Imprecision as an Account of Violations of Independence and Betweenness

David Butler
University of Western Australia

Graham Loomes
University of Warwick

ABSTRACT

In an earlier paper we put forward a model of imprecise preferences which accounted for various forms of preference reversal. In this paper we show that the same model can also explain the best-known violations of expected utility theory's axioms of independence and betweenness. It appears that a simple model of imprecise preferences can account for a broader range of anomalies than any one of the more elaborate alternative theories developed to date.

Keywords: Imprecise preferences, independence, betweenness, experiment
(JEL: C91, D81)

Corresponding Author: David J. Butler, Economics Program, UWA Business School, University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia Tel: 61 8 6488 7160 Fax: 61 8 6488 1035 (e-mail: david.butler@uwa.edu.au);
Graham Loomes, Department of Economics, University of Warwick, Coventry, UK
(e-mail: g.loomes@warwick.ac.uk).

Standard competitive markets do not exist for some important goods and services such as many environmental improvements, or a range of health and safety benefits. As a result, some government bodies draw on 'stated preference' methods to provide data about people's values and preferences which may inform policy in these domains. But such survey data are liable to exhibit substantial uncertainty, imprecision and patterns which fail to satisfy various standard assumptions (see, for example, Bateman et al. (2002, Chapter 12) and Loomes (2006)). If such data are to be used for policy decisions which may impact upon people's risks of death, injury and illness, it is important to gain a better understanding of the imprecision in stated preferences and any susceptibility to bias that may be involved.

Experimental research has also shown that even when the 'goods' are relatively familiar and straightforward, many intelligent and numerate individuals find it hard to know their own preferences precisely and may systematically transgress basic axioms of rationality that underpin standard consumer theory and the policy principles that flow from it. If this is true, and given the growing variety and sophistication of goods and services, are there implications for consumer protection? If consumers' preferences *are* imprecise and susceptible to manipulation, this may be used against their own best interests. But it is hard to reach firm conclusions about the right policies to pursue until we have a better model of imprecise preferences. This paper explores some basic issues about the way that people deal with risk, with a view to contributing towards the construction of better models.

During the last three decades, hundreds of studies have been published which report a variety of seemingly systematic violations of expected utility (EU) theory. The breadth and weight of this evidence has inspired more than a dozen alternative decision theories (for a review see Starmer, 2000). However, no single model has been able to accommodate more than a subset of these patterns. For example, different rank-dependent expected utility (RDEU) models can account for violations of betweenness and independence, but cannot explain systematic intransitivity or the preference reversal phenomenon, while other models such as regret theory allow standard preference reversals and certain patterns of cyclical choice but cannot explain many of the violations of betweenness and independence[1].

Third-generation prospect theory (Schmidt, Starmer, & Sugden 2008) is a generalization of prospect theory which, unlike earlier versions, can account for standard selling preference reversals as well as the other anomalies predicted by prospect theory (pp.212-220). Despite this important advance it still cannot account for buying preference reversals, nor choice cycles. It would also struggle to find plausible parameter values that can capture the strong reversals and non-standard reversals that are reported in Butler and Loomes (2007).

More recently, and partly as a response to those data, Blavatskyy (2009) has produced a model that provides a possible probabilistic choice account of the directions of the asymmetries reported in Butler and Loomes (2007). Blavatskyy argues that his model can also account for some violations of independence and betweenness, although this requires additional assumptions about the non-homogeneity of the probabilistic function. As such, it is in the tradition of taking some deterministic 'core' theory – in this case, EU – and embedding it in some particular stochastic specification to account for seemingly systematic deviations from core principles (for more discussion and examples of this literature, see the special issue of Experimental Economics in late 2005, Wilcox's major review chapter in Cox and Harrison (2008) and chapter 7 in Bardsley et al., (2009)).

However, our strategy is different. Rather than focusing on a probabilistic model revolving around some specific set of principles or axioms, we consider what insights may be gained from the fairly minimal structure suggested in an unpublished paper by MacCrimmon and Smith (1986) – henceforth, M&S. In Butler and Loomes (2007), we showed how their model of imprecise preferences, though quite loosely specified, not only accounted for the ‘standard’ preference reversal phenomenon involving a systematic disparity between choice and certainty equivalent valuation but also predicted the *opposite* asymmetry involving choice and probability equivalents. In the present paper, we consider whether an extension of that same simple model can also accommodate systematic violations of betweenness and independence, and we report experimental data suggesting that this is indeed a possibility.

I. The Model

Consider the Marschak-Machina triangle diagram in Figure 1a.

FIGURE 1a HERE

This diagram enables us to depict the kinds of lotteries most often deployed in tests of independence and betweenness: namely, lotteries involving combinations of up to three payoffs – x_1, x_2 and x_3, where $x_3 > x_2 > x_1$ (and where, in most cases – and in our experiment – x_1 is set at 0). In this diagram, the vertical axis shows the probability of x_3 and the horizontal axis shows the probability of x_1, with the probability of x_2 being given by 1-pr(x_3)-pr(x_1). Hence the point on the hypotenuse labelled M_1 depicts a lottery offering a 0.8 chance of x_3 and a 0.2 chance of x_1, while M_4 represents a lottery involving a 0.2 chance of x_3 and a 0.8 chance of x_1.

Machina (1982) showed that an EU maximizer’s preferences over such lotteries can be represented by indifference curves within the triangle that are linear and parallel, each sloping up from the south-west to the north-east and with the slope reflecting the individual’s risk attitude (the steeper the slope, the more risk averse the individual). The most frequent violations of independence can be represented by curves that are not parallel but ‘fan out’ as if from some point of convergence to the south-west of the right angle of the triangle. Could behaviour taking on this appearance be consistent with the intuition behind the M&S model of imprecise preferences?

We start with the case (which was one we investigated experimentally) where x_3 = \$40, x_2 = \$20 and x_1 = 0, and suppose that a respondent is asked to identify a lottery L_1^* on one of the other edges of the triangle which she regards as equivalent to M_1, where M_1 can be denoted by (40, 0.8; 0, 0.2)[2]. In line with M&S, we suppose that the respondent recognises and respects transparent dominance. Thus she recognises that L_1^+ (= (40, 0.8; 20, 0.2)) and all lotteries above it on the vertical edge strictly dominate M_1, and therefore realises that L_1^* must lie below L_1^+. On the other hand, L_1^-(= (20, 0.8; 0, 0.2)) and all lotteries to the right of it on the horizontal edge are strictly dominated by M_1, so that L_1^* must lie somewhere to the left of L_1^-.

So the 'permissible' range for equivalents to M_1 consists of all points below L_1^+ on the vertical edge and to the left of L_1^- on the horizontal edge. A conventional deterministic model would entail that each respondent could identify a single one of those points as the lottery which she regards as exactly as good as M_1. But the M&S notion of imprecise preferences allows for the possibility that, although the respondent should be able to eliminate some parts of the range, she may end up not being able to say exactly which lottery she regards as equivalent to M_1.

Purely to illustrate the idea, consider one such individual who is fairly confident that she would prefer an L-lottery to M_1 if it were higher on the vertical edge than, say, (40, 0.6; 20, 0.4) and who is also confident that she would prefer M_1 to any L-lottery located lower down the vertical edge than, say, (40, 0.1; 20, 0.9). However, this leaves an interval on the vertical edge between 0.6 and 0.1, depicted in Figure 1a by the bracket, where she is less than sure about her preferences but from within which she is required to identify a point of equivalence. To keep the example simple, suppose that if a sample of people of this kind were each asked to pick some single point from the interval, they would, between them, generate a distribution of points, the median of which (let us say) happens to be the mid-point[3] of the interval, i.e. (40, 0.35; 20, 0.65). For the purposes of the current exposition, let us take this as the 'representative' response and label it L_1* in Figure 1a. If we connect this representative equivalence to M_1 by a straight dashed line, we have a depiction of one indifference curve in the triangle.

Now consider Figure 1b and a corresponding account of trying to identify an equivalent for M_4 (= (40, 0.2; 0, 0.8)).

FIGURE 1b HERE

We can see that L_4^+ and every point above it on the vertical edge dominate M_4, while L_4^- and every point to the right of it are dominated by M_4, so the lottery regarded as equivalent to M_4 must lie somewhere along the stretch of edges between those two points. The typical individual is, let us say, sure she would prefer every L lottery on the permissible section of the vertical edge and anywhere to the left of (20, 0.8; 0, 0.2) and is also sure she would prefer M_4 if the alternative L lottery were worse than (20, 0.3; 0, 0.7). Suppose once again that the representative response is the mid-point of this 'imprecision interval' – in this case, (20, 0.55; 0, 0.45), which we label L_4* in Figure 1b. The dashed line in Figure 1b depicts the relevant representative indifference curve.

These are only examples, of course, and their purpose is simply to illustrate how the intuitions behind the M&S model might be extended to the Marschak-Machina framework; and in so doing, to indicate the *potential* for predicting behaviour within that environment which violates EU in the way typically reported. Whether actual behavior operates broadly along the lines of the example is what the experiment set out to investigate. However, to stay with the example for a moment, a comparison between Figures 1a and 1b shows that the gradient of the slope joining M_1 and L_1* is 2.25, considerably steeper than the slope of 0.57 joining M_4 and L_4*, and consistent with the 'fanning out' pattern characteristic of many data sets.

To be as faithful as possible to the M&S notion, rather than focus exclusively upon the 'representative' point in the middle of the interval, we may allow that a respondent might, on any particular occasion, opt for a point anywhere within the interval. Without imposing some probability distribution over the interval, this entails observing gradients for the indifference curve between M_1 and L_1* drawn from the range between 1 and 3.5, while gradients for the indifference curve between M_4 and L_4* would, in this example, be drawn from the range between 0.33 and 2. So while there is some overlap in these ranges, and there could be pairs of observations where the M_4-L_4* gradient would actually be *steeper* than the M_1-L_1* gradient (which would look like 'fanning-in'), there is clearly more scope for the opposite inequality typical of fanning-out.

For those readers who find the above examples too *ad ho*c, a more general way of thinking about the model may be helpful. Consider first the lottery on the vertical edge which has the same *expected value* as M_1 and so marks the boundary between risk aversion and risk seeking. That lottery is (40, 0.6; 20, 0.4) – call it L_1^{EV}. Four-fifths of the interval between L_1^+ and L_1^- lie below L_1^{EV}, indicating the scope for imprecision to favour equivalences for M_1 that show up as risk *averse*. For M_4, the equivalent lottery on the horizontal edge, L_4^{EV}, is (20, 0.4; 0, 0.6). In this case, four-fifths of the interval between L_4^+ and L_4^- lie *to the left* of that lottery, so that imprecision would be more likely to pull equivalences for M_4 in the direction of risk *seeking*. The specific examples shown in Figures 1a and 1b are just particular cases of the general tendency, consistent with the body of past evidence, for individuals to give responses which look risk averse in the middle and upper part of the triangle but appear to be risk seeking in the bottom right hand corner.

The way in which the same notion of imprecision might also account for violations of betweenness can be demonstrated in conjunction with Figure 2.

FIGURE 2 HERE

Consider the case where the values of x_3, x_2 and x_1 are such that the representative equivalent L_2* for M_2 is the lottery (x_2, 0.9; x_1, 0.1) on the bottom edge, so that the straight line connecting L_2* to M_2 passes through M_5. Betweenness entails that the representative equivalent for M_5 should also be L_2*. However, transparent dominance constrains the interval for M_5 to lie somewhere inside the range from L_5^+ (= (x_3, 0.2; x_2, 0.8)) to L_5^-, (= (x_2, 0.8; x_1, 0.2)). Thus three times as much of the 'permissible' range lies to the left/above L_2* as lies to the right of it. Of course, this does not necessarily mean that the imprecision interval for L_5* will reflect those exact proportions; but on the other hand, there is clearly much more scope for the bulk of that interval to lie to the left of L_2*, in which case M_5 would appear to be preferred both to M_2 and to L_2*, a result which would violate betweenness predominantly in the direction of convexity, as has often been reported – see, for example, Camerer (1995).

Those who are accustomed to thinking in the way that economists are trained to think, whereby individuals are supposed to make their choices on the basis of reasonably stable, well-articulated and self-contained preferences, might be uncomfortable with the implication that the limits and range of the permissible interval may play an influential role in shaping patterns of response. However, such an implication is consistent with an established body of psychological evidence.

For example, Parducci and Wedell (1986) discussed range-frequency effects, whereby people's judgments of the values of items could be influenced by their ranking in whatever range and/or set of other items they were embedded. If respondents had clear and precise preferences, their judgments would be much less susceptible to changes in the range or distribution of other alternatives. But if preferences are somewhat imprecise, it may be a reasonable strategy for decision makers to look for points of reference – such as alternatives which dominate or are dominated – that eliminate some possibilities and facilitate a greater focus. Thus it may not be surprising to find those 'sure' points of reference exerting some influence.

Inconvenient though it may be from a normative perspective or for the enterprise of building general all-purpose models, such influences do appear to play a role even in the controlled conditions of laboratory experiments involving lotteries with no more than two or three modest monetary sums and relatively straightforward probabilities. For example, Bateman et al (2007) found clear evidence of range-frequency effects when certainty equivalents were inferred from ranking exercises. They describe (p.52) how two lotteries, labelled I and J, were ranked in two separate sets, one of which was composed of other lotteries that were generally more attractive than I and J, while the other set contained lotteries that were mostly less attractive. The same sure sums were included in both sets, but the differences in the distributions of other lotteries resulted in the inferred certainty equivalents of I and J being some 50% higher when included in the less attractive set than when the majority of other lotteries were more attractive. In a somewhat different task, Blavatskyy and Kohler (2009) found strong evidence of range effects when using the Becker, DeGroot & Marshak (BDM) valuation mechanism for lotteries by comparing a restricted and unrestricted interval from which the valuations could be drawn.

To summarise, then, the main implications and issues of interest in our application of the M&S model to the present context are as follows:

a) Most individuals are liable to manifest a degree of imprecision in their statements of their preferences, reflected by an 'imprecision interval' for each M-lottery.

b) In the vicinity of the L_i^+ lottery that dominates M_i and also in the vicinity of the L_i^- lottery that is dominated by M_i, individuals will express high confidence in their choices, but for most individuals there will typically be a non-trivial imprecision interval.

c) While the average size of the imprecision intervals might be strongly influenced by the ranges between L^+ and L^-, we were interested to see whether other features might exert some influence. For example, while the lengths of the lines that constituted the 'permissible' range were the same for all of M_1-M_4, the distance between hypotenuse and edge varied, being greater for M_2 and M_3 than for M_1 and M_4. If these distances are related to the (dis)similarity between lotteries, and if dissimilarity adds to uncertainty about preferences (see Buschena and Zilberman (1999) for a discussion of this possibility), we might see this reflected in the widths of the imprecision intervals. Likewise, increasing the x_3 payoff while holding x_2 and x_1 constant may be regarded as generating greater dissimilarities, and this might also correlate with imprecision intervals.

d) The L_i-lottery giving the same expected value as an M_i-lottery may lie much closer to L_i^+ for some M_i, and closer to L_i^- for other M_i. This asymmetry may cause choices to appear more (less) risk averse in some evaluations than in others.

The next section describes the experimental design intended to investigate how far the various possibilities outlined above are manifested in actual behaviour. Section III reports the results and concludes with a discussion of their interpretation and possible implications.

II. Design and Implementation of the Experiment

The design was built around two Marschak-Machina triangles. One of these was described in the previous section; the other was the same in every respect except that x_3 was set at \$60 rather than \$40. Respondents were allocated at random to one or other of the two triangles.

Our first objective, in Stage 1 of the experiment, was to get respondents to compare each of the fixed lotteries M_1-M_5 with a series of alternative L lotteries located on the vertical and horizontal edges of the triangle, and to identify the point at which they switched between the M lottery and the L alternative.

To illustrate how we did this, take the case where the fixed lottery was M_2 in the \$60 sub-sample: that is, it offered a 0.60 chance of \$60 and a 0.40 chance of 0. This lottery was presented on a computer screen as option A. The alternative, option B, was a lottery on the vertical or horizontal edge. For half of each sub-sample (again, determined at random), B was initially located on the vertical edge at L_2^+: so in this case, for that half of the sub-sample, B initially offered a 0.60 chance of \$60 and a 0.40 chance of \$20. For the other half of the sub-sample, B was initially located on the horizontal edge at L_2^-, offering a 0.60 chance of A\$20 and a 0.40 chance of 0.

Because we were interested in the role of imprecision in explaining behaviour, we set out not only to identify respondents' stated preferences but also aimed to obtain some measure of the confidence with which those preferences were recorded. To that end, respondents were asked to respond in one of four ways, 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. (The instructions, available on request, explained the terms "definitely prefer" and "probably prefer" in more detail.)

To illustrate how this worked, consider first a respondent initially presented with a choice between M_2 (i.e., A) and L_2^+ (i.e., B). Since B here dominates A, almost every respondent signified a definite preference for B, coded as 4. Once the initial response had been recorded, the computer program changed B, making it two points worse: that is, displaying a lottery which offered ($60, 0.58; $20, 0.42) instead of the initial ($60, 0.6; $20, 0.4). The respondent was then asked again to state their preference and the confidence with which they held it. Thereafter, B was made progressively worse, so that it moved steadily down the vertical edge, reducing the chances of $60 and increasing the chances of $20, until B reached the corner (the certainty of $20), after which point B moved along the horizontal edge until it eventually became L_2^-, where the procedure came to an end.

So for those starting on the vertical edge and initially recording 4's, there came a point at which they indicated that they still chose B but no longer felt so sure, coded as 3. As B was degraded further, there came a point at which the respondent switched from B to A: if this was initially a 'probable' preference for A, it was recorded as 2; when, after further degradation of B, it became a definite preference for A, it was recorded as 1.

We refer to the treatment where B initially dominated A, and then was progressively degraded, as 'iterating down'. For the other half of the subsample, B was initially set at L_2^- and the program progressively improved it, in effect moving B leftwards along the horizontal edge towards the corner, then up the vertical edge until it became L_2^+. We refer to this treatment as 'iterating up'. Within a sub-sample, the same direction of iteration was used for all five fixed lotteries, the only difference being that for M_5 the iteration involved decrements or increments of one point at a time, rather than the two-point changes used for each of the four lotteries on the hypotenuse[4].

In this way, in the course of Stage 1 of the experiment, we elicited from each respondent their implied point of indifference between the M and L lotteries (the 2↔3 switchpoints) and also some indication of the intervals (between the 1↔2 switchpoint and the 3↔4 switchpoint) over which they considered themselves to be less than sure about their preference[5]. We neither claim nor require that this represents the same level of confidence for different respondents. It is necessary only that whatever a particular respondent regarded as the point of transition between a 'definite' and a 'probable' preference in the case of one pair of lotteries would correspond with that same respondent's judgment of their own confidence for the other pairs.

The main objective of these questions in the experiment was to explore how far our extension of the M&S model might be able to account for violations of betweenness and independence. However, the bulk of the existing body of evidence has taken the form of pairwise choice data, so we wanted to see how the patterns yielded in Stage 1 by iterating through a succession of very similar pairwise choices would compare with the usual approach of asking respondents to make a number of separate one-off choices between a variety of predetermined pairs. Note that these pairs were the same for all subjects and fixed in advance, so subjects could not affect their future choices by their responses in Stage 1.

To this end, Stage 2 of the experiment involved presenting each subsample with a set of 20 pairwise choices: that is, 4 B's for each of M_1-M_5, with each B chosen to produce a particular gradient of the line connecting it to A, as shown in Figure 3 for the \$60 triangle. These gradients, which we shall denote by g_1 ... g_4, were as follows:

	g_1	g_2	g_3	g_4
When x_3 = \$40, the gradients were:	**1,**	**1⅔,**	**2½,**	**5**
When x_3 = \$60, the gradients were:	**½,**	**1,**	**2,**	**4**

Because a risk-neutral subject's indifference curves would have a gradient of 1 when x_3 = \$40 but a gradient of ½ when x_3 = \$60, it was necessary to use a different set of gradients in each triangle to obtain useful data from each.

FIGURE 3 HERE

We know of no way of making incentive compatible the distinction between stating a 'definite' preference and stating a 'probable' preference. 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. 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 questions carefully. So is it right 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 incentives in decision experiments is beyond the scope of this paper (though see Bardsley et al., 2009, chapter 6), so we address our remarks to the specific question of the usefulness of our imprecision data for understanding the phenomena central to 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, 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 which 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 numbers of steps in the iterative procedure, irrespective of the nature of the lottery: for example, the 1↔2 switch-point for M_1-M_4 was typically 25% (in the $40 triangle) and 30% (in the $60 triangle) of the distance from the bottom of the iterative range, while the 3↔4 point was 50% of the way along for the $40 triangle and 55% for the $60 triangle.

The questions relating to betweenness and independence alternated with questions investigating preference reversals, as reported in Butler and Loomes (2007). So we are able to examine whether switchpoints were sensitive to the differences between the 'triangle' questions and the preference reversal questions. They were: the 1↔2 switch-point for the certainty equivalent of the $-bet was typically between 15% and 20% of the distance from the bottom of the iterative range, while the 3↔4 point was just over 40% of the way along (i.e. less than halfway); while 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%-80% of the way along (depending on the direction of iteration). Given the different probabilities of winning offered by these bets, 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. 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. (For further discussion of this issue, see Butler and Loomes (2007, pp.293-294)).

We *were* of course able to make the Stage 2 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, and they would each be paid according to the way their decision in that particular question worked out. There were 43 such questions, (twenty of direct relevance to this paper) so 1 in 43 choices was played for real. This is entirely within the range of usual practice regarding the random lottery incentive system (see Starmer and Sugden, 1991). Average earnings in our study were $26, ranging from a low of $0 to a high of $160. Moreover, as will become apparent in the next section, the data from these incentive-compatible questions exhibited essentially the same overall patterns of behaviour as displayed by responses to the Stage 1 elicitations of switchpoints and imprecision intervals.

A total of 89 individuals drawn from a broad cross-section of students and staff at the University of Western Australia took part. Verbal and on-screen explanations plus on-screen practice questions for both stages introduced the experiment. 45 participants were allocated at random to the $40 triangle (of whom 23 iterated down and 22 iterated up in Stage 1) and 44 to the $60 triangle (with equal numbers iterating in each direction).

As listed at the end of Section I, the main issues we hoped that the data would illuminate were as follows. First, do people typically have non-trivial imprecision intervals (i.e. between 1↔2 and 3↔4)? And if so, what determines the widths of these imprecision intervals? And can their positioning within the permissible range offer an account of systematic violations of betweenness and independence?

III. Results

We begin with the Stage 2 pairwise choice data. Table 1 reports, for each triangle and each gradient, the numbers of respondents who chose the riskier M lottery.

TABLE 1 HERE

Reading *down* the rows, things were much as virtually every model would lead us to expect: as the gradient increased, the L lottery became less favourable, and more respondents chose M.[6]

However, when we read *along* the rows from left to right, we find the kind of patterns consistent with the 'usual' departures from EU. There is not much to see in the top row for each triangle, but for the other three gradients in each triangle, there was a clear trend for the numbers of M choices to increase as the lotteries move towards the bottom right hand corner. A within-subject analysis of those choices shows that the numbers choosing L_1 and M_4 outnumbered those choosing M_1 and L_4 to an extent that was significant at the 1% level in all six comparisons (using a McNemar exact binomial test). This is consistent with a significant degree of fanning out. Indeed, at the level of the individual, 45 of the 89 subjects satisfied *strict* fanning out, while just two satisfied strict fanning in.

Violations of betweenness were also in evidence, directly and indirectly. In the case of the $60 triangle, there were two direct tests. When the gradient was ½, the straight line joining M_3 to its L counterpart passed through M_5, and when the gradient was 2, the straight line joining M_2 to *its* L counterpart also passed through M_5. In the first of these cases, there were very few M choices at all, so the fact that 10 out of 44 chose M_5 over L as opposed to 6 choosing M_3 over the same L is consistent with convex indifference curves but is not a statistically significant difference. However, as Table 1 reports, when the gradient was 2, 32 out of 44 respondents chose M_5 compared with just 21 who chose M_2, and this asymmetry in the direction consistent with convex indifference curves was significant at the 1% level.

In the \$40 triangle, the tests were less direct, but the results pointed to the same conclusions. Consider first M_2 and the straight line with gradient 1⅔ linking it to $L_2 = (x_2, 0.96; 0, 0.04)$. This line passes just to the left of M_5 – it goes through $(x_3, 0.2; x_2, 0.64; 0, 0.16)$ – but there is little space between it and the straight line of the same gradient joining M_5 to $L_5 = (x_2, 0.92; 0, 0.08)$. Yet there were 22 respondents choosing M_5 over that L_5 as opposed to just 9 choosing M_2 over the corresponding L_2, with the within-subject asymmetry registering as significant at the 1% level. In case this might be attributed to some very acute fanning out in that thin slice of the triangle, consider M_2 and the gradient 2½ which links it to $(x_2, 0.84; 0, 0.16)$. This straight line passes to the *right* of M_5 through $(x_3, 0.2; x_2, 0.56; 0, 0.24)$ – that is, by the same distance to the right that the previous line passed it to the left – so that any fanning out effect while maintaining linearity might be expected to favour M_2 more than M_5. But once again Table 1 shows that M_5 was chosen much more often – by 31 as opposed to 17 respondents; and once again the asymmetry was significant at the 1% level.

Overall, then, when viewed from the perspective of deterministic models, the patterns of choice in Stage 2 appear entirely consistent with a model of convex indifference curves fanning out as if from some point to the south-west of their respective triangles. But how far do such patterns also show up in the Stage 1 data? And to what extent were they predicted by the sorts of propositions about imprecision discussed earlier?

Table 2 shows the data analogous to those in Table 1, but this time drawn from individuals' responses to the Stage 1 iterative procedure. The one additional complication is that with the Stage 1 procedure we may occasionally observe the 2↔3 switching point coincide with the relevant pre-set L lottery. In such cases, we have counted this as half a choice of each option. Of necessity, the data in Table 2 must be more regular than those in

Table 1 when it comes to reading down the columns, with at least as many choices of M at steeper gradients as at shallower ones. However, the important issue is the pattern reading along the rows. And as far as fanning out is concerned, the picture here is even sharper than it was in Table 1: for all four gradients in both triangles, the differences between M_1 and M_4 patterns of choice are significant at the 1% level.

TABLE 2 HERE

The picture is not quite so sharp with respect to violations of betweenness. Making the same comparisons as in Table 1, all four disparities were in the direction consistent with convex indifference curves near the bottom edge. However, once again there were relatively few M choices in the \$60 triangle when the gradient was ½, so that the difference (8.5 of 44 against 7 of 44) was not statistically significant. By contrast, when the gradient was 2, the asymmetry (35.5 of 44 against 20 of 44) was again significant at the 1% level. Meanwhile in the \$40 triangle, the two comparisons between M_2 and M_5 when the lines from M_2 with gradients 1⅔ and 2½ pass either side of M_5 produced one difference that was significant at 10% and another that just failed to be significant at that level.

However, with the iterative procedure we are not confined to looking just at preferred choices: with these data we can not only examine the behaviour of the 2↔3 switch-points, but also the widths and locations of the intervals of imprecision around those points. When reporting these results, we shall refer to the mean index values of the various switching points. These index numbers can be best understood with reference to the triangle as follows: the top left point has the value 100 and the numbers fall to 0 at the right angle, and then become progressively more negative as we move along the bottom edge, with the bottom right corner taking the value -100.

Table 3 reports the mean switch-points for all M lotteries in both triangles, as well as (in bold) the implied gradients of straight lines connecting the M lotteries to their respective mean switch-point L lotteries. The mean intervals between the 3↔4 and the 1↔2 switch-point L lotteries are also computed. What do these data show?

TABLE 3 HERE

We begin by considering the lotteries on the hypotenuse of both triangles. The gradients from these M's to their 2↔3 switch-point L lotteries get progressively less steep as we go from M_1 to M_4: in the A$40 triangle, the gradients fall from 3.25 to 0.68, while in the $60 triangle, the corresponding fall is from 2.89 to 0.56. This pattern of strict fanning-out corresponds with the patterns of choice reported in Tables 1 and 2.

The data also enable us to see that, holding the permissible range constant, the distance between the fixed lottery and the edge upon which the equivalence response is recorded does not seem to systematically influence the width of the imprecision interval. If we use the length of the straight lines connecting M_1-M_4 to their respective 2↔3 switch points as a rough estimate of that distance[7], then in both triangles M_4 would be closest to the relevant edge, followed by M_3, then M_1, with M_2 furthest away, these last distances being between 75% and 100% greater than those for the respective M_4's. But the widths of the intervals between 1↔2 and 3↔4 did not follow that pattern. Rather, as Table 3 shows, within a given triangle all four interval widths were very similar and there were no significant differences between any two of them. Moreover, in relation to the intervals between L^+ and L^-, the positions of the imprecision intervals were remarkably stable, as Table 4 shows.

TABLE 4 HERE

For M_1-M_4 in the \$40 triangle, the range over which the M lottery is definitely preferred lies between 20.4 and 27.6 points of the corresponding dominated L, while the range over which the L lottery is definitely preferred lies between 50.7 and 56.3 points of the L which dominates M. So as we go from M_1 to M_4 and as the positions of L^+ and L^- and the 100-point ranges between them shift, so too do the positions of the intervals between 1↔2 and 3↔4 and the 2↔3 switch-points. The \$60 triangle exhibits similar behaviour, except that, with x_3 being larger, the ranges over which M is definitely preferred are wider and the ranges over which L is definitely preferred are narrower. So increasing x_3 while keeping x_2 and x_1 constant had the effect of increasing the *widths* of all of the imprecision intervals and systematically shifted the location of those intervals. At the same time, the *positions* of those intervals within each L^+ to L^- range were as stable for the \$60 triangle as for the \$40 triangle.

The finding that the widths of the imprecision intervals are more a function of the L^+ to L^- range than of the distance from an M to the equivalence edge is given further support by the data relating to M_5. Table 4 reports the actual intervals in the bottom row; but just above, in the row labelled (M_5), these are converted to percentages to make them comparable with the M_1-M_4 data. This shows that, as proportions of the relevant L^+ to L^- range, all of the imprecision intervals within the same triangle are of much the same magnitude: 21-23 percentage points for the \$40 triangle, 25-28 percentage points for the \$60 triangle.[8]

What the M_5 row also shows is a tendency for the position of the imprecision interval to be shifted somewhat relative to its position for M_1-M_4: with M_5, a relatively larger proportion of the range is associated with a definite preference for M and a correspondingly smaller proportion represents a definite preference for L. This is in line with our conjecture that for M_5 the imprecision interval would be pushed in a clockwise direction, producing an effect that looks like a violation of betweenness consistent with convex indifference loci near the bottom edge.

Figures 4 and 5 depict the overall patterns of responses in terms of lines from the M lotteries to their respective 2↔3 switch point L lotteries. These figures are based on the mean values of 45 and 44 subjects respectively. For simplicity, these lines are drawn straight, but the fact that the M_5 line has a shallower slope than might be extrapolated from its position relative to M_2 and M_3 – and indeed, the fact that in the $60 triangle the M_5 line actually crosses the M_2 line – suggests that if one were to wish to impose an indifference map of the kind typical of deterministic theories, the best fit for an 'average' subject would be one which involved curves that are convex near the bottom edge and fanning out from the south-west of the triangle: that is, the kind of configuration which some RDEU models are able to generate.

IV. Concluding Remarks

It might be argued that the violations of independence and betweenness discussed in this paper can be explained just as well by one of the class of RDEU models as by the model of imprecision we have investigated. So what are the arguments for and against adding this imprecision model to the existing catalogue of theories?

The first argument *for* so doing is that although RDEU models may be *technically* capable of accommodating these patterns, they could be regarded as behaviourally implausible: in particular, the process of converting probabilities into the decision weights required to fit the data is one which is quite complex (as anyone who has tried to teach the notion to students will know). By contrast, the imprecision model is behaviourally very simple – indeed, for some theorists' tastes it has, if anything, too *little* structure. Despite its modest assumptions, it can make refutable predictions.

Second, the imprecision model can explain other phenomena – most notably, the two opposite forms of the preference reversal phenomenon described in Butler and Loomes (2007) – which RDEU models simply cannot deal with. While it may be that ultimately no single model can be expected to account for *all* behaviour in all contexts, it does seem reasonable to expect one model to capture the key phenomena generated by the same subjects performing similar tasks in the course of a single experiment. A relatively simple descriptive model of imprecision is able to accommodate a variety of 'anomalies' that have defied capture by any one of the many alternative deterministic models developed to date.

There are limitations, of course. For example, it is not (yet) obvious how to apply the model to decisions involving more than three payoffs, or where there are larger choice sets. Nevertheless, what the model and the data presented above clearly suggest are that imprecision *is* a feature of many people's preferences and that there *is* some potential for explaining regularities in behaviour in terms of such imprecision. There may be a good deal more work to be done to investigate the scope and limitations of such models, but the present paper, in conjunction with Butler and Loomes (2007), gives grounds for believing that this may be a useful line of enquiry to pursue.

Figure 1a

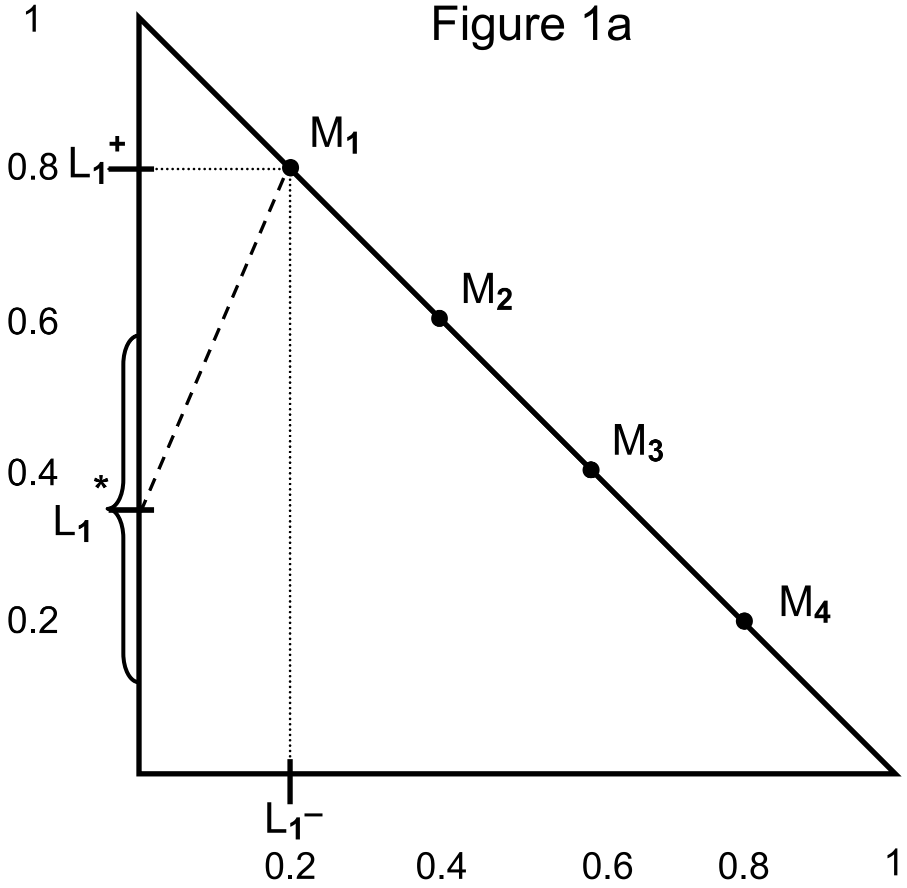

Figure 1b

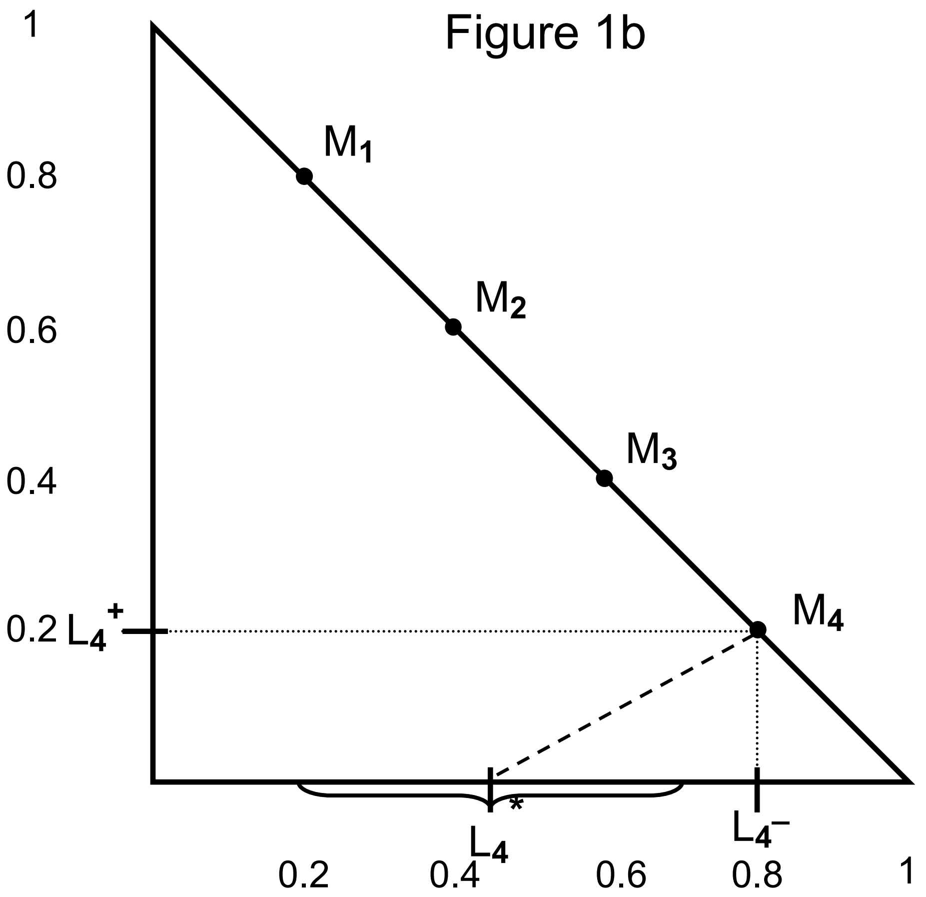

Figure 2: Illustration of Violation of Betweenness

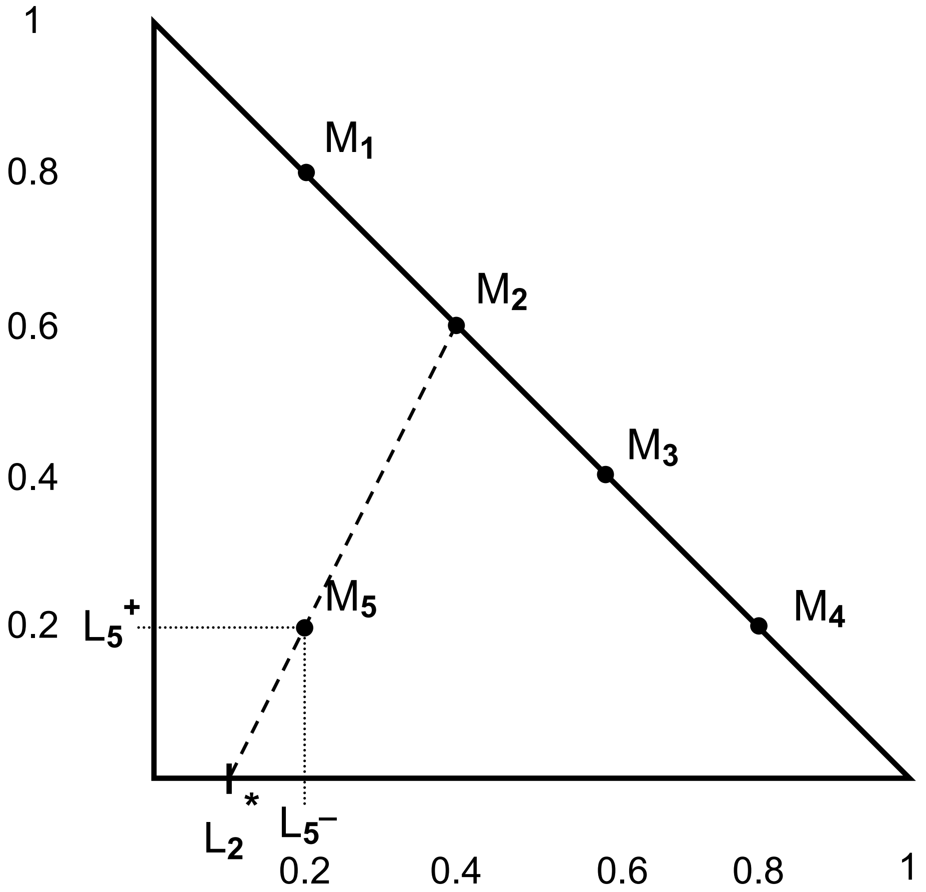

Figure 3: The Gradients for Pairwise Choices in the $60 Triangle ($M_5$ omitted for clarity)

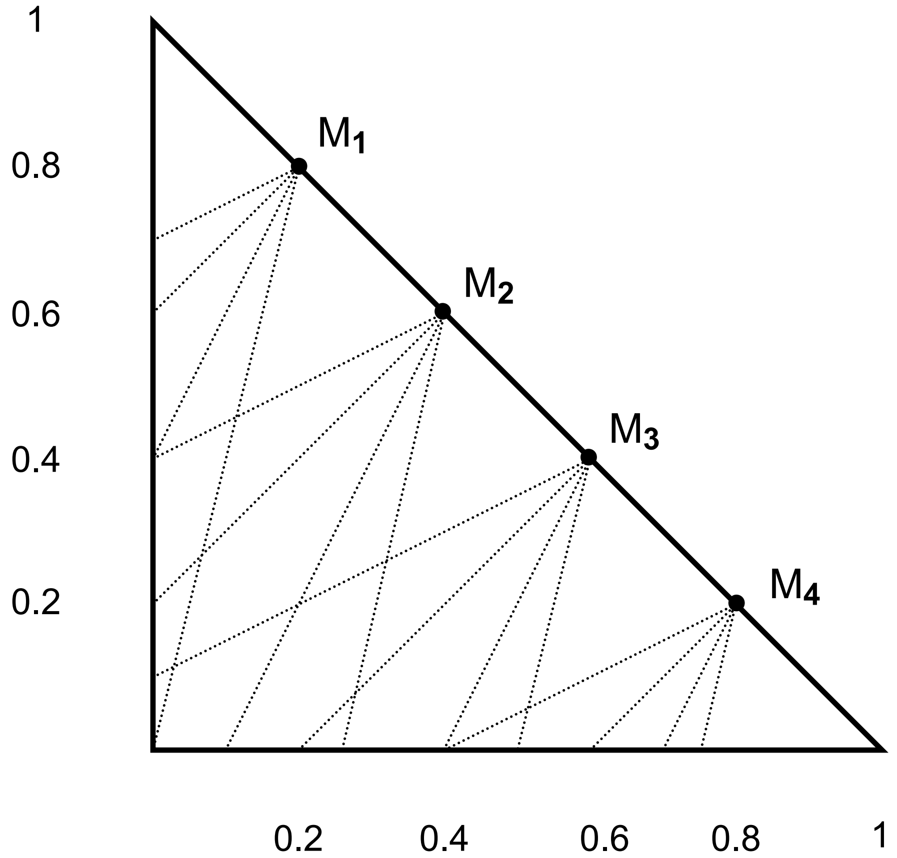

Figure 4: The Fitted Lines from the 2↔3 Switch-points in A$40 Triangle

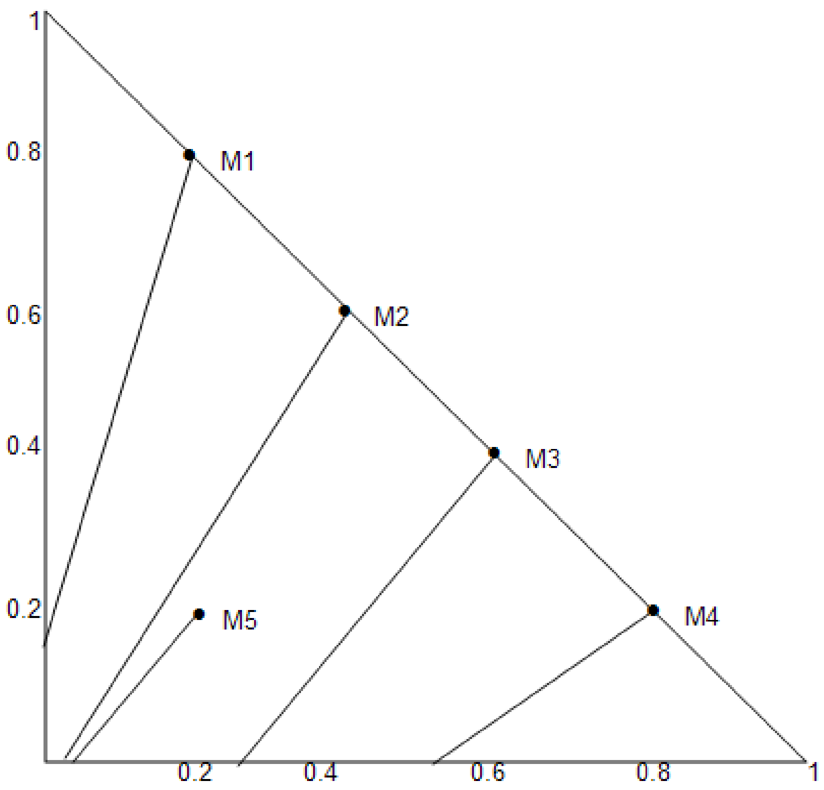

Figure 5: The Fitted Lines from the 2↔3 Switch-points in A$60 Triangle

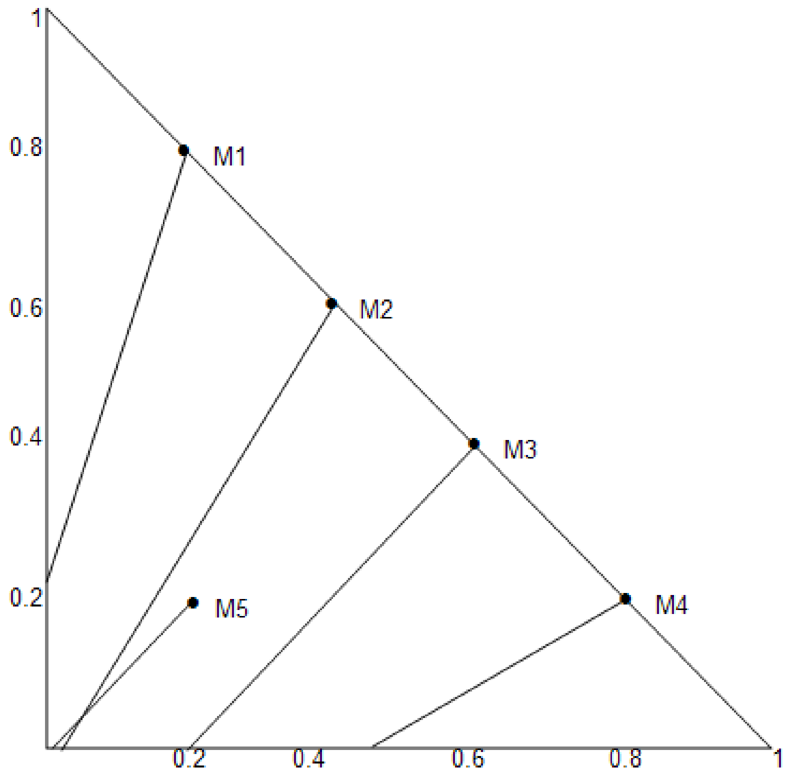

Table 1: Numbers of M Choices in Stage 2					
	M_1	M_2	M_3	M_4	M_5
x_3 = A$60					
g_1 (=½)	5	2	6	4	10
g_2 (=1)	8	11	18	25	24
g_3 (=2)	17	21	34	36	32
g_4 (=4)	26	35	36	41	33
x_3 = A$40					
g_1 (=1)	12	5	8	8	13
g_2 (=$1^{2/3}$)	8	9	18	25	22
g_3 (=$2^{1/2}$)	11	17	29	36	31
g_4 (=5)	26	34	38	43	36

Table 2: Numbers of M Choices Inferred from Stage 1					
	M_1	M_2	M_3	M_4	M_5
x_3 = A$60					
g_1 (=½)	1.5	3	7	18	8.5
g_2 (=1)	4.5	13	20.5	34	20
g_3 (=2)	10.5	20	36.5	40	35.5
g_4 (=4)	31.5	36	41	43	41
x_3 = A$40					
g_1 (=1)	3.5	12	14	27	15.5
g_2 (=1⅔)	7	18.5	30	38	26
g_3 (=2½)	13	25	35	41	32
g_4 (=5)	30.5	40	43.5	44	36.5

Table 3: Mean Switch-points, Gradients and Imprecision Intervals					
	M_1	**M_2**	**M_3**	**M_4**	**M_5**
x_3 = A$40					
3↔4 switch-point	27.07	9.33	-13.07	-36.27	1.60
3↔4 **gradient**	**2.65**	**1.27**	**0.85**	**0.46**	**0.92**
2↔3 switch-point	15.02	-3.29	-26.04	-50.58	-3.96
2↔3 **gradient**	**3.25**	**1.63**	**1.18**	**0.68**	**1.25**
1↔2 switch-point	5.91	-12.36	-34.18	-59.60	-7.69
1↔2 **gradient**	**3.70**	**2.17**	**1.55**	**0.98**	**1.62**
1↔2 to 3↔4 interval	21.16	21.69	21.11	23.33	9.29
x_3 = A$60					
3↔4 switch-point	35.18	15.18	-4.73	-27.32	4.80
3↔4 **gradient**	**2.24**	**1.12**	**0.72**	**0.38**	**0.76**
2↔3 switch-point	22.27	-1.55	-19.50	-44.59	-1.11
2↔3 **gradient**	**2.89**	**1.56**	**0.99**	**0.56**	**1.06**
1↔2 switch-point	9.36	-10.09	-31.95	-54.77	-6.64
1↔2 **gradient**	**3.53**	**2.01**	**1.43**	**0.79**	**1.50**
1↔2 to 3↔4 interval	25.82	25.27	27.23	27.45	11.43

Table 4: Widths of 'Definite' and 'Imprecise' Intervals							
	x_3 = A$40				**x_3 = A$60**		
	L^- to 1↔2	1↔2 to 3↔4	3↔4 to L^+		L^- to 1↔2	1↔2 to 3↔4	3↔4 to L^+
M_1	25.9	21.2	52.9		29.4	25.8	44.8
M_2	27.6	21.7	50.7		29.9	25.3	44.8
M_3	25.8	21.1	53.1		28.1	27.2	44.7
M_4	20.4	23.3	56.3		25.2	27.5	47.3
(M_5)	(30.8)	(23.2)	(46.0)		(33.4)	(28.6)	(38.0)
M_5	12.3	9.3	18.4		13.4	11.4	15.2

(All figures rounded to one decimal place)

References

Bardsley, N., Cubitt, R., Loomes, G., Moffatt, P. and Starmer, C. (2009), *Experimental Economics: Rethinking the Rules*, Princeton University Press.

Bateman, I., Carson, R., Day, B., Hanemann, M., Hett, T., Hanley, N., Jones-Lee, M., Loomes, G., Mourato, S., Ozdemiroglu, E., Pearce, D., Swanson, J. and Sugden, R., (2002), *Economic Valuation with Stated Preference Techniques: A Manual*, Edward Elgar, Cheltenham.

Bateman, I., Day, B., Loomes, G. and Sugden, R. (2007) Can Ranking Techniques Elicit Robust Values? *Journal of Risk and Uncertainty*, **34**, 49-66.

Blavatskyy, P. (2009) Preference Reversals and Probabilistic Decisions, *Journal of Risk and Uncertainty*, **39**, 237-250.

Blavatskyy, P. and Kohler, R. (2009) Range Effects and Lottery Pricing, *Experimental Economics*, **12**, 332-349.

Buschena, D. and Zilberman, D. (1999) Testing the Effects of Similarity on Risky Choice: Implications for Violations of Expected Utility, *Theory and Decision*, 46, 253-80.

Butler, D. and Loomes, G. (2007) Imprecision as an Account of the Preference Reversal Phenomenon, *American Economic Review,* **97**, 277-297.

Camerer, C. (1995) Individual Decision Making. In *Handbook of Experimental Economics*, eds Kagel, J. and Roth, A., Princeton University Press.

Kahneman, D. and Tversky, A. (1979) Prospect theory: An Analysis of Decision under Risk, *Econometrica*, **47**, 263-291.

Loomes, G. (2005) Modelling the Stochastic Component of Behaviour in Experiments: Some Issues for the Interpretation of Data, *Experimental Economics*, **8**, 301-323.

Loomes, G. (2006) (How) Can We Value Health, Safety and the Environment? *Journal of Economic Psychology*, **27**, 713-36.

Loomes, G. and Sugden, R. (1982) Regret Theory: An Alternative Theory of Rational Choice under Uncertainty, *Economic Journal*, **92**, 805-824.

Machina, M. (1982) Expected Utility Theory without the Independence Axiom, *Econometrica*, **50**, 277-323.

McCrimmon, K. and Smith, M. (1986) Imprecise Equivalences: Preference Reversals in Money and Probability, *Working Paper 1211*, University of British Columbia.

Parducci, A. and Wedell, D. (1986) The Category Effect with Rating Scales: Number of Categories, Number of Stimuli, and Method of Presentation, *Journal of Experimental Psychology: Human Perception and Performance*, **12**, 496-516.

Starmer, C. (2000) Developments in Non-Expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk, *Journal of Economic Literature*, **38**, 332-82.

Tversky, A. and Kahneman, D. (1986) Rational Choice and the Framing of Decisions, *Journal of Business*, **59**, 251-278.

Wilcox, N. (2008) Stochastic models for binary discrete choice under risk: A critical primer and econometric comparison. In J. C. Cox and G. W. Harrison, eds., Research in Experimental Economics Vol. 12: Risk Aversion in Experiments pp. 197-292. Bingley, UK: Emerald.

Footnotes

---

1 Regret theory *can* account for some violations of independence such as the 'common ratio effect', but only by assuming statistical independence between the alternatives (see Loomes and Sugden, 1982). However, many experiments have shown that even when the juxtaposition of payoffs is such that regret effects are controlled for, the common ratio effect continues to be manifested to a degree that regret theory cannot account for.

2 All payoffs were in Australian dollars.

3 The data in fact show a slight asymmetry in the average size of the imprecision intervals above and below the 2↔3 switch point, of between 0 and 3 percentage points, so preferences are slightly more risk-averse than the mid-point of the interval would imply. However this does not affect the thrust of the argument given here.

4 This was intended to reduce the difference between the procedure used for M_5 and that used for the other four lotteries, although some disparity remained: i.e. the M_5 procedure involved only 40 changes of B rather than the 50 changes involved in the course of eliciting responses for M_1-M_4.

5 A respondent who felt no sense of uncertainty could, of course, switch from 4 to 1 (or vice-versa) without ever recording either 2 or 3. A few (male) subjects consistently did just this.

6 The only exception to this was M_1 in the A$40 triangle, where more respondents chose M_1 when the gradient was 1 than when the gradient was either 1⅔ or 2½. This is a case for which we have no explanation except chance aberrations. Further evidence that the top left cell *was* aberrant comes from comparing that whole g_1 row with the g_2 row when x_3 = A$60. Since the gradient was 1 in both cases, the M lotteries should have been chosen by more respondents in the A$60 sub-sample. This was what happened for M_2-M_5, where in each case the number of M choices was about two or three times higher; but not for M_1.

7 We are not ruling out that the indifference loci might, in fact, be convex; but moderate convexity would not alter the general result.

9 It also turns out that these intervals are of a similar magnitude to those found for the preference reversal study: 24 points for the certainty equivalents of the $-bet and P-bet; 28 points for the probability equivalents, when converted to percentages of the non-dominated range.

Acknowledgements

This research was funded by a grant from the University of Western Australia Business School, and by the University of Newcastle, UK. We wish to thank Peter Hill for writing the software for the experiment and Andrew Williams for help in running the sessions. We also thank seminar participants at various universities for their comments on earlier versions.