CERE Working Paper, 2018:2

CERE

Understanding the Preference Imprecision

Oben K. Bayrak

Understanding the Preference Imprecision

Oben K. Bayrak

Abstract

The idea of preference imprecision challenges an underlying assumption in economics: individuals make choices between options confidently and they can articulate their subjective valuations for goods in terms of single precise amounts. In this paper, I review different strands of literature related to preference imprecision. Besides distilling conclusions from the existing literature, I also initiate a discussion on modelling the imprecision from a deterministic perspective by introducing two new frameworks.

JEL classification: D0, D80, D81

Keywords: preference imprecision, stochastic models, stochastic specifications, incommensurability, anomalies, valuation gap, preference imprecision, strength of preference, SoP, economic preferences, interval valuations

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

This work was supported by Tore Browaldh and the Jan Wallanders Foundations.

Understanding the Preference Imprecision

1. Introduction

What percentage of the choices do you make with full confidence? Without a part of you leaning towards the alternatives? Do you always articulate your subjective valuations for goods as precise monetary amounts or come up with ranges? The idea of preference imprecision challenges an underlying assumption in economics: individuals make choices between options confidently and they can articulate their subjective valuations for goods in terms of single precise amounts.

There are two different strands of literature illuminating the concept of preference imprecision: i. experimental studies eliciting range of subjective valuations; and, ii. stochastic models of choice interpreting the imprecision as noise around the predictions of deterministic theories such as Expected Utility Theory (EUT). These two strands differ in their methods and interpretation of the imprecision. Although this state of the literature enriches our perspective, it may also cause confusion and misapprehension of the concept of preference imprecision. The literature lacks of a study which synthesizes the results of the existing results. The objective of this paper is to distill the findings of the related studies to provide a guidance for the prospective studies.

Evidence from the experimental studies suggests that most of the subjects prefer to state their subjective valuations in terms of ranges when they are allowed (Håkansson, 2008, Benarjee and Shogren, 2014; Cubit et al., 2015; Bayrak and Kriström, 2016). These studies reveal important insights about the nature of the preference imprecision: For example, Cubitt et al. (2015) found that imprecision is not a temporary phenomenon; it does not decay with repetition and experience. Butler and Loomes (1988) observed that there is a positive relationship with the size of the imprecision and the dispersion of the lotteries used in their experiments. The focus of the literature so far is not merely to understand the nature of the preference imprecision, but also to investigate the possible connection of the imprecision with the anomalies of standard economic theory. For example, Butler and Loomes (2011) claims that preference imprecision can be an explanation for the prominent anomalies of standard economic theory: For valuation gap, except the only study employing an incentive compatible method for eliciting imprecision intervals (Bayrak and Kriström, 2016), the results does not support this claim (Morrison, 1998 and Dubourg et al., 1994). On the preference reversals side, Butler and Loomes (2007) found subjects state wider intervals for $-bet and the intervals for the two type of bets often overlap.

The second strand of the literature views the imprecision as an error component attached to core deterministic theories. The imprecision is interpreted as a random error reflecting the sways from the predictions of deterministic models. Hey and Orme (1994) find a sizable noise, i.e., the estimated deterministic models cannot explain a significant portion of the observed behavior. Harless and Camerer (1994) find that the unexplained part left as noise having a pattern instead of solely being random, can be interpreted as a call for new deterministic theories. Table 1 lists the distilled conclusions (C hereafter) together with the related section numbers to guide the reader throughout the paper.

Table 1: Conclusions distilled from the literature and the relevant sections

Distilled Conclusions	**Section**
C1.More than half of the subjects in the experiments exhibit imprecision by stating a range of valuations.	2.2.1.
C2. Results suggest a positive relationship between the imprecision range and the dispersion of the lotteries.	2.2.1.
C3.There seems to be a persistent part of the imprecision that does not decay with repetition.	2.2.1.
C4.Except the only incentive compatible experimental study, results suggest that the imprecision alone is not a sufficient explanation of the valuation gap.	2.2.2.
C5. Evidence support the idea that imprecision might be an explanation for the preference reversals.	2.2.3.
C6.Results strongly support the importance of the errors: the deterministic core models do not describe the significant portion of the observed behavior.	3.1.
C7.The variation in the behavior that is not predicted by the core deterministic theories is systematic rather than solely being random errors.	3.2.
C8. Best fit is achieved by Rank Dependent Utility Theory under Random Preference Model with trembles.	3.3.

The rest of the paper is organized as follows: Section 2 reviews the experimental studies related to preference imprecision, starting with the elicitation methods used (Section 2.1). In Section 2.2, the results of these studies providing insights about the size, prevalence and stability of the imprecision, and also its account for explaining the valuation gap and preference reversals are examined. Section 3 focuses on the stochastic models of choice such as random error approach (Section 3.1), constant error approach (Section 3.2.) and finally random preference approach and a comparison of the three (Section 3.3). Section 4 reviews the only deterministic model for preference imprecision: Imprecise Expected Utility (IEU) model (Bayrak and Hey, 2016). In Section 5, I introduce new deterministic frameworks to model the decision under preference imprecision. The motivation for suggesting a deterministic part is explained in Section 4 by relying on C6 and C7.

2. Experimental Studies Related to Preference Imprecision

In this section, I review the experimental studies which use direct elicitation methods of imprecision intervals mostly relying on the subjects' self-reporting (except Bayrak and Kriström (2016) reviewed in Section 2.1). Self-reporting is often used in environmental valuation studies; however, it is unconventional in experimental economics to rely on unincentivized methods. As a principle; 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 e.g. their offers (See Camerer and Hogarth (1999) for a detailed discussion). The results of the unincentivized experimental studies does not provide ultimate proofs but they can be seen as suggestive evidence for further research. I start with the most prevalent elicitation methods used in experimental studies on preference imprecision (Section 2.1.), and proceed with reviewing the results of the experimental studies (Section 2.2).

2.1. Methods of Elicitation

There are mainly three methods developed in the literature: i. *Response Table* (Cohen et al., 1987; Cubitt et al., 2015), ii. *Iterative Process* (Butler and Loomes, 2007, 2011; Dubourg et al., 1997, 1994), and, iii. *Buyer-Seller Uncertainty Mechanism* (Bayrak and Kriström , 2016).

Regarding the first method, subjects are asked to respond a series of binary choice questions between a risky prospect and a sure amount of money by filling a *response table* similar to Table 2. For the example depicted in Table 2, subject prefers the risky prospect for the certain amounts 2 and below, whereas risky prospect is preferred for the amounts 5 and above. Subject's imprecision interval corresponds to the values 3 and 4, for which the subject cannot confidently state a preference between the bet and these amounts. Cohen et al., (1987) included a fourth column which provides subjects the option of stating equivalence (indifference) between the two.

Table 2. Example Response Table

Certain Amounts	I definitely prefer the good	Not sure	I definitely prefer the certain amount
0	✓		
1	✓		
2	✓		
3		✓	
4		✓	
5			✓
6			✓

However, due to the misunderstandings detected among subjects, authors combined the imprecision and equivalence column while analyzing the data. More recently, Cubitt et al. (2015) also used the reduced version of the response table discarding the equivalence statement. Another difference between the two studies is related to payoff determination: For the resulting values inside the imprecision interval, Cohen et al., (1987) randomly determined which of the two options picked; whereas Cubitt et al. (2015) left the choice to the subjects by asking them to determine a switching point inside the imprecision interval.

The second method relies on an *iterative process*: For example, Dubourg et al. (1994) used a numbered disk, which has a small window shows only 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 was “definitely willing to pay”, interviewer rotates the disk to reveal a higher value through the window, whereas if the answer is “definitely not willing to pay”, interviewer reveals a lower amount. The experiment continues until a maximum amount that 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 used a type of iteration procedure. For each two outcome lottery in Table 3, subjects answered a series of binary choice questions where the second option was a sure amount of money. If a subject chooses the risky option, the sure amount is increased in the following question; on the other hand, if a subject choose the sure amount, the value updated downwards in the following question. Additionally, subjects are asked to use a cursor to state their confidence about their decision or strength of preference (SoP). The cursor can be moved to 51 different positions, corresponding the feeling of confidence between "very confident" to "very unsure".

Table 3. 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) elicited the valuations for risky prospects using a similar method: "incremental choice method". They focused on preference reversal phenomenon[1] 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 included 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" (See Section 5 for a detailed discussion about SoP).

Overall, the existing methods in the literature are slightly modified versions of the ones mentioned above. The main problem of first two is not being properly incentivized. On the other hand, the third approach, *Buyer-Seller Uncertainty* (BSU) mechanism is a modified version of standard Becker-DeGroot-Marschak mechanism in which subjects are free to state their subjective valuations for a good either as an interval or a precise amount. Yet, the crucial point is that a random mechanism assigns them as buyers or sellers after they record their valuations. Moreover, in the buyer role the trade occurs, if the market price is inside or below the stated range, and in the seller role the trade occurs, if the market price is inside or above the stated range. Subjects are informed about these procedures at the beginning of the experiment so they should consider while they are stating their valuations. Bayrak and Kriström

[1] 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 the same or insignificantly different. 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).

(2016) use this mechanism to investigate the imprecision account of the Willingness-to-Pay (WTP) and Willingness-to-Accept (WTA) gap[2]. To see how the incentives work under this mechanism, they assume that there can be only three possible individual types:

i. Individuals having precise preferences and complying with standard economic theory: In other words WTA and WTP are precise values and equal to each other; $WTA = WTP$. These type of individuals are predicted to state their values as single amounts.

ii. Individuals having a precise estimate of WTA and WTP but exhibiting loss aversion, i.e., $WTA > WTP$: For these type of subjects, the expected payoff under this mechanism is negative. So what they can do is to minimize their expected loss which might be seen as a weakness of this design: If they state WTA as their valuation, there is a 0.5 probability of assigning to buyer role, and in that case they might end up buying the good at an undesirable high price (P): $WTP < P < WTA$. On the other hand, stating WTP is also not optimal for them because they might end up being a seller with a probability of 0.5 and give away the good for undesirable low prices for them. Thus, the optimal response for a subject is to state the weighted average of 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)[3]. The stated response is predicted to be a single point as in type i.

iii. The final possible type of subjects are the ones who have imprecise preferences, in other words, they can come up with a range of subjective valuations for a good. Bayrak and Kriström (2016) hypothesize that valuation gap simply occurs because subjects who have range of subjective valuations simply state a value closer to lower (upper) bound of the range as their WTP (WTA) when they are required to state a precise amount in all other conventional

---

[2] 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 (Hanemann, 1991). The dominating explanation in the literature seems to be the "endowment effect" related to the loss aversion notion of prospect theory (Thaler, 1980).

[3] More formally, let $u^{-1}(X)$ denote the monetary equivalent for the utility of the good X in a buying task and $u^{-1}(\lambda \cdot X)$ is the one for the selling task where $\lambda > 0$ is the loss aversion parameter. The expected payoff from the experiment, $E[Y]$ can be written as: $E[Y] = 0.5 \cdot [u^{-1}(X) - SV] + 0.5 \cdot [SV - u^{-1}(\lambda \cdot X)]$, where SV stands for the subject's stated value. This mechanism is unfortunate for the subjects exhibiting loss aversion because they lose in all cases, what they can do is to minimize their loss. $E[Y]$ is negative for all values except the average of WTP and WTA. So the optimal SV i.e., $SV^* = 0.5 \cdot (WTP + WTA)$ achieves the maximum possible $E[Y]$ of zero. (For a detailed proof see the Appendix in Bayrak and Kriström, 2016).

experimental studies. Thus they let their subjects to be free between stating a precise value and a range under BSU mechanism. They assume that subjects who are in this group cannot distinguish the difference between the psychological satisfaction of the values in their range, i.e., they are indifferent between the values in the range and the good. They show that it is weakly dominant strategy for these subjects to state their true range of subjective valuations (for a detailed proof see Appendix in Bayrak and Kriström, 2016).

Considering the three possible types of subjects, one can conclude that subjects who prefer to state their valuations are the ones with imprecise preferences i.e., have a range of subjective valuation for a good (See Bayrak and Kriström (2016) for details).

The first two methods discussed above are not incentive compatible, but BSU is the mechanism so far developed in the literature as being the closest to be incentive compatible, except the weaknesses mentioned above.

2.2. Results from Experimental Studies

This section summarizes the results of the experimental studies which employ one of the three elicitation methods presented in Section 2.1. The results provide insights about the prevalence, size and the stability of the imprecision (Section 2.2.1), as well as the imprecision account of explaining the prominent anomalies such as the valuation gap (Section 2.2.2) and preference reversals (Section 2.2.3).

2.2.1. Some Characteristics of Preference Imprecision: Prevalence, Size and Stability

C1.More than half of the subjects in the experiments exhibit imprecision by stating a range of valuations.

Cohen et al. (1987) which is one of the early studies used response table method and observed that 10% of the subjects exhibit imprecision, the lowest ratio observed in the literature. On the other hand, Cubit et al. (2015) found that 87% of the subjects exhibit imprecision in their preferences. Bayrak and Kriström (2016) found that more than half of the subjects prefer to state their valuations as intervals when they are allowed, in line with the studies such as Håkansson (2008) and Benarjee and Shogren (2014).

C2. Results suggest a positive relationship between the imprecision range and the dispersion of the lotteries.

Butler and Loomes (1988) used the iterative elicitation procedure and elicited the certainty equivalents of the four binary outcome lotteries in Table 3: They found that from A4 to A1, as the dispersion of the lotteries increases, the size of the imprecision range also increases. More Recently, Cubitt et al. (2015) find support for the positive relationship between the size of the imprecision range and the lottery's distance to certainty which is analogous to dispersion.

C3. There seems to be a persistent part of the imprecision that does not decay with repetition.

Cubitt et al. (2015) provide tests for the stability of the imprecision as well, i.e. whether the size of the intervals change with repetition or not: If imprecision is merely a result of errors or unfamiliarity with the experimental mechanisms, one should expect it to disappear with repetition and experience. They found no evidence for imprecision declining with repetition. The analysis supports that imprecision is stable and not temporary; instead it seems to be an inherent part of the preferences. This is in line with the findings of Loomes et al. (2002) reviewed in Section 3.1.3.

2.2.2. Imprecision Account of the Valuation Gap

C4. Except the only incentive compatible experimental study, results suggest that the imprecision alone is not a sufficient explanation of the valuation gap.

Morrison (1998) focuses on the imprecision account of the valuation gap. Three responses for each WTA or WTP question were elicited: a lower-bound, an upper-bound, and subjects' "best estimate". Morrison tested for a significant overlap between the ranges for WTA and WTP. The results reject the imprecision as an explanation for valuation gap, because the lower bound of WTA is significantly higher than the upper bound of WTP. Similarly, Dubourg et al. (1994) elicited WTP and WTA values for changes in the risk of nonfatal road injuries using an iterative process. They found that individuals exhibit a significant amount of imprecision, however this imprecision alone is not insufficient to explain the observed disparity between WTA and WTP.

On the other hand, Bayrak and Kriström (2016) provide supporting evidence for the imprecision account of the valuation gap: "…individuals cannot intrinsically determine precise single points, but are able to identify a range of values for their personal valuation of the good. If the experiment forces them to state a point, they employ a heuristic: buyers state the lower bound while sellers state the upper bound in their admissible range". To test this, they simply compared the results of BSU mechanism with the treatments in which they elicited the subjective valuations in the conventional manner i.e., requiring subjects to state precise amounts as their WTA and WTP. Statistical tests confirm that WTP elicited in conventional method and lower bound of the offers in the BSU group come from identical distributions. Similarly, WTA elicited in the conventional manner and upper bound of the BSU group come from the identical distributions.

2.2.3. Imprecision Account of the Preference Reversals

C5. Evidence support the idea that imprecision might be an explanation for the preference reversals.

Butler and Loomes (2007) investigate the imprecision account of the preference reversals, using a similar iterative mechanism to Butler and Loomes (1988). Theoretical background of their study is based on an unpublished but influential paper; MacCrimmon and Smith (1986). MacCrimmon and Smith (1986) conjecture that individuals might have interval values rather than precise amounts for the risky prospects and claim that preference reversal phenomenon can be explained by $-bets having a wider interval than P-bets. Authors refrain themselves from presenting a formal structure about how individuals form these interval valuations, but they suggest that as the risky prospect become more dissimilar to the certainty, the width of the imprecision interval increases. Butler and Loomes (2007) found that the imprecision argument can be seen as one of the explanations of the preference reversals, since the intervals elicited for the $-bet is significantly higher than the P-bet and more importantly they overlap, which is in line with the conjectures of MacCrimmon and Smith (1986)[4].

3. Modelling Imprecision as Stochastic Deviations from Deterministic Theories

The ideas go back to the literature on "noisy" preferences based on the early works in 1950s and 1960s in the form of probabilistic choice and random preference models (Becker et al., 1963; Georgescu-Roegen, 1958; Luce, 1959; Luce and Suppes, 1965).

Beginning with 1990s, the idea of imprecision began to receive attention by researchers in the form of modelling it as the stochastic component of a deterministic core model such as EUT. (Harless and Camerer, 1994; Hey and Orme, 1994; Loomes and Sugden, 1998, 1995; Sopher

[4] Butler and Loomes (2011) tested the preference imprecision as an explanation for the observed violations of independence and betweenness axioms. This study can be seen as a continuation study of Butler and Loomes (2007) which focusses on the imprecision account of the preference reversals. Their results confirm the fanning out hypothesis and favor the preference imprecision as an explanation of the violations of EUT.

and Gigliotti, 1993). The common approach employed by these studies is to incorporate the imprecision as a stochastic component i.e. the random error part of a deterministic model. They rely on an assertion that the choice under risk might not be predictable in each and every choice problem by the core theories, instead, for some problems, individual might sway from the psychological satisfaction calculated by the deterministic model and make an unpredicted decision among the options. This idea is also supported by the experimental literature such as Mosteller and Nogee (1951), Hey and Orme (1994) and Hey (2001): subjects seem to behave inconsistently in repeated trials of the same choice problems in the same or different days. The most prominent approaches in the literature for modelling the imprecision as stochastic preferences are random error (Section 3.1), trembling hand (Section 3.2) and random preference models (Section 3.3)[5].

3.1. Random Error

Hey and Orme (1994) inspired by Fechner's (1860/1966) ideas of individuals' imprecise judgements of the stimuli modelled as white noise (iid), normally distributed with zero-mean. The reason for such an error might be the subjects' misunderstanding the nature of the experiment or operational mistakes during the experiment e.g. pressing the wrong key by accident. Moreover, subjects' inattentiveness such as being in a hurry to complete the experiment and/or having another motivation rather than maximizing their welfare from participating in the experiment might be the reasons behind those errors.

Hey and Orme's (1994) idea is that the preferences can be represented by a core theory plus a random error term:

$$V(f)-V(g)+\varepsilon>0 \Leftrightarrow f \succ g \qquad (1)$$

where $V(.)$ is the preference functional of a deterministic core theory, f and g are the two options and ε is the stochastic component with a constant variance and mean of zero. Individual prefers f over g if the difference between the utility of the two options plus some random error is positive. When $\varepsilon=0,$ the choice solely depends on the core deterministic theory part of the model. Moreover, higher the difference in deterministic part; $V(f)-V(g)$, it is less likely that the preferences predicted by the core theory will be reversed by the error term.

[5] See Wilcox (2008) for a detailed discussion on stochastic models.

The data they use composed of 100 pairwise choice questions answered by 80 subjects. They estimated eleven different preference functionals including: risk neutrality (expected value), EUT, Disappointment Aversion Theory, Prospective Reference Theory, Quadratic Utility Theory, Regret Theory with dependence and independence, rank dependence with the power weighting function and "Quiggin" weighting function, Weighted Utility Theory and Yaari's Dual Model.

Their results provide insights about the winner and loser theories. For example, risk neutrality is rejected in favour of EUT, on the other hand, at 1% level, EUT is rejected in favour of the remaining 9 preference functionals. Overall, for approximately 39% of the subjects, EUT does not perform worse than any of the alternative models. For the remaining portion of the subjects, Rank Dependent Utility Theory and Quadratic Utility Theory seem to be the strongest models. Additionally, they find that Regret Theory with independence performs better than the one with dependence, which suggests that the subjects perceived the two lotteries as being statistically independent. Among the remaining nine models, Yaari's Dual Model and Disappointment Aversion Theory are the poorest.

C6. Results strongly support the importance of the errors: the deterministic core models do not describe significant portion of the observed behavior.

Hey and Orme (1994) report the significance of the errors by calculating a vacillation/determination index for the subjects. Subject prefers the left hand gamble if $V(f)-V(g)+\varepsilon>0$ with a proportion of r and chooses the right hand one if $V(f)-V(g)+\varepsilon\leq 0$ with a proportion of p. A value of p or r equals to unity implies that subject is certain about the choice, whereas a value of 0.5 shows subject is completely uncertain. Based on these, the vacillation index is calculated as the following: if $100|p-r|=100$, then the subject is completely determined; if $100|p-r|=0$, then the subject is completely vacillating. The results show that the vacillation index for the estimated preference functionals is observed as below 50 frequently and for some functionals, the figures are as low as 0 implying sizable errors.

3.2. Constant Error (Tremble)

Harless and Camerer (1994) suggest a simpler error generating mechanism which is analogous to the game theoretic term of "trembling hand". It is assumed that individuals have true underlying preferences characterized by a core deterministic theory, but making the wrong choice with a fixed probability of w . One problem with this assumption is that, intuitively it might be an insufficient way to incorporate the stochastic nature of human behavior, since the likelihood of making an error is expected to increase as the expected satisfaction given by the options become dissimilar (See Loomes et al. (2002) for further discussion). On the other hand, in tremble approach, the probability of making erroneous decision is independent of the features of the options: individuals choose the less preferred option with a probability of w, no matter how much the difference between the utility of the options according to a core deterministic theory is.

They analyzed 23 data sets consisting of approximately 8,000 choices and the results suggest that none of the existing theories perform significantly better than others. Yet, the analysis identify some dominated and dominant patterns: the dominated theories are generally the ones assuming betweenness rather than independence and fanning in. The dominants are mixed fanning, Prospect Theory, EUT, and expected value. Interestingly, EUT is never dominated, but it is never selected as the best model according to Schwarz criterion.

Another important observation that theories like EUT and Weighted EUT can be improved by further generalizations to incorporate commonly observed patterns. Moreover, the alternative models such as Rank Dependent Utility Theory seem to allow patterns which are rarely observed. Thus, the results suggest not to completely abandon EUT but to improve it.

C7. The variation in the behavior that is not predicted by the core deterministic theories is systematic rather than solely being random errors.

All theories are rejected by a chi-squared test. This implies that the variation which is not predicted by the existing core theories can be explained by another theory to be developed. In other words, for every model, the “trembling part” is systematic variation rather than being an error. So, the “noise” has an underlying pattern, rather than being random. This is an important result, because it simply means that there is a room for developing a new deterministic model. In particular, a theory predicting parallel and linear indifference curves inside the Machina triangle, and non-conventional patterns on the boundaries would fit the data better.

3.3. Random Preference Model and a Comparison of the Three

The final approach that I will discuss is known as the *random preference model (RPM),* first developed by Becker et al. (1963), then generalized by Loomes and Sugden (1995). Central to the RPM, the assumption is that individuals decide according to a core theory, but the parameters of the theory are determined randomly for each action. Suppose the core theory is EUT with a simple power utility function, $u(x) = x^a$, and parameter a, determining the curvature of the utility function i.e. the risk attitude of the individual. Now suppose a equals 0.8, implying that the individual exhibiting risk aversion. Conventionally, the value of a is an inherent characteristic of the individual and assumed to be stable for all types of tasks such as buying, selling or choice, and also for all goods. Whereas in *RPM*, the value of a is randomly drawn with replacement for each decision task, i.e. the risk profile of the individual is not assumed to be stable. Intuitively, it views the individual as a collection of multiple selves behaving in accordance with the same core theory, but which self that is deciding for each task is randomly chosen.

One problem with these three approaches is related to the violation of dominance: The frequency of cases which exhibit violation of dominance is over-predicted by *random error* approach of Hey and Orme (1994), where *RPM* fails to predict any violation of dominance (Loomes et al., 2002).

Consider two binary outcome prospects f and g, which give the same amount X and zero, but first one has slightly higher probability of winning X such as 0.25 and 0.20, respectively. The expected utility of the prospects are:

$$EU(f) = 0.25 \cdot u(X) + 0.75 \cdot u(0) \tag{2}$$

$$EU(g) = 0.20 \cdot u(X) + 0.80 \cdot u(0) \tag{3}$$

Clearly, expected utility of f is higher than g, since the first one stochastically dominates other. The choice problem can be represented as:

$$0.05 \cdot [u(20) - u(0)] > 0 \tag{4}$$

Thus, regardless of which parameters are drawn randomly for $u(.)$, if the core theory behaves according to the dominance notion, so does *RPM* (Loomes and Sugden, 1998). On the other hand, random error model incorporates the error term separately:

$$0.05 \cdot [u(20) - u(0)] + \varepsilon > 0 \tag{5}$$

If the error term is negative and sufficiently high, the inequality will be reversed and model predicts the dominated option, g to be preferred over f .

We know from the previous experimental literature that individuals seldom violate dominance at least when it is transparent, i.e. they frequently choose the stochastically dominant option (Loomes, 2005). For example, Loomes et al. (2002) analyze the data presented by Loomes and Sugden (1998), in which the binary choices of 92 subjects for 45 lottery pair are collected. The distinct feature of the data is that each pair is presented twice in different orders. Among the 45 different lottery ticket pairs, in 5 of them, one option stochastically dominates the other such as offering a slightly higher chance of winning the same amount or lower chance of losing the same amount. When they include the dominance cases, Fechner models perform poorly: They predict 10-15% of subjects to violate dominance, the observed ratio was less than 1.5%.

C8. Best fit is achieved by Rank Dependent Utility Theory under Random Preference Model with trembles.

Subsequently, Loomes et al. (2002) implemented the trembling modification to *random error* model and *RPM*. Authors compared *EUT* and *Rank Dependent Utility Theory* using different stochastic specifications such as *RPM* and *random error model*; with and without tremble. Results show that the trembling modification significantly increases the explanatory power of the two stochastic specifications. The best fitting menu seems to be the Rank Dependent Utility Theory together with *RPM* with trembles[6]. Moreover, they find that the trembles disappear as subjects gain experience i.e. towards the completion of 90 choice questions. This draws interesting conclusion that the tremble part can be seen as a type of error due to e.g., misunderstandings which capture the temporary mistakes made by subjects, i.e., disappear with experience and repetition, but the variation incorporated by the *RPM* captures the stable and inherent part of the imprecision. This result is in line with C3 in Section 2.2.1.

4. Modelling the Systematic Part of Imprecision: Imprecise Expected Utility Model

As reviewed in Section 3, the literature mostly comprises the studies which model the imprecision in a stochastic manner. There are at least three important conclusions that can be

[6] More recently, Blavatsky and Pogrebna (2010) estimated seven deterministic theories using different stochastic specifications including the three discussed in this section and also their variations. They found that the estimated parameters of the core theories differ significantly across different stochastic specifications. The best fit of the core theories is achieved under Fechner model of heteroscedastic truncated errors or random utility model.

drawn from the existing studies providing motivation for focusing on the issue from a deterministic perspective: As C6 underlines, existing deterministic theories do not explain a significant portion of the observed behavior when coupled by a stochastic component (See Section 3.1.1). Moreover, the observed variation seems to be systematic rather than being random. One way to interpret this observation as a room for a deterministic model of preference imprecision is yet to be developed (C7 in Section 3.2). Finally, the imprecision seems to be an inherent part of the preferences, not a temporary phenomenon which disappears with experience (C3 in Section 2.2.1).

IEU model developed by Bayrak and Hey (2017) is the only model employing a deterministic approach in modelling preference imprecision: They assume that the imprecision arises due to the individuals' imperfect 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). For example, in the experiment reported by Budescu et al. (1988): subjects stated almost identical valuations for the same lotteries under different representations of the probabilities such as numerical, graphical and verbal (e.g. less likely). This suggests that individuals perceive the objective numerical probabilities in a similar vagueness as they understand the verbal correspondences of the probabilities, if we interpret the concept of verbal probabilities as interval of probabilities (Wallsten and Budescu, 1995). Zimmer (1984) also suggests that numerical probabilities are not natural to people, appearing recently as the 17th century as a mathematical concept, instead, individuals process and communicate the uncertainty in a verbal manner.

The model developed by Bayrak and Hey (2017) is applicable only for two-outcome lotteries: $L:\{x_1, p; x_2\}$, where x_i denotes the monetary payoff such that $x_1 > x_2$ and p is the probability of winning x_1. Each probability is perceived as an interval: $[p-\beta(p,\psi), p+\beta(p,\psi)]$ where $\beta(p,\psi)$ is a function of objective probability p and individual subjective sophistication level ψ. Sophistication level is contingent to individual's familiarity and knowledge in uncertainty. Therefore, imprecision is higher for a less sophisticated individual in probability concept. Furthermore, the imprecision level is minimum for the probabilities 0 and 1, and reaches its highest level when the probability is 0.5. This assumption is in line with the C3 (See Section 2.2.1) suggesting that the imprecision increases with the dispersion of a lottery.

Using the imprecise judgement of the probabilities individuals calculate a range of expected utilities for the good:

$$EU_L(L) = [p - \beta(p,\psi)] \cdot u(x_1) + [1 - p + \beta(p,\psi)] \cdot u(x_2) \tag{6}$$

$$EU_U(L) = [p + \beta(p,\psi)] \cdot u(x_1) + [1 - p - \beta(p,\psi)] \cdot u(x_2) \tag{7}$$

where EU_L and EU_U are the lower and the upper bounds of the expected utility range. Next, individual calculates a precise estimate of the range by simply taking the expectation of it. Bayrak and Hey (2017) suggest that individuals calculate the weighted average of the range depending on their pessimism level, since this stage of the decision problem resembles the decision under ambiguity i.e., individuals do not have any information about which expected utility value from the range reflects his or her true preference for lottery L. Thus, individual weights the worst case with pessimism level, α and the best case with $1-\alpha$. The crucial point for explaining the anomalies of standard theory such as valuation gap and preference reversals is that depending on the task type such as buying, selling and choice; individual sees the different bound of the range as the worst and the best case. In a buying and choice task, the lower bound of the expected utility range is seen as the worst case, because the good is an incoming one. On the other hand, in a selling task, seller perceives the upper bound of the range as the worst case since he or she is giving the good away. Bayrak and Hey (2017) demonstrate how the new model can explain the preference reversals with a numerical example. They show that under imprecision i.e., $\beta(p,\psi) > 0$; a sufficient level of pessimism results in preference reversals whereas optimism causes non-standard preference reversals.

Although the model can account for preference reversals and valuation gap, it is only applicable for two-outcome lotteries. Extending it to the risky prospects having more than two outcomes is not straightforward. Another point worth mentioning is that the level of pessimism is assumed to be exogenous, i.e., independent of the task type such as buying and selling, and the features of bets such as the outcomes and the associated probabilities. Yet, the nature of the task and the elements of a risky prospect might also affect the level of pessimism.

5. New Frameworks for Modelling the Imprecision in a Deterministic Manner

Suppose a subject can come up with a range of valuations for a risky prospect but and asked to state a precise WTP for it. If the experiment requires the individual to determine a precise amount, which value will the subject state? The lower bound or upper bound of the range or something between? What is the underlying behavioural story behind pinning the range down to a single amount? Stochastic modelling attempts in Section 3 suggest that it is a random draw around a core deterministic theory, whereas IEU model assumes that individual first forms an imprecision range, then calculates the weighted average of the lower and the upper bounds using a pessimism parameter. Yet, these are not the only ways of modelling. The views introduced in this section provide alternative deterministic ways of modelling the process of selecting a single amount for the psychological value of a good under preference imprecision[7].

To provide the alternative stories, I shall begin by viewing the preference imprecision as a collection of *multiple selves*: Suppose decision maker (DM) ends up having a range of certainty equivalents for a Lottery L as between \$5 and \$7. For simplicity, I assume that the smallest monetary increment is \$1. Thus, the range implies that there are three selves, disagreeing with each other about L's certainty equivalent. The most generous self values the good as \$7, whereas the stingiest one thinks that L is worth \$5.

Returning back to the question stated in the beginning of this section about the underlying behavioural story behind pinning the range down to a single amount: One way to model decision in the presence of preference imprecision is to assume DM acting as ***intrapersonal planner***, analogous to social planner concept of welfare economics, i.e. DM states the certainty equivalent from the range which maximizes the total welfare of multiple selves:

$$W = f\left(U\left(v_i, \pi_i, x\right)\right); i = \{1, \ldots, n\}, \tag{8}$$

where π_i is the weight attached to utility of i^{th} inner self by DM. Let $v_i \in V$ be set consisting of the selves' subjective valuations for the good, i.e. imprecision range. Finally, $x \in V$ is the DM's choice to be stated from the set V, which is e.g. in buying task is the WTP, or in selling task it is WTA to be stated for the good in question. Thus, the objective of the DM is to maximize the total welfare by choosing $x \in$V.

For the example above, the elements of V are \$5, \$6 and \$7. For simplicity, assume that the initial wealth is zero and DM attaches equal weights to each self, i.e., $\pi_i = 1/3$ for $i = \{1,2,3\}$. Thus, in a buying task, DM chooses an x from V for WTP according to the following:

[7] See Section 4 for the motivation behind modelling the imprecision in a deterministic manner.

$$WTP = \underset{x \in V}{\arg\max}\ 1/3 \cdot \left[u(5-x) + u(6-x) + u(7-x) \right] \tag{9}$$

It is straightforward to see that WTP equals \$5, which is the lower bound of the imprecision range. Similarly, in a selling task, DM's problem is to find the value of x that maximizes the total welfare of the selves:

$$WTA = \underset{x \in V}{\arg\max}\ 1/3 \cdot \left[u(x-5) + u(x-6) + u(x-7) \right] \tag{10}$$

In this case, the x stated from V as WTA will be \$7 corresponding to the upper bound of the imprecision range.

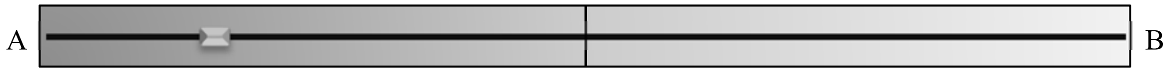
A second alternative story can be constructed using a ***voting scheme***: The stated point value from the imprecision range is the one which attains a threshold of 'Yes' votes. More formally, suppose in a buying task, i^{th} self votes 'Yes' for a particular x from V , if and only if:

$$u(v_i - x) \geq 0 \tag{11}$$

DM's decision problem is to state an x from V, which can get a certain number of 'Yes' votes from multiple selves. I define threshold percentage of 'Yes' votes as *required level of willingness*, w^*. For instance, if w^* equals 0.7, it implies that DM states a value from V, which can get the consent of 0.7 of the inner selves. The required level of wilingness does not need to be the same for all type tasks and goods: an individual having bad experiences in buying goods might have a higher required level of willingness, whereas the same individual might be tranquil in selling tasks and have a relatively lower w^*. Additionally, an individual might act more meticulous in buying important goods such as a house and have e.g. w^* of 1, relative to the case of buying a pair of shoes, e.g., 0.7. In other words, for some situations, DM might want to feel totally confident about his or her decision by convincing all inner selves by employing a unanimity rule, and for some situations, DM might prefer a majority rule which is to convince a certain portion of the selves.

Voting scheme also helps us to interpret the concept of *'strength of preference' (SoP)*, which is a measure elicited in economic experiments including choices between pair of lotteries: In a typical experiment eliciting SoP, subjects are not only asked for their choice between options, but also for their judgements about how confident they feel about their choice either using a numerical scale as in Figure 1 or selecting the appropriate phrase from a list of given phrases as shown in Figure 2. Subjects use the scroll bar instrument in Figure 1 to exhibit how confidently they are close to choosing Option A or B.

Figure 1: Scroll bar instrument used in Butler et al., (2014)



In Figure 1, subject seems to be choosing Option A, but not with full confidence. Another example is used in an experiment reported in Butler and Loomes (2007). Subjects are asked to complete a series of binary choice questions: in each question the first option was a lottery which was fixed, whereas the second option was a sure amount of money which was updated for each question. Additionally, for each choice, subjects stated their SoP by choosing the appropriate phrase as depicted in Figure 2.

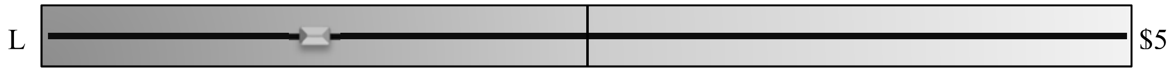
Figure 2: An Example Instrument for Eliciting SoP. Reprinted from Butler and Loomes (2007)

I definitely prefer lottery A	I think I prefer lottery A but I am not sure	I think I prefer lottery B but I am not sure	I definitely prefer lottery B

Consider Figure 1 as an example, and suppose Option A is a degenerate lottery which gives \$5 with certainty and Option B is a lottery; and finally assume that the scrollbar can be placed to 100 different positions but placed on the thirtieth point from the left. Thus, the elicited data point looks something similar to the following: subject prefers \$5 with certainty over the lottery with a confidence level of 70%. Although there is a growing literature on SoP, it is not clear how to incorporate this type of data in decision models. Existing models of choice is incapable of interpreting this concept formally, because it is assumed that a subject is either willing to or not willing to accept the e.g. the certain amount of money against the lottery. However, SoP extends the willingness concept from a discrete yes-no situation to a continuous measure of willingness.

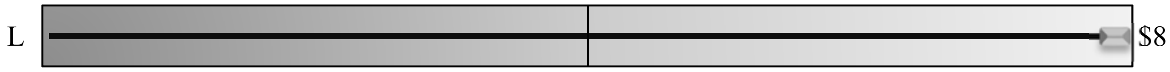
One way to interpret SoP judgements is to view it as analogous to the *level of willingness concept* (percentage of the multiple selves voting ‘Yes’) introduced above. Suppose a subject ends up having a range of certainty equivalents for lottery L between \$5 and \$8 and asked to choose between the lottery and a series of sure amounts of money. For a choice between L and \$5, subject’s *level of willingness* is 25%, since only for one of the four multiple selves value L as \$5. Subject will position the scroll bar in a similar way depicted in Figure 3, if one views the bar having a 100 point scale, increasing from left to right. The utility will be nonnegative resulting in a ‘Yes’ vote, however the other selves—who value L as \$6, \$7 and \$8—prefer lottery L instead of \$5 with certainty. Similarly, for a binary choice question in which the options are lottery L and \$7, all of the inner selves will be convinced and the level of willingness will be 1.

Figure 3: Subject stating SoP for a choice between L and \$5



Similarly, for a choice between L and \$8, all of the multiple selves are convinced, so subject exhibits 100% confidence and level of willingness as shown in Figure 4.

Figure 4: Subject stating SoP for a choice between L and \$8



SoP as a new measure might be a useful input for the prospective choice models, yet the literature lacks the theoretical background and a clear definition of the concept: It is not clear how individuals use—if they use—SoP in their judgements and decision making. Perhaps, it would be a better strategy for the literature to advance in setting a theoretical background for SoP rather than eliciting it without knowing what it means, and how and whether it is used in decision making. In this lack—with the aim of initiating a new discussion—I suggested above to interpret SoP as analogous to the level of willingness that individual uses to calculate a single amount from the imprecision range, which is just one way to interpret the concept of SoP.

6. Discussion and Conclusion

Although the aforementioned studies provide insights about the nature of the imprecision, the concept itself is imprecisely defined: experimental studies elicit ranges of subjective valuations and view imprecision as the incommensurability of the values inside the stated range. On the other hand, stochastic models of choice define imprecision as a noise, i.e., the random sways or mistakes attached to deterministic models. To construct a solid understanding of the preference imprecision, as a first step, a precise definition of the concept should be provided. In light of C6 (Section 3.2), and after reviewing the studies related to preference imprecision, I can now in this part of the paper—after reviewing the related studies—suggest a definition for the imprecision as having two components: *i. noise*, i.e., the random mistakes that people make; and, *ii. incommensurability*, i.e., the difficulty for articulating one's psychological satisfaction from a good precisely. The two components of imprecision can exist simultaneously: an individual might make mistakes calculating the psychological satisfaction of a good, but also might end up having a haziness for the result which can be represented as a range of e.g. expected utilities.

For the first component, the literature on stochastic models of choice provide a rich menu of methods, and has so far been successful in comparing the existing deterministic models and suggesting common patterns of behaviour that can be incorporated in prospective deterministic models. However, the experimental studies focusing on the second component have a key limitation: relying on merely subjects' self-reporting except one study; Bayrak and Kriström (2016), yet it should be developed further to solve its flaws mentioned in Section 2.1. I conjecture and to some extent hope that the new contributions to the literature will develop incentive compatible elicitation methods for the imprecision, and construct new models by considering the distinction between the noise and the incommensurability concepts.

References

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.

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.

Blavatskyy, P. R., & Pogrebna, G. (2010). Models of stochastic choice and decision theories: Why both are important for analyzing decisions. Journal of Applied Econometrics, 25(6), 963-986.

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.

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.

Fechner, G. (1966). Elemente der Psychophysik. Leipzig: Breitkopf and Hartel, 1860. Translated by Helmut E. Adler as Elements of Psychophysics (New York: Holt, Reinhart and Winston).

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

Butler, D., Isoni, A., Loomes, G., & Tsutsui, K. (2014). Beyond choice: investigating the sensitivity and validity of measures of strength of preference. Experimental Economics, 17(4), 537-563.

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

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

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.

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

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

Loomes, G., & Sugden, R. (1998). Testing different stochastic specificationsof 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.

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

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 Handbook of Mathematical Psychology Volume III, 249–410. New York: Wiley.

Sopher, B., & Gigliotti, G. (1993). Intransitive cycles: Rational Choice or random error? An answer based on estimation of error rates with experimental data. Theory and decision, 35(3), 311-336.

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.

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.

Wilcox, N. T. (2008). Stochastic models for binary discrete choice under risk: A critical primer and econometric comparison. In Risk aversion in experiments (pp. 197-292). Emerald Group Publishing Limited.

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