S_2)$. In the student data, 3 of 30 (10%) subjects chose R_1S_2 and 13 of 30 (43%) subjects chose S_1R_2, providing strong evidence of the common consequence effect, and inconsistent with the reduction conjecture. This result is consonant with Starmer and Sugden's (7) results using University of East Anglia (Norfolk, U.K.) students. Among traders, however, 9 of 54 (17%) chose R_1S_2 and only 7 of 54 (13%) chose S_1R_2. This difference, which is in the opposite direction of the student results and the common consequence prediction, is not statistically significant, and therefore we cannot reject the null hypothesis that traders behave according to the reduction principle. Also, the amount of "switching" for dealers (16 of 54) is less than that for students (16 of 30), suggesting again that the trader population falls prey to the Allais paradox less than the student population.

Concluding Remarks

In this study, we recruited futures and options floor traders from the CBOT to participate in an experiment involving choice over risky outcomes. Making use of undergraduate students as our experimental control group, we find some evidence that both students and professional traders behave in accordance with the Allais paradox. However, we find that the trader population falls prey to the Allais paradox slightly less frequently than the student population. A few caveats are in order. First, it is important to note that traders were making

ECONOMIC SCIENCES

decisions over higher financial stakes. Second, we should highlight that our findings do not necessarily conflict with those of previous authors who have found that experienced marketers do not fall prey to behavioral anomalies. In many of those cases, subject decision-making was observed in natural settings in which the agents have experience with the good, the task, the institution, the environment, etc. In our experimental design, it may be the case that all of the important design parameters, even the task, is foreign to the professional traders, leading to the observed results. A next step in our research agenda is to explore the comparative static effects of adding/removing context and observing behavior in/out of the natural setting.

Marc Nerlove and three sharp reviewers provided astute comments on an earlier version of this paper. We thank John Di Clemente, former Managing Director of Research at the Chicago Board of Trade, for authorizing the study. Chicago Board of Trade officials Frederick Sturm, Dorothy Ackerman Anderson, and Keith Schap also provided incredible support on site. We thank Jonathan Alevy for research assistance and the University of Maryland for funding this research.

1. Schoemaker, P. (1982) *J. Econ. Lit.* **20,** 529–563.
2. Bernoulli, D. (1738) *Commentarii Acad. Scientiarum Imperialis Petropolitanae* **5**, 175–192; English translation (1954) *Econometrica* **22,** 23–36.
3. von Neumann, J. & Morgenstern, O. (1944) *Theory of Games and Economic Behavior* (Princeton Univ. Press, Princeton).
4. Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M. & Combs, B. (1978) *J. Exp. Psychol.* **4,** 551–578.
5. Tversky, A. & Kahneman D. (1981) *Science* **211,** 453–458.
6. Machina, M. (1987) *J. Econ. Perspect.* **1,** 121–154.
7. Starmer, C. & Sugden, R. (1991) *Am. Econ. Rev.* **81,** 971–978.
8. Baron, J. (1992) *Thinking and Deciding* (Cambridge Univ. Press, New York).
9. Thaler, R. (1992) *The Winner's Curse* (Free Press, New York).
10. Plott, C. R. (1996) in *The Rational Foundations of Economic Behavior*, eds., Arrow, K. J. Colombatto, E., Perlman, M. & Schmidt, C. (MacMillan, Basingstoke, U.K.).
11. List, J. A. (2002) *Am. Econ. Rev.* **92,** 1636–1643.
12. List, J. A. (2003) *Q. J. Econ.* **118,** 41–71.
13. Christensen-Szalanski, J. J. & Beach, L. R. (1984) *Am. Psychol.* **39,** 75–78.
14. Bonner, S. & Pennington, N. (1991) *J. Acc. Lit.* **10,** 1–50.
15. Burns, P. (1985) in *Research in Experimental Economics* (JAI Press, Greenwich, CT), pp. 139–157.
16. Holt, C. A. & Villamil, A. P. (1986) in *Laboratory Market Research*, ed. Moriarity, S. (Univ. of Oklahoma Press, Norman), pp. 80–85.
17. List, J. A. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 15827–15830.
18. Allais, M. (1953) *Econometrica* **21,** 503–546.
19. Machina, M. (1982) *Econometrica* **50,** 277–323.
20. Holt, C. A. (1986) *Am. Econ. Rev.* **76,** 508–515.
21. Quiggin, J. (1982) *J. Econ. Beh. Org.* **3,** 323–343.
22. Yarri, M. (1987) *Econometrica* **55,** 95–115.
23. Kahneman, D. & Tversky, A. (1979) *Econometrica* **42,** 263–291.
24. Starmer, C. (1992) *Rev. Econ. Stud.* **59,** 813–830.
25. Marschak, J. (1951) in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* (Univ. of California Press, Berkeley), pp. 493–506.