
AEA Journals Copyright and Reprint Permissions


Permission to make digital or hard copies of part or all of American Economic Association publications for personal or classroom use is granted without fee provided that copies are not distributed for profit or direct commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation, including the name of the author. Copyrights for components of this work owned by others than AEA must be honored. Abstracting with credit is permitted.

The author has the right to republish, post on servers, redistribute to lists and use any component of this work in other works. For others to do so requires prior specific permission and/or a fee. Permissions may be requested from the American Economic Association Business Office.
Imprecision as an Account of the Preference Reversal Phenomenon

By David J. Butler and Graham C. Loomes*

Many individuals’ choices and valuations involve a degree of uncertainty/imprecision. This paper reports an experiment designed to obtain some measure of imprecision and to examine the extent to which it can explain preference reversals of two opposite forms, one of which appears not to have been reported previously. The model of imprecision we examine not only predicts both patterns but also provides an account of earlier results that are otherwise not well explained. The results suggest that any successful descriptive theory of choice and valuation will need to allow in some way for the imprecision surrounding people’s decisions. (JEL C91, D11, D81)

The preference reversal “phenomenon” occurs if individuals reverse their stated preference orderings over two goods in a predictable manner when different procedures are used to elicit them, even though the different elicitation procedures are formally equivalent and incentive compatible.

The large number of papers discussing this phenomenon (for a review, see Christian Seidl 2002) show that it is easy to produce, but much harder to explain. Yet finding an explanation is an issue of both theoretical and practical significance. This phenomenon is not just an experimental curiosity with little relevance outside the laboratory. There are many important areas of the political economy—including health, safety, and the environment—where policy decisions may be strongly influenced by “stated preference” measures of value elicited from the public.1 If different ways of eliciting those preferences can produce substantially different results, which may have radically different implications for resource allocation and public welfare, it is important to understand the processes people are using to generate their responses. In the absence of such an understanding, it is difficult—and potentially dangerous—to make judgments about the appropriate elicitation method(s) to use or about the degree of confidence we can have in the measures they produce.

This paper tries to move us a step closer to such an understanding. In what follows, we describe a novel elicitation instrument used to explore the ways in which a model of reasonable but imprecise preferences might give rise to various patterns of preference reversal. That

* Butler: Economics Program, UWA Business School, University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia (e-mail: david.butler@uwa.edu.au); Loomes: Department of Economics, University of East Anglia, Norwich, NR4 7TJ, UK (e-mail: g.loomes@uea.ac.uk). We would like to thank the School of Economics and Commerce, UWA, and the Department of Economics, University of Newcastle, UK, for funding this work. We thank Peter Hill for the programming of the experiment and Andrew Williams for assistance in running the experiment. The first author also thanks the Department of Economics at the University of Arizona, where he was visiting for part of the duration of this work, and Peter Walley for early discussions on the structure of preferences. We are also grateful to seminar participants at the Universities of Arizona, East Anglia, Western Australia, and Warwick for comments on earlier versions, as well as participants at FUR XII in Paris, 2004. We also thank three anonymous referees for many detailed and constructive comments on earlier drafts. Thanks are also due to Bob Sugden, Terry Connolly, Ken McCrimmon, and Maxwell Smith for their comments. All opinions and errors in this paper remain our responsibility.

1 For examples and a fuller discussion of the range of “stated preference” techniques and the controversies associated with them, see Ian Bateman et al. (2002).
model also organizes data that earlier researchers were unable to explain satisfactorily.

I. Background and Conceptual Framework

The idea that people’s judgments and preferences may to some degree be imprecise and “noisy” is not new (see, for example, Gustav Fechner 1860, 1966), and in the 1950s and 1960s there was considerable interest in models of probabilistic choice and random preferences (e.g., Frederick Mosteller and Philip Nogee 1951; Nicholas Georgescu-Roegen 1958; Duncan Luce 1959; Gordon Becker, Morris De Groot, and Jacob Marschak 1963; Luce and Patrick Suppes 1965). It was known then—and there has been a considerable accumulation of supporting evidence since—that when presented with exactly the same decision on two separate occasions, an individual’s response might differ from one occasion to the other. So not all of an individual’s observed responses could be consistent with a single stable utility function.

In the last quarter of the twentieth century, however, those interested in explaining various robust and frequently replicated departures from expected utility theory (EUT) did not look to noise or random error for an answer. Instead, the main focus was upon developing alternative deterministic models of individual choice, most of which relaxed one or more of the axioms of EUT in ways that might accommodate the “violations.”

It was not until the mid-1990s that the issue of the stochastic component of decision behavior began once again to receive serious attention (see, for example, Barry Sopher and Gary Gigliotti 1993; David Harless and Colin Camerer 1994; John Hey and Chris Orme 1994; Loomes and Robert Sugden 1995, 1998). Although those papers formulated the stochastic specification in somewhat different ways, they shared what might be regarded as the standard economists’ “top-down” approach: that is, they took some highly articulated deterministic theory to be at the “core” of people’s preferences and then proposed some method of adding randomness and/or error to that core. By so doing, it was possible to show that certain patterns of behavior that might appear to be systematic violations of a certain core theory could in fact be consistent with that same core theory plus a particular model of the stochastic term. It remained the case, however, that no single combination of deterministic core and particular stochastic specification could accommodate more than a subset of the best-known “regularities” observed in decision experiments.²

By contrast with those “top-down” models, our focus in this paper will be upon a more “bottom-up” approach. We take as our starting point individuals who are typical of those who participate in experiments and surveys: that is, those who are somewhat uncertain and imprecise about their preferences in the face of the kinds of tasks they are presented with in decision experiments or stated preference surveys. We then take a model of imprecise preferences which Kenneth MacCrimmon and Maxwell Smith (1986) proposed as an account for preference reversals and we implement an experiment designed to explore and develop that model.

MacCrimmon and Smith—henceforth M&S—considered two forms of preference reversal: in money and in probability. The first of these is the best-known and most frequently studied format for a preference reversal experiment, and operates as follows. Respondents are presented with two binary lotteries. One offers a relatively large chance of a modest sum of money and a residual chance of zero; the other offers a rather smaller chance of a considerably bigger prize and a larger chance of zero.³ After Sarah Lichtenstein and Paul Slovic (1971), these have come to be known respectively as the P-bet and the $-bet. Respondents are then asked to undertake three tasks (not necessarily in this order): to place a “certainty equivalent” money value on the P-bet; to place a certainty equivalent money value on the $-bet; and to make a straight choice between the two bets.

² For a review of the main permutations of “core + error” models, and the extent to which different “regularities” can or cannot be accommodated by them, see Loomes (2005). For an example of a specific attempt to apply one such model to preference reversals, see Ulrich Schmidt and John Hey (2004), who found that even when both pricing and choice errors of a particular kind were excluded, a systematic deviation from standard theory still remained.

³ A few experiments—especially earlier ones—used small losses rather than zero payoffs.
Most decision theories entail that if an individual strictly prefers one bet over the other, she will reveal that preference both by placing a higher money value on the preferred bet and by choosing that bet when offered a straight choice between the two. There are now dozens of studies, however, showing that a substantial proportion of respondents (often the mode, sometimes the majority) violate that expectation and instead exhibit a “standard” reversal by choosing the P-bet while placing a higher money value on the $-bet. The opposite “nonstandard” reversal—choosing the $-bet but placing a higher money value on the P-bet—is much less frequently observed.\(^4\) The fact that standard reversals usually outnumber nonstandard reversals so heavily has led almost all observers to conclude that this phenomenon cannot be accounted for in terms of random noise, however specified.

Other types of reversal are possible. One of these is the second form considered by M&S, involving probability equivalents rather than money equivalents. Taking the same two bets, one task, as before, is to make a straight choice between them. But now, instead of stating a money value for each, respondents are asked to consider a third “yardstick” lottery that offers some given payoff \(x\), but that leaves unspecified the probability \(q\) of receiving that payoff. Respondents are then required to set the level of \(q\) so that they are indifferent between that lottery and the P-bet; and then, separately, to set the level of \(q\) again so that they are indifferent between this yardstick lottery and the $-bet. These values of \(q\) are the “probability equivalents” of the two bets; and, again, the usual assumption is that an individual will set a higher \(q\) for whichever of the two bets she prefers, and that this will correspond with the bet she picks in a straight choice between the two.

There is much less evidence regarding the relationship between choice and probability equivalents for bets. Indeed, apart from MacCrimmon and Smith’s own experiment, the only substantial body of data about “probability reversals” is the one reported in the paper by Robin Cubitt, Alistair Munro, and Chris Starmer (2004). What they found is that although the total numbers of reversals (that is, standard plus nonstandard) were roughly the same for both money equivalents and probability equivalents, the distributions were rather different: whereas the money equivalents exhibited the usual marked asymmetry, the probability equivalence method produced approximately equal numbers of both kinds of reversal, rather than any consistently asymmetrical pattern in either direction. This is broadly in line with what M&S had themselves found (although their sample size was much smaller and their experiment involved only one pair of bets).

Cubitt, Munro, and Starmer (2004) concluded that none of the economic or psychological theories they had set out to test could explain the patterns they observed. They then discussed other possible explanations, including stochastic preference models, but could not identify any variant of such a model that they considered likely to be able to account for their data. They did not, however, consider the approach that M&S thought might provide an explanation for their own broadly similar results.

To explain the nub of the M&S approach, while at the same time setting the scene for our experiment, consider Figures 1A and 1B. In both figures, the basic elements are identical: the vertical dimension represents probability, the horizontal dimension represents payoffs, and $ and P depict the two bets we used in our experiment—respectively, a 0.25 chance of 80 Australian dollars (A$) and a 0.70 chance of A$24. For both bets, the “losing” payoff was zero.

Figure 1A illustrates the way that M&S viewed the certainty equivalence task. Essentially, an individual has to identify a sum of money to be received with certainty—that is, a point on the north side of the rectangle—which she regards as equally as good as the bet being considered. In a conventional deterministic model, there would be just one such point for a given individual. But what if the individual does not have such precise preferences? Starting with the P-bet, all that we can be fairly sure about is that transparent dominance will be respected, so that the individual will be sure that the amount must lie somewhere between 0 and A$24. Of

\(^4\) An exception to this was reported by Jeff Casey (1991), who discovered a manipulation that produced a majority of nonstandard reversals, which he interpreted as suggesting that a contingent decision process, rather than a hard-wired information processing limitation, underlay the phenomenon.
course, most people can say more than that. Any particular individual might be able to narrow their personal range down somewhat further: perhaps she feels quite confident that she’d rather have the P-bet than A$7 (or anything less), and, on the other side, is quite confident that if offered A$15 or more, she would always take the sure sum.

There may, however, be some range—between A$7 and A$15 in this example—where this individual finds it difficult to be so confident. Should it be A$8, A$10, A$12, or A$14? This range of values, within which the person finds it hard to be sure whether one is a better reflection of their preferences than another, is what M&S referred to as an imprecise equivalence. To avoid possible confusion with “thick” indifference curves, we think the term “imprecision interval” might be more appropriate. By this, we mean the range where an individual finds it hard to say whether any particular value is preferable to or less preferred than the stimulus, and therefore the range where no one value is obviously better or worse than another

---

5 See Richard Quandt (1956, 311–12) for the difficulties associated with thick indifference curves.
as an expression of equivalence for the stimulus being evaluated. For the individual in our example, that interval is shown in Figure 1A as the range on the north side of the rectangle encompassed by the two lines connected to the $bet. The lines are drawn straight purely for clarity.

For any particular lottery, both the width and the location of such an interval may vary from one individual to another; and for some, the interval may not exist at all—for example, someone who has convinced himself that an “expected value” (EV) rule is the right one to use may be quite sure he would choose the lottery rather than any sure sum less than its EV, and be equally sure that he would choose any sure sum that was bigger than the EV. But M&S conjectured that, for many people, there would be such an interval, and that the existence of such intervals might help explain patterns of observed response—such as preference reversals—which appear to depart systematically from the implications of standard theory.

M&S refrained from putting a great deal of formal structure on these intervals, and they left open the possibility that particular features of different elicitation procedures might influence their width and location. Their main proposition was that as a bet gets farther away from the north side of the rectangle—that is, as the bet becomes more dissimilar from a certainty—the interval is liable to grow wider. Thus, we might expect the interval for the typical $bet to be wider. In the case of the $bet in Figure 1A, dominance requires the interval to lie between 0 and AS80, although a particular individual may be able to identify, say, AS5 and AS30 as the lower and upper bounds of her personal interval.

Thus, Figure 1A depicts a hypothetical case where the P-bet interval is entirely contained within the $bet interval, but where the range of values above the P-bet upper bound that could be assigned to the $bet is much broader than the corresponding range of values below the P-bet’s lower bound. Of course, the likelihood that this individual will actually assign a higher value to the $bet than to the P-bet would depend on some model of how particular values are selected from within those intervals, and M&S are agnostic on that question, simply pointing to the considerable scope for picking a higher value from the $bet interval than from the P-bet interval, particularly if the elicitation procedure cues the respondent to pick a response more toward the upper end of each of the two intervals.

Having outlined the idea with respect to a certainty equivalence task, consider now the implications for the probability equivalence task, as depicted in Figure 1B. For this task, a particular payoff is identified—in our case, AS160—and the respondent is asked to state the probability of receiving AS160, which she regards as equally as good as the bet under consideration.

Again, transparent dominance sets some bounds on the intervals: for the $bet, the stated probability must be less than 0.25, while for the P-bet it must be less than 0.7. Figure 1B shows the case of an individual who feels confident that she would prefer the $bet if the alternative offered less than a 0.1 chance of AS160, and feels confident that she would prefer the alternative if it offered more than a 0.2 chance of AS160, but is unsure about her preference when the chance of AS160 lies between 0.1 and 0.2. For the more dissimilar P-bet, the imprecision interval spans a broader range—from 0.05 to 0.35. As depicted in Figure 1B, the probability equivalence task allows more scope for the P-bet to be “valued” more highly than the $bet. We wanted to see whether actual behavior corresponded with this depiction. By using the parameters shown in Figures 1A and 1B, our experiment was designed to give plenty of room for standard reversals to outnumber nonstandard reversals in the certainty equivalence task.

---

6 A similar-sounding, but essentially different, notion of “indecisiveness” has been formalized by Kfir Eliaz and Efe Ok (2006). Their approach also offers an account of some cases of the standard preference reversal phenomenon, but it cannot accommodate much of the evidence we present in this paper.

7 The lines are drawn straight purely for clarity.

8 This is not a requirement of the analysis: it just avoids complicating the diagram.

9 For example, if the task is framed, as is the case in quite a few preference reversal experiments, as an open-ended question along the lines of—“What is the smallest sum you would be willing to sell this bet for?”—the respondent may be cued to “anchor” on the positive payoff and adjust cautiously downward; or may adopt a strategic/bargaining stance where the opening ask is on the high side. Either of these ways of generating a response may increase the chances of picking a value from the upper end of the interval.
while at the same time allowing considerable scope for the opposite asymmetry in the probability equivalence task.

Neither the original experiment by M&S nor the more recent experiment by Cubitt, Munro, and Starmer gave their respondents this kind of scope in the probability equivalence task. In both of those cases, the payoff in the “reference” lottery was set somewhere between the payoffs in the $-bets and the P-bets, giving a situation much more like the one shown in Figure 1C (where the broken line signifies a reference payoff of A$50). On the basis of Figures 1A and 1B, it seems quite plausible that for many respondents the two intervals might largely coincide, or at least—as in this example—show no particular asymmetry between the regions that do not coincide. Thus, it would not be surprising to find approximately equal propensities for reversals in both directions—which is more or less what both papers report.

In that sense, this model does seem to have some potential for organizing the Cubitt, Munro, and Starmer results. However, because a reference payoff between the $-bet and P-bet payoffs could in principle cause some responses to be truncated, and because the potentially more dramatic contrast allowed by the Figure 1B design had not been investigated by the other studies, we selected the parameters shown in Figures 1A and 1B.

II. Design and Objectives of the Experiment

In the previous section, we set out the background and broad conceptual framework motivating the experiment and indicated the basic parameters to be used: a P-bet offering a 70-percent chance of A$24; a $-bet offering a 25-percent chance of A$80; and a reference lottery with a payoff of A$160 (US$1 = A$1.40 at PPP). In this section, we describe our design in more detail, and in particular the methods used to elicit responses.

In order to focus upon the issue of imprecision, we wished to control as far as possible for other procedural effects and to make the elicitation of certainty equivalent and probability equivalent values as similar to each other, and to the elicitation of straight choice, as possible.

Consider, for example, one of the $-bets offered by Cubitt, Munro, and Starmer: a 31-percent chance of £32. The expected value of this bet is £9.92, so a risk-neutral respondent would set the probability of the £10 reference payoff at 0.99. There is thus nowhere for any risk-seeking respondents to go, except to set the reference probability at 1.00, which for many such individuals would not really be an equivalent, and they would, in effect, be forced to understate their value of the bet.
We therefore adopted an incremental choice method.

Figure 2A shows an example of the computer display used to elicit the money equivalent of the P-bet. A respondent was asked to consider option A (in this case, the P-bet) and option B (a sure sum of money), with the payoffs shown in the middle two rows and their respective probabilities shown above (for A) and below (for B). Half of the respondents were allocated at random to a treatment where the first sure payoff they were presented with was A$1, while the other half initially saw a value equal to the positive payoff offered by option A (i.e., A$24 when option A was the P-bet, or A$80 when option A was the $-bet).

Respondents were then asked to respond in one of four ways,\(^\text{12}\) which we recorded on a 1–4 scale: if they “definitely preferred” option A, we coded it as 1; if they “probably preferred” A, a 2 was recorded; 3 signified “probably preferring” B; and a definite preference for B was coded as 4.

For the example displayed, it was always the case that a respondent initially signified either a “definite” or “probable” preference for option A. Once the response had been recorded, the computer program changed the amount offered by option B—either raising it by A$1 if the starting point had been low, as in Figure 2A, or reducing it by A$1 if the starting point had been high—and the next response was requested, and so on, for the full range of sure amounts.\(^\text{13}\) We shall refer to these treatments respectively as “iterating up” (U) and “iterating down” (D). The software did not permit subjects to make glaring inconsistencies, such as recording a lesser preference for an option when its attractiveness is increased. However, if they wished, they could use the “back” button to undo their earlier selections, before recording a new set of choices.

Note that this incremental choice technique is not a “valuation” method as commonly understood in preference reversal experiments. All preference elicitation via this procedure rely on expressions of (strength of) choice preferences. Thus, for each respondent we collected not only

\(^{12}\) The instructions are available on the AER Web site (http://www.e-aer.org/data/mar07/20050175_/data.zip).

\(^{13}\) Mosteller and Nogee (1951, 374) described a somewhat related procedure for a wager offering an 80-percent chance of losing 5c combined with a 20-percent chance of winning x, where x is raised incrementally from 1c to $5. They conjectured that “... as \([x]\) increases from 1c towards $5, vacillation sets in, the bet is taken occasionally, then more and more often, until finally, the bet is taken nearly all the time. There is not a sudden jump from no acceptances to all acceptances at a particular offer.”
the certainty equivalent (the $2 \leftrightarrow 3$ switchpoint) for each bet, but also some indication of the interval (between $1 \leftrightarrow 2$ and $3 \leftrightarrow 4$) over which they considered themselves to be less than sure about their preference. This is in contrast to M&S, who did not attempt to measure imprecision in preferences.

Figure 2B gives an example of the display used to elicit the probability equivalent, this time for the $-$bet. In this format, half of the respondents were allocated at random to a treatment where the B option offered A$160 with the probability set at its minimum value of 0.01, while the other half initially saw a display (as illustrated in Figure 2B) with the B option offering A$160 with the same probability as the A option offered for its positive payoff. As with the money equivalence task, respondents were asked to respond to a succession of incremental (0.01) changes to reveal the probability equivalent ($2 \leftrightarrow 3$) point, and any interval of uncertainty around it.

Of course, once we allow the possibility that individuals may be somewhat hazy or uncertain about their preferences, it would be rather paradoxical to expect that they should be absolutely sure about the exact points at which they start and stop being uncertain. Moreover, given the considerable literature about people’s susceptibility to all sorts of procedural effects, it was conceivable that the incremental choice procedure might induce some systematically different patterns of response when compared with a one-off straight choice procedure.

In particular, it seemed quite plausible that the direction of iteration might have an effect—hence the randomization of respondents between U and D in order to check for any such effect. What we had in mind was this. In the U treatment, where respondents are presented with either the P-bet or the $-$bet as option A, and an option B starting at either a very low sure sum or a very low probability of receiving the reference payoff, the great majority can be expected to start by stating a definite preference for option A. This may then become their “reference” option, and the subsequent increases in the values in option B may be seen as attempts to persuade them to give up what they started with. If, however, they are uncertain about the precise point of parity between the two options, they might be inclined to hold on to the perceived status quo for a little longer, so that their $1 \leftrightarrow 2$ and $3 \leftrightarrow 4$ (and possibly $3 \leftrightarrow 4$) switchpoints are all raised somewhat.

The opposite effect might occur for those randomized to the D treatment. Their starting point is very likely to be a strong preference for option B, with successive reductions in the value of that option being seen as an attempt to induce them to give it up. It is true that an option that is continually changing may have less “reference” status than one that stays constant.

---

14 A respondent who felt no sense of uncertainty could switch from 1 to 4 (or vice versa) without ever recording 2 or 3. In fact, 6 of our 89 respondents did this for both the $-$bet and the P-bet.
throughout; but if, in the face of uncertainty, there is any inclination to stick with what has initially been chosen, the effect would be to lower the 3 ↔ 4 and 2 ↔ 3 (and possibly the 1 ↔ 2) switch-points.

To examine that issue, we not only randomized respondents between U and D, but also asked them to make pairwise choices between each of the bets and various predetermined sure amounts and chances of the reference payoff. So we asked each respondent to make: five one-off choices between the $-bet and five different sure amounts—8, 12, 16, 20, and 24 dollars; five one-off choices between the $-bet and five different probabilities of receiving A$160—0.10, 0.12, 0.15, 0.18, and 0.20; five one-off choices between the P-bet and five different sure amounts—4, 8, 12, 16, and 20 dollars; and five one-off choices between the P-bet and five different probabilities of receiving A$160—0.10, 0.15, 0.20, 0.25, and 0.30. In addition, using the same basic display, we asked respondents to make a straight choice between the $-bet and the P-bet on three separate occasions.

In order to try to keep each preference-reversal-related decision somewhat separate, we alternated them with questions similar in format but using different parameters and designed to investigate whether imprecision might also account for violations of independence and betweenness. Space does not permit a comprehensive analysis and report of those data, but there is one aspect that is particularly pertinent to the present paper, so a brief outline of certain salient features is in order.

Figure 3 shows a Marschak-Machina triangle diagram of the kind often used to explain designs intended to test for independence and betweenness. Any lottery involving no more than three payoffs \( x_3 > x_2 > x_1 \) can be depicted as a point in the triangle, where the vertical axis represents the probability of \( x_3 \), the horizontal axis represents the probability of \( x_1 \), and the probability of \( x_2 \) is given by \( 1 - pr(x_1) - pr(x_3) \). Thus, in Figure 3, lottery \( M_1 \) offers a 0.6 chance of \( x_3 \) and a 0.4 chance of \( x_1 \), while \( M_4 \) offers a 0.2 chance of \( x_3 \), a 0.6 chance of \( x_2 \), and a 0.2 chance of \( x_1 \).

The questions that alternated with the preference reversal tasks each involved taking one of the lotteries \( M_1 - M_5 \) as the fixed option A in Figure 2B and then eliciting a series of strength of preference responses as option B moved progressively along the vertical and horizontal edges of the triangle, either from a starting point which strictly dominated the M lottery to a finishing point which was strictly dominated by the relevant M, or in the opposite direction, with half of the sample being assigned to each direction of iteration. Within each of those treatments, half were assigned to lotteries where \( x_3 = A$60, x_2 = A$20, and x_1 = 0 \), while for the other half the payoffs were \( A$40, A$20, and 0 \) respectively.

Thus, for example, in the question where the fixed lottery was \( M_2 \), a respondent who was assigned to the subsample that iterated up (from dominated to dominating) with lotteries where \( x_3 = A$60, x_2 = A$20, and x_1 = 0 \), while for the other half the payoffs were \( A$40, A$20, and 0 \) respectively.

A more comprehensive report can be found on the AER Web site.
of A$20—that is, moving steadily up the vertical edge, finishing at the point level with M2, where lottery B now dominates M2 by offering a 0.6 chance of A$60 and a 0.4 chance of A$20.

As with the preference reversal questions, this iterative procedure was intended to identify not only the 2 ↔ 3 switch-point for each M but also the interval between 1 ↔ 2 and 3 ↔ 4 which we are taking to be an indication of imprecision. It is these latter measures that are pertinent to the present paper, since they provide additional information relating to the M&S conjecture that these intervals get wider as the fixed lottery gets farther from the edge where its equivalence is being judged. In terms of the Marschak-Machina triangle, M2 and M3 will generally be more dissimilar from their counterparts on the horizontal or vertical edges than will be the case for M1, M4, and M5. But will the imprecision intervals reflect this? In addition, we might expect that, all other things being equal, broader ranges of payoffs will increase uncertainty about where the equivalences lie: so will the intervals be wider for the lotteries where the range is A$60 than for those where the range is A$40?

Just as with the preference reversal questions, there were also a number of one-off straight-choice questions based on the Marschak-Machina triangle; and as with the preference reversal questions, these were intended to provide a basis for examining how far the direction of iteration might affect the patterns of choices.

We could find no way of making the distinction between “definitely preferring” an option and “probably preferring” that same option incentive compatible. We doubt that such a mechanism can be devised—at least, not in a form simple and transparent enough to work without creating additional uncertainty. So we relied upon respondents making the distinction simply because we asked them to do so and because they found that distinction meaningful. In Section IV we shall consider arguments for and against taking these data seriously.

However, we were able to make their straight choices between A and B incentive compatible, and it was explained that all these choices were made on the basis that, at the end of the session, one question would be selected at random for each respondent, who would each be paid according to the way their decision in that particular question worked out. Average earnings across the sample were A$26, ranging from a low of A$0 to a high of A$160.

Thus, by the end of the experiment we had obtained from each respondent the following data relating to preference reversal: point estimates of certainty equivalents for the $-bet and P-bet (denoted by CE$ and CEP), together with an indication of the interval of imprecision around those points; corresponding probability equivalent estimates (PE$ and PEP) and intervals; three straight choices between the two bets; and a total of 20 straight choices between the bets and various prespecified sure sums or preset chances of receiving A$160. We had also obtained additional relevant data from the Marschak-Machina component of the design.

III. Results

A total of 89 individuals took part in the experiment, 44 allocated at random to D, 45 to U. The first issue is whether the patterns of response were systematically affected by the direction of iteration in the manner hypothesized above.

Table 1 reports summary statistics of responses to the four incremental tasks, in each case showing the means and standard errors of the 1 ↔ 2, 2 ↔ 3, and 3 ↔ 4 switch-points. It is clear that in three tasks out of four, the direction of iteration made a significant difference to all switch-points in the hypothesized direction, while in the other task—the certainty equivalent of the $-bet—the difference is in that same direction, although not to an extent which registers as significant.

Of course, even though respondents were allocated to D and U at random, it is just possible that those in the D subsample happened to have lower values for the $- and P-bets than their U subsample counterparts. To examine this,
Table 2 shows the data from the straight choices. The first column shows whether option A was the $- or P-bet while the second column shows the various sure amounts or chances of A$160 offered by option B. The next two columns show, respectively, the numbers in the D subsample and the numbers in the U subsample who actually chose option A rather than whatever option B was offering. In general, the numbers are very similar, and only one comparison (for the $-bet when option B offers the certainty of A$8) exhibits a difference that is significant at the 5-percent level—but since that is just one of the 20 comparisons, it could have occurred by chance.

The fifth and sixth columns in Table 2 show for D and U, respectively, the numbers inferred from their iterative responses to prefer option A rather than whatever option B was offering. In keeping with expectations and with the data from Table 1, inferred choices of A are higher for the U subsample than the D subsample in 19 out of 20 cases, and in many cases the differences are pronounced.

We can also examine any patterns of differences between responses to the iterative questions (which were not incentive-linked) and responses to the straight choice questions (which were linked to incentives in the conventional manner).

In the case of the U subsample, these differences are systematic and marked: here, too, the direction of difference goes the same way in 19 out of 20 cases, with the numbers inferred as preferring option A being strictly higher than the numbers actually picking option A in the corresponding straight choice. The eighth column of Table 2 reports significance levels for these differences, and in the majority of cases they are significant at the 5-percent level.
For the D subsample, however, the majority of differences (11/20) are in the opposite direction, and the pattern is much more mixed. For the certainty equivalence of the P-bet and the probability equivalence of the $-bet, inferred choices of option A undershoot the actual choices for the three highest values of option B, but the differences fade thereafter. In the case of the probability equivalence of the P-bet, there is not much to choose between actual and inferred. And for the certainty equivalence of the $-bet, inferred choices of the $-bet exceed actual choices at every level, although this effect is weaker for the D subsample than for the U subsample.

The divergence between the responses elicited via the Up (U) treatment and those from both the Down (D) treatment and the straight choices is entirely compatible with the notion that many respondents are to some extent unsure about their preferences and therefore vulnerable to certain procedural effects: in particular, a reluctance to switch away from an unchanging “reference” option. However, the design aimed to control for this, since respondents either iterated up for all four preference reversal equivalence tasks, or iterated down for all four.

We turn next to the issue of the size of intervals. Recall that although M&S made no strong assumptions about the nature/extent of imprecision, their hypothesis was that the intervals tended to become wider, the more dissimilar is the fixed lottery from the edge on which the equivalence is being expressed. Our proxy measure for these intervals is the range between the $\frac{1}{2}$ and the $\frac{3}{4}$ points.

For each respondent and for each of the four tasks, we computed these intervals, denoting them as follows: CEPint and CE$\text{S}_\text{int}$ are, respectively, those intervals for the certainty equivalence for the P-bet and $-bet; and PEPint and PES$\text{int}$ are the corresponding measures for the probability equivalence tasks. Figures 4A and 4B depict the average intervals for each subsample. The points representing the P-bet and $-bet
$\text{-bet}$ are joined (for ease and simplicity) by straight lines to the means of the 1 ↔ 2 and 3 ↔ 4 responses.

Although the position of these lines is affected by the direction of iteration—the greater degree of kink around each bet for the U subsample being consistent with a stronger reference point effect—the width of the intervals is fairly constant across subsamples: tests for the difference between the D and U mean intervals—reported in the upper part of Table 3—show no significant difference for any of the four tasks.

To examine the issue at the level of individuals, we computed for each respondent two measures of the differences between their intervals: for the certainty equivalence task we computed $\text{CEintdiff} = \text{CE$int$} - \text{CEPint}$, and for the probability equivalence task, $\text{PEintdiff} = \text{PE$int$} - \text{PEPint}$. We then tested the null hypotheses that $\text{CEintdiff} = 0$ and $\text{PEintdiff} = 0$ against the alternative hypotheses that $\text{CEintdiff}$ is positive and $\text{PEintdiff}$ is negative. As Table 3 reports, for both cases and both subsamples, the null hypotheses were clearly rejected in favor of the alternatives. The breakdown at the
level of the individual was as follows: 73 of the 89 individuals had \( CE_{\text{Pint}} > CE_{\text{Pint}} \), with 10 having the same interval and 6 having the opposite inequality; at the same time, 71 of the 89 individuals had \( PE_{\text{Pint}} > PE_{\text{Pint}} \), with 12 having the same interval and 6 having the opposite inequality. In both cases, if the null hypothesis were true, the likelihood of such asymmetry occurring by chance is vanishingly small.\(^{17}\)

These data could be interpreted as supporting M&S’s conjecture that interval widths might be an increasing function of the distance between the position of the bet in the Figure 1 rectangle and the side of that rectangle upon which the response is recorded. For example, the \$-bet is roughly 2.5 times farther away from the top side of the rectangle than the P-bet, and the average CES intervals are 3.15 and 3.53 times bigger than the average CEP intervals for the D and U subsamples, respectively. Similarly, the P-bet is about 1.7 times farther from the right-hand side of the rectangle than the S-bet, and the ratios of \( PEP_{\text{Pint}} > PE_{\text{Pint}} \) are 2.70 and 3.29 for D and U.

There is, however, another possibility that fits the data at least as well, and is no less consistent with the spirit of the model: namely, that the width of each interval is positively correlated with the ranges within which dominance is not transgressed. For example, the CE of the P-bet could lie anywhere between \$24 and 0 without violating dominance, while the allowable range for the CE of the \$-bet (\$80 to 0) is 3.33 times bigger—a figure that neatly bisects the CESint:CEPint ratios of 3.15 and 3.53. Similarly, the PE of the P-bet could lie between 0.70 and 0, while the allowable range for the \$-bet is between 0.25 and 0: the 2.8:1 ratio of these two ranges also lies squarely between the observed PEPint:PEPint ratios of 2.70 and 3.29.

The data from the lotteries in the Marschak-Machina triangle, reported in Table 4, provide further support for this possibility. Although the distances between lotteries \( M_1 \) to \( M_4 \) and their equivalents on the vertical/horizontal edges vary a good deal, the imprecision intervals within each triangle are strikingly similar, arguably reflecting the fact that the allowable range for each of \( M_1 \) to \( M_4 \) is the same—that is, 100 percentage points along the vertical/horizontal edge.

\(^{17}\) Despite the similarity of the two breakdowns in aggregate, it was not the same individuals who constituted the six observations of opposite differences: there was only one respondent for whom CESint < CEPint and PEPint < PEint.

| Table 3—Imprecision Intervals for P- and \$-Bets |
|-----------------|----------------|----------------|----------------|
|                | D              | U              |                |
| \( CE_{\text{Pint}} \) | mean 5.98      | s.e. 0.59      |                |
| \( CE_{\text{Sint}} \) | mean 18.82     | s.e. 2.06      |                |
| \( PEP_{\text{Pint}} \) | mean 19.50     | s.e. 2.11      |                |
| \( PES_{\text{Pint}} \) | mean 7.23      | s.e. 0.75      |                |
| \( CE_{\text{intdiff}} \) | mean 12.84     | s.e. 1.78      |                |
| \( P_{\text{intdiff}} \) | mean -12.27    | s.e. 1.88      |                |

| Table 4—Imprecision Intervals for M Lotteries |
|-----------------|----------------|----------------|
| Lottery         | \( x_2 = \$40 \) | \( x_2 = \$60 \) |
| \( M_1 \)       | 21.16           | 25.82           |
| \( M_2 \)       | 21.69           | 25.27           |
| \( M_3 \)       | 21.11           | 27.23           |
| \( M_4 \)       | 23.33           | 27.45           |
| \( M_5 \) Percent | (23.22)       | (28.57)       |

(M5 Percent)
edges.\textsuperscript{18} For \( M_5 \), which might be regarded as more like \( M_1 \) and/or \( M_4 \) in terms of distance from its edge equivalent, the imprecision interval is narrower in absolute terms; but as the figures in brackets in the bottom row of Table 4 show, that interval represents much the same proportion of the smaller allowable range as we observe for \( M_1 \) to \( M_4 \) in the same triangle. In addition, there does appear to be some effect of payoff range: as tentatively conjectured, the imprecision intervals are all higher for the lotteries where payoffs span A$60 as compared with the corresponding lotteries in the A$40 triangle (although no single difference registered as statistically significant).

Thus, there might be grounds for modifying the M&S conjecture about the width of the imprecision interval: it may not simply be a function of the degree of (dis)similarity between the fixed lottery and the locus on which the equivalence is being expressed, but may also be influenced by the width of the range of possible answers on that locus which are not disallowed by dominance.

Of course, it is not just the widths of these intervals but their location relative to one another that is crucial to accounting for preference reversals. As Figures 4A and 4B show, on average the CE$ intervals range over higher money values than the CEP intervals: in fact, for each subsample, the two intervals only barely intersect. The PE intervals show a mirror image of this pattern, with the PEP interval involving much higher probabilities than the PE$ interval (and not intersecting at all in the case of subsample U). Such a pattern is even more extreme than those envisaged in Figures 1A and 1B, and clearly has the potential to accommodate both the standard preference reversal phenomenon involving certainty equivalents and the opposite asymmetry involving probability equivalents, as M&S conjectured.

The patterns were as follows. Recall that we asked respondents to make straight choices between the P-bet and the $-bet on three separate occasions. Of the 89 respondents, 51 chose the P-bet on all three occasions, 17 chose the P-bet twice and the $-bet once, 6 chose the P-bet once and the $-bet twice, and 15 chose the $-bet on all three occasions. In Tables 5 and 6, we report the relationships between choices and equivalences for each occasion of choice and each type of equivalence.

Table 5 shows a pattern to which we have become accustomed: setting aside those who stated CEP = CE$ (and who would therefore be weakly consistent whichever bet they chose), between 32 and 36 respondents gave strictly consistent answers to both tasks, while between 43 and 47 (that is, between 48.3 percent and 52.8 percent of the sample) reversed their stated preferences, with the least asymmetric split relating to choice 1, where the ratio of standard:nonstandard reversals was 42:1. The usual null hypothesis—that both types of reversal are equally likely to occur—is emphatically rejected in favor of the alternative hypothesis consistent with imprecise preferences. In case it might be thought that the phenomenon relies heavily on those who are least consistent in their

\textsuperscript{18}For example, for \( M_2 \) the lowest 60 percent of the vertical axis is in the allowable range, as is the leftmost 40 percent of the horizontal axis; for \( M_5 \), the allowable range consists of the lowest 20 percent of the vertical edge and the leftmost 20 percent of the horizontal edge.
choices, the lowest panel in Table 5 focuses exclusively on those who made the same choice on all three occasions. Exactly half of these 66 exhibited reversals, all of which were in the standard direction.

Table 6 displays the corresponding data for the probability equivalence task. Here, as predicted by the model of imprecision, reversals displayed the opposite asymmetry; and even in the least asymmetrical instance, relating to choice 2, the probability of the 14:6 split occurring by chance if the null hypothesis were true is 0.058 (exact binomial test), while for the more asymmetric instances relating to choices 1 and 3, the corresponding probabilities are 0.002 and 0.006. Once again, to check whether this pattern was largely attributable to those who were least consistent in their choices, the lowest panel of Table 6 reports the data for the 66 who made the same choice on all three occasions. Of the 15 who consistently chose the $-bet, 9 gave a strictly higher probability equivalent to the P-bet, while only 2 of the 51 who consistently chose the P-bet exhibited the opposite reversal: the likelihood of that 9:2 split occurring by chance is 0.033.

The patterns of imprecision depicted in Figures 4A and 4B not only are consistent with all of the above, but also help us to make sense of another aspect of the preference reversal phenomenon which has not been well explained by earlier theories: namely, what Peter Fishburn (1988) called “strong reversals.” In the context of certainty equivalents, a strong reversal is said to occur when the P-bet is chosen even though the certainty equivalent of the $-bet is strictly greater than the positive payoff offered by the P-bet. (Fishburn referred to cases where CE$ is less than the positive payoff of the P bet as “weak” reversals.) It turns out that even models such as regret theory (David Bell 1982; Loomes and Sugden 1982, 1987) and SSB utility theory (Fishburn 1982), which can accommodate weak preference reversals by relaxing transitivity, cannot accommodate strong reversals. And although Fishburn did not address the question of reversals in probability, it also turns out that strong reversals of this kind—where the $-bet is chosen even though the probability equivalent of the P-bet is strictly greater than the chance offered by the $-bet of receiving its positive payoff—are not compatible with regret or SSB theories in their deterministic forms.19

However, preference imprecision allows for the possibility of strong reversals of both kinds. Almost half of the width of subsample D’s CE$ interval, and slightly more than half of subsample U’s CE$ interval, occupy values that are strictly higher than the A$24 payoff. At the same time, more than half of subsample D’s PEP interval, and all of subsample U’s PEP interval, involve probabilities that are strictly higher than the $-bet’s 0.25 probability. Of course, those intervals are based on averages and do not represent every respondent. Nevertheless, it turned out that 40.3 percent of the

19 In fact, the restrictions on regret theory’s $\phi(\cdot, \cdot)$ function that allow the classic pattern of CE reversals also entail reversals in that same direction when probability equivalents are elicited—that is, regret theory is incompatible not only with strong probability reversals but also with weak PE reversals of the kind that predominate here.
standard money reversals were strong reversals, while a similar proportion—43.8 percent—of the relevant probability reversals were also strong ones.

**IV. Discussion**

We set out to explore the degree of imprecision as indicated by people’s statements of their preferences, taking as a general framework a model along the lines sketched out by M&S. We devised an instrument that not only elicited the points at which respondents switched from one option to another, but also attempted to get some measure of the intervals around those points where respondents were less than sure about their preferences. We conducted what we believe to be the first test of the implications of the M&S approach for PE reversals when the yardstick lottery offers a payoff higher than the payoffs of either bet, as well as replicating the usual CE reversal phenomenon. It appears that the model of imprecision discussed in this paper not only accounts for our data, but also organizes the data that M&S (1986) and Cubitt, Munro, and Starmer (2004) reported, and provides a framework that can accommodate the “strong” CE reversals observed in many studies and remarked upon by Fishburn (1988), as well as strong PE reversals never previously reported but found in our data.

These data suggest that in the absence of very precise “true” preferences, respondents faced with equivalence tasks may be liable to pick one value from an imprecision interval, with their perception of the range of this interval, and their selection of a particular value from within it, both liable to be influenced by various “cues” or “anchors.” In our data, it appears that one anchor was the starting point in the iterative procedure, with the difference between the D and U treatments, as reported in Table 1 and displayed in Figures 4A and 4B, suggesting that this had a discernible influence both upon the location of the interval and the switch-points within it.

In addition, there was evidence suggesting that the width of the imprecision interval was liable to be influenced by the width of the range of possible responses that did not violate transparency dominance. The implication is that imprecision not only allows starting point effects to influence people’s responses, but also makes them susceptible to range effects sufficient to accommodate a number of strong reversals as well as a great many weak ones.

Someone skeptical of our interpretation might question the status of responses to procedures not directly linked to financial incentives. It might be suggested that respondents really have fairly precise preferences which they reveal with reasonable accuracy when offered the appropriate financial incentives, but that in the absence of such incentives they have no motivation to engage properly with the tasks and answer carefully. So is it safe to rely on data from the iterative procedures to inform us about behavior when the stakes are real?

A comprehensive discussion of the general importance (or otherwise) of financial (or other) incentives in decision experiments is beyond the scope of this paper, so we address our remarks to the specific question of the usefulness of our imprecision data for understanding the preference reversal phenomena that are the subject of this paper. We suggest that this issue might be judged on the basis of two criteria: first, whether the data show reasonable signs of being the product of engagement and deliberation, as opposed to being generated haphazardly, with little thought or effort; and second, 20 Starting point effects and range effects have been widely reported in the psychology literature and in a number of studies seeking to elicit values for health, safety, or environmental goods—again, see Bateman et al. (2002) for a review. One striking example of both effects in operation can be found in Richard Dubourg, Michael Jones-Lee, and Loomes (1997), where values were elicited for reductions in the risks of road accident injuries. Respondents were asked first to identify amounts they were sure about—the one hand, amounts they were sure they would pay, and on the other hand, larger amounts they were sure they would not pay. There was often some interval in between where they felt unsure, but they were asked to identify a single value in that interval as their “best estimate” of their maximum willingness to pay. As in our experiment, both the position of the intervals and the distributions of best estimates were influenced significantly by the starting point in an iterative bidding procedure and also by the range of values presented on a payment card.
whether they tell a story that is broadly consistent with patterns in the incentive-linked responses.

Regarding the first question, the great majority of our respondents expressed definite preferences over some ranges and more tentative preferences over other ranges on either side of the point where they switched from one option to the other, and did so in ways that showed considerable and systematic responsiveness to the characteristics of the different questions. Respondents were clearly not changing from “definite” to “probable” preference, or vice versa, after much the same number of steps in the iterative procedure, irrespective of the nature of the lottery: for example, the $1 → 2 switch-point for the certainty equivalent of the $-bet was typically between 15 percent and 20 percent of the distance from the bottom of the iterative range, while the $3 → 4 switch-point was just over 40 percent of the way along (i.e., less than halfway); by contrast, the $1 → 2 switch-point for the certainty equivalent of the P-bet was typically between a third and a half way along the range, with the $3 → 4 switch-point lying roughly 60 percent to 80 percent of the way along (depending on the direction of iteration). Given the different probabilities of winning offered by the two bets in conjunction with the kind of risk aversion typically exhibited in decision experiments, this seems entirely consistent with the proposition that respondents were attending to the parameters of the lotteries and trying to reflect their feelings about them.

Moreover, most respondents varied the widths of their imprecision intervals quite substantially from one question to another: as noted in the text, more than 80 percent of respondents gave wider intervals in the expected direction, even though each interval was elicited in a question separated from its counterpart by at least one other question (taken from the Machina triangle set) involving iteration in the opposite direction; and as reported in Table 3, the wider intervals were typically at least three times wider. All this suggests that most participants had at least some intuitive feel for the distinction between definite and probable preference and, having been asked to do so, reported those feelings as best they could and in a manner that was broadly responsive to the varying parameters of the lotteries presented to them.

Nevertheless, a skeptic might still be troubled by any systematic divergence between hypothetical and incentive-linked data: in particular, the clear disparity between the actual and inferred choices of the U subsample reported in Table 2 might be a source of concern.

On that specific issue, we anticipated that this particular direction of iteration might encourage a reference effect, and it seems that it did. But that is not something peculiar to hypothetical questions: such effects have been reported in many studies using real—and sometimes quite substantial—incentives, and they are often quite resistant to attempts to counteract them. Having anticipated this possibility, however, we controlled for it with respect to the main focus of the study (by using the same iteration direction for the set of preference reversal questions), so the patterns of reversals are not attributable to this effect.

On a more general point, it appears that even among the responses that were linked to incentives in a standard way, there were disparities. To see this, consider the “demand” for the $-bet and the P-bet inferred from the choices where four different sure sums—A$8, A$12, A$16, and A$20—were offered as alternatives to both bets. The total number of respondents choosing the $-bet at those “price” levels were 58, 48, 26, and 10, while the corresponding figures for the P-bet were 66, 50, 23, and 5. On this basis, there would appear to be little difference between the demands for each bet.

By contrast, if we use the choices between each of the bets and the three levels of yardstick lottery against which they are both compared—that is, 0.10, 0.15, and 0.20 probabilities of receiving A$160—the patterns of demand are very different: against those three levels, the P-bet was chosen, respectively, by 75, 70, and 51 respondents, whereas the corresponding figures for the $-bet were 66, 43, and 12.

So the different sets of choice problems produced systematically different pictures of preference, even though the incentive mechanism was exactly the same for both. With respondents’ preferences displaying the kind of imprecision reported in this paper, it may be that all methods of eliciting those preferences, with or without incentives, are vulnerable to procedural effects of one kind or another. The claim that
incentives produce "truth" and/or that any one procedure represents the gold standard—either in the laboratory or in the field—may need to be treated with caution.21

Having said all that, the key question for the present study is whether the most striking feature of the data from the iterative procedure—namely that the asymmetry found with certainty equivalents is reversed when probability equivalents are elicited—is also evident in the incentive-linked choice data. The short answer is: yes, it is.

Of course, asking pairwise choices involving preset parameters is inevitably a more coarse-grained procedure and is liable to miss cases. For example, suppose an individual prefers P to $ in a straight choice and would place certainty equivalents of $15 on the $-bet and $13 on the P-bet. Eliciting those equivalences would reveal a preference reversal. However, if that same individual is asked simply to choose between each bet and a couple of preset certainties—say $12 and $16—the reversal may be missed: she prefers $16 for sure to each bet, and prefers each bet to $12, and no reversal is observed because the "action" lies between the two sure amounts selected by the experimenter.

Nevertheless, if the phenomena revealed by the iterative procedures have their counterparts in straight choices, we should expect them to be at least some indications. Denoting a sure amount in a straight choice by C, the choice analogue of a standard preference reversal is the intransitive cycle P ≻ C, C ≻ P, with the nonstandard reversal translating to $ ≻ P, P ≻ C, C ≻ $. If the certainty equivalent asymmetry reported in Table 5 also manifests itself in straight choices, we should expect cycles of the first type to outnumber the second type. And, indeed, taking the choices where C was set at $8, $12, $16, and $20, there were a total of 69 cases of the "standard" cycle as compared with 19 cases of the "nonstandard" cycle. Of course, given that each individual made numerous choices, it was possible for the same individual to record more than one version of the same type of cycle. Taking the individual as the unit of analysis, there were 22 respondents who exhibited the standard cycle on at least one occasion and never exhibited the nonstandard cycle, as opposed to just 8 who exhibited the nonstandard cycle on at least one occasion. Using the analogue of the null hypothesis applied earlier—namely that individuals exhibit cycles as a result of random error and are equally likely to exhibit either form—the probability of this asymmetry occurring by chance is 0.008.

Correspondingly, let Q represent the preset lottery offering $160, with the probability taking one of the values 0.10, 0.15, or 0.20 against which both bets were compared in straight choices. Here the choice analogue of the reversal most often observed in Table 6 is $ ≻ P, P ≻ Q, Q ≻ $, while the less frequent form translates to P ≻ $, $ ≻ Q, Q ≻ P. The choice data show a total of 44 of the first type and 17 of the second type, with 18 individuals exhibiting only the first type, 6 exhibiting only the second type, and 2 individuals each exhibiting 2 of the first type and 1 of the second type. Again, the null hypothesis is rejected, either with probability 0.038 if the two "mixed" individuals are included on the side of their more frequent cycles, or with probability 0.011 if those individuals are omitted. A more general null hypothesis, which does not suppose that both types of cycle are equally likely to be manifested, but only that there should be no difference between the asymmetries whether preferences are elicited via choices involving C or via choices involving Q, is also very firmly rejected ($ \chi^2 = 10.15, p < 0.001$).

So although the choice data are necessarily coarser, and although there may be more noise due to the presentation in quick succession of a number of rather different pairs, the mirror-image asymmetries revealed by the iterative procedures are also manifested to a less visible, but still significant, extent in the choice data.

Our conclusion, then, is as follows. We con-
sider there are grounds for supposing that many respondents’ preferences are to some degree imprecise, and that such imprecision may be a factor in explaining the occurrence of preference reversals and in accounting for the opposite asymmetric patterns predicted by MacCrimmon and Smith (1986) and reported here. We do not suggest that our iterative instrument is immune from procedural effects, nor do we claim that it does more than give some proxy measure of the degree of imprecision. But it does appear to produce relevant additional information about behavior in decision experiments and about the role that imprecision may play in systematic departures from standard models of deterministic preferences. The implication is that no decision theory is likely to be descriptively adequate unless it makes appropriate allowance for the imprecision in people’s expressed preferences; and that further exploration of the nature of imprecision—whether by means of the kind of instrument we have used, or by other methods—may prove a fruitful line of future inquiry.

REFERENCES


This article has been cited by:


