Economics and Philosophy, **18** (2002) 287–302

LEVI'S ACCOUNT OF PREFERENCE REVERSALS

Erik Angner
University of Pittsburgh

Abstract

This paper argues that Isaac Levi's account of preference reversals is only a limited success. Levi succeeds in showing that an agent acting in accord with his theory may exhibit reversals. Nevertheless, the specific account that Levi presents in order to accommodate the behavior of experimental subjects appears to be disconfirmed by available evidence.

1. INTRODUCTION

In an early volume of this journal, Isaac Levi (1986b) argues that his theory of decision can account for a wide range of anomalous experimental results, including the Allais paradox, the Ellsberg paradox, and the preference reversal phenomenon. By claiming that the theory, which is primarily intended as a normative theory, can accommodate the paradoxical behavior patterns exhibited in empirical studies, Levi wants to rescue the notion that choices made by experimental subjects are perfectly rational. While more recent volumes of *Economics and Philosophy* have seen a rather intense discussion about Levi's account of the Allais and Ellsberg paradoxes (see Maher, 1989; Levi, 1989; Maher and Kashima, 1991; Levi, 1991), his treatment of reversals has received little attention.

In the present paper, I assess Levi's account of preference reversals. Preference reversal occurs when an agent chooses one alternative over another, yet prices the rejected alternative higher. According to Levi, such choice patterns can be accommodated by assuming that agents have indeterminate probabilities or utilities, and that they resort to other considerations when expected utility reasoning fails. Specifically, Levi

Thanks to Horacio Arló-Costa, Isaac Levi and Teddy Seidenfeld for helpful criticism on earlier drafts. I also want to register my gratitude to the late Tamara Horowitz for enthusiastically encouraging me to explore these issues in the first place. Errors, obviously, remain my own.

https://doi.org/10.1017/S0266267102002067 Published online by Cambridge University Press

suggests that agents who exhibit reversals choose bets so as to minimize losses. Here, I evaluate his proposal by confronting it with new evidence gathered from various published sources.

At the end of the day, I claim, Levi's treatment of reversals is only a limited success. It is true that the theory permits rational agents to exhibit preference reversals. Thus, Levi is right to assert that on his theory, preference reversal is not necessarily irrational. Still, the specific account he defends – according to which agents who reverse their preferences choose bets so as to minimize losses – appears to be disconfirmed by available evidence. In fact, it is compatible with no more than about one half of recorded reversals. I also consider a number of other accounts consistent with Levi's theory, but conclude that no one of them is likely to accommodate the behavior of experimental subjects.

2. THE PREFERENCE REVERSAL PHENOMENON

The preference reversal phenomenon was first demonstrated in the early seventies.[1] Lichtenstein and Slovic (1971) presented their subjects with pairs of gambles, or ''bets'', of roughly equal expected value. One of the bets, which they called the ''P-bet'', had a higher probability of winning a smaller amount; the other, which they called the ''\$-bet'', had a lower probability of winning a larger amount. First, the subjects were asked which bet they would prefer to play. Second, they were presented with the bets one by one and prompted to place a price on each. In Experiment I, subjects were asked to name the minimum price at which they would be willing to sell the right to play the bet. In Experiment II, subjects were asked to name the maximum price they would be willing to pay in order to buy the right to play the bet. In Experiment III, a device known as the Becker–DeGroot–Marschak mechanism was used to elicit prices,[2] and subjects played for real money. Lichtenstein and Slovic predicted that when the P-bet was chosen, the \$-bet would often be priced higher. Such reversals were termed ''predicted''. Reversals in which the \$-bet is chosen but the P-bet is priced higher were termed ''unpredicted''.

The results conformed to the experimenters' expectations. In Experiment I, the responses of 173 subjects to six pairs of bets were analyzed. The proportion of predicted reversal was high; 73% priced the \$-bet higher every time they chose the P-bet. Unpredicted reversals were uncommon, however; 83% of the subjects never exhibited an unpredicted

[1] Useful overviews of the research, and more references, can be found in Slovic and Lichtenstein (1983), Tversky *et al.* (1990), Tversky and Thaler (1990), Hausman (1992, Chapter 13), Camerer (1995, Section III.I), and Guala (2000).

[2] The mechanism, which was introduced in Becker *et al.* (1964), is explained in Lichtenstein and Slovic (1971, pp. 51–2).

reversal. In Experiment II, 74 subjects was presented with 49 pairs of bets. The proportion of predicted reversals was somewhat lower than in Experiment I, and the proportion of unpredicted reversals somewhat higher. In Experiment III, 14 subjects were presented with six pairs of bets. Eleven subjects exhibited reversals, and six reversed their preference every time they chose the P-bet. Meanwhile, there were few unpredicted reversals. Thus, a significant proportion of subjects exhibited reversals, and there was a clear asymmetry in the kind of reversals exhibited; predicted reversals were much more common than unpredicted ones.

The preference reversal phenomenon has been demonstrated in a number of other studies. Lichtenstein and Slovic (1973) replicated the results at a Las Vegas casino, where a croupier served as experimenter, professional gamblers served as subjects, and winnings and losses were paid in real money. The economists Grether and Plott (1979) set out with the explicit intention to ''discredit the psychologists'' works as applied to economics'' (p. 623), but to their evident surprise merely confirmed the original results. Others have demonstrated reversals in different cultures, when subjects are more carefully informed, when incentives are increased, when expected values of the gambles are negative as well as positive, and when diverse elicitation mechanisms are employed. Though some studies succeed in decreasing the frequency of reversals (Bostic, Herrnstein and Luce, 1990), or even eliminating them (Chu and Chu, 1990), the phenomenon appears on the whole both common and robust.

Preference reversals are often considered evidence that expected utility theory is descriptively inadequate, and that irrationality is systematic and widespread. According to Slovic, Griffin and Tversky (1990), for example, phenomena like preference reversals ''represent deep and sweeping violations of classical rationality'', and therefore ''it may not be possible to construct a theory of choice that is both normatively acceptable and descriptively adequate'' (p. 26). Some have gone so far as to argue that subjects in these studies do not act on the basis of underlying beliefs and preferences at all (see e.g., Kahneman *et al*., 1999).

3. LEVI'S THEORY OF DECISION

Levi agrees that expected utility theory is descriptively inadequate, but resists drawing such radical conclusions. He aspires to show that his theory of decision, which is a generalization of Bayesian decision theory, ''can give account of responses of experimental subjects in a systematic manner'' (1997 [1986b], p. 185).[3] Since Levi's theory is a normative one,

[3] I will consistently refer to the reprinted version of Levi's (1986b) paper, Chapter 10 of

this amounts to showing that subjects may be perfectly rational, and that their choices are consistent with having beliefs and preferences. In this section I give a brief account of Levi's theory of decision.

According to ordinary Bayesian expected utility theory, each agent has a unique probability distribution over states of the world, and a unique utility function over outcomes.[4] The probability distribution reflects the subjective probabilities that the agent assigns to the occurrence of each state, and the utility function reflects the agent's valuations of the various outcomes. Given a probability distribution and a utility function, both with the appropriate properties, the agent calculates the expected utility of each alternative course of action. She then chooses the optimal alternative with respect to expected utility, if there is a uniquely optimal one, or one of the optimal alternatives, if there are several.

On Levi's more general theory, each agent has a set of probability distributions, called *permissible* probability distributions, and a set of utility functions, called *permissible* utility functions. Indeterminacy in probability and utility, on Levi's account, reflects doubt or suspense with respect to belief and value respectively (1997, pp. xi, 185). In Levi's terms, an alternative *A* is E-*admissible* if and only if there is some permissible probability distribution and some permissible utility function such that *A* is optimal relative to the probability distribution and the utility function. When making a decision, the rational agent first determines which alternatives are *E*-admissible. If there is only one *E*-admissible alternative, the agent chooses it.

In cases when there are several *E*-admissible alternatives, the agent uses some other criterion to break ties.[5] For example, she may choose an alternative that maximizes security according to a maximin criterion, or what Levi refers to as a lexicographical maximin (leximin) criterion. That is, if the worst possible outcome of the one is better than the worst possible outcome of the other, the former will be chosen. If the worst possible outcomes of the two alternatives are equivalent, the agent will consider the second worst outcome, etc.[6] An alternative is called *S-admissible* if and only if it is optimal according to the security criterion among the *E*-admissible options. There are many other criteria that can be used to break ties, however. The agent may choose an alternative that

Covenant of Reason (1997, pp. 185–216). In the more recent version, a small number of errors have been corrected.

4 This paragraph and the next draw on Levi's "On indeterminate probabilities" (1997 [1974], Chapter 6).

5 This paragraph draws on *The Enterprise of Knowledge* (1980, pp. 156–62).

6 What alternative is favored by security criteria may depend on how the space of possibilities is partitioned. If so, the choice of partition is up to the agent (see 1980, pp. 156–61).

https://doi.org/10.1017/S0266267102002067 Published online by Cambridge University Press

maximizes minimum expected value. Indeed, she may even choose randomly or flippantly among the *E*-admissible options. What secondary criterion to use is up to the agent. On Levi's view, rationality does not favor one over the other, and the choice of secondary criterion should be seen as a value commitment on the part of the agent.[7]

4. LEVI'S ACCOUNT OF REVERSALS

Although Levi's theory is mainly intended as a prescriptive theory (see e.g., 1997, p. 185), the issue concerning its descriptive adequacy – which I take to be the main issue raised in Levi's paper – is interesting in its own right. In the present section, I estimate the extent to which the theory can accommodate the data on preference reversal.

According to Levi's account, people exhibit reversals because both bets are *E*-admissible, and subjects invoke security considerations so as to minimize losses. To illustrate, Levi invites us to consider two gambles. One involves a 90% chance of winning a million dollars, and the other involves an 80% chance of winning five million dollars (1997, p. 211). Formally, the two gambles can be represented as follows:

G_1: (.9, \$1,000,000; .1, \$0)
G_2: (.8, \$5,000,000; .2, \$0)

Next, Levi asks us to consider an agent whose set of permissible utility functions is the convex hull – i.e., the set of all weighted averages – of the following two functions (1997, p. 211):

$U_1(x) = \ln(x+1)$
$U_2(x) = \ln(0.1x+1)$

Given the utility functions, we can calculate expected utilities as well as certainty equivalents for the gambles, where the certainty equivalent of a gamble is defined as that amount of money *X* such that the agent is indifferent between receiving the gamble and receiving *X* for sure. Tables 1a and 1b give the values for expected utilities and certainty equivalents (see 1997, p. 211).

Table 1a Expected utilities

	U_1	U_2
G_1	12.44	10.36
G_2	12.34	10.49

Table 1b Certainty equivalents

	U_1	U_2
G_1	\$252,000	\$310,000
G_2	\$230,000	\$359,000

[7] Though Levi's earlier work (e.g., 1974) assumes that agents are rationally required to break ties using a security criterion, his later writings (e.g., 1986a) permit agents to use any criterion they like.

Which gamble does our agent choose? Levi writes: ''Both options are *E*-admissible because they come out optimal according to some permissible utility function in the set. Leximin favors the gamble on the $5 million (gamble 2)'' (1997, p. 211). Since no gamble is optimal according to all permissible utility functions, the agent resorts to security reasoning. As leximin favors gamble 2, the agent chooses the 5 million gamble.[8]

Which gamble does the agent price higher? According to Levi's account:

> [If] asked for the smallest price at which he would sell gamble 1, the decision maker should answer $252,000, and for gamble 2 he should answer $230,000. For lower prices, retaining the gamble would be uniquely *E*-admissible. For prices of $252,000 and $230,000 respectively, both retaining and selling becomes *E*-admissible and selling is the maximin solution. (1997, p. 211)

The argument can be understood as follows. Suppose the agent has to choose between the right to play gamble 1 and some fixed amount *X* of money. Now, there are three cases. In the first case, *X* is higher than the highest certainty equivalent assigned to the gamble, in this case $310,000. Here, the money is uniquely *E*-admissible and, therefore, rationally required. In the second case, *X* is lower than the lowest certainty equivalent assigned to the gamble, in this case $252,000. Now, the gamble is uniquely *E*-admissible. In the third case, *X* is no lower than the lowest but no higher than the highest certainty equivalent assigned to the gamble. Thus, *X* is higher than or equal to $252,000, but lower than or equal to $310,000. In this last case, both the money and the gamble are *E*-admissible. However, the fixed amount would be favored by security considerations, and so the agent should take the money. It follows that the agent should choose the gamble if *X* is lower than $252,000, and the money otherwise. That is, she should price gamble 1 at (approximately) $252,000. In general, the agent should choose the gamble so long as *X* is no higher than the lowest certainty equivalent assigned to the gamble. Accordingly, the agent should price gamble 2 at $230,000. The same considerations apply when the agent is asked for the highest price at which she would buy the gambles. Thus, whether asked for selling prices or buying prices, the agent should price gamble 1 higher.

Levi also considers a modified pair of gambles, in which the agent is charged $10 for playing gamble 1 (1997, p. 211).[9] The modified pair of gambles can be described as follows:

[8] Levi implicitly assumes that the agent adopts probabilities equal to those provided by the experimenter. I follow Levi in this regard.

[9] Actually, the text tells us to charge $10 for gamble 2 rather than gamble 1 (1997, p. 211). This suggestion is difficult to reconcile with Levi's subsequent assertion that leximin still favors gamble 2, but the figures have been accidentally switched. Levi in fact intended the penalty to apply to gamble 1 (personal communication).

G_1*: (.9, $999,990; .1, $-10)
G_2: (.8, $5,000,000; .2, $0)

Now, unfortunately, the expected utility for gamble 1 is undefined, as ln(x) is undefined for non-positive numbers. However, this problem is easily solved, if we redefine the utility functions slightly:

$U_1(x) = \ln(x+1)$ for $x \geq 0$; $x/100$ otherwise
$U_2(x) = \ln(0.1x+1)$ for $x \geq 0$; $x/100$ otherwise

Given these small modifications, the same argument applies; in fact, the figures remain roughly the same. The agent should choose gamble 2, but price gamble 1* (i.e., the modified gamble 1) higher.

Even so, however, the account does not quite accommodate the behavior of most experimental subjects. In the original experiments of Lichtenstein and Slovic, for example, a majority of subjects who exhibited reversals chose the P-bet but priced the $-bet higher. That is, they exhibited what Lichtenstein and Slovic (1971) called "predicted" reversals. In Levi's example, gamble 1 is the P-bet and gamble 2 the $-bet. Since Levi's agent chooses gamble 2, she exhibits a case of unpredicted reversal.

It is possible to modify Levi's example in order to accommodate predicted reversals as well. Suppose we modify gamble 2 instead of gamble 1, so as to produce the following two gambles:

G_1: (.9, $1,000,000; .1, $0)
G_2*: (.8, $4,999,990; .2, $-10)

Suppose, moreover, that the agent's set of utility functions is the convex hull of the following two functions:

$U_1(x) = x/100{,}000{,}000 + 99/100$ for $x \geq 1{,}000{,}000$;
$x/1{,}000{,}000$ for $0 \leq x < 1{,}000{,}000$; and
$x/100$ for $x < 0$.
$U_2(x) = x/2{,}000{,}000 + 1/2$ for $x \geq 500{,}000$;
$3x/2{,}000{,}000$ for $0 \leq x < 500{,}000$; and
$x/100$ for $x < 0$.

Given these utility functions, expected utilities and certainty equivalents are given by Tables 2a and 2b.

Table 2a Expected utilities

	U_1	U_2
G_1	.90	.90
G_2	.81	2.4

Table 2b Certainty equivalents

	U_1	U_2
G_1	$900,000	$800,000
G_2	$812,000	$3,760,000

Again, both gambles are *E*-admissible. Since leximin favors gamble 1, the agent chooses it. Since the agent prices gambles at a price corresponding to the lowest certainty equivalent assigned to it, gamble 1 should be priced at $800,000, and the modified gamble 2 at $812,000. While the agent chooses gamble 1, he prices gamble 2 higher. Thus, Levi's theory can accommodate predicted reversals as well.

In sum, Levi is correct to say that an agent acting in accordance with his theory may exhibit preference reversals. Note that these agents strictly speaking neither "reverse" nor misstate their preferences, but make choices that reflect security considerations as well as expected utility calculations.[10] Levi concludes:

> As far as I can see, the situation I have just described reproduces the phenomenon reported by Grether and Plott (1979) and others (Reilly, 1982) on so-called preference reversal. On my account, of course, there is no preference reversal – only indeterminacy in preferences. (1997, p. 212)

5. AN EMPIRICAL TEST OF LEVI'S ACCOUNT

Levi's last, rather upbeat paragraph has a footnote in which he adds that he cannot conclusively confirm the account. He writes:

> I cannot cite decisive evidence that my scheme models the Grether-Plott phenomenon because the reports of the experimental design and results provided by Grether and Plott do not furnish information relevant to this matter. They invite experimental subjects to compare pairs of bets of types 1 and 2 (so-called P-bets and $-bets). Counter to the impression given on p. 623 of Grether and Plott (1979), each bet in a given pair always incurs a risk of loss. However, in some pairs, the losses are greater for P-bets and in some pairs, the losses are greater for $-bets. Grether and Plott do not report, however, the percentages of experimental subjects who choose members of pairs in a way that minimizes such losses (and hence maximizes security). Hence, although their results and those of Reilly appear compatible with my model, there is an unsettled empirical question about their experiments, which is relevant to the empirical adequacy of the model I have proposed. (1997, p. 212)

In effect, this passage suggests an empirical test of Levi's account – or "model" – according to which subjects who exhibit reversals choose bets so as to minimize their losses. On this account, subjects look at the worst possible outcome for each bet, and choose between bets accordingly. Let us call a bet that minimizes losses the *safe* bet, and a reversal in which the agent chooses the safe bet a *safety* reversal. Then, Levi's account implies that all reversals are safety reversals. The question of whether this implication is true is an empirical one, and can be settled by experiment.

[10] I will continue to use "preference reversal" to refer to the characteristic choice pattern exhibited in these studies. I trust this will not cause confusion.

If actual subjects who reverse their preferences choose the safe bet, Levi's account is confirmed. In what follows, I will argue that the data in fact do not confirm it, and that no more than about half of recorded reversals can be accounted for in this way.

All told, I have found only two published articles that report the data required to compute the frequency of subjects who minimize losses. Reilly (1982), quoted in Levi's original article, in fact, provides some relevant data. In the second stage of his experiment, Reilly presented his subjects with six pairs of bets. Sometimes the $-bet is the safe one, sometimes the P-bet is, and it is unclear what proportion of subjects who reversed their preferences chose the safe bet. However, Reilly's paper also presents data from pairs 2 and 6 only, in which it so happens that the $-bet is the safe bet (1982, pp. 580–81). For example, consider pair 6: P-bet (34/36, $3.00; 2/36, $-2.00) vs. $-bet (18/36, $6.50; 18/36, -1.00). In this case, Levi's account implies that all agents who reverse their preferences over these two bets should choose the $-bet but price the P-bet higher.

Reilly had two groups of subjects, with 45 and 41 members respectively. Each subject made two choices. In the first group, there were 18 cases of subjects who chose the P-bet but priced the $-bet higher. At the same time, there were 13 cases of subjects who chose the $-bet but priced the P-bet higher. Thus, there were 31 reversals; of those only 42% were of the kind implied by Levi's account. In the second group, there were 12 cases of subjects who chose the P-bet but priced the $-bet higher. Meanwhile, there were 11 cases of subjects who chose the $-bet but priced the P-bet higher. Thus, there were 23 reversals; of those only 48% were of the kind implied by Levi's account. This is a slight improvement with respect to the first group, but still a far cry from the 100% we would expect if Levi's account adequately describes the behavior of the subjects.

Clearly, we should not expect any study to provide data that are perfectly compatible with the theory under investigation. We should expect some subjects to misunderstand the question, or to misstate their answers, even though it appears prima facie unlikely that subjects would do so on a large scale in this case, since Reilly went to great lengths to assure that subjects understood the questions and that payoffs were clearly presented (see 1982, Appendix, pp. 582–84). The question we should ask is which competing hypothesis is most closely compatible with the data. What would be an appropriate alternative hypothesis in this case? Let me suggest, as an alternative, the hypothesis that reversals are randomly distributed over safety and non-safety reversals. This would be true if subjects chose bets on the basis of features uncorrelated with safety. If my alternative hypothesis is true, we should expect approximately 50% of each. This prediction is rather close to the actual results of 42% and 48%, and significantly better than Levi's suggestion of 100%.

MacDonald *et al*. (1992), who explored reversals in the domain of losses, report results similar to Reilly's (1982). In MacDonald *et al*.'s stage 2 experiments, given the partition, security considerations would favor gambles A, C, and E (A′, C′, and E′). Here, a typical pair of bets is their bet C (16/32, $-8.00; 16/32 $-1.00) and D (12/32, $-14.00; 20/32, $-1.00), where C minimizes losses. Thus, agents who chose A, C, and E (A′, C′, and E′) chose the safe bet. According to Table 3 there were 55 such reversals (MacDonald *et al*. 1992, p. 123). Agents who chose B, D, and F, (B′, D′, and F′), however, chose the non-safe bet. According to the table, there were 33 such reversals (p. 123). All in all, there were 133 reversals in which the subject chose a safe bet, and 123 reversals in which the agent chose a non-safe bet. Thus, the proportion of safety reversals was 52%. Again, this result appears far too low to be accounted for by subjects' mistakes; again, the result is compatible with the hypothesis that reversals are randomly distributed over safety and non-safety reversals.

In summary, it appears that Levi's empirical test fails. Subjects who exhibit reversals do not in general minimize losses when choosing between bets. In the three studies reviewed here, the proportion of safety reversals is distributed rather closely around 50% (see Table 3). In contrast, Levi's account would make us expect a figure close to 100%. The data are most compatible with the hypothesis that reversals are randomly distributed over safety and non-safety reversals.[11]

Table 3 Summary of results

Study (year)	Total number of reversals	Number of safety reversals	Number of non-safety reversals	Proportion of safety reversals
Reilly (1982) group 1	31	13	18	42%
Reilly (1982) group 2	23	11	12	48%
MacDonald *et al*. (1992)	256	133	123	52%
Total:	310	157	153	51%

[11] In the light of the alternative hypothesis, none of the results are statistically significant even at the .32 level ($Mean_1 = 15.5$, $SD_1 = 2.8$; $Mean_2 = 11.5$, $SD_2 = 2.4$; $Mean_3 = 128$, $SD_3 = 8.0$; $Mean_4 = 155$, $SD_4 = 8.8$).

6. DISCUSSION

Available evidence, then, appears to disconfirm the account that Levi proposes. It is important to note, however, that this conclusion does not by itself show that his theory cannot account for the behavior of experimental subjects. For, there may be other accounts consistent with the theory that produce the same patterns of choice. In this section, I consider other manners in which one might model reversals within the framework of Levi's theory.

Levi is right, I believe, to suggest that subjects need to invoke secondary criteria in order to exhibit reversals in pair-wise choice. For a proof by contradiction, suppose that an agent who exhibits a reversal chooses one of the gambles because it was favored by expected utility considerations, that is, because it was uniquely *E*-admissible. Without loss of generality, we can let gamble 1 be the chosen gamble, and gamble 2 be the one that was priced higher. Let the utilities and certainty equivalents of the two gambles be defined as in tables 4a and 4b (as before, the agent's set of utility functions is the convex hull of U_1 and U_2). Since G_1 is uniquely *E*-admissible, it must be the case that $U_i(G_1) > U_i(G_2)$ for all permissible utility functions U_i. Thus, in particular, $a > c$ and $b > d$. Since G_1 is priced lower, the lowest certainty equivalent associated with G_1 must be lower than that associated with G_2. Hence, either $A < C$ or $B < D$. Monetary consistency (or monotonicity), the assumption that more money is preferred to less, implies that $U_1(A) < U_1(C)$ or $U_2(B) < U_2(D)$.[12] But that is to say that $a < c$ or $b < d$, by the definition of certainty equivalents, which is a contradiction. Thus, an agent acting in accordance with Levi's theory can exhibit reversals only if her utilities and probabilities are sufficiently indeterminate to warrant the use of secondary criteria.

Table 4a Expected utilities

	U_1	U_2
G_1	*a*	*b*
G_2	*c*	*d*

Table 4b Certainty equivalents

	U_1	U_2
G_1	*A*	*B*
G_2	*C*	*D*

Even so, there are many different accounts, consistent with Levi's theory, which could account for reversals. One such account assumes that agents choose flippantly or randomly between *E*-admissible options. As indicated above, Levi's theory permits agents to act in this way. Yet, this account is hard to reconcile with the experimental data in the

[12] Monetary consistency says that for fixed amounts *X* and *Y*, if $X > Y$ then $U_i(X) > U_i(Y)$ (Tversky *et al.*, 1990, p. 205). I take monetary consistency to be an uncontroversial assumption in these contexts; at any rate, Levi himself adopts it (1989, p. 85).

literature. First, the account does not explain why predicted reversals are more common than unpredicted ones. Second, it does not explain why some subjects are so adamant about their choices. In debriefing sessions after their third experiment, Lichtenstein and Slovic (1971) informed subjects of the fact that they had exhibited reversals, that their behavior could be characterized as ''inconsistent'' and ''irrational'', and that they would be vulnerable to a money-pump game. That is to say, if the agents persist in their choices, they can be guided through a series of exchanges that leave them strictly worse off. Many subjects were quite determined; out of eleven subjects, in spite of rather heavy pressure to change their responses, three changed only after the money-pump game had been explained to them, and two refused to change at all (p. 53). One of the interviews ended with the following exchange. Experimenter: ''Well, I think I've pushed you as far as I know how to push you short of actually insulting you.'' Subject, who just lost part of his winnings by being turned into a money pump: ''That's right.'' (Lichtenstein and Slovic, transcript.) The fact that subjects feel so strongly about their choices disconfirms the hypothesis that subjects choose randomly or flippantly.

An alternative account assumes that agents choose so as to maximize minimum expected value. However, this account too is disconfirmed by the data. In the case of both pairs of gambles from Reilly (1982), the $-bet has a higher expected value than the P-bet (p. 580). Thus, maximizing minimum expected value would favor the $-bet, but as we have seen (see Table 3), less than half recorded reversals are of this kind. Moreover, when Reilly explained the concept of expected value to his subjects, and provided them with information about the expected value of each bet, the number of $-bets chosen actually decreased. While the first group did not receive this extra information, the second group did (p. 577), and the proportion of $-bets chosen went from 48% to 42% (see Table 3).

Another account assumes that agents apply a leximin criterion, but that they partition the space of possibilities in some other way. In his account of reversals, Levi implicitly assumes that agents partition the space of possibilities in the straightforward manner – ''either I win or I lose''. As we have seen above, however, Levi permits agents to assess security levels as they wish. It is possible, at least in theory, that there is another partition of the space of possibilities such that security considerations would favor only one bet (perhaps, even, the P-bet). If we could find such a partition, we could rely on Levi's theory to explain reversals after all. In actuality, however, it appears difficult to find such a partition; at the very least, I have not succeeded.

In his reply to Maher (1989), Levi (1989) suggests two other ways in which one might model individual behavior in order to accommodate robust empirical results. First, he indicates that agents may neglect small differences in payoffs when assessing security levels (p. 82). Thus, for the purposes of calculating security levels, the agent may consider a payoff of, say, \$0.10 equivalent to a payoff of zero. Second, Levi suggests that agents may rely on a criterion of *probabilistic* S*-admissibility* instead of standard *S*-admissibility (pp. 83–4). The set of probabilistically *S*-admissible options consists of ''[those] *S*-admissible options whose security levels have minimum probability or probability negligibly different from the minimum probability'' (p. 83). So, if two gambles are *S*-admissible, but gamble 1 yields the worst payoff with probability .01, while gamble 2 yields it with probability .50, only the former is probabilistically *S*-admissible.

However, these proposals do not seem sufficiently radical to accommodate the data, even though it would be up to the agent to decide what constitutes a small difference. In the pair of gambles from MacDonald *et al*. (1992) considered above, the one involves a risk of losing \$14 and the other a risk of losing eight dollars. In the context of such gambles, it seems unlikely that subjects would ignore the difference between 14 and eight. Besides, even if they did, security would not explain why many subjects choose the first gamble, since the second-worst outcome is the same in both cases, namely, losing one dollar. The second proposal is equally unlikely to accommodate the data on reversals, since it is equivalent to the original proposal when the worst payoff of one gamble is significantly lower than the worst payoff of the other (as in the gamble from MacDonald *et al*.). In the latter case, assuming that a loss of \$14 is significantly different from a loss of eight, there is only one *S*-admissible alternative.

A more promising approach is that of Martha Lacey (1996), who suggests that agents who exhibit reversals choose bets so as to maximize the probability of winning (pp. 7–8). This hypothesis would explain why predicted reversals – in which agents choose the bet with the high probability of winning a small amount – tend to be more common than unpredicted ones. Thus, Lacey's suggestion has a distinct advantage over the other accounts considered so far. However, the data she reports throw doubt on the hypothesis as a general account of reversals. Let C_P denote the price assigned to the P-bet, and $C_\$$ denote the price assigned to the \$-bet. In Lacey's experimental design, after going through the ordinary choice and pricing tasks, subjects face a four-way choice between the P-bet, the \$-bet, an amount of C_P for sure, and an amount of $C_\$$ for sure. If subjects use the security criterion Lacey suggests, those who reverse their preferences should choose $C_\$$ in the four-way choice.[13] If they choose the P-bet and exhibit reversal, $C_\$$ must be greater than C_P, and therefore by monetary consistency they prefer $C_\$$ to C_P. Also, the

[13] Here, I assume receiving $C_\$$ for sure is an *E*-admissible alternative for subjects who exhibit reversals.

probability of winning in the case of a sure amount (i.e., one) is better than the probability of winning in the case of either bet. Unfortunately, in cases when subjects exhibited reversals, no more than 50.7% went on to choose $C_\$$ in the four-way choice (1996, p. 11). Thus, much like Levi's own account, Lacey's can accommodate no more than about half of subjects' reversals.

All told, it appears that no one account consistent with Levi's theory can accommodate the behavior of experimental subjects. Even the most promising accounts are compatible with only about half of the recorded reversals. While Lacey's proposal has the distinct advantage of explaining why predicted reversals are so much more common than unpredicted reversals, it is inconsistent with almost half the reversals exhibited in her experiment. It is possible, of course, that I have overlooked *the* secondary criterion that allows us to explain subjects' behavior. Ignoring this possibility for now, however, I conclude that no one account consistent with Levi's theory is able to accommodate the behavior of experimental subjects.

7. CONCLUSION

Levi's account of reversals must be pronounced only a limited success. He is right to say that his theory can accommodate reversals, in the sense that agents acting in accord with his theory may exhibit reversals. Thus, someone who accepts Levi's theory must admit that the preference reversal phenomenon is not necessarily irrational. Nevertheless, the specific account or model that Levi suggests, according to which agents who exhibit reversals choose bets so as to minimize their losses, seems disconfirmed by the data. In fact, no more than about half of all recorded reversals can be accommodated in this way. Worse, it appears that no one alternative account does much better than Levi's original one.

This state of affairs leaves us with two possibilities. Either, subjects' choices simply fail to conform to Levi's theory. If so, the theory is descriptively inadequate and the behavior of actual subjects must be accounted for by some other theory. Or, subjects do act in accordance with the theory, but apply different secondary criteria. Most certainly, Levi's theory does not require different people to apply the exact same criterion to break ties. And if they do not use the same one, we should not expect this fact to be apparent in data reported in the aggregate. Since the disaggregated data that would allow us to explore this issue is not currently available, it remains possible that there exists an adequate account consistent with Levi's theory.

Either way, the current argument should not be taken as a critique of the normative status of Levi's theory. As indicated above, the theory is primarily intended to be a normative or prescriptive theory. In my view, nothing that has been said here threatens its normative status. Although experimental subjects in psychological experiments may fail to act in accordance with the theory, it can still serve as a correct and useful guide for someone aspiring to make more rational decisions.

REFERENCES

Becker, Gary M., Morris H. DeGroot and Jacob Marschak. 1964. Measuring utility by a single-response sequential method. *Behavioral Science*, 9:226–32

Bostic, Raphael, R. J. Herrnstein and R. Duncan Luce. 1990. The effect on the preference-reversal phenomenon of using choice indifferences. *Journal of Economic Behavior and Organization*, 13:193–212

Camerer, Colin. 1995. Individual decision making. In *The Handbook of Experimental Economics*. John H. Kagel and Alvin E. Roth (eds.). Princeton University Press

Chu, Yun-Peng and Ruey-Ling Chu. 1990. The subsidence of preference reversals in simplified and marketlike experimental settings. *American Economic Review*, 80:902–11

Grether, David M. and Charles R. Plott. 1979. Economic theory of choice and the preference reversal phenomenon. *American Economic Review*, 69:623–38

Guala, Francesco. 2000. Artefacts in experimental economics: preference reversals and the Becker–DeGroot–Marschak mechanism. *Economics and Philosophy*, 16:47–75

Hausman, Daniel M. 1992. *The Inexact and Separate Science of Economics*. Cambridge University Press

Kahneman, Daniel, Ilana Ritov and David Schkade. 1999. Economic preferences or attitude expressions? An analysis of dollar responses to public issues. *Journal of Risk and Uncertainty*, 19:203–35

Lacey, Martha. 1996. Preference reversal in non-binary choice. Carnegie Mellon University Master Thesis

Levi, Isaac. 1974. On indeterminate probabilities. *The Journal of Philosophy*, 71:391–418. Reprinted in Levi (1997), pp. 117–44

Levi, Isaac. 1980. *The Enterprise of Knowledge: An Essay on Knowledge, Credal Probability and Chance*. MIT Press

Levi, Isaac. 1986a. *Hard Choices: Decision Making Under Unresolved Conflict*. Cambridge University Press

Levi, Isaac. 1986b. The paradoxes of Allais and Ellsberg. *Economics and Philosophy*, 2:23–53. Reprinted in Levi (1997), pp. 185–216

Levi, Isaac. 1989. Reply to Maher. *Economics and Philosophy*, 5:79–90

Levi, Isaac. 1991. Reply to Maher and Kashima. *Economics and Philosophy*, 7:101–3

Levi, Isaac. 1997. *The Covenant of Reason: Rationality and the Commitments of Thought*. Cambridge University Press

Lichtenstein, Sarah and Paul Slovic. Transcript. Excerpted from a half-hour post-experimental interview with an *S* who is being questionned [sic] about the reversal inherent in his bids and choices

Lichtenstein, Sarah and Paul Slovic. 1971. Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology*, 89:46–55

Lichtenstein, Sarah and Paul Slovic. 1973. Response-induced reversals of preference in gambling decisions: an extended replication in Las Vegas. *Journal of Experimental Psychology*, 101:16–20

MacDonald, Don N., William L. Huth and Paul M. Taube. 1992. Generalized expected utility analysis and preference reversals: some initial results in the loss domain. *Journal of Economic Behavior and Organization*, 17:115–30

Maher, Patrick. 1989. Levi on the Allais and Ellsberg paradoxes. *Economics and Philosophy*, 5:69–78

Maher, Patrick and Yoshihisa Kashima. 1991. On the descriptive adequacy of Levi's decision theory. *Economics and Philosophy*, 7:93–100

Reilly, Robert J. 1982. Preference reversal: further evidence and some suggested modifications in experimental design. *American Economic Review*, 72:576–84

Slovic, Paul, Dale Griffin and Amos Tversky. 1990. Compatibility effects in judgment and choice. In *Insights in Decision Making: A tribute to Hillel J. Einhorn*, pp. 5–27. Robin Hogarth (ed.). Chicago University Press

Slovic, Paul and Sarah Lichtenstein. 1983. Preference reversals: a broader perspective. *American Economic Review*, 73:596–605

Tversky, Amos, Paul Slovic and Daniel Kahneman. 1990. The causes of preference reversal. *American Economic Review*, 80:204–17

Tversky, Amos and Richard H. Thaler. 1990. Anomalies: preference reversals. *Journal of Economic Perspectives*, 4:201–11