Understanding Preference Imprecision

Oben K. Bayrak and John D Hey

Abstract

The term 'preference imprecision' seems to have different meanings to different people. In the literature, one can find references to a number of expressions. For example: vagueness, incompleteness, randomness, unsureness, indecisiveness and thick indifference curves. Some of these are theoretical constructs, some are empirical. The purpose of this paper is to survey the various different approaches and to try to link them together: to see if they are all addressed to the same issue, and to come to some conclusions. In the course of this survey, we report on evidence concerning the existence of preference imprecision, and its impact on theoretical and empirical work.

JEL classification: D0, D80, D81

Keywords: preference imprecision, stochastic models, anomalies, incompleteness, equivalence interval, vagueness, strength of preference.

Oben K. Bayrak: Centre for Environmental and Resource Economics, Swedish University of Agricultural Sciences 90183, Umeå, Sweden, obenbayrak@sabanciuniv.edu.

John D Hey: Department of Economics and Related Studies, University of York, Heslington, York Y10 5DD, UK, john.hey@york.ac.uk

Acknowledgements

We have benefited from the helpful comments of Ilke Aydogan, Glenn Harrison, Eline van der Heijden, Scott Lemos, Yudistra Permana and Nathaniel Wilcox. This work was supported by the Tore Browaldh and the Jan Wallanders Foundations.

1. Introduction

It may seem odd, in a paper allegedly surveying economists' contributions on preference imprecision, to start by saying that it is not clear what is meant by it: different people have different perceptions and definitions. Yet, that is where we start: by looking at different interpretations provided by different strands of literature. Hopefully, by the end, precise definitions and interpretations will be clear. In addition, we hope to show the connections between them, where they exist.

Let us begin by giving examples of the use of the word ‘imprecision’ when used in conjunction with the word ‘preferences’. Butler and Loomes (2007) write “Many individuals' choices and valuations involve a degree of uncertainty/imprecision”. Cubitt *et al* (2015) write, in a paper entitled *On Preference Imprecision*, that “Recent research invokes preference imprecision to explain violations of individual decision theory “and “[that there is] convincing evidence that individuals will often express imprecision in their preferences, when allowed to do so”.

Take the quotation from Butler and Loomes. What do they mean by this? Choices presumably are just that – choices. People make a choice. What does it mean that the choice “involves a degree of uncertainty/imprecision”. One interpretation is that, if the person was asked to make exactly the same choice on another occasion, then that second choice might be different from the first choice. A second interpretation is that the person, when making the choice, feels uncertain as to what is the ‘correct’ choice (correct from the point of view of his or her preferences). The first interpretation is observable, the second is not – we cannot measure the uncertainty present in his or her choice (unless we have some sophisticated neuro-imaging technology and can use it to detect unsureness). The same is true of valuations – but much depends on how people are ‘allowed’ to express their choice or their valuation. This brings us to the second quote – that from Cubitt *et al*: what does “when allowed to do so” mean?

This creates a conundrum for experimental economists – particularly those who believe in appropriate incentives. One technique that has been used is to ‘allow’ subjects, not only to choose some option A or some option B, but to choose a 50-50 mixture of the two (for example Cettolin and Riedl 2016). The problem with this is that it essentially elicits a preference for randomisation (though this, of course, could be driven by imprecise preferences). Other experimental economists (for example, Cubitt *et al* 2015) have allowed subjects to state “I prefer A”, or “I prefer B” or “I am not sure”, but without implications for this latter statement.

At this point, we should bring in theorists. Standard economic theory does not allow for choice correspondences that include ‘unsureness’ or ‘imprecision’, since the standard neoclassical assumption is that preferences are complete, and that people should either prefer A, or prefer B, or be indifferent between them. Of course, this latter provides a reason for subjects not to be sure which they prefer, but the chances of it being the case are low, and it is experimentally difficult to detect indifference. One way to try to incorporate preference imprecision, therefore, is to drop the *completeness axiom*. An important discussion of doing this can be found in Aumann (1962). He shows that dropping the completeness axiom implies that

> "We still get a utility function *u* that satisfies the expected utility hypothesis … and *u* still "represents" the preference order … but now in a weaker sense: as before, if *x* is preferred to *y* then *u(x)* > *u(y)*, but the opposite implication is no longer true. … Furthermore, we no longer have uniqueness of the utility."

In an early but unpublished paper, MacCrimmon and Smith (1986) introduce the idea of equivalence intervals instead of the conventional precise certainty equivalent concept. The concept can be seen as a thick form of indifference meaning that individual is unable to state one of the values from the equivalence interval for a lottery confidently as the precise certainty equivalent of the lottery. Yet, the idea of thick indifference curves is problematic due to monotonicity, that is the 'more is better' assumption.

Eliaz and Ok (2006) focus on the meaning of incompleteness from a revealed preference perspective, their central focus is to distinguish indifference from indecisiveness by merely observing choice sets of individuals and articulating imprecise choice correspondences using preference relations. Moreover, they show that such preferences can be represented by utility functions by using a weaker version of (rather than drop) one of the key axioms of revealed preference theory, namely the Weak Axiom of Revealed Preference, replacing it by a *Weak Axiom of Revealed Non-Inferiority*. As the authors say, "The resulting choice theory, thus, allows a 'rational individual' to remain *indecisive* at times."

Manzini and Mariotti (2004) similarly propose a theory, which they call *Vague Expected Utility,* in which preferences are not complete and decisions must be made by a lexicographic process. Essentially the theory proposes that preferences are clear if the alternatives A and B are 'sufficiently far apart', but if they are 'close' (and hence the individual cannot distinguish between them) some kind of heuristic is invoked to resolve the indecisiveness.

This leaves us with a number of questions:

1. Are preferences imprecise? How can we have proof that they are?
2. Can we measure the degree of imprecision?
3. Does decision theory help us to understand what imprecise preference might mean, where they come from, and how a decision-maker with such preferences might reach a single decision?
4. What other stories provide a route to choice from imprecise preferences?
5. Can preference imprecision explain some of the prominent anomalies of standard theory?

In the next five sections, we turn to possible answers.

2. Are preferences imprecise? How can we have proof that they are?

There seems to be what one might call *direct but unincentivised* evidence, and also *indirect* evidence. The former comes from experiments in which subjects were allowed to state 'I am not sure', but in which there are no financial implications of so doing. If there were financial implications, for example saying 'I am not sure' leads to a random payoff, then indicating this implies a preference for randomisation (though, this, possibly, could result from imprecise preferences). Indirect evidence comes in a variety of forms, but mainly in the form of variability of choices. It is widely accepted that, in fitting preference functionals to experimental data, one needs to add in some kind of random element. This could take the form of a tremble (Harless and Camerer 1994), a Fechner error added to the utility difference (Hey and Orme 1994), or randomness added to one or more of the parameters of the preference functional (Loomes *et al* 2002). A tremble does not seem to be a sign of preference imprecision (more a matter of the decision-maker making a mistake), while a Fechner error and random preferences seem to be possible examples of preference imprecision (see Section 3 and Appendix A for a detailed discussion).

2.1 Experimental Studies Related to Preference Imprecision

We now examine experimental studies trying to find direct but unincentivised evidence of preference imprecision. These experiments mainly rely on subjects' self-reporting. Self-reporting is often used in environmental valuation studies and in psychology; however, it is unconventional in experimental economics to use such an unincentivised method. This is an important principle of experimental economics: unlike psychology, intrinsic motivation is not seen as sufficient to elicit true preferences, since, it is not a costly action for subjects to lie about, for example, their offers (see Camerer and Hogarth (1999) for a detailed discussion), if there is no punishment for so doing. Butler and Loomes (2007, 2011) state their doubt that an incentive compatible mechanism can be devised for eliciting imprecise preferences—"*at least, not in a form simple and transparent enough to work without creating additional uncertainty*". A skeptic may argue that subjects might not engage in enough effort to pin their preferences down to a more precise degree in the absence of suitable monetary incentives. Butler and Loomes (2007, 2011) address this argument by further investigating their subjects' "considerable and systematic responsiveness" to the characteristics of the decision problems in their experiment. The results of these experimental studies does not provide ultimate proof but they can be seen as suggestive evidence for further research[1].

Section 2.2 summarises the elicitation methods used in the experimental studies, while Section 2.3 reviews the main findings concerning the characteristics of preference imprecision — such as its prevalence, size and stability.

2.2. Methods of Elicitation

There are mainly three methods used in the literature:

(1) a *Response Table* (Cohen *et al*, 1987; Cubitt *et al*, 2015)

(2) an *Iterative Process* (Butler and Loomes, 2007, 2011; Dubourg *et al*, 1997, 1994)

(3) a *Buyer-Seller Uncertainty Mechanism* (Bayrak and Kriström, 2016).

In (1), subjects are asked to respond to a series of binary choices between a risky prospect and a sure amount by completing a *response table* similar to Table 1 below.

Table 1. Example Response Table

Sure Amount	I definitely prefer the sure amount	Not sure	I definitely prefer the risky prospect
\$0			✓
\$1			✓
\$2			✓
\$3		✓	
\$4		✓	
\$5	✓		
\$6	✓		

For the example in the table, the subject states that he or she prefers some specified risky prospect for the certain amounts 2 and below, whereas the certain amount is preferred for certain amounts 5 and above. The subject's imprecision interval corresponds to the values 3 and 4, for which the subject cannot confidently state a preference between the risky prospect and the sure amount[2]. More recently, Cubitt *et al* (2015) used the same form of the response table. One key difference between the two studies is the payoff rule. For values inside the imprecision interval, Cohen *et al* (1987) randomly determined which of the two options was picked. In contrast, Cubitt *et al* (2015) left the choice to the subjects by asking them to determine a switching point inside the imprecision interval; in fact, in this latter study there were no financial implications for the subjects of stating that they were 'not sure'. In the Cohen *et al* (1987) experiment, subjects were expressing a desire for randomisation[3].

Method (2) relies on an *iterative process*[4]: for example, Dubourg *et al* (1994) uses a numbered disk, which has a small window showing a single value at a time. For each value, subjects state their preference by choosing one of the three phrases: "definitely willing to pay", "definitely not willing to pay", or "not sure". If the response is "definitely willing to pay", the interviewer rotates the disk to reveal a higher value through the window, whereas if the answer is “definitely not willing to pay”, the interviewer reveals a lower amount. The experiment continues until a maximum amount that the subject is definitely willing to pay is reached. The amounts for which the subject chooses the phrase “not sure” corresponds to the imprecision interval.

Butler and Loomes (1988) also uses a type of iterative process[5]. For each two-outcome lottery in Table 2, subjects answered a series of binary choice questions where the second option was a sure amount of money. If a subject chose the risky option, the sure amount is increased in the following question; on the other hand, if a subject chose the sure amount, the sure amount is decreased in the following question. Additionally, subjects were asked to use a cursor to state their confidence about their decision – that is their Strength of Preference (*SoP*). The cursor can be moved to 51 different positions, corresponding to feelings of confidence between “very confident” and “very unsure”. At the moment, the literature is not clear, and indeed mostly silent, about how to incorporate such *SoP* measures in a decision model.

Table 2. Lotteries used by Butler and Loomes (1988)

Lottery	p_1	x_1	p_2	x_2
A1	0.2	30 GBP	0.8	0 GBP
A2	0.4	15 GBP	0.6	0 GBP
A3	0.6	10 GBP	0.4	0 GBP
A4	0.8	7.5 GBP	0.2	0 GBP

Butler and Loomes (2007) elicits valuations for risky prospects using a similar method, which they call the *incremental choice method*. They focus on the preference reversal phenomenon by eliciting value and probability equivalents for a series of P-bets and $-bets. The procedure is very similar to the method described before with a small modification, they include four categories instead of three to describe the subjects’ confidence in their choice: “definitely preferring A”, “probably preferring A”, “probably preferring B” and “definitely preferring B”.

Method (3) improves the incentive compatibility of the elicitation of preference imprecision in a Willingness-to-Pay (WTP) and Willingness-to-Accept (WTA) gap experiment, by using the “*Buyer-Seller Uncertainty*” (BSU) mechanism[6]. This is a modified version of the standard Becker-DeGroot-Marschak mechanism, modified so that subjects are free to state their subjective valuations for a good as either an interval or a precise amount. The crucial point is that a random mechanism assigns them as buyers or sellers after they have recorded their valuations. If they are designated as a buyer, then they buy if the randomly drawn market price is inside or below the stated range, if they are designated as a seller, then they sell if the market price is inside or above the stated range. Subjects are informed about these procedures at the beginning of the experiment. The objective of this experimental design is to identify those subjects who have precise and those who have imprecise preferences.

Under this mechanism, subjects with precise preferences are predicted to state precise valuations, whereas the ones with imprecise preferences are predicted to state a range of subjective valuations. Bayrak and Kriström assume that subjects with precise preferences belong to one of two groups: (1) those whose preferences comply with standard theory, so that their WTA equals their WTP; for these it is optimal to state a precise amount equal to their WTP and WTA; (2) those whose preferences do not comply with standard theory, that is their WTA is higher than their WTP; for these subjects, their expected payoff is negative under this mechanism, and what they should do is to minimise their loss from the experiment (see the appendix in Bayrak and Kriström, 2016). They can either state a range which is between WTP and WTA or a single amount which is the average of these two measures. For example, if they state WTA as their valuation, there is a 0.5 probability of being assigned to a buyer role, and in that case they might end up buying the good at an undesirably high price (P): $WTP < P < WTA$. On the other hand, stating WTP is also not optimal for them because subjects might end up being a seller with a probability of 0.5, and hence give away the good for an undesirably low price. Thus, the optimal response for a subject is to state a precise valuation that is, the weighted average of his or her WTP and WTA, where the weights are the probabilities of being a buyer and a seller (0.5 for each since there is equal chance).

Bayrak and Kriström assume that subjects with imprecise preferences are not able to articulate precise subjective valuations for goods; instead, they can come up with a range of valuations. The authors interpret a range of values as an *equivalence interval* or *thick form of indifference,* as in MacCrimmon and Smith (1986). Therefore, subjects with imprecise preferences are indifferent between the values in the range and the good. Bayrak and Kriström show that it is a weakly dominant strategy for these subjects to state their true range of subjective valuations. The flaw of this design is the following: (i) since it is a weakly dominant strategy for these subjects to state a range of values, some of them might state a precise value from their range. Therefore, subjects who state precise valuations might also include at least a part of the subjects with imprecise preferences (for details see the appendix in Bayrak and Kriström, 2016); (ii) similarly, subjects with precise preferences but having a WTA higher than their WTP might prefer to state a range of values, between WTP and WTA.

So, (i) might cause experimenters to observe lower number of responses in range form than the unknown true amount, whereas (ii) might lead a higher number of range responses.

2.3. Results from Experimental Studies

This section summarises the results of the experimental studies employing one of the three elicitation methods described in Section 2.2. These results provide insights about the prevalence, the size and the stability of the imprecision.

More than half of the subjects in the experiments exhibit imprecision by stating a range of valuations. Cohen *et al* (1987) is one of the early studies that used the response table method. They observed that 10% of the subjects exhibit imprecision. On the other hand, Cubit *et al* (2015) found that 87% of the subjects exhibit imprecision *somewhere*. Bayrak and Kriström (2016) found that more than half of the subjects prefer to state their valuations as intervals when they are allowed to, in line with studies such as Håkansson (2008), Banerjee, and Shogren (2014).

Interestingly the results suggest a positive relationship between the size of the imprecision range and the dispersion of the lotteries. Butler and Loomes (1988) used an iterative elicitation procedure and elicited the certainty equivalents of the four lotteries in Table 2: They find that going from A1 to A4, as the dispersion of the lotteries decreases, the size of the imprecision range also decreases. More recently, Cubitt *et al* (2015) found support for the positive relationship between the size of the imprecision range and the lottery's distance from certainty, which can be seen as analogous to dispersion. This result is intuitively appealing since stating a certainty equivalent for a degenerate lottery is straightforward, but articulating preferences for risky prospects in a precise way is relatively difficult. In addition to distance from certainty, Butler and Loomes (2007, 2011) and Cubitt *et al* (2015) found that the imprecision range is approximately a constant proportion of the difference between the worst and the best outcome (around 25-27%).

However, there seems to be a persistent part of the imprecision that does not decay with repetition. It is reasonable to argue that imprecision might be a transient phenomenon and might decrease as level of deliberation increase as subjects gain experience. Cubitt *et al* (2015) provided tests for the stability of the imprecision, that is, whether the size of the intervals changes with repetition or not; if imprecision is merely a result of errors or unfamiliarity with the experimental mechanism, one should expect it to disappear with repetition and experience. They find no evidence for the imprecision declining with repetition. Their analysis supports the idea that imprecision is stable and not temporary; instead, it seems to be an inherent part of the preferences.

3. Can we measure the degree of imprecision?

If decision-makers have a constant-absolute-risk-averse or a constant-relative-risk-averse utility function, we can define a single measure of risk aversion. It would be useful if we could do the same for a measure of preference imprecision. However, as with the measure of absolute risk aversion, or the measure of relative risk aversion, the measure would rely on the story that is being told, the restrictions on the preference functional, and possibly on the problem to which it was applied. Some approaches, as we will see, might provide a single measure. Other approaches provide multiple measures.

One clear indirect (but not incentivised) measure can be obtained from those experiments who add an additional column ('I am not sure') to the standard Holt-Laury price list. One can simply count the number of times that a subject ticks the 'I am not sure' column. This, as has been shown by Cubitt *et al* (2015) and Butler and Loomes (2007), is not constant but varies across price lists (in an apparently systematic way).

More direct (though clearly not indisputable) measures concern the magnitude of the noise of the stochastic term incorporated when estimating preference functionals, though it would seem that one cannot use the magnitude of the tremble in Harless and Camerer (1994) since this seems to be an indicator of *error* not preference imprecision. One could use the standard deviation of the Fechner error in the strong utility model (used by Hey and Orme (1994) amongst others), or the standard deviation(s) of the preference parameter(s) in the random preference story. For these (assuming a homoscedastic Fechner error or a homoscedastic parameter distribution) one can come up with a single number, which will vary across individuals.

Theories come up with other possible measures. The Aumann (1962) approach would suggest measuring the *degree of incompleteness* of the preference ordering, as would Eliaz and Ok (2006). However, these are, unfortunately, difficult things to measure. The Manzini and Mariotti (2004) might be simpler – one could parameterise and estimate the parameters of their 'vagueness function'. This we discuss in Section 4.

4. Does decision theory help us understand what imprecise preferences might mean, where they come from, and how a decision-maker with such preferences might reach a single decision?

This section assumes that a DM with imprecise preferences has, necessarily, difficulties in choosing a single item from some choice menu. He or she might think of first deciding on a *choice set* from the menu, and then choosing a single item from this choice set. If we could observe the choice set, we could then begin to understand the nature of his or her imprecision, and, for example, understand whether it is really imprecision or simply indifference. This might also help us to answer how we might represent such preferences.

If DMs with imprecise preferences do proceed as discussed above, we need to understand how a DM forms his or her non-singular choice set. For example, how does he or she come up with the imprecision *range* (a non-singular choice set) elicited in the experimental studies as in Cubitt et al (2015). We now see how decision theory might help us answer these questions.

Studies that might provide answers are from a literature expressing doubt on the *necessity of completeness axiom for defining rationality*, such as Eliaz and Ok (2006) and Mandler (2005)[7]. From an empirical perspective, the completeness axiom simply means that individuals are not allowed to say “I am not sure” and “I cannot make a decision”. Yet, there might be some decision problems for which the DM might find it difficult to make a precise judgement and find the options incomparable. Readers should note that although these studies do not use the word *imprecision* explicitly, they use *indecisiveness*, which essentially refers to identical choice correspondences that can be inferred from DMs’ statements such as “I am not sure” or “I cannot make a decision”. Thus, for the rest of the paper we refer to indecisiveness and imprecision implying identical choice correspondences which implies that DM finds the options difficult to compare. In such correspondences, choice sets include all these options for which individual feels indecisive or unsure.

Of course, choice sets might include more than one element in the case of indifference as well. Conventionally indifference is defined as both $x \succsim y$ and $y \succsim x$ are true. Incomparability correspondences are defined by neither $x \succsim y$ nor $y \succsim x$ being true.

Eliaz and Ok (2006), from a revealed preference perspective, propose a solution to the problem of distinguishing indifference from incomparability. Let us quote[8] from Eliaz and Ok (2006) on page 67, where they suppose that the

> “choice correspondence … satisfies:
>
> $c\{x,y\} = (x,y)$, $c\{x, z\} = (x, z)$, $c\{y, z\} = (y)$ and $c\{x, y, z\} = (x,y)$”

They remark that:

> “It is readily checked here that both *(x, y)* and *(x, z)* are […] incomparable pairs, while *(y,z)* is not.”

It will be noted that first and last of these include both *x* and *y*. This can be either because the DM is indifferent between *x* and *y* or finds them incomparable. The check for *x* and *y* incomparability is done by noting that, if we take *x* out of the choice set (*x,z*) from the choice menu {*x,z*} we get left with *z*, while, if we take *y* out of the choice set (*y*) from the choice menu {*y,z*} we get left with nothing. The intuition is as follows: in order to find out whether the DM is indifferent between *x* and *y*, we need choice set data from two other menus: (i) one menu should include *x* and some good(s) *z*, and the other should include *y* and the same other good(s) *z*. Their criterion suggests that if the individual is indifferent between *x* and *y*, the role of *x* in the choice set of the first menu should be the same as the role of *y* in the second; otherwise as in the example above, we say *x* and *y* are incomparable. A similar check can be done for *x* and *z* incomparability. The basic issue here is that the DM does not have complete preferences over *x, y* and *z.*

A natural question arises in decision theory: can we represent such preferences with a preference function? Eliaz and Ok (2006) solves the representation problem by replacing the Weak Axiom of

Revealed Preferences (WARP) with a Weak Axiom of Revealed Non-Inferiority (WARNI) and show that WARNI is sufficient for a choice correspondence to be rationalized by unique regular preference[9].

The contribution of Eliaz and Ok (2006) is ingenious and theoretically important, yet it is difficult to see how it could be implemented experimentally in an incentive compatible way (as discussed in Section 2), as the experimenter would have to infer the choice set from the choice, or observe the choice set. If the latter is not observable, how can the experimenter determine whether the individual is 'tossing up' between several options – except possibly by repeating the question several times? Even then, different choices on different repetitions might result from indifference, or from randomness unconnected with imprecision. Unfortunately, experimental economists seem currently unable to devise an incentive compatible way of eliciting the underlying choice set[10].

Another solution provided in the literature, but again one that relies on the observability of choice sets, is to make a distinction between *psychological* and *revealed* preferences: the former implies individuals' judgements about their utility, the latter is determined by their choice behaviour. Mandler (2005) suggests psychological preferences may not be complete, yet revealed preferences are. This argument is in line with Sen (1997) who writes

> "A chooser, who may have to balance conflicting considerations to arrive at a reflected judgement, may not, in many cases, be able to converge on a complete ordering when the point of decision comes. If there is no escape from choosing, a choice decision will have to be made even with incompleteness in ranking."

Mandler's focus is on sequential choice. He extends choice functions and allow current choices to depend on prior choice sets and decisions. He shows that transitive but incomplete psychological preferences can lead to intransitive but rational revealed preferences in which dominated options are never chosen. One way to relate this idea to the experiments reviewed in Section 2 is perhaps by viewing stated imprecision for a choice between a lottery and a series of sure amounts of money as shown in Table 1: in that example, the subject's imprecision on the 4th and 5th rows can be seen as a reflection of underlying psychological preferences. On the other hand, in Cubitt *et al* (2015) subjects had to refine their choice further by selecting a single line from their unsureness interval. This single line can be seen as revealed preference in Mandler's terminology[11].

How might DMs with imprecise preferences reach a single decision? We start with two models that assume that the individual acts according to the standard view of preferences, *if* there is a sufficient level of utility difference between the options, and otherwise employs some heuristic or rule of thumb. Rubinstein (1988) proposed a stepwise model for decision-making based on the similarity relation between outcomes and probabilities of risky prospects. A similarity relation is defined in a similar way to semi-orders in Luce (1956): we say a and b (denoted by $a \sim b$) are similar if:

$$a \sim b \text{ if } 1/\lambda \leq a/b \leq \lambda \tag{1}$$

where $\lambda > 1$. Let *S* and *R* be lotteries giving *s* and *r* with a probabilities *p* and *q*, respectively, and zero otherwise. Rubinstein's model is summarised with the following two stages:

Stage 1: The DM compares the outcomes and probabilities of the two options, if both $s > r$ and $p > r$ then DM chooses *S*. If step one is not decisive, DM proceeds to the second stage.

Stage 2: DM decides according to the non-similar elements of the pair. If the outcomes are similar, then decides according to the probabilities and *vice versa*.

However, Rubinstein does not tell us what happens next if the second stage is also not decisive.

Vague Expected Utility Theory (VEUT) proposed by Manzini and Mariotti (2004) suggests an answer to how DMs with imprecise preferences might reach a single decision[12]. The first departure of VEUT from Rubinstein's model is by focusing on similarity between risky prospects rather than between the elements of risky prospects. VEUT proposes a primary and secondary criterion for decision-making. The primary criterion is used when there is a sufficient difference between expected utilities. We say:

$$\begin{aligned} R \succ S \ \text{ if } \ EU(R) > EU(S) + \sigma(S,R) \\ S \succ R \ \text{ if } \ EU(R) < EU(S) - \sigma(S,R) \end{aligned} \tag{2}$$

where $\sigma(S,R)$ is a 'vagueness function', which, in the simplest version of VEUT, is assumed to be constant over all lottery pairs, that is, $\sigma(S,R) = \sigma$ for all lottery pairs. The above condition gives us a partial order:

$$EU(S) - \sigma(S,R) \leq EU(R) \leq EU(S) + \sigma(S,R)\,, \tag{3}$$

VEUT assumes that the DM employs a secondary criterion to reach a complete and precise judgement. This secondary criterion looks either at outcomes or at probabilities.

VEUT can also provide insights about how subjects in experiments, where unsureness intervals are elicited, construct their subjective ranges. To illustrate, let *R* be the lottery paired with a series of degenerate lotteries in an experiment similar to the ones discussed in Section 2. In such an experiment, using the lower and upper bound of the unsureness interval for a lottery *R* corresponds to certainty equivalents of $EU(R) - \sigma(S,R)$ and $EU(R) + \sigma(S,R)$, respectively. In other words, according to VEUT, the size of the imprecision range depends on the vagueness function and the characteristics of the pair.

The final model that we discuss is Imprecise Expected Utility Theory (IEUT) of Bayrak and Hey (2017). This assumes that the imprecision arises from the vagueness in individuals' perception of the numerical objective probabilities. The support for this assumption comes from the psychophysics literature (see Budescu *et al*, 1988; Budescu and Wallsten, 1990; Wallsten *et al*, 1986 and Bisantz *et al*, 2005; Wallsten and Budescu, 1995; Zimmer, 1984).

IEUT is at the moment applicable only for two-outcome lotteries: $L:\{x_1, p; x_2\}$, where x_1 and x_2 denote monetary payoffs ($x_1 > x_2$) and p is the probability of winning x_1. The probability is perceived as an interval: $[p-\beta(p,\psi), p+\beta(p,\psi)]$ where $\beta(p,\psi)$ is a function of the objective probability p and the individual's *subjective sophistication level* ψ . The sophistication level is dependent on the individual's familiarity and knowledge with uncertainty. Therefore, imprecision is higher for a less sophisticated individual. Further, the imprecision level is minimum for probabilities 0 and 1, and reaches its highest level when the probability is 0.5. This assumption is in line with findings summarised in Section 2.3; suggesting that imprecision increases with the dispersion of a lottery.

Using the imprecise judgement of the probabilities, the DM calculates a range of expected utilities for the lottery:

$$EU_L(L) = [p - \beta(p,\psi)] \cdot u(x_1) + [1 - p + \beta(p,\psi)] \cdot u(x_2) \tag{4}$$

$$EU_U(L) = [p + \beta(p,\psi)] \cdot u(x_1) + [1 - p - \beta(p,\psi)] \cdot u(x_2) \tag{5}$$

where EU_L and EU_U are the lower and the upper bounds of the expected utility range. To reach a final decision, IEUT assumes that the DM calculates a precise value by taking a weighted average of the lower and upper bounds, the weights depending on the DM's pessimism level.

5. What other stories provide a route to choice from imprecise preferences?

These are essentially stochastic stories, in which observed behaviour is viewed as a stochastic departure from the prediction of a deterministic model. The primary motivation for this specification is not to explain the range of certainty equivalents observed in the experiments reviewed in Section 2, but to provide an explanation for choice variability observed in experimental studies[13], though this specification does shed some light on the formation of intervals. However, building a connection between preference imprecision and stochastic models is not a straightforward task. One possibility is that stated imprecision in experiments might be a *reflection of anticipated noise in preferences* (see, for example, Loomes (2005)). In other words, subjects might be aware of their potential choice variability, and state a range of valuations to express the variability.

Several ways of adding a stochastic component to a deterministic model have been proposed. A detailed review of these models can be found in Appendix A. Adding a stochastic component to a deterministic model also enables an econometric implementation of such models and opens up a method for comparing their goodness-of-fit. Such studies mainly but not exclusively[14] use pairwise choice data (either embedded in Holt-Laury price lists or independently listed) and employ maximum likelihood estimation to fit and compare deterministic preference functionals (for example Expected Utility Theory, Rank Dependent Expected Utility Theory) embedded in a stochastic model. Comparison is done by using likelihood ratio tests for nested models and Vuong likelihood ratio tests for non-nested models (Vuong, 1989). Table B1 in the online Appendix B lists the datasets employed in these studies. Table B2, also in the online Appendix B, is a snapshot of the literature, showing the deterministic and stochastic models estimated in each study. Early studies focused on finding the best deterministic model, so they focused on a single error story. As the table shows, recently the literature has turned its attention to comparing stochastic models combined with, for example, EUT and RDUT. Early works in the literature employed a tremble[15] or a homoscedastic strong utility model to compare deterministic models (Hey and Orme, 1994; Harless and Camerer, 1994)[16]. These studies were stimulated by the development of a large number of non-expected utility theories. Hey (1995) claimed that fit of EUT can be further improved when coupled with heteroscedastic stochastic specifications compared to non-EUT models coupled with homoscedastic errors. Buschena and Zilberman (2000) comprehensively investigates and confirms Hey's conjecture. They found that when all models are embedded in homoscedastic formulation, EU is inferior to other models. Yet, when all deterministic models are embedded in a heteroscedastic formulation, non-EUT models do not perform significantly better than EUT. In addition, heteroscedastic error specifications of models significantly outperform the homoscedastic versions in terms of fit to data. This suggests that imprecision is not constant across problems.

Loomes *et al* (2002) is the first study making a distinction between errors and imprecision. The former, according to Loomes *et al* (2002), is possibly a temporary phenomenon. It might diminish as subjects gain experience. For example, calculation mistakes, making the wrong choice by a slip of hand, resulting from limited cognitive capabilities, inattentiveness to the experiment, misunderstanding experimental procedures. However, imprecision seems to be an inherent part of preferences: it is the inability to articulate one's own preferences precisely. Loomes *et al* (2002) added trembles to the *strong utility model* and to the *Random Preference Model* (RPM), focusing on EUT and RDUT. Here, trembles are for the temporary mistakes, and either strong utility or RPM is for the inherent imprecision. Results show that trembles disappear as subjects gain experience that is, towards the completion of 90 choice questions, but noise incorporated by secondary models remains relatively stable.

The online Appendix B presents details of studies examining stochastic choice behaviour. Some of these shed light on preference imprecision.

6. Can preference imprecision explain some of the prominent anomalies of standard theory?

The two prominent anomalies are the *valuation gap* and *preference reversals.*

Morrison (1998) focused on the *valuation gap*. Three responses for each WTA or WTP question were elicited: a lower bound, an upper bound, and the subject's "best estimate". Morrison tested for a significant overlap between ranges for WTA and WTP. The results reject imprecision as an explanation for the valuation gap, because the lower bound of WTA is significantly higher than the upper bound of WTP. Similarly, Dubourg *et al* (1994) elicits WTP and WTA values for changes in risk of non-fatal road injuries using an iterative process. Results show that individuals exhibit a significant amount of imprecision, but this imprecision alone is not sufficient to explain the observed disparity between WTA and WTP.

In contrast to these studies, Bayrak and Kriström (2016), using the BSU mechanism (Section 2.2), provides evidence in support of the imprecision account of the valuation gap: individuals cannot intrinsically determine a precise subjective valuation for a good but a range. When they are compelled to state a precise point they state the lower (upper) bound of the range in a buying (selling) task. In a between-subject design the authors compare the responses elicited using the BSU mechanism with the WTA and WTP values elicited using the conventional method. Results suggest that the WTP elicited using the conventional method and the lower bound of the offers in the BSU group come from identical distributions. Similarly, the WTA elicited in the conventional manner and the upper bound of the BSU group come from identical distributions.

As far as *preference reversals* are concerned, there have been many experimental investigations, one of the earliest being Butler and Loomes (2007). They used an iterative mechanism similar to that in Loomes (1988)[17]. The theoretical background to their study is based on an unpublished but influential paper by MacCrimmon and Smith (1986). This conjectured that individuals might have interval values rather than precise amounts for risky prospects, and claims that the preference reversal phenomenon can be explained by $-bets having a wider interval than P-bets. Butler and Loomes (2007) found that the imprecision argument can be seen as one of the explanations of preference reversals, since value intervals elicited for the $-bet are significantly larger than those for the P-bet and more importantly they overlap. However, further investigation by Butler and Loomes (2007) shows a puzzling finding: although subjects exhibit greater imprecision for $-bets in pricing tasks, they exhibit greater imprecision for P-bets (see Butler *et al* (2012) for a detailed discussion).

7. Summary

This section summarises what we have discussed during this paper. We start with evidence on preference imprecision.

There are two sources of data for preference imprecision: direct and indirect evidence. *Direct evidence* comes from experimental studies that rely on self-reported unsureness in binary choice experiments. These studies allow subjects to state “I’m not sure” while their preferences are being elicited. The main problem of these experimental studies is that appropriate incentives are not (cannot be?)

provided for the reporting of imprecision. *Indirect evidence* comes mainly from studies that fit deterministic functionals coupled with a random element to account for noise, such as random utility or random preference models. However, not all kinds of noise can be interpreted as imprecision; for example, the tremble story seems to be a more appropriate story to represent the mistakes that subjects might make when taking decisions.

One can measure the degree of imprecision in experimental studies by counting the statements of "I am not sure". For the indirect studies, we can look at the standard deviation of the error term such as in the strong utility model.

We then examined theoretical studies, particularly those focussing on: (1) ways to represent imprecise preferences; (2) functionals that can explain the formation of imprecision ranges elicited in experimental studies; and (3) theories focussing on how individuals reach a precise decision when they have imprecise preferences. Starting from the representation problem, we see that imprecise preferences imply preference incompleteness since complete preferences do not allow individuals to state "I'm not sure" for any choice problem (unless they are indifferent). Accordingly, we looked at a strand of literature that argues that the completeness axiom is not a necessary element for rationality. There are two ways of representing such preferences, while making a distinction between indifference and imprecision. The first is to employ an 'axiom of revealed non-inferiority' instead of the weak axiom of revealed preference. A second way is to make a distinction between psychological and revealed preferences. The former can be incomplete, while the latter is always complete, because real life situations usually compels decision-makers to make a precise decision.

We then focused on studies explaining how imprecision intervals are formed. VEUT connects the formation of imprecision intervals to a vagueness function which gives the least perceivable utility difference that an individual need to see to reach a precise judgement between two goods. IEUT proposes that preference imprecision arises from individuals' imprecise judgements about the probabilities of outcomes.

The final issue is modelling how individuals with imprecise preferences reach a precise decision when they are compelled to do so. VEUT suggests individuals use a type of lexicographic heuristic whereas IEUT assumes that they calculate the weighted average of the imprecision bounds. The weights in IEUT are a function of optimism.

In addition to these deterministic stories, there are also stochastic models that imply an indirect relationship between models and preference imprecision. It is indirect because stochastic models talk about errors and noise not about imprecision. However, it is possible to establish a link if we view the imprecision range elicited in a single setting as a collection of errors that a subject will exhibit in a repeated setting. In other words, subjects might be aware of their potential choice variability, and state a range of valuations to express the variability. For example, the distribution of a risk aversion parameter used in RPM might be linked to an imprecision range.

Finally, we focused on the possibility of preference imprecision explaining the prominent anomalies of standard theory, such as the valuation gap and preference reversals. For the former, the evidence is mixed and scarce, yet for the latter there is evidence for the argument that preference imprecision might be an explanation for preference reversals.

8. Conclusion

The term 'imprecise preferences', when used in economics, appears to be differently interpreted by different writers. Appropriately, then, there does not appear to be a widely accepted precise definition. This survey tries to give an overview of the various interpretations, and the relationships, if any, between them. It starts with various examples of the use of the expression in the literature, and evidence for its existence. If it exists (which some in the neoclassical mainstream might not admit) it is obviously useful to measure its extent, and to see if it is small (and hence not of concern to the mainstream) or it is large. We report on ways of measuring it. We then examine the work of those theorists who admit its existence and want to find ways of modelling it in a formal way (while not departing too far from the mainstream) and exploring the implications. Empiricists are also concerned about preference imprecision, and its implications for decision-making, so we examine ways that they have tried to model and incorporate imprecision into their empirical studies. Finally, we examine whether preference imprecision might tell us something about well-known violations of the accepted wisdom. We conclude that preference imprecision does exist and is important.

REFERENCES

Agranov, M., & Ortoleva, P. (2017). Stochastic choice and preferences for randomization. Journal of Political Economy, 125(1), 40-68.

Alberini, A., Boyle, K., & Welsh, M. (2003). Analysis of contingent valuation data with multiple bids and response options allowing respondents to express uncertainty. Journal of environmental economics and management, 45(1), 40-62.

Aumann, R. J. (1962). Utility theory without the completeness axiom. Econometrica: Journal of the Econometric Society, 445-462.

Ballinger, T. P., & Wilcox, N. T. (1997). Decisions, error and heterogeneity. The Economic Journal, 107(443), 1090-1105.

Banerjee, P., & Shogren, J. F. (2014). Bidding behavior given point and interval values in a second-price auction. Journal of Economic Behavior & Organization, 97, 126-137.

Bayrak, O. K., & Kriström, B. (2016). Is there a valuation gap? The case of interval valuations. Economics Bulletin, Volume 36, Issue 1, pages 218-236.

Bayrak, O. K., & Hey, J. D. (2017). Expected utility theory with imprecise probability perception: explaining preference reversals. Applied Economics Letters, 24(13), 906-910.

Becker, G. M., DeGroot, M. H., & Marschak, J. (1963). Stochastic models of choice behavior. Systems Research and Behavioral Science, 8(1), 41-55.

Bewley, T. (1986). Knightian uncertainty theory: Part I (No. 807, pp. 79-110). Cowles Foundation Discussion Paper.

Bisantz AM, Marsiglio SS and Munch J (2005). Displaying Uncertainty: Investigating the Effects of Display Format and Specificity. Human Factors: The Journal of the Human Factors and Ergonomics Society, 47, 777–96.

Budescu D.V., Weinberg S. and Wallsten T.S. (1988). Decisions Based on Numerically and Verbally Expressed Uncertainties. Journal of Experimental Psychology: Human Perception and Performance, 14, 281.

Budescu D.V. and Wallsten T.S. (1990). Dyadic Decisions with Numerical and Verbal Probabilities. Organizational Behavior and Human Decision Processes, 46, 240–63.

Buschena, D., & Zilberman, D. (2000). Generalized expected utility, heteroscedastic error, and path dependence in risky choice. Journal of Risk and Uncertainty, 20(1), 67-88.

Butler, D., & Loomes, G. (1988). Decision difficulty and imprecise preferences. Acta Psychologica, 68(1-3), 183-196.

Butler, D., Loomes, G. (2007). Imprecision as an account of the preference reversal phenomenon. American Economic Review 97, 277–297.

Butler, D., & Loomes, G. (2011). Imprecision as an account of violations of independence and betweenness. Journal of Economic Behavior & Organization, 80(3), 511-522.

Butler, D., Isoni, A., & Loomes, G. (2012). Testing the 'standard' model of stochastic choice under risk. Journal of Risk and Uncertainty, 45(3), 191-213.

Camerer, C. F. (1989). An experimental test of several generalized utility theories. Journal of Risk and uncertainty, 2(1), 61-104.

Camerer, C. F., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. Journal of risk and uncertainty, 19(1-3), 7-42.

Cettolin, E. and Riedl, A. (2016) "Revealed incomplete preferences under uncertainty: Evidence for Bewley preferences." Working paper.

Champ, P. A., Bishop, R. C., Brown, T. C., & McCollum, D. W. (1997). Using donation mechanisms to value nonuse benefits from public goods. Journal of environmental economics and management, 33(2), 151-162.

Cohen, M., Jaffray, J. Y., & Said, T. (1987). Experimental comparison of individual behavior under risk and under uncertainty for gains and for losses. Organizational behavior and human decision processes, 39(1), 1-22.

Cubitt, R., A. Munro, & C. V. Starmer. (2004). "Testing Explanations of Preference Reversal," Economic Journal 114, 709–726.

Cubitt, R. P., Navarro-Martinez, D., & Starmer, C. (2015). On preference imprecision. Journal of Risk and Uncertainty, 50(1), 1-34.

Danan, E. (2008). Revealed preference and indifferent selection. Mathematical Social Sciences, 55(1), 24-37.

Dubourg, W. R., Jones-Lee, M. W., & Loomes, G. (1994). Imprecise preferences and the WTP-WTA disparity. Journal of Risk and Uncertainty, 9(2), 115-133.

Dubourg, W.R. Jones–Lee, M.W., and Loomes, G. (1997) "Imprecise Preferences and Survey Design in Contingent Valuation", Economica, 64, 681–702.

Dubra, J., Maccheroni, F., & Ok, E. A. (2004). Expected utility theory without the completeness axiom. Journal of Economic Theory, 115(1), 118-133.

Eliaz, K., & Ok, E. A. (2006). Indifference or indecisiveness? Choice-theoretic foundations of incomplete preferences. Games and economic behavior, 56(1), 61-86.

Evren, Ö., & Ok, E. A. (2011). On the multi-utility representation of preference relations. Journal of Mathematical Economics, 47(4-5), 554-563.

Fishburn, P. C. (1985). "Nontransitive Preference Theory and the Preference Reversal Phenomenon," Rivista Internazionale di Science Economiche e Commerciali 32, 39–50.

Galaabaatar, T., & Karni, E. (2013). Subjective expected utility with incomplete preferences. Econometrica, 81(1), 255-284.

Georgescu-Roegen, N. (1958). Threshold in Choice and the Theory of Demand. Econometrica: Journal of the Econometric Society, 157-168.

Giarlotta, A., & Greco, S. (2013). Necessary and possible preference structures. Journal of Mathematical Economics, 49(2), 163-172.

Gilboa, I., Maccheroni, F., Marinacci, M., & Schmeidler, D. (2010). Objective and subjective rationality in a multiple prior model. Econometrica, 78(2), 755-770.

Hanemann, W.M., (1991). Willingness to Pay and Willingness to Accept: How Much Can They Differ? The American Economic Review 81 (3): 635–47.

Hanemann, M.W., Kriström, B., Li, C.Z. (1996). Nonmarket Valuation Under Preference Uncertainty: Econometric Models and Estimation. Berkley: University of California, Department of Agricultural and Resource Economics, Working Paper 794.

Harless, D. W., & Camerer, C. F. (1994). The predictive utility of generalized expected utility theories. Econometrica: Journal of the Econometric Society, 1251-1289.

Hey, J. D., & Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. Econometrica: Journal of the Econometric Society, 1291-1326.

Hey, J. D. (1995). Experimental investigations of errors in decision making under risk. European Economic Review, 39(3-4), 633-640.

Hey, J. D. (2001). Does repetition improve consistency?. Experimental economics, 4(1), 5-54.

Holt, C.A. (1986). “Preference Reversals and the Independence Axiom,” American Economic Review 76, 508–515.

Håkansson, C. (2008). A new valuation question: analysis of and insights from interval open-ended data in contingent valuation. Environmental and resource economics, 39(2), 175-188.

Karni, E. and Z. Safra. (1987). “Preference Reversals and the Observability of Preferences by Experimental Methods,” Econometrica 55, 675–685.

Loomes, G. (2005). Modelling the stochastic component of behaviour in experiments: Some issues for the interpretation of data. Experimental Economics, 8(4), 301-323.

Loomes, G., & Sugden, R. (1983). A rationale for preference reversal. The American Economic Review, 73(3), 428-432.

Loomes, G. (1988). Different experimental procedures for obtaining valuations of risky actions: Implications for utility theory. Theory and Decision, 25(1), 1-23.

Loomes, G., & Sugden, R. (1998). Testing different stochastic specifications of risky choice. Economica, 65(260), 581-598.

Loomes, G., Moffatt, P. G., & Sugden, R. (2002). A microeconometric test of alternative stochastic theories of risky choice. Journal of risk and Uncertainty, 24(2), 103-130.

Luce, R. D. (1956). Semiorders and a theory of utility discrimination. Econometrica, Journal of the Econometric Society, 178-191.

Luce, R. D. (1959). Individual Choice Behavior. New York: John Wiley & Sons.

Luce, R. D., and Suppes, P. (1965). Preference, utility and subjective probability. In: R. D. Luce, R. R. Bush & E. Galanter (Eds), Handbook of mathematical psychology (Vol. III, pp. 249–410). Wiley: New York.

MacCrimmon, K., & Smith, M. (1986). Imprecise equivalences: Preference reversals in money and probability. University of British Columbia working paper, 1211.

Mandler, M. (2005). Incomplete preferences and rational intransitivity of choice. Games and Economic Behavior, 50(2), 255-277.

Manzini, P., & Mariotti, M. (2004). A theory of vague expected utility. Topics in Theoretical Economics, 4(1).

Manzini, P., & Mariotti, M. (2006). A vague theory of choice over time. Advances in Theoretical Economics, 6(1), 1-27.

McFadden, D. (1981), Econometric Models of Probabilistic Choice, in C. Manski & D. McFadden, eds, 'Structural Analysis of Discrete Data with Econometric Applications', MIT Press, Cambridge, pp. 198–272.

Morrison, G. C. (1998). Understanding the disparity between WTP and WTA: endowment effect, substitutability, or imprecise preferences?. Economics Letters, 59(2), 189-194.

Mosteller, F., & Nogee, P. (1951). An experimental measurement of utility. Journal of Political Economy, 59(5), 371-404.

Nishimura, H., & Ok, E. A. (2016). Utility representation of an incomplete and nontransitive preference relation. Journal of Economic Theory, 166, 164-185.

Nishimura, H. and E. A. Ok. (2018). Preference Structures, mimeo , NYU.

Ok, E. A. (2002). Utility representation of an incomplete preference relation. Journal of Economic Theory, 104(2), 429-449.

Ok, E. A., Ortoleva, P., & Riella, G. (2012). Incomplete preferences under uncertainty: Indecisiveness in beliefs versus tastes. Econometrica, 80(4), 1791-1808.

Ready, R. C., Whitehead, J. C., & Blomquist, G. C. (1995). Contingent valuation when respondents are ambivalent. Journal of environmental economics and management, 29(2), 181-196.

Ribeiro, M., & Riella, G. (2017). Regular preorders and behavioral indifference. Theory and Decision, 82(1), 1-12.

Riella, G. (2015). On the representation of incomplete preferences under uncertainty with indecisiveness in tastes and beliefs. Economic Theory, 58(3), 571-600.

Rubinstein, A. (1988). Similarity and decision-making under risk (Is there a utility theory resolution to the Allais paradox?). Journal of economic theory, 46(1), 145-153.

Schmidt, U., & Hey, J. D. (2004). Are preference reversals errors? An experimental investigation. Journal of Risk and Uncertainty, 29(3), 207-218.

Segal, U. (1988). "Does the Preference Reversal Phenomenon Necessarily Contradict the Independence Axiom?" American Economic Review 78, 233–236.

Sen, A. (1997). Maximization and the Act of Choice. Econometrica: Journal of the Econometric Society, 745-779.

Starmer, C., & Sugden, R. (1989). Violations of the independence axion in common ratio problems: An experimental test of some competing hypotheses. Annals of Operations Research, 19(1), 79-102.

Starmer, C. (2010). Preference reversals. In Behavioural and Experimental Economics (pp. 206-211). Palgrave Macmillan UK.Thaler, R., 1980. Toward a Positive Theory of Consumer Choice. Journal of Economic Behavior & Organization 1 (1): 39–60.

Thaler, R. (1980). Toward a positive theory of consumer choice. Journal of Economic Behavior & Organization, 1(1), 39-60.

Tversky, A., and Russo, E. J. (1969). Substitutability and similarity in binary choices. Journal of Mathematical Psychology 6, 1-12.

Tversky, A., Slovic, P., & Kahneman, D. (1990). The causes of preference reversal. The American Economic Review, 204-217.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica: Journal of the Econometric Society, 307-333.

Wallsten T.S., Fillenbaum S. and Cox J.A. (1986). Base Rate Effects on the Interpretations of Probability and Frequency Expressions. Journal of Memory and Language, 25, 571–87.

Wallsten T.S. and Budescu D.V. (1995). A Review of Human Linguistic Probability Processing: General Principles and Empirical Evidence. The Knowledge Engineering Review, 10, 43–62.

Zimmer A. C. (1984). A Model for the Interpretation of Verbal Predictions. International Journal of Man-Machine Studies, 20, 121–34.

Appendix A: Stochastic Models

We need to distinguish between three different kinds of stochastic models: *Tremble Models, Random Utility Models* and *Random Preference Models.* As we have noted before, the first of these – being related to trembles – cannot be really thought of as related to preference imprecision. However the second and third can be: if utility is random, then it would seem that preference must be random; similarly if a preference parameter is random; though we should note that not all would agree that randomness is the same as imprecision. If, however, one is allowed to make this judgement, then the precision of the stochastic term in these models is an indicator of the precision of the preferences.

Tremble Models were the earliest to be proposed (Harless and Camerer, 1994), and simply posit that there is a correct decision, but, in implementing it, the DM 'trembles' and probabilistically makes the wrong decision(s). This model is rather simplistic and is now rarely used in isolation, though it is frequently used in combination with one of the other models (when they, in isolation, are unable to explain behaviour[18]).

Random Utility Models (RUM) are the most frequently used. They posit that the randomness enters through the utility component of the preference functional. In the simplest earliest formulations, the noise was a simple, zero mean, constant variance, additive term to the utility, but recent variants introduce heteroscedasticity and other embellishments. A problem with many RUMs is that dominance is not necessarily respected.

Random Preference Models (RPM), which do not allow dominance to be violated, are perhaps more acceptable to theorists[19]. They introduce the randomness through the parameter(s) of the preference functional. This the seventh model we describe below; the first six are Random Utility models.

Throughout the rest of this section, we focus[20] on *binary choice problems* between two lotteries S and R . $V(.)$ denotes the subjective value calculated according to a deterministic model such as EUT. Therefore, $V(S)-V(R)$ denotes the difference between the two lotteries; we will call this the *V-distance.* Usually it is assumed that the noise is centered around zero (so there is no bias) and is symmetrical, so, if *F(.)* denotes the *cdf* of the noise, then $F(0)=0.5$ and $F(x)=1-F(-x)$. In most (but not all) stochastic models there is a precision/scale parameter; let us denote this by λ; the larger is this precision parameter, the lower is the variance of noise.

1. The homoscedastic strong utility model

$$P\left(S \succ R\right)=F\left(\lambda[V\left(S\right)-V(R)]\right) \tag{6}$$

This, as its name implies, is homoscedastic; all the remaining RUM models are heteroscedastic.

2. The contextual utility model (CU)

The difference between this model and the homoscedastic strong utility model is the denominator:

$$P(S \succ R) = F\left(\lambda \cdot \frac{V(S)-V(R)}{V(z^{\max})-V(z^{\min})}\right) \tag{7}$$

Here $V(z^{\max})$ and $V(z^{\min})$ are the values of the highest and lowest monetary outcomes in the lottery pair S and R. This formulation implies that the standard deviation of noise is proportional to the range of outcome utilities in a pair (Wilcox, 2011). Contextual utility reduces to homoscedastic strong utility if all lottery pairs have the same range.

3. The strong utility model with heteroscedastic and truncated errors (THes)

Blavatskyy (2007) introduced this error specification by first applying it to EUT and naming this 'stochastic expected utility theory' (StEUT, hereafter). The truncation is done in a way to ensure that the StEUT value of a lottery is between the utility of the lowest and the utility of the highest outcomes – in order to rule out transparent errors, such as valuing a risky prospect less (more) than its lowest (highest) outcome. Another departure from strong utility is that StEUT attaches an error term to the expected utility of each lottery, instead of attaching it to the difference between the expected utilities of the two lotteries.

Unlike EUT, StEUT explains the fourfold pattern of risk attitudes, that is, risk seeking for unlikely gains or probable losses and risk averse behavior for probable gains and unlikely losses. Intuitively, a StEUT person is more likely to overvalue a lottery with an expected utility closer to its lower bound, and *vice versa.* Blavaskyy's approach is important because literature traditionally has been trying to incorporate the anomalies of EUT by developing deterministic models. StEUT, on the other hand, tries to explain these anomalies by incorporating a stochastic component within EUT.

4. Blavatskyy's "Model 1" (M1)

In this model, the probability of S being chosen over R can be written as:

$$P(S \succ R) = \frac{\varphi\left[U(S)-U(S \wedge R)\right]}{\varphi\left[U(S)-U(S \wedge R)\right]+\varphi\left[U(R)-U(S \wedge R)\right]} \tag{8}$$

where $U(.)$ is the von Neumann-Morgenstern EU function, φ is a non-decreasing function with $\varphi(0)=0$ and $S \wedge R$ is the greatest lower bound, defined as a lottery which is dominated by both lotteries in the pair and there exists no other lottery which is dominated by the pair and dominates $S \wedge R$. Notice that, when S stochastically dominates R, then $S \wedge R = R$, and $\varphi[U(R)-U(S \wedge R)]=0$, so the probability of choosing S over R equals 1; dominance is respected.

This formulation looks similar to a model that Luce developed. This is called the 'Luce Model' in McKelvey and Palfrey (1995) and Camerer and Ho (1999):

$$P(S \succ R) = \frac{\varphi[\lambda \cdot V(S)]}{\varphi[\lambda \cdot V(S)] + \varphi[\lambda \cdot V(R)]}, \text{ where } \varphi(x) = e^x \tag{9}$$

Model 1 uses deterministic values of lotteries $(V(.))$ with respect to a 'reference lottery' that is, the greatest lower bound of the pair, $S \wedge R$. The Luce Model is a special case putting $\phi(x) = e^x$.

5. Stronger utility model (Ser)

This model is a modified version of the strong utility model designed to respect dominance. To accomplish that, Blavatsky (2014) first defines two options such as the least upper bound $(S \vee R)$ and the greatest upper bound $(S \wedge R)$ using the characteristics of the pair that it is composed of safe (S) and risky (R) lotteries. $S \vee R$ stochastically dominates both lotteries but is itself dominated by options which dominate both lotteries. Similarly, $S \wedge R$ is stochastically dominated by both lotteries, but it dominates every other lottery which are dominated by both options.

The probability of choosing S over R can be written as:

$$P(S \succ R) = H\left(\frac{V(S) - V(R)}{V(S \vee R) - V(S \wedge R)}\right) \tag{10}$$

where $H : [-s, s] \rightarrow [0,1]$ is the cumulative distribution function of the error. For the H in equation, Blavatsky (2010) suggests using the cumulative distribution function of the raised cosine distribution after trying seven other forms (See Blavatsky, 2014, p. 270). This has interesting implications:

$$H(x) = \begin{cases} 0, & x < -s \\ \frac{1}{2} \cdot \left[1 + \frac{x-\mu}{s} + \frac{1}{\pi} \cdot \sin\left(\frac{x-\mu}{s} \cdot \pi\right)\right], & x \in [-s, s] \\ 1, & x > s \end{cases} \tag{11}$$

where $x = [V(S) - V(R)]/[V(S \vee R) - V(S \wedge R)]$ and $s > 0$. The distribution is assumed to be symmetric around 0 (i.e. $\mu = 0$), thus the distribution function is bounded on the interval of $[-s, s]$.

6. Decision Field Theory (DFT) and Boundedly Rational Expected Utility Theory (BREUT)

These models focus on the decision *process* by taking into account the limited attention and cognitive capabilities of DMs. These models are known as accumulator or sequential-sampling frameworks (for reviews, see Ratcliff and Smith, 2004; Otter *et al*, 2008). For example, DFT developed by Busemeyer and Townsend (1993) describes the decision process for a pair of lotteries with following elements:

i. DFT uses attention weights instead of objective probabilities. During the decision process, attention weights vary like a random walk as a function of objective probabilities.

ii. Using attention weights, the DM calculates several *V-distances* during the process, and accumulates them according to some rule. For example, Busemeyer and Townsend assume that the DM puts more weight on the current *V-distance* than previous ones.

iii. The process of deliberation is terminated once the accumulated V-distance crosses a threshold.

A simplified version of DFT following Rieskamp (2008) is

$$P(S \succ R) = F\left(2 \cdot \frac{V(S) - V(R)}{\sigma_{V-\text{distance}}} \cdot \theta\right) \tag{12}$$

Here the choice probability is a decreasing function of the standard deviation of the *V-distance* ($\sigma_{V-\text{distance}} = \sqrt{\sigma_R^2 + \sigma_S^2 - 2 \cdot \sigma_{R,S}}$). The smaller the variance of *V-distance*, the easier it is to make a judgement between the two lotteries. The covariance part implies that noise decreases with the similarity between the options, which is reminiscent of the wandering vector model and some heteroscedastic strong utility models. The decision threshold parameter, θ can be seen to be identical to the precision parameter, λ. Thus, the threshold parameter has identical implications: the higher the threshold, the lower the standard deviation of the noise. This implies that the DM behaves closer to the predictions of a deterministic model embedded in a stochastic structure. Busemeyer and Towsend (1993) formulates the decision threshold as $\theta^{*} = \theta / \sigma_{\text{V-distance}}$.

BREUT developed by Navarro-Martinez *et al* (2017) is similar in spirit to DFT, but has differences. First, BREUT assumes fluctuations in the subjective values of outcomes rather than in the probabilities. Second, in BREUT, the differences in the CEs of the lotteries are accumulated instead of differences in the utilities of options. Finally, in BREUT, the threshold is not arbitrary; instead, it reflects the DM's desired level of confidence, which is represented as the probability that the DM picks the option that she would choose after unlimited deliberation.

7. Random Preference Models

These are fundamentally different from Random Utility Models, first developed by Becker et al. (1963), then generalized by Loomes and Sugden (1995). They posit that the randomness comes through the parameters of the preference functional. So, for example, when fitting EU with a CRRA utility function, the risk-aversion parameter *r* is assumed to be stochastic with a given mean and variance. The idea here is that the DM, on every decision problem, chooses a risk-aversion parameter from some distribution and applies it to that decision problem. This implies that dominance is never violated. More formally RPM can be articulated as follows:

$$P(S \succ R) = \mathrm{P}(\beta \in B) \text{ such that } B = \left\{\beta \middle| V(S|\beta) - V(R|\beta) \geq 0\right\} \tag{13}$$

For each choice task, individual draws parameters of a deterministic model (β) randomly. Intuitively, it views the individual as a collection of multiple selves behaving in accordance with the same core theory, yet which self that is deciding for each task is randomly chosen.

References for Appendix A

Birnbaum, M. H., & Navarrete, J. B. (1998). Testing descriptive utility theories: Violations of stochastic dominance and cumulative independence. Journal of Risk and Uncertainty, 17(1), 49-79.

Blavatskyy, P. R. (2007). Stochastic expected utility theory. Journal of Risk and Uncertainty, 34(3), 259-286.

Blavatskyy, P. R. (2010). Modifying the mean-variance approach to avoid violations of stochastic dominance. Management Science, 56(11), 2050-2057.

Blavatskyy, P. R. (2014). Stronger utility. Theory and Decision, 76(2), 265-286.

Becker, G. M., DeGroot, M. H., & Marschak, J. (1963). Stochastic models of choice behavior. Systems Research and Behavioral Science, 8(1), 41-55.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field-theory—a dynamic cognitive approach to decision-making in an uncertain environment. Psychological Review, 100, 432–459.

Camerer, C., & Hua Ho, T. (1999). Experience-weighted attraction learning in normal form games. Econometrica, 67(4), 827-874.

Harless, D. W., & Camerer, C. F. (1994). The predictive utility of generalized expected utility theories. Econometrica: Journal of the Econometric Society, 1251-1289.

McKelvey, R. D., & Palfrey, T. R. (1995). Quantal response equilibria for normal form games. Games and economic behavior, 10(1), 6-38.

Loomes, G., & Sugden, R. (1995). Incorporating a stochastic element into decision theories. European Economic Review, 39(3), 641-648.

Navarro-Martinez, D., Loomes, G., Isoni, A., Butler, D., & Alaoui, L. (2017). Boundedly rational expected utility theory.

Otter, T., Johnson, J., Rieskamp, J., Allenby, G. M., Brazell, J. D., Diedrich, A., Hutchinson, J. W., MacEachern, S., Ruan, S. and Townsend, J. (2008). Sequential sampling models of choice: Some recent advances. Marketing Letters, 19, 255–267.

Ratcliff, R. and Smith, P.L. (2004). A Comparison of Sequential Sampling Models for Two-Choice Reaction Time. Psychological Review, 111, 333–367.

Rieskamp, J. (2008). The probabilistic nature of preferential choice. Journal of Experimental Psychology: Learning, Memory, and Cognition, 34(6), 1446.

Wilcox, N. T. (2011). ‘Stochastically more risk averse: A contextual theory of stochastic discrete choice under risk. Journal of Econometrics, 162(1), 89-104.

NOTES

[1] Source of evidence is not limited to economic experiments. Since 1990s, a related concept known as "preference uncertainty" has long been discussed in the stated preference literature. To inform policy makers, contingent valuation studies collect data from respondents about their willingness-to-pay and willingness-to-accept measures for non-market goods such as endangered species and recreational areas by presenting them with hypothetical scenarios. These studies assume that individuals can articulate their preferences precisely (Hanemann *et al*, 1996). Yet, empirical evidence suggests that this assumption might not be realistic (see for example, Ready *et al*, 1995; Champ *et al*, 1997; Alberini *et al*, 2003).

[2] Cohen *et al* (1987) included a fourth column, which provided subjects the option of stating equivalence (indifference) between the two. However, due to misunderstandings detected among subjects, subsequent authors combined the imprecision and equivalence columns.

[3] Which some theorists would say is a way to take a decision when preferences are imprecise (see Section 4).

[4] Which is manipulable in the sense that the questions that the subject gets depends upon his or her previous answers.

[5] Which is also manipulable in the same sense.

[6] An individual's WTA for a good defined as the minimum amount that an individual is willing to accept to give away a good, whereas WTP is the maximum amount that an individual is willing to pay to acquire a good. According to neoclassical economic theory, WTP and WTA should be similar if the goods in question have close substitutes and the income effects are small, but experimental literature has documented that the WTA is higher than the WTP (Hanemann, 1991). The widely-accepted explanation in the literature seems to be the "endowment effect" related to the loss aversion notion of prospect theory (Thaler, 1980).

[7] Pioneers are Aumann (1962) and Bewley (1986). Unfortunately, the absence of completeness creates a problem for representing preferences by standard utility functions. However, some literature has developed solutions for this problem. For example, Ok (2002) suggests multi-utility representation theorems. Dubra *et al* (2004) presents an extended version of EUT in which potentially incomplete preferences are represented with a set of utility functions which still satisfy the basic axioms of EUT. Other related studies are Evren and Ok (2011), Ok *et al* (2012), Galaabaatar and Karni (2013), Riella (2015), Nishimura and Ok (2016).

[8] Note that we have slightly changed the notation, so that curly brackets indicate a choice menu (that is, the set of items from which the DM can choose) and round brackets indicate the choice set (that is, the set of items which the DM chooses).

[9] See Riberio and Riella (2017) for a recent refinement of the work by Eliaz and Ok (2006).

[10] One could, as some experimental economists have done, tell the subject that if he or she had several items in the choice set (or a range of valuations) then the experimenter would choose from them at random, but then this could imply a preference for randomness, rather than preference imprecision.

[11] For other studies which model individual preferences by means of two binary relations see Danan (2008), Gilboa *et al* (2010), Giarlotta and Greco (2013), Nishimura and Ok (2018).

[12] Manzini and Mariotti (2006) presents a model for choice over time based on the same core idea of VEUT but the axiomatic structure is separate.

[13] It is a fact that subjects seem to behave inconsistently in repeated trials of the same choice problems in the same or different days (Mosteller and Nogee, 1951; Tversky and Russo, 1969; Starmer and Sugden, 1989; Camerer, 1989; Hey and Orme 1994; Ballinger and Wilcox, 1997; Loomes and Sugden, 1998; Hey, 2001; Agranov and Ortoleva, 2017).

[14] Allocation data and data using the Becker-Degroot-Marschak mechanism have also been used.

[15] As we have already noted, not a manifestation of imprecision.

[16] Roots of such works goes back to studies such as Mosteller and Nogee (1951), Becker *et al* (1963), Georgescu-Roegen (1958), Luce (1959), Luce and Suppes (1965) and McFadden (1981).

[17] In a typical preference reversal experiment, subjects are asked to make a choice between two lotteries and in another task, they are asked to state their selling prices. The two binary outcome gambles in the preference reversals experiments have distinct features: one of them typically called the 'P-bet' offers a relatively better chance of winning a modest prize, whereas the other bet, the '$-bet', offers a relatively small chance of winning a larger prize. Moreover, those two bets are constructed such that their expected values are close. The results show that a significant proportion of subjects choose the P-bet in the choice task but value the $-bet more. However, preferences are expected to be independent of the method that we elicit them (see Starmer (2010) for a review). Other prominent explanations for the reversals are violations of transitivity (Loomes and Sugden, 1983; Fishburn, 1985); violations of the independence axiom (Holt, 1986; Karni and Safra, 1987); violations of the reduction axiom (Segal, 1988). Tversky *et al* (1990) finds that the existent three explanations for preference reversals can only account for a small portion of the reversals. Additionally, there is also a fourth explanation: violations of procedure invariance related to scale compatibility. However, Cubitt *et al* (2004) and Schmidt and Hey (2004) found that it only accounts for a part of reversals.

[18] For example, when there are empirical violations of dominance.

[19] RPM does not allow dominance to be violated if it is used with a deterministic model that does not allow for such violations. This, however, is a problem from a descriptive perspective since there is

experimental evidence suggesting that subjects do violate stochastic dominance (see Birnbaum and Navarrete (1998)).

[20] Though they can be extended to be used in allocation problems and complete ranking problems.