Preference Reversals

James C. Cox
University of Arizona

Study of preference reversals originated with cognitive psychologists and has spread to experimental economics because it is directly relevant to the empirical validity of economic theories of decision-making under uncertainty. A preference reversal experiment involves paired choice and valuation responses, usually over simple two-outcome gambles. Subjects are asked to choose which of a pair of gambles they want to play. They are also asked to place minimum selling prices on the gambles in an experimental context in which telling the truth is a dominant strategy. A preference reversal occurs when a subject places a lower selling price on the gamble that he/she chooses than on the other gamble in a pair.

Preference reversals call into question the empirical validity of economic theory because they provide support for the conclusion that the preferences that subjects reveal vary with the response mode (choice or valuation) that is used to elicit the preferences. If the preference reversal phenomenon is robust, then standard economic decision theory is on shaky ground as an empirically useful positive theory of decision-making. Robust preference reversals would be even more of a problem for normative economics: consider the implications for cost-benefit analysis of preferences over alternatives that reverse with a change in the response mode used to elicit the preferences.

Binary lotteries that were used in the early experiments by Lichtenstein and Slovic (1971) and many subsequent papers are reported in Table 1. There are six pairs of lotteries. Each pair consists of a probability (or P) bet with a relatively large probability of a relatively small win state payoff and a dollar (or $ bet) with a relatively small probability of a relatively large win state payoff. Psychologists explain preference reversals as response mode effects on decisions. For example, when asked to place a value (minimum selling price) on a bet, a subject may first “anchor” on the values of the win state payoffs in two bets and then make an insufficient “adjustment” for the difference in probability of winning. In contrast, when asked to choose between two bets a subject may first anchor on the probabilities of the win state payoffs and then

make an insufficient adjustment for the difference in dollar values of the win state payoffs. This explanation implies that preference reversals will be asymmetric: it will be much more frequent that subjects will (a) place a higher value on the $ bet and choose the paired P bet, than (b) place a higher value on the P bet and choose the paired $ bet. Hence, preference reversals of type (a) are called "predicted reversals" and those of type (b) are called "unpredicted reversals."

1. Seminal Experiments

Grether and Plott (1979) explained several design features of earlier preference reversal experiments that called into question their implications for economics. They developed an experimental design that was more appropriate for economics and were surprised that the results confirmed the earlier findings. Figure 1 presents results from experiments by Lichtenstein and Slovic (1971, 1973) and Grether and Plott (1979). The reported experiments have the following characteristics: (a) Lichtenstein and Slovic III is a monetary payoff experiment with psychology students as subjects; (b) Lichtenstein and Slovic P is a monetary payoff experiment with positive expected payoff gambles run on the floor of a Las Vegas casino with adult gamblers as subjects; (c) Grether and Plott 1H is a hypothetical payoff experiment with undergraduate students as subjects; and (d) Grether and Plott 1M is a monetary payoff experiment with undergraduate students as subjects. Results from all of these experiments are similar. About 1/3 of all decisions involve preference reversals and predicted reversals are much more common than unpredicted reversals. Furthermore, there is no notable difference between the results from Grether and Plott's hypothetical payoff and monetary payoff experiments.

Thus the preference reversal phenomenon was robust to the changes in experimental procedures introduced by Grether and Plott. They attributed preference reversals to subjects' violations of transitivity.

2. Independence Axiom Treatments

Grether and Plott's paper motivated several authors to critique their design and question how robust were their results. However the experiments reported by these other authors produced results similar to those reported by Grether and Plott.

Holt (1986) and Karni and Safra (1987) questioned Grether and Plott's interpretation of their results and offered alternative interpretations as follows. In order to eliminate changes in subjects' wealth during an experiment, and thus remove a "wealth effects" easy explanation of preference reversals, Grether and Plott randomly selected a single decision of each subject for monetary payoff at the end of the experiment. Holt explained that this procedure requires the independence axiom of expected utility theory in order for the experimental results to be interpreted as preference reversals. Thus, Grether and Plott's conclusion that the preference reversals in their experiments were violations of transitivity was called into question. Karni and Safra examined another feature of the preference reversal experiments reported by Grether and Plott and most others, the use of the Becker, DeGroot, and Marschak (1964) procedure for eliciting selling prices. They explained that this procedure requires the independence axiom in order for the results to be interpreted as preference reversals.

Cox and Epstein (1989) and Tversky, et al. (1990) designed experiments that did not use either the Becker-DeGroot-Marshak (BDM) or the random decision selection procedure. They both used price elicitation procedures ("ordinal pricing tasks") that gave subjects an incentive to reveal sales prices with the correct relative values, but not the true values, and that did not require the independence axiom. Cox and Epstein paid their subjects after every decision in order to avoid the random decision selection procedure. They then used econometric analysis of the data to check for any significant wealth effects on decisions and found none. The experiments of Tversky, et al. used hypothetical payoffs (except in one treatment in which 15% of the subjects were randomly selected for small monetary payoffs at the end of the experiment).

Results from the experiments by Cox and Epstein and Tversky, et al. are reported in Figure 2, along with those from Grether and Plott's experiments 1M (with monetary payoffs) and 1H (with hypothetical payoffs) for comparison. The overall reversal rate in Cox and Epstein 1 is almost as high as in Grether and Plott 1M, but the former does not exhibit the asymmetric pattern of the latter. The different reversal pattern in Cox and Epstein 1 may be evidence that it was transparent to the subjects in their experiment that an ordinal pricing task is, in fact, a choice task framed as valuation.

Tversky, et al. 1 has a higher overall reversal rate than Grether and Plott 1H and both exhibit the asymmetric pattern. The asymmetry of reversals in Tversky, et al. 1 may be evidence that in their experiment, which is more complicated than Cox and Epstein's, the subjects did not realize that the ordinal pricing task is a choice task framed as valuation. Alternatively, it may instead be the case that the asymmetry resulted from subjects' lack of motivation in the complicated experiment. (See the next section.)

Together, the Cox and Epstein and Tversky, et al. experiments make clear that preference reversals cannot simply be attributed to violations of the independence axiom of expected utility theory.

3. Incentive Treatment

As noted above, Grether and Plott (1979) found that the use of monetary payoffs did not affect their conclusions about preference reversals. The results reported by Cox and Grether (1996) were also invariant to cash versus hypothetical payoffs with one striking exception, the English clock auction treatments. The English clock auction is a market sequential choice task that can be used to elicit selling prices. Figure 3 reports some results that make clear the pattern found. The results for the hypothetical payoff treatments, BDM 1H and ECA 5H, are quite similar. This would suggest the conclusion that after five replications in the market environment of the English clock auction the frequency and asymmetry of preference reversals are essentially

the same as for the BDM mechanism. But now consider the results for the monetary payoff experiments, BDM 1M and ECA 5M. Results for BDM 1M are very similar to BDM 1H. In contrast, results for ECA 5M are very different than those for all other experiments in Figure 3; in particular, the asymmetry of reversals is strikingly reversed.

The effects of financial incentives in this market environment are quite intuitive if one considers the decision procedure that it is involved. The sooner the subject chooses to play the bet rather than remain in the market, and thus remain eligible to sell it, the less time and effort that is expended watching the computer screen and making decisions. This effect is quite pronounced for $ bets with their high win state payoffs and high starting prices for the clock. A subject could spend less time watching the computer screen by pressing the key for choosing the bet and dropping out early. This is exactly what many subjects did when no money was at stake in the hypothetical payoff experiments.

Together, the Grether-Plott and Cox-Grether papers make clear that many results from preference reversal experiments are the same for hypothetical and monetary payoffs but that some results differ dramatically.

4. Transitivity Treatments

Tversky, et al. (1990) designed their hypothetical payoff experiments to be able to discriminate between violations of transitivity and other causes of preference reversals. Cox and Grether (1996) adopted this design feature in their monetary payoff experiments. Figure 4 reports the rates of intransitivity (IT), predicted reversals (PR), and unpredicted reversals (UR) from some of the experiments in these two papers. Note that the rates of intransitivity are far lower than the overall rates of preference reversal.

These experiments support the interpretation of the preference reversal phenomenon as a response mode effect, not a result of preference intransitivity.

5. Risk Neutrality Treatment

Most of the preference reversal experiments used gambles like those in Table 1 that involve pairs in which the P bet has essentially the same expected value as its paired $ bet. If the subjects were risk neutral then they would be essentially indifferent between the two bets in any such pair; in that case, preference "reversals" might simply reflect some convenient rule for resolving indifference. To check on this possibility, Cox and Epstein ran a second experiment in which there was a 50% difference between the expected payoffs of the two bets in each pair. In addition, in one-half the pairs the P bet had the higher expected payoff and in the other one-half the $ bet did. Figure 5 reports results from Cox and Epstein's experiment 1M using some of the P bets and $ bets in Table 1 and results from their experiment 2M using bet pairs with the 50% difference in expected payoff. The ordinal pricing task used in these experiments did not produce the asymmetric pattern of reversals characteristic of preference reversal experiments; hence the results are called choice reversals.

The observed choice reversals in the two experiments are similar; hence these results indicate that such reversals do not result from the resolution of indifference by risk neutral subjects.

6. Market Treatment

The preference reversal phenomenon has troubling implications for the applicability of expected utility theory to non-market, non-repetitive choice and valuation decisions. But the central concern of economics is market behavior, much of which is repetitive decision-making. Cox and Grether (1996) examined preference reversals in paired market and non-market environments with choice and valuation response modes. Results from some of their treatments are reported in Figure 6 for comparison with those of Grether and Plott. First note that the results from Cox and Grether BDM 1M are quite similar to those for Grether and Plott 1M. Thus, Cox and Grether's computerized experiments with the BDM mechanism involving monetary payoffs after each decision replicated Grether and Plott's manual experiments with BDM involving

random selection of one decision for monetary payoff. Cox and Grether SPA 1M is the first round of market experiments in which the second-price sealed-bid auction is the valuation task used to elicit selling prices rather than the BDM mechanism. The preference reversals for SPA 1M are comparable to BDM 1M; hence the preference reversal phenomenon is robust to the market environment. Cox and Grether SPA 5M is the fifth round of experiments with the second-price auction. Here, the results are very different from those for BPA 1M and SPA 1M. In SPA 5M, the frequency of reversals has markedly decreased and, more importantly, the asymmetry of reversals has disappeared.

The preference reversal phenomenon is not robust to five repetitions in the second-price sealed-bid auction market environment with monetary incentives for the subjects.

Acknowledgement

Todd Swarthout provided valuable assistance by preparing the table and figures.

References

Becker, G. M., Morris H. DeGroot and Jacob Marshak. “Measuring Utility by a Single-Response Sequential Method,” *Behavioral Science*, 9, July 1964, pp. 226-32.

Cox, James C. and Seth Epstein. “Preference Reversals Without the Independence Axiom,” *American Economic Review*, 79, June 1989, pp. 408-26.

Cox, James C. and David M. Grether. “The Preference Reversal Phenomenon: Response Mode, Markets, and Incentives,” *Economic Theory*, 7, no. 3, 1996, pp. 381-405.

Grether, David M. and Charles R. Plott. "Economic Theory of Choice and the Preference Reversal Phenomenon," *American Economic Review*, 69, Sept. 1979, pp. 623-38.

Holt, Charles. "Preference Reversals and the Independence Axiom," *American Economic Review*, 76, June 1986, pp. 508-515.

Karni, Edi and Zvi Safra. “‘Preference Reversal’ and the Observability of Preferences by Experimental Methods,” *Econometrica*, 55, May 1987, pp. 675-85.

Lichtenstein, Sarah and Paul Slovic. “Reversals of Preference Between Bids and Choices in Gambling Situations,” *Journal of Experimental Psychology*, 89, Jan. 1971, pp. 46-55.

Lichtenstein, Sarah and Paul Slovic. “Response-induced Reversals of Preference in Gambling: An Extended Replication in Las Vegas,” *Journal of Experimental Psychology*, 101, Nov. 1973 pp. 16-20.

Tversky, Amos, Paul Slovic, and Daniel Kahneman. “The Causes of Preference Reversal,” *American Economic Review*, 80, March 1990, pp. 204-17.

Figure Legends

Figure 1. A preference reversal occurs when a subject places a lower selling price on the lottery that he/she chooses than on the other lottery in the pair. Using data from the seminal experiments, this figure shows the percentage of choices of P bets for which the paired $ bets had higher selling prices ("predicted reversals," shown in blue) and the percentage of choices of $ bets for which the paired P bets had higher selling prices ("unpredicted reversals," shown in red). Lichtenstein and Slovic III is a monetary payoff experiment with psychology students as subjects. Lichtenstein and Slovic P is a monetary payoff experiment with positive expected payoff gambles run on the floor of a Las Vegas casino with adult gamblers as subjects. Grether and Plott 1H is a hypothetical payoff experiment with undergraduate students as subjects. Grether and Plott 1M is a monetary payoff experiment with undergraduate students as subjects.

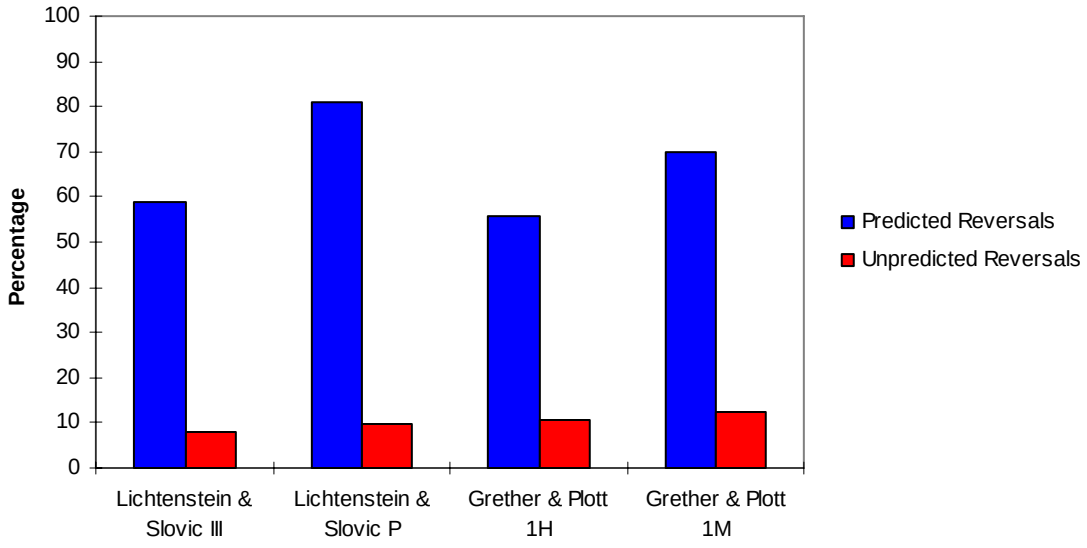

Figure 2. This figure compares results from Grether and Plott's experiments, that require the independence axiom for interpretation, with results from the ordinal pricing experiments of Cox and Epstein and Tversky, et al. that do not require this axiom to interpret the data as preference reversals. The Cox and Epstein experiment produced about the same rate of reversals as the Grether and Plott experiment (1M) with monetary payoffs but the latter did not replicate the asymmetry of reversals. The Tversky, et al. experiment with hypothetical payoffs replicated both the rate and asymmetry of the preference reversals in the Grether and Plott experiment (1H) with hypothetical payoffs. Thus preference reversals cannot be attributed to violations of the independence axiom of expected utility theory.

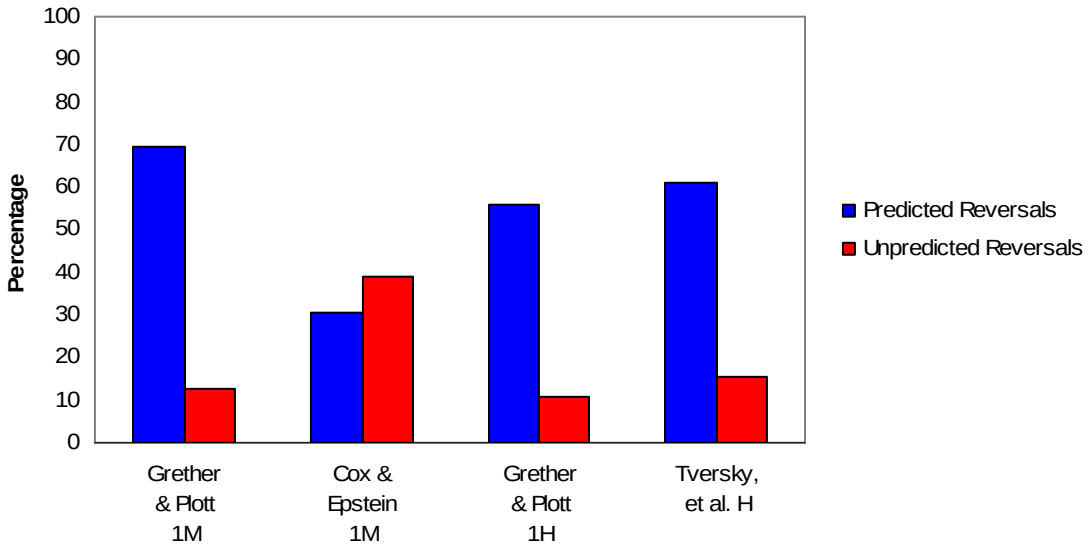

Figure 3. Results from some preference reversal experiments have been robust to use of hypothetical or monetary payoffs. This figure reports results from an experiment in which monetary payoffs produced fundamentally different preference reversals than did hypothetical payoffs. BDM 1H (respectively, M) shows the preference reversals that occurred with selling prices produced by the first repetition of the Becker-DeGroot-Marshak mechanism with hypothetical (respectively, monetary) incentives. ECA 5H (respectively, M) shows the preference reversals that occurred with selling prices produced by the fifth repetition of a sequential choice task using the English clock auction with hypothetical (respectively, monetary) payoffs.

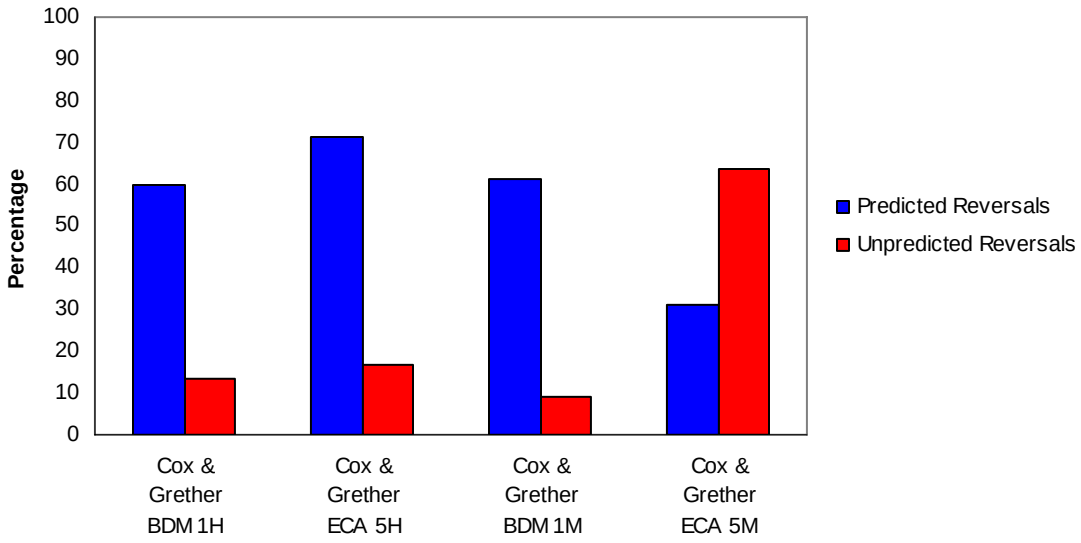

Figure 4. Tversky, et al. and Cox and Grether report results from experiments with designs that can discriminate between violations of transitivity and other causes of preference reversals. This figure reports the rates of intransitivity (IT), predicted reversals (PR), and unpredicted reversals (UR) from some of the experiments in these two papers. Note that the rates of intransitivity are far lower than the overall rates of preference reversal. These experiments support the interpretation of the preference reversal phenomenon as a response mode effect, not a result of preference intransitivity.

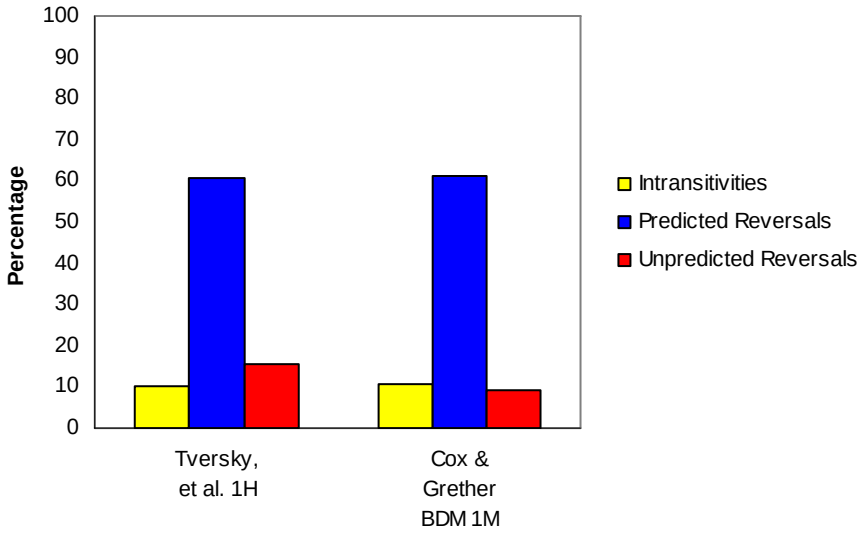

Figure 5. Most preference reversal experiments used pairs of P and $ bets with approximately equal expected values. Thus it is necessary to ascertain whether preference "reversals" simply reflect some convenient rule used by risk neutral subjects to resolve indifference. This figure reports results from Cox and Epstein's experiment, 1M using the gambles in Table 1, and results from their experiment 2M, using bet pairs with a 50% difference in expected payoff. The ordinal pricing task used in these experiments did not produce the asymmetric pattern of reversals characteristic of preference reversal experiments; hence the results are called choice reversals. The observed choice reversals in the two experiments are similar; therefore these results indicate that such reversals do not result from the resolution of indifference by risk neutral subjects.

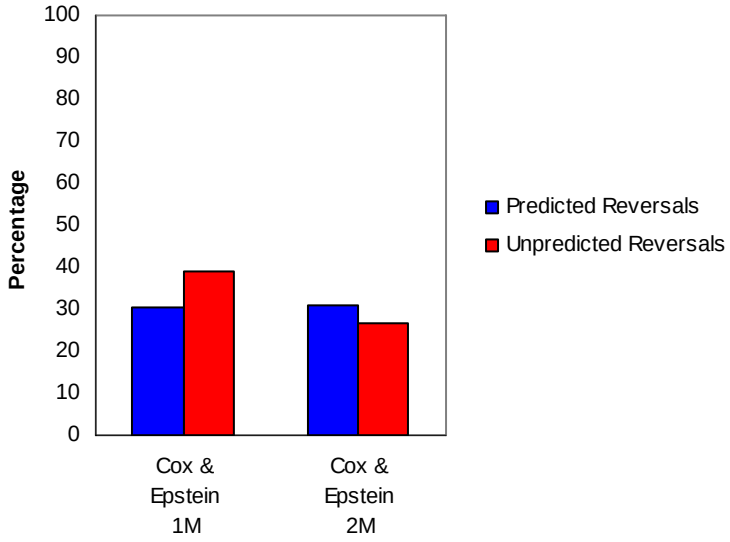

Figure 6. This figure presents comparisons of preference reversals from Grether and Plott's monetary payoff experiment, using the BDM mechanism, with Cox and Grether's monetary payoff experiments using the BDM mechanism and the second price auction. The first round of responses in the second price auction experiment, SPA 1M, produce about the same level and asymmetric pattern of reversals as in both BDM experiments. But by the fifth round of responses in the second price auction experiment, SPA 5M, the frequency of reversals has markedly decreased and, more importantly, the asymmetry of reversals has disappeared. Thus the preference reversal phenomenon is not robust to five repetitions in the second-price sealed-bid auction market environment with monetary incentives for the subjects.

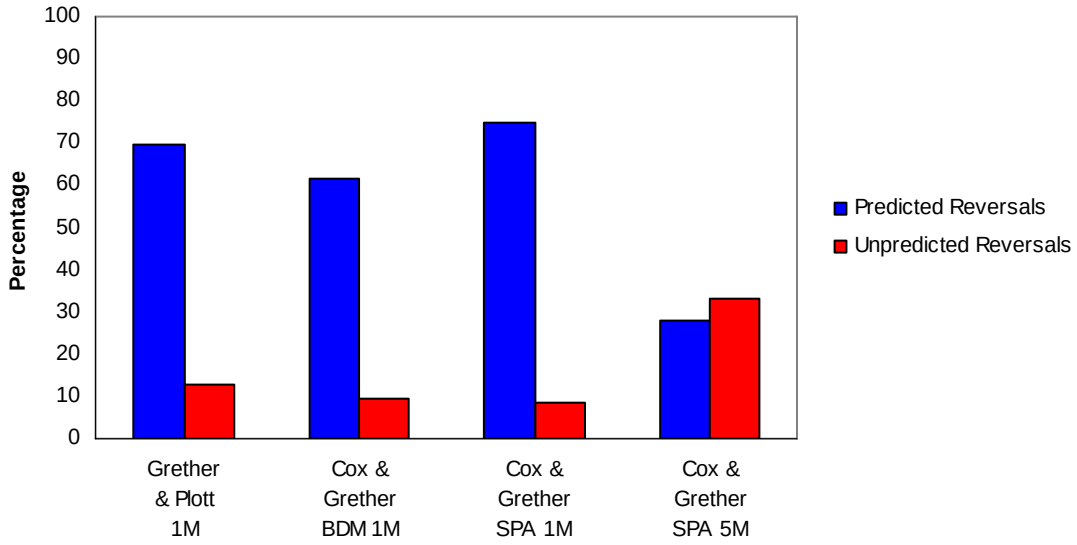

Table 1. Typical Pairs of Binary Lotteries Used in Experiments.
A preference reversal experiment involves choice and valuation responses, usually over pairs of simple two-outcome lotteries such as the ones in this table. Each lottery pair contains a "P bet," with a relatively high probability of a relatively low win state payoff, and a "$ bet," with a relatively low probability of a relatively high win state payoff.

Pairs	Type	Probability of Winning	Amount if Win	Amount if Lose	Expected Value
1	P	35/36	$4.00	-$1.00	3.86
	$	11/36	$16.00	-$1.50	3.85
2	P	29/36	$2.00	-$1.00	1.42
	$	7/36	$9.00	-$0.50	1.35
3	P	34/36	$3.00	-$2.00	2.72
	$	18/36	$6.50	-$1.00	2.75
4	P	32/36	$4.00	-$0.50	3.50
	$	4/36	$40.00	-$1.00	3.56
5	P	34/36	$2.50	-$0.50	2.33
	$	14/36	$8.50	-$1.50	2.39
6	P	33/36	$2.00	-$2.00	1.67
	$	18/36	$5.00	-$1.50	1.75