Article:
Bayrak, Oben and Hey, John Denis orcid.org/0000-0001-6692-1484 (2020) Decisions under Risk: Dispersion and Skewness. Journal of Risk and Uncertainty. pp. 1-24. ISSN 0895-5646

https://doi.org/10.1007/s11166-020-09333-6

Decisions under Risk

Dispersion and Skewness

Oben K. Bayrak and John D. Hey

Abstract

When people take decisions under risk, it is not only the expected utility that is important, but also the *shape* of the distribution of utility: clearly the *dispersion* is important, but also the *skewness*. For given mean and dispersion, decision-makers treat positively and negatively skewed prospects differently. This paper presents a new behaviourally-inspired model for decision making under risk, incorporating both dispersion and skewness. We run a horse-race of this new model against six other models of decision-making under risk and show that it outperforms many in terms of goodness of fit and shows a reasonable performance in predictive ability. It can incorporate the prominent anomalies of standard theory such as the Allais paradox, the valuation gap, and preference reversals, and also the behavioural patterns observed in experiments that cannot be explained by Rank Dependent Utility Theory.

JEL classification: D81.

Keywords: Allais Paradox; Anomalies; Decision under Risk; Dispersion; Expected Utility; Experiments; Non-Expected Utility; Pairwise Choice; Preference Functionals; Preference Reversals; Skewness; Stochastic Specifications; Valuation Gap.

Oben K. Bayrak: Centre for Environmental and Resource Economics (Umeå, Sweden) and Department of Economics, University of Gothenburg (Gothenburg, Sweden), obenbayrak@sabanciuniv.edu. Corresponding author.

John D. Hey: Department of Economics and Related Studies, University of York, Heslington, York, YO10 5DD, UK, john.hey@york.ac.uk.

Acknowledgements: We thank an anonymous referee, Glenn W. Harrison, Bo Ranneby, Gábor J. Székely, Nathaniel T. Wilcox and Wenchao Zhou for suggestions and/or comments. All errors are ours. The financial support from Tore Browaldh and the Jan Wallanders Foundations is also acknowledged.

1. Introduction

We present a new model of decision-making under risk. Crucial to our story is that the decision-maker (henceforth DM) considers not only the expected utility of a lottery, but also the dispersion and skewness of the utilities. Our new theory explains how an individual values lotteries and hence takes decisions under risk. The theory is based on a behavioural description of the evaluation process.

In this theory, evaluation is thought of as a two-stage process: first, the DM formulates an interval for the value of each lottery; secondly, the DM takes a weighted average of the extremes of this interval. Crucially, the interval depends upon the dispersion of the lottery, while the weights in the weighted average depend upon the skewness of the lottery and the optimism/pessimism of the individual.

Let us break this down into its two stages. As to the first stage, the literature suggests that many individuals find it difficult to state a precise Willingness-to-Pay (WTP) or Willingness-to-Accept (WTA) for a good (Bayrak and Kriström, 2016; Dubourg *et al.*, 1994, 1997; Morrison, 1998). Studies show that if subjects are given the option of stating their subjective valuations in terms of a single amount or an interval, more than half of subjects prefer to state their valuations in terms of an interval (Banerjee and Shogren, 2014, Håkansson, 2008; Bayrak and Kriström, 2016). However, because of the problems in incentivising the true revelation of intervals (if they exist), this evidence does not *prove* that people think in terms of an interval, but only *suggests* it. But this seems a natural phenomenon: if asked to state their valuation for some lottery, individuals usually find it difficult to specify a precise number. Of course, this depends upon the lottery: if it is a certainty, then there is no difficulty; if however, the lottery is risky then there is, and it becomes more difficult the more dispersed is the lottery. This is consistent with findings of Butler and Loomes (1988) and Cubitt *et al.* (2015), who conclude that, on the basis of their experimental evidence, the higher the variance of a lottery, the broader the imprecision range for a lottery[1].

Consider a lottery with pays either $x-d$ or $x+d$ each with probability one-half. The individual might, for example, say that the value is between $x-ad$ and $x+ad$ where $a<1$, and

[1] See Bayrak and Hey, 2020 for a review of relevant literature.

where *a* depends upon the confidence of the decision-maker. We formulate this more precisely shortly.

After the formation of an interval, the second stage sees the individual selecting a single value from the interval. This is done by taking a weighted average of the extremes of the interval, where the weights depend upon the *skewness* of the lottery and upon the optimism/pessimism of the decision-maker. We shall explain in more detail in the next section.

Readers should note that we are not presenting a new normative theory, instead we focus on the descriptive side of the problem: we distill our new model from the accumulated experimental findings in the literature to explain observed behaviour.

The paper is structured as follows: section 2 formalises our theory; section 3 describes how our model explains some typical 'anomalies' of EUT and RDUT found in the literature. Section 4 describes the 'horse race' that we conducted, comparing our model to six others familiar in the literature, with our methodology and stochastic assumptions described in section 5. Section 6 details the results of the 'horse race'. Section 7 concludes.

2. Model

Let *X* be the set of outcomes (consequences) with elements denoted by x_i, $i=1...I$. The outcome set consists of real numbers designating amounts of money. The objects of choice are lotteries, which are probability distributions over the set *X*. A lottery is denoted by $z=\{x_1,p_1;...;x_I,p_I\}$, where $x_1<...<x_I$ and $p_1,...,p_I$ are the associated probabilities such that $p_i \geq 0$ and $\sum_{i=1}^{I} p_i = 1$. Since this paper focusses on decisions under risk, the probabilities are taken as given by the DM. Let us denote by $\tilde{z}$ the *utilities* of the outcomes in the lottery and their associated probabilities: $\tilde{z}=\{u(x_1),p_1;...;u(x_I),p_I\}$. For notational convenience, we denote the expected utility of the lottery by $E(\tilde{z})=\sum_{i=1}^{I} p_i u(x_i)$.

We consider first the situation when the DM is choosing between two lotteries; later we shall consider the changes necessary when the individual owns a lottery and is considering selling it, or when the individual does not own a lottery and is considering buying it.

In the first stage, the DM is thought of as formulating an interval for the value of each lottery, perceiving it as between *LEU(z)* and *HEU(z)*, which are the lowest and highest expected utilities. This captures the idea that the DM is unable to attach a precise number to the value of a lottery, instead coming up with an interval, saying that the utility of the lottery is somewhere between some lower number and some higher number. The bounds are given by:

$$LEU(z) = E(\tilde{z}) - kD(\tilde{z}) \tag{1}$$

$$HEU(z) = E(\tilde{z}) + kD(\tilde{z}) \tag{2}$$

These are centered on the expected utility of the lottery with the distance between them depending upon the dispersion $D(\tilde{z})$ of the utilities of the outcomes of the lottery, and upon a parameter k – which adds the individual heterogeneity to the model and reflects the DM's uncertainty about his or her evaluations. $D(\tilde{z})$ denotes the standard deviation of the utilities, defined by $D(\tilde{z}) = \sqrt{\sum_{i=1}^{I} p_i (u(x_i) - E(\tilde{z}))^2}$.

Our theory now posits that at the second stage the DM evaluates the lottery by taking a weighted average of the Worst and the Best. If we denote this valuation by *V(z),* it is given by:

$$V(z) = \alpha_{S(\tilde{z})} \cdot WEU(z) + [1 - \alpha_{S(\tilde{z})}] \cdot BEU(z) \tag{3}$$

Here $V(z)$ is a weighted average of the Worst and the Best expected utilities. $\alpha_{S(\tilde{z})} \in [0,1]$ is defined as the pessimism/optimism level of the individual, and this is a function of the skewness of the utilities in the lottery. Skewness, $S(\tilde{z})$, is defined by the Pearson Measure[2].

Notice that in equation (3), the weights are attached to $WEU(z)$ and $BEU(z)$, instead of $LEU(z)$ and $HEU(z)$. Because the designation of the bounds as the $WEU(z)$ and $BEU(z)$ is done in a task/situation contingent way: for a choice and buying task the lower bound and the upper bound are perceived as the worst case (WEU) and the best case (BEU), respectively. However, for a selling task the reverse is the case: for a buyer the upper bound of the imprecision range is the best thing that can happen; for a seller the opposite is the case, since the lottery is going to be given away.

For expositional simplicity, we initially restrict our analysis to pairwise choice problems, so equation (3) becomes $V(z) = \alpha_{S(\tilde{z})} \cdot LEU(z) + [1 - \alpha_{S(\tilde{z})}] \cdot HEU(z)$.

---

[2] $S(\tilde{z}) = \mu^3 / \sigma^3$, where μ^3 is the third central moment and σ is standard deviation. If σ is zero (so that the lottery is degenerate and the Pearson measure undefined), we naturally put the skewness equal to zero.

When we substitute (1) and (2) into (3) we get the following:

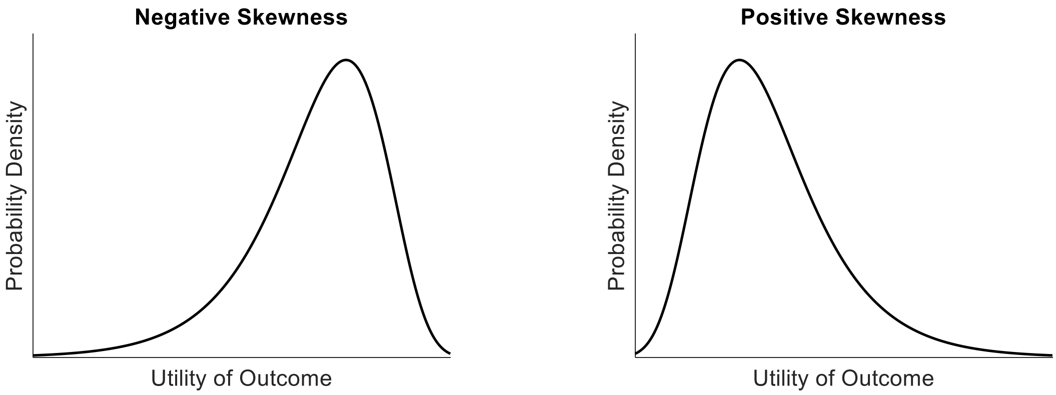
$$V(z) = E(\tilde{z}) + (1 - 2\alpha_{S(\tilde{z})}) \cdot k \cdot D(\tilde{z}) \tag{4}$$

Here all variables are measured in units of utility. Let us simplify this by putting $(1 - 2\alpha_{S(\tilde{z})}) = S(\tilde{z})$ thus getting the final form of our model[3]:

$$V(z) = E(\tilde{z}) + kD(\tilde{z})S(\tilde{z}) \tag{5}$$

There are three key components[4] in the model: First the *dispersion* of (the utilities of) the lottery, $D(\tilde{z})$. Second the *skewn*ess of (the utilities of) the lottery, $S(\tilde{z})$. And there is the parameter *k*. Note that $D(\tilde{z})$ is necessarily positive while $S(\tilde{z})$ can be either positive or negative. The parameter *k* is individual-specific and could be either positive or negative. We discuss the implications below.

Figure 1: Skewness Examples



If k is positive: Dispersion hurts when skewness is negative (left-skewed) and is a desirable property when the skewness is positive (right-skewed) [5]. As $D(\tilde{z})$ is positive, then how

[3] Note that in the most general form of DS given in (3), $\alpha_{S(\tilde{z})}$ is introduced as a function of skewness, but here we adopt a simplification. Alternatively, pessimism can have a more complicated functional form that, for example, includes the DM's experience in the past; we leave this issue for future work.

[4] We note that Hagen (1991) also proposes a model involving dispersion and skewness. Hagen incorporates these in an additive manner in the preference functional; they are seen as the source of extra utility and disutility, respectively. Instead our theory has behavioural motivations for the way that dispersion and skewness affect the preference functional.

[5] The literature on the relationship between the skewness and preferences dates back to the 1990s: The common finding is that individuals favour positively/right skewed lotteries more than the negatively/left skewed lotteries (Ebert, 2015). One explanation for this behaviour is that people feel excited and hopeful for the large gains that come with a low probability such as national lotteries. For example, national lotteries can be interpreted as one is actually "buying a dream" which includes for example imagining how one can spend the prize and the joy of quitting one's job (Forrest *et al.*, 2002; Garrett and Sobel, 1999). Possibly the most direct evidence comes from a study which employs neuroimaging measures: Wu et al. (2011) found that positive skewed gambles increased positive arousal, but negatively skewed gambles increased negative arousal and

skewness affects the valuation depends upon the sign of the skewness. See the figure above. If the skewness is positive our theory implies that the lottery is valued more than it would be by its expected utility. Moreover, an increase in dispersion *in*creases the value of the lottery. This can represent the behaviour of an *optimistic* person whose attention is drawn to the possible high outcomes of the lottery. If the skewness is negative our theory implies that the lottery is valued less than it would be by its expected utility. Moreover, an increase in dispersion *de*creases the value of the lottery. This can represent the behaviour of a *pessimistic* person whose attention is drawn to the possible low outcomes of the lottery. *If k is negative* we get the reverse.

In summary, our model has three key components: dispersion, skewness and optimism/pessimism[6].

3. Anomalies

In Section 3.1, we look at what can DS can tell us about the prominent 'anomalies' of standard economic theory[7]. Most of the 'anomalies' are situations in which behaviour is not consistent with that of EU. With EU, indifference curves in the Marschak-Machina Triangle (MMT) are parallel straight lines. This is not the case with non-expected utility theories, DS included. In Section 3.2, we focus on behavioral patterns observed in experiments that cannot be explained by RDUT and CPT, but possible under DS.

3.1. Anomalies of EUT

3.1.1. The Common Consequence and the Common Ratio effects

perceived risk. Subjects preferred positively skewed and high dispersion lotteries more than the negatively skewed gambles. Tversky and Kahneman (1992) found that subjects exhibit risk-loving preferences for positively skewed lotteries and risk-averse preferences for negatively skewed lotteries. Golec and Tamarkin (1998) find that people tend to favour the long-shot options in horse races with high prizes but low probabilities. Grossman and Eckel (2015) developed a new protocol consisting of 3 tasks. In task 1 subjects make a choice between 6 lotteries which are zero-skewed lotteries but differ in dispersion. In task 2, subjects are asked to keep the preferred lottery in task 1 or to choose one of the new 6 lotteries. The new 6 lotteries are the modified versions of Task 1 lotteries in a way to keep their dispersion same but have a skewness level of 1. Finally, in task 3, a new set of 6 lotteries are presented which have a skewness level of 2. There are two important results from that study: firstly more 88.2% of the subjects prefer the skewed lotteries over the zero-skewed lotteries. Secondly, and more importantly is that an increase in skewness leads 37.6% of the subjects to take on greater risk/seek for higher dispersion in their choice for lotteries.

[6] We could insert a fourth component – an editing phase, in which the DM looks for (first-order stochastically) dominating lotteries in the lottery pair and simply chooses them without evaluating them first. We discuss this in Appendix G.

[7] For a reader unfamiliar with the anomalies see online for a brief summary.

A Common Consequence problem is described by two choice problems between pairs of lotteries, constructed in the specific way presented in Table 1. Here p_1, p_2 and p_3 are the associated probabilities of the outcomes of the lotteries; such that $x_1 < x_2 < x_3$. S2 and R2 are derived from S1 and R2 by moving probability *cc* (the 'common consequence') from x_2 to x_1. An individual whose preferences are compatible with EU would choose either 'S' or 'R' in **both** choice problems; common consequences added or subtracted to the two prospects should have no effect on the desirability of one prospect over the other, because the probabilities are incorporated in a linear way in EU[8]. Figure 2 superimposes the common consequence lotteries on a DS indifference map. Here S1 would be preferred to R1 and R2 to S2.

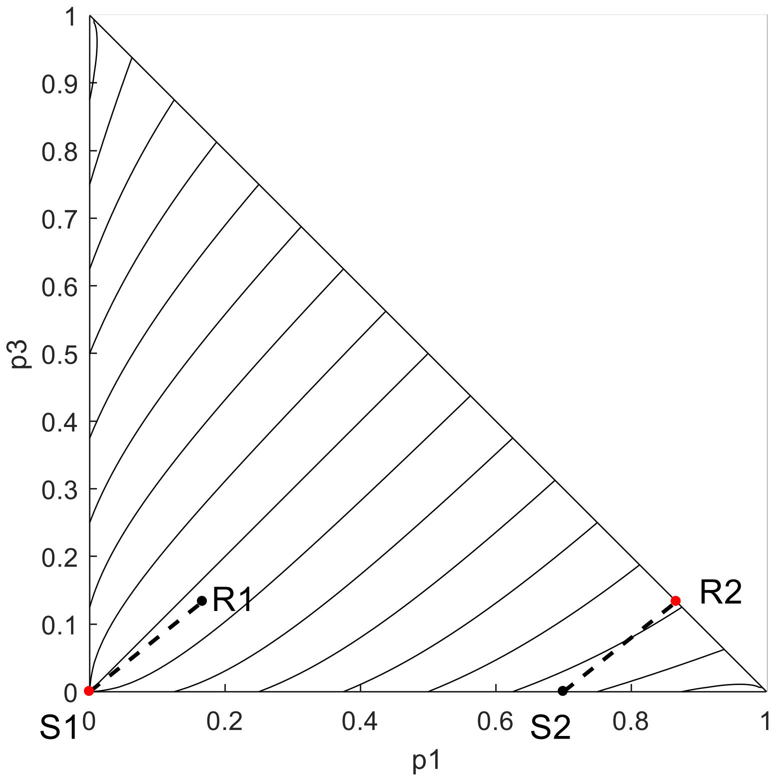
Table 1: Common Consequence Lotteries

Lottery	**p_1**	**p_2**	**p_3**
S1	*0*	*1*	*0*
R1	*a*	*cc*	*1-a-cc*
S2	*cc*	*1-cc*	*0*
R2	*a+cc*	*0*	*1-a-cc*

Notes: The table presents the generic structure of four lotteries used in common consequence problems. There are two binary choice tasks between lotteries S1 and R1 and between S2 and R2. In the first column are the lottery labels. All lotteries can have at most three outcomes with associated probabilities, p_1, p_2 and p_3 presented in the last three columns. *a* is a scaling parameter whereas *cc* stands for common consequence.

[8] Formally, under EU, putting $u(x_1)=0$, $u(x_2)=u$ and $u(x_3)=1$, S1 is preferred to R1 if and only if $u > cc \cdot u + (1-a-cc)$, while S2 is preferred to R2 if and only if $u \cdot (1-cc) > 1-a-cc$.

Figure 2: Common Consequence Effect



A related phenomenon is the Common Ratio effect: there are two choice tasks and each task includes a pair of lotteries as shown in Table 2.

Table 2: Common Ratio Lotteries

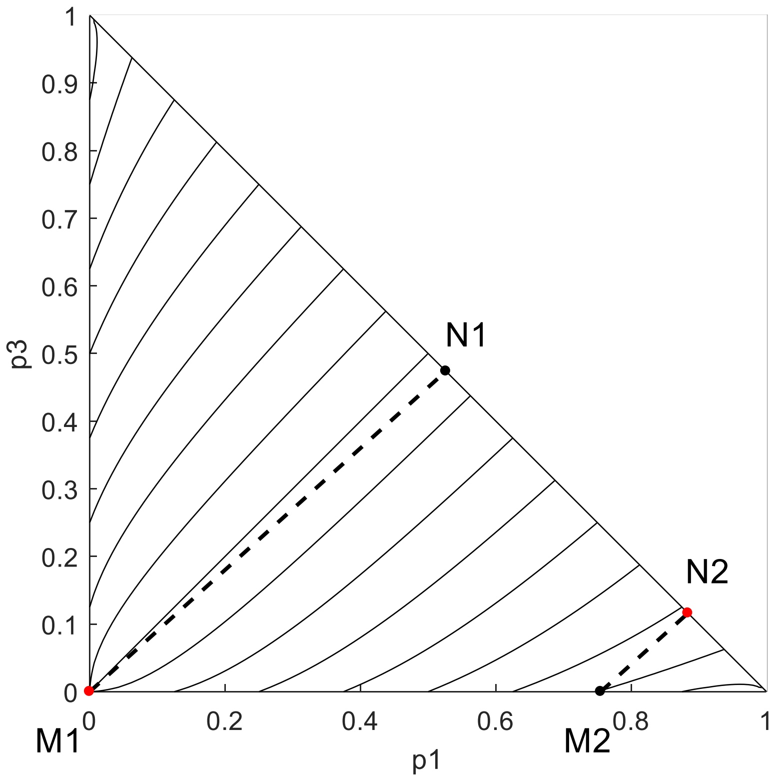
Lottery	p_1	p_2	p_3
M1	*0*	*1*	*0*
N1	*1-a*	*0*	*a*
M2	*1-cr*	*cr*	*0*
N2	*1-a.cr*	*0*	*a.cr*

Notes: The table presents the generic structure of four lotteries used in common ratio problems. There are two binary choice tasks between lotteries M1 and N1 and between M2 and N2. In the first column are the lottery labels. All lotteries can have at most three outcomes with associated probabilities, *p1, p2* and *p3* presented in the last three columns. a is a scaling parameter whereas *cr* stands for common ratio.

The common choice pattern of choosing M1 and N2 is inconsistent with the predictions of EU[9]. Figure 3 shows an example of such lotteries superimposed on DS indifference curves. Here M1 would be preferred to N1 and N2 to M2.

[9] Formally, under EU, putting $u(x_1)=0$, $u(x_2)=u$ and $u(x_3)=1$, M1 is preferred to N1 if and only if $u > a$, while M2 is preferred to N2 if and only if $u \cdot cr > a \cdot cr$.

Figure 3: Common Ratio Effect



3.1.2. The Valuation Gap

Our model implies that Willingness To Pay (WTP) may differ from Willingness To Accept (WTA) and this can be shown immediately from equation (3). As explained in Section 2, the key point that allows DS to incorporate the valuation gap is the *task contingent designation* of the bounds of the imprecision range as the worst and the best cases: in a buying task *LEU(z)* and *HEU(z)* are perceived as the worst and the best cases, respectively; whereas in a selling task the perceptions are reversed.

So, equation (3) has the following forms for buying and selling tasks respectively as below:

$$V(z) = \alpha_{S(\tilde{z})} LEU(z) + [1 - \alpha_{S(\tilde{z})}] HEU(z) \quad (6)$$

$$V(z) = \alpha_{S(\tilde{z})} HEU(z) + [1 - \alpha_{S(\tilde{z})}] LEU(z) \quad (7)$$

Inspection of (6) and (7) shows that valuations will change unless $\alpha_{S(\tilde{z})}$ = 0.5 or the dispersion is zero. So WTP may differ from WTA.

DS ascribes the gap to uncertainty. For a gap to be observed a necessary condition is that the good should have an uncertain nature; this implies that there exists a dispersion of possible utilities. This conjecture is consistent with the findings in Plott and Zeiler (2005) which reports a significant gap for lotteries, but not for the ordinary goods used in their experiments. Additionally, we note that uncertainty also can be associated with *unfamiliarity* of goods. In other words, having an uncertain nature regarding the utilities is not a specific feature of the standard lotteries used in experiments. DMs might view unfamiliar goods as if they are lotteries giving different utilities for different states of the world with subjective probabilities attached to those states. Contraiwise, a DM might view familiar goods that are regularly purchased as close to degenerate lotteries –since the DM can be sure about the utility that (s)he will attain from them. For example, Bateman et al., (1997) observed the gap both for a familiar (Coke) and an unfamiliar (luxury chocolate), yet the gap is larger for the latter (for a detailed discussion see Teitelbaum and Zeiler, p.376, 2018).

3.1.3. Preference Reversals

A preference reversal occurs when the DM prefers a lottery A to lottery B, but values B more highly. This possibility is immediately apparent from our discussion of the valuation gap above.

3.2. Anomalies of RDUT and CPT

3.2.1. The Reverse Common Ratio Effect

Blavatskyy (2010) discovered that the exact opposite of the classical common ratio pattern is more frequently observed than the classical one when in the first question, the expected value of risky lottery is above the sure payoff: 57.1% of the subjects chose the risky lottery in the first question and switch to choosing a safer lottery when the probabilities of winning non-zero prizes are scaled down by a common ratio. The lotteries used by Blavatskyy (2010) are presented in Table 3: the ones in question 2 were simply constructed by dividing the probability of winning nonzero prizes of the lotteries in question 1 by 3.

Table 3: Blavatskyy (2010) pairs for reverse common ratio effect

Question 1		**Question 2**		
Sure Payoff	**Risky Lottery**	**Safer Lottery**	**Riskier Lottery**	**k range**
\$60	3/4 chance of \$100	1/3 chance of \$60	1/4 chance of \$100	k< -0.2
\$50	3/4 chance of \$100	1/3 chance of \$50	1/4 chance of \$100	k< -0.3
\$40	3/4 chance of \$100	1/3 chance of \$40	1/4 chance of \$100	k< -0.3

Note: Questions 1 and 2 are binary choice questions. Labels of the lotteries in Question 1 are "Sure Payoff" which implies a degenerate lottery and "Risky Lottery", and in Question 2 are "Safer Lottery" and "Riskier Lottery".. Risky, Safer and Riskier lotteries are two outcome lotteries whose second outcome is getting zero. Lotteries in question 2 were constructed by dividing the probability of winning nonzero outcomes of the corresponding lotteries in question 1 by 3. Last column reports the range of values for parameter k which allows DS to incorporate reverse common ratio effect.

Neither EUT nor non-EUT theories which were designed to account for the Allais paradox (including RDUT and CPT), predict a Reverse Common Ratio effect to be observed in two or all three simultaneously (See Blavatskyy, 2010, p.228). On the other hand, DS can account for the reverse common ratio effect in all three pairs even for the simplest case of linear utility function for the range of k values presented in the final column of Table 3 calculated using the model in (5).

3.2.2. Violation of Ordinal Independence/Upper Tail Independence

DS allows for violations of ordinal independence (OI) or upper tail Independence, yet these must be satisfied by RDUT and CPT (Green and Jullien, 1988, Quiggin, 1982). Wu (1994) reports systematic violations of OI which requires that replacing a common right tail of two lotteries/cumulative distributions with another common right tail should not change the original preference order. Wu's experiment includes 25 OI questions. For 12 of them, the results are significant at 1%. Consider the two pairs from Wu (1994, p.45) presented in the table below.

Table 4: Two pairs of lotteries for testing Ordinal Independence in Wu (1994, p.45)

Outcomes	R	S	R'	S'
$750	65%	65%	-	-
$240	30%	-	95%	65%
$200	-	34%	-	34%
0	5%	1%	5%	1%
EV	559.5	555.5	228.0	224.0
Standard Deviation	264.3	265.8	52.3	29.4
Skewness	-0.8	-0.7	-4.1	-4.5

Notes: In the upper part of the table, the outcomes of the 2 pairs of lotteries used in OI shown under column 1 and the associated probabilities for each lottery is under the remaining columns. Lottery S and R have a common tail which is the best outcome (750$), the remaining lotteries are constructed by transferring the probabilities of this outcome to the second-best outcome. In the bottom part of the paper EV, standard deviation and skewness of the lotteries are reported.

All the lotteries designed for testing OI in the experiment have the same structure: R and S pairs have a common right tail, i.e. winning $750 with a probability of 65%. R' and S' are constructed simply by transferring the probability of the common tail to the next best outcome in the pair's context, i.e. winning \$240. The typical finding in the experiments is that subjects choose R in the first pair and S' in the second pair.

For the simplest case when the utility function is linear, the expected utility, standard deviation and skewness of the lotteries are shown in the bottom part of the table above. An individual chooses R over S, using (5) we have:

$$559.5 + k \cdot 264.3 \cdot -0.8 > 555.5 + k \cdot 265.8 \cdot -0.7 \tag{8}$$

$$k < 0.14 \tag{9}$$

Similarly, when an individual chooses S' over R', we have:

$$228.0 + k \cdot 52.3 \cdot -4.1 > 224.0 + k \cdot 29.4 \cdot -4.5 \tag{10}$$

$$k > 0.05 \tag{11}$$

Both choice correspondences can be simultaneously accommodated under DS for k values between 0.05 and 0.14.

3.2.3. Violation of Stochastic Dominance

Violation of stochastic dominance contradicts theories such as RDUT and CPT. The experimental literature has documented several cases showing significant violations of stochastic dominance (see, for example Tversky and Kahneman, 1986; Carbone and Hey, 1995; Loomes and Sugden, 1998; Hey, 2001 and Birnbaum, 2005). Birnbaum and Navarrete (1998) used the four lotteries shown in Table 5 to test for stochastic dominance. The fourth column shows that a considerable portion of the subjects chose the stochastically dominated lottery over the dominating one. The proportion of subjects violating stochastic dominance is significantly higher than 50% for all lotteries as shown in the third column in table 5. Using the model in (5), the last column solves for the parameter k that allows for such violations for the simplest case when the utility function is linear.

Table 5: Lotteries used to test stochastic dominance in Birnbaum and Navarrete (1998, p.61)

Dominant	Dominated	% Violations	k range
\$12, 5%; \$14, 5%; \$96, 90%	\$12, 1%; \$90, 5%; \$96, 85%	73	k< -1.5
\$3, 6%; \$5, 6%; \$97, 88%	\$3, 12%; \$92, 4%; \$97, 84%	61	k< -0-7
\$6, 2%; \$8, 3%; \$99, 95%	\$6, 5%; \$91, 3%; \$99, 92%	73	k< -0.7
\$4, 1%; \$7, 1%; \$95, 98%	\$4, 2%; \$89, 2%; \$95, 96%	73	k< -0.2

Notes: The first two columns present the outcomes and associated probabilities of the stochastically dominant and dominated lotteries respectively. The third column reports the percentage of people violating the dominance in Birnbaum and Navarrete (1998); the final column lists the range of values for parameter k which allows DS to incorporate the observed violations of stochastic dominance when the utility function is linear.

3.2.4. Lower Distribution Independence: 3-LDI and 3-2 LDI

Lower distribution independence (LDI) is a property that is implied by EUT, yet CPT (for gains only)/RDUT predicts the violation of it if the employed de-cumulative weighting function is nonlinear (see Birnbaum, 2005, p.1349). The first type of lower distribution independence, that is 3-LDI, is defined as follows:

$$\begin{aligned} &S=(x,p;y,p;z,1-2p)\succ R=(x',p;y',p;z,1-2p)\Leftrightarrow \\ &S2=(x,p';y,p';z,1-2p')\succ R2=(x',p';y'p';z,1-2p') \end{aligned} \qquad (12)$$

where $x>y>z>0$. Both lotteries in the first pair give the lowest outcome *z* with a probability of *1-2p*. The remaining probability is split equally between the other outcomes in each lottery. The second pair is simply constructed by assigning a different probability to the common outcome (*z*), and then the remaining probability is distributed equally between the outcomes in each lottery. Birnbaum (2005) uses the four lotteries in Table 6 and concludes that predicted violations by RDUT and CPT do not materialise in his experiments.

Table 6: Lotteries for testing 3-LDI (Birnbaum, 2005)

Lotteries	\$2	\$4	\$40	\$44	\$96	EV(.)	D(.)	S(.)
S	0.80		0.10	0.10		10.0	16.0	1.51
R	0.80	0.10			0.10	11.6	28.1	2.66
S2	0.10		0.45	0.45		38.0	12.1	-2.54
R2	0.10	0.45			0.45	45.2	46.0	0.20

Notes: In the left part of the table, the lotteries used in 3-LDI experiments are shown; the values in cells are the probabilities associated with the five monetary outcomes reported as the column labels. The right part of the table includes the expected value, standard deviation and skewness of the lotteries, respectively.

Let us now investigate how DS behaves for different values of the parameter *k* in the simplest case of a linear utility function. Using the model in (5) we get the following inequalities:

$$S \succ R \Leftrightarrow 10 + 24.2 \cdot k > 11.6 + 74.9 \cdot k \quad (13)$$

$$S2 \succ R2 \Leftrightarrow 38 + (-30.9) \cdot k > 45.2 + 9 \cdot k \quad (14)$$

Solving these leads to the following constraints on the parameter *k* that can be simultaneously satisfied: $k < -0.03$ and $k < -0.17$.

The second type of lower distribution independence that Birnbaum (2005) examines is 3-2 LDI and defined as follows:

$$\begin{aligned} &A = (x, 0.5; y, 0.5) \succ B = (x', 0.5; y', 0.5) \Leftrightarrow \\ &C = (x, p; y, p; z, 1-2p) \succ D = (x', p; y', p; z, 1-2p) \end{aligned} \quad (15)$$

3-2 LDI simply states that cancelling a common part of two lotteries (*z* with a probability of *1-2p*) and allocating its probability equally to other outcomes should not reverse the preference order. Birnbaum (2005) shows that RDUT might violate the property when the commonly estimated parameters used in the functional. Yet, in his experiments, the data show no systematic violation of the property. In table 7, the first five columns describe the four lotteries used to test 3-2 LDI and last three columns show the three measures that DS uses in its formula.

Table 7: Lotteries for testing 3-2 LDI (Birnbaum, 2005)

Lotteries	**\$2**	**\$4**	**\$40**	**\$44**	**\$96**	**EU**	**D(.)**	**S(.)**
S	0	0	0.5	0.5	0	42.0	2.0	0.00
R	0	0.5	0.0	0.0	0.5	50.0	46.0	0.00
S2	0.04	0	0.48	0.48	0	40.4	8.1	-4.25
R2	0.04	0.48	0	0	0.48	48.1	46.0	0.08

Notes: In the left part of the table, the lotteries used in 3-2 LDI experiments are shown; the values in cells are the probabilities associated with the five monetary outcomes reported as the column labels. The right part of the table includes the expected value, standard deviation and skewness of the lotteries, respectively.

Let us now investigate values of the parameter *k* for which DS conforms to 3-2 LDI. For the simplest case of a linear utility function, using the model in (5), we see that *R* is chosen over *S* since, both lotteries have zero skewness, that is, they are symmetric around the mean. Thus, only the EU part of our model becomes relevant for this comparison: $42 < 50$. For the second pair, R2 has a higher EU than S2. Even at a first glance it is easy to see that individual will choose R2 over S2, therefore will not reverse his or her preferences that we articulated for the first pair, if *k* is sufficiently high:

$$S2 \prec R2 \Leftrightarrow 40.40 + (-34.4) \cdot k > 48.08 + 3.7 \cdot k \qquad (16)$$

DS satisfies this choice correspondence, in other words 3-2 LDI if $k > -0.2$.

4. A horse race

We compare the goodness-of-fit and the predictive ability of our theory with six other prominent theories in the literature. We use data from Hey (2001) which contains the pairwise choice responses of 53 individuals for the same 100 pairs of lotteries on five different days presented in different orders. The four monetary outcomes for the lotteries were -£25, £25, £75 and £125 respectively[10].

We consider the following preference functionals: Expected Utility theory (EU), Disappointment Aversion theory (DA) (Gul, 1991), Prospective Reference theory (PR) (Viscusi, 1989), Rank dependent expected utility theory with a Prelec weighting function (RL) (Quiggin, 1982; Prelec, 1998), Salience Theory (ST) (Bordalo et al., 2012) and Weighted Utility theory (WU) (Chew, 1983; Dekel, 1986). We test these against our Dispersion and Skewness theory (DS). Details of the preference functionals can be found in Hey (2001) and Appendix A, though we should comment briefly on our implementation of Salience Theory, as this was not considered in Hey (2001) but is now popular in the literature.

In the Hey (2001) experiment subjects were presented with two lotteries side by side and not juxtaposed as in Salience theory. So, we have to make some assumption as to how subjects did the juxtapositioning. What we have assumed is the following. If the two lotteries are $X = \{x_1, p_1; x_2, p_2; x_3, p_3; x_4, p_4\}$ and $Y = \{y_1, q_1; y_2, q_2; y_3, q_3; y_4, q_4\}$, then we have assumed that the subjects consider the choice problem as over 16 'states of the world' leading to outcomes either x_i or y_j with probabilities $p_i q_j$ (for *i=1,2,3,4* and *j=1,2,3,4).* Now we can apply Salience Theory[11].

Our procedure is to estimate all seven models by maximum likelihood using GAUSS. We do this using the data in a variety of ways. We have 500 observations, collected in batches of 100 on 5 separate days. We do the following:

[10] There was a participation fee of £25. See Appendix B for the lotteries used in the experiment.
[11] Clearly this is not the only way that we can posit how subjects do the juxtapositioning.

1. Estimate using the first 100 observations (“1st 100”).
2. Estimate using the second 100 observations (“2nd 100”).
3. Estimate using the third 100 observations (“3rd 100”).
4. Estimate using the fourth 100 observations (“4th 100”).
5. Estimate using the fifth 100 observations (“5th 100”).
6. Estimate using all 500 observations (“All 500”).
7. Estimate using the first 400 observations and predict on the last 100, using the estimates of the parameters from the first 400 (“1st 400”).
8. Estimate using the first 300 observations and predict on the last 200 using the estimates of the parameters from the first 300 (“1st 300”).
9. Estimate using the first 200 observations and predict on the last 300 using the estimates of the parameters from the first 200 (“1st 200”).
10. Estimate using the first 100 observations and predict on the last 400 using the estimates of the parameters from the first 100 (“1st 100”).

5. Methodology and stochastic assumptions

As noted above we fit the various models by maximum likelihood. To do this we need some assumptions about the stochastic nature of the data since it is abundantly clear that subjects make mistakes in experiments. We follow first what Wilcox (2008) calls “a strong utility model”. The particular form of strong utility that we use first is what is sometimes called the Luce Model. In addition, since Wilcox (2008) reports that the stochastic specification may be more important than the preference functional, we also investigate the White Noise or Fechner story[12]; the differences are small.

To explain what we have done, we need to give more detail. In the experiment there were four possible outcomes: -£25, £25, £75 and £125. All the models involve a utility function over the various outcomes. Such a function involves two normalisations. We normalise the utility of -£25 to be 0; the second normalisation comes through our strong utility story. In the Luce Model, in a pairwise choice between A and B where the value of A is V_A and the value of

[12] Further results are reported in Appendices C and F: The distribution of estimated values of the parameter k can be found in Appendix C, results with White Noise specification are presented in Tables F2a and F3b in Appendix F.

B is V_B, then the probability of A being chosen over B is given by $\frac{\exp(\lambda V_A)}{\exp(\lambda V_A)+\exp(\lambda V_B)}$ where λ is a parameter inversely related to the level of error; the smaller is λ the noisier is the subject. We put *$\lambda=1$;* this is our second normalisation. The estimated values of the utilities of £25, £75 and £125 are therefore relative to this normalisation. The smaller are the estimated values of the utilities of £25, £75 and £125, the noisier is the subject. The probability that A is chosen over B is thus $1/(1+\exp(V_B-V_A))$. In contrast, in the White Noise (Fechner) story this probability is given by 1-cdf(V_B-V_A) where cdf(.) is the cumulative distribution function of the unit normal. We apply this specification for all the decision problems. We call this specification Version C.

Clearly, this is just one of many possible stochastic specifications. Problems where one option *dominates* the other might be considered different from non-dominating problems. DS assumes an editing phase, where domination is recognised, but the DM might tremble in its implementation. So, in Appendix D, we introduce two different specifications, both with the editing of dominated problems for DS, and with a tremble in its implementation: Version A where the tremble is exogenous; Version B where it is endogenous (and estimated from the data): the results from them are in Appendices E and F. We note that the results for DS with these other specifications are better than the results for DS reported below. These versions obviously favour DS.

6. Results

This section summarises our results. We measure performance both by goodness-of-fit and by predictive ability.

All these comparisons involve *ranking* the models in some way. We note that there is no general agreement on the 'best' ranking, nor even on what that might mean. So, we present a set of different rankings and leave it up to the reader to judge.

We first count the percentage of times that each model comes first, either on the Akaike or on the Bayes information criterion, or in terms of predictive ability. The first two both penalise the goodness-of-fit – the maximised log-likelihood – by the number of parameters involved in the preference functional. EU has three parameters, WU and RL have five and the others have four. The results using the Akaike criterion are in the top part of Table 8. We mark the 'winners' in **bold**. It will be seen that RL dominates under the Akaike criterion.

Considering the nine rows under the Akaike criterion, DS is ranked as the best model in one case, second best in five cases, third best in two cases and fifth in one case. Using the Bayesian criterion, which penalizes the number of estimated parameters (degrees of freedom) more heavily compared to the Akaike Criterion, EU wins the race for six cases in which DS is ranked as the second best, and DS is ranked as the best model in three cases.

We get a different picture on predictive ability shown in the final four rows of Table 8. We measure this by fitting the models on a subset of the observations and using the estimated parameters to predict decisions on the remaining observations. We measure this by fitting the models on a subset of the observations and using the estimated parameters to predict decisions on the remaining observations. We measure predictive ability by the log-likelihood of the prediction set using the parameters from the estimation set, **without** penalizing models for the number of parameters estimated. We note that this gives an unfair advantage to *more complicated* models (such as RL in our case). (See Busemeyer and Wang (2002) for a detailed discussion). RL is ranked as the best model, and DS is the second in one case and third in two cases and the last model in one case.

An alternative way of looking how 'good' models are is to look at the *average* ranking rather than at the number of times each model comes first. Why one might prefer to do this is that if one model comes first for half the subjects and last for the other half, while a second model is always second, one might prefer the latter. We start again with the Akaike criterion. Note here that 'first' is ranked 1 and 'last' is ranked 7, so that the lower the average ranking the better. Table 9 gives the detail.

Table 8: % of the time that each model comes first; Luce Model

	EU	DA	DS	PR	RL	ST	WU
Akaike Criterion							
1st 100	13	11	**23**	8	19	15	13
2nd_100	17	9	23	6	**28**	6	15
3rd 100	17	11	15	13	**26**	8	11
4th 100	15	9	11	13	**25**	6	21
5th 100	13	8	17	11	**34**	4	15
all 500	9	6	13	9	**38**	4	21
1st 400	11	6	15	11	**42**	4	11
1st 300	9	13	13	8	**40**	8	9
1st 200	9	9	17	8	**38**	6	13
Bayesian Criterion							
1st 100	**36**	9	21	8	9	13	6
2nd_100	**51**	6	23	4	13	6	2
3rd 100	**32**	13	19	11	11	8	8
4th 100	**38**	9	17	11	11	4	9
5th 100	**43**	4	21	9	11	6	8
all 500	17	6	**28**	8	26	6	9
1st 400	19	8	**25**	9	23	6	11
1st 300	19	11	**26**	9	17	9	8
1st 200	**34**	8	25	9	15	8	2
Predictive Ability							
1st 400	4	11	19	11	**30**	8	19
1st 300	0	15	17	15	**23**	9	21
1st 200	6	17	15	9	**26**	15	11
1st 100	15	9	8	9	**26**	19	13

Notes: The table reports the number of times a model comes first in descriptive ability (using the Akaike and Bayesian Information Criteria) and predictive ability. The explanation for row labels can be found in Section 4. Model abbreviations from left to right: EU: Expected Utility; DA: Disappointment Aversion DS: Dispersion Skewness; PR: Prospective Reference; RL: Rank Dependent Utility with Prelec Function; ST; Salience; WU: Weighted Utility.

DS does well throughout. Under the Akaike criterion, DS is the best in five out of nine cases and second best in the remaining four cases for which RL is the winner. When we look at the Bayesian Criterion, DS is ranked as the best in eight out of the nine cases, and the second best in the one case in which EU is the winner. For predictive ability, we see that RL is the winner and DS follows it as being the second best with a small difference in average rankings.

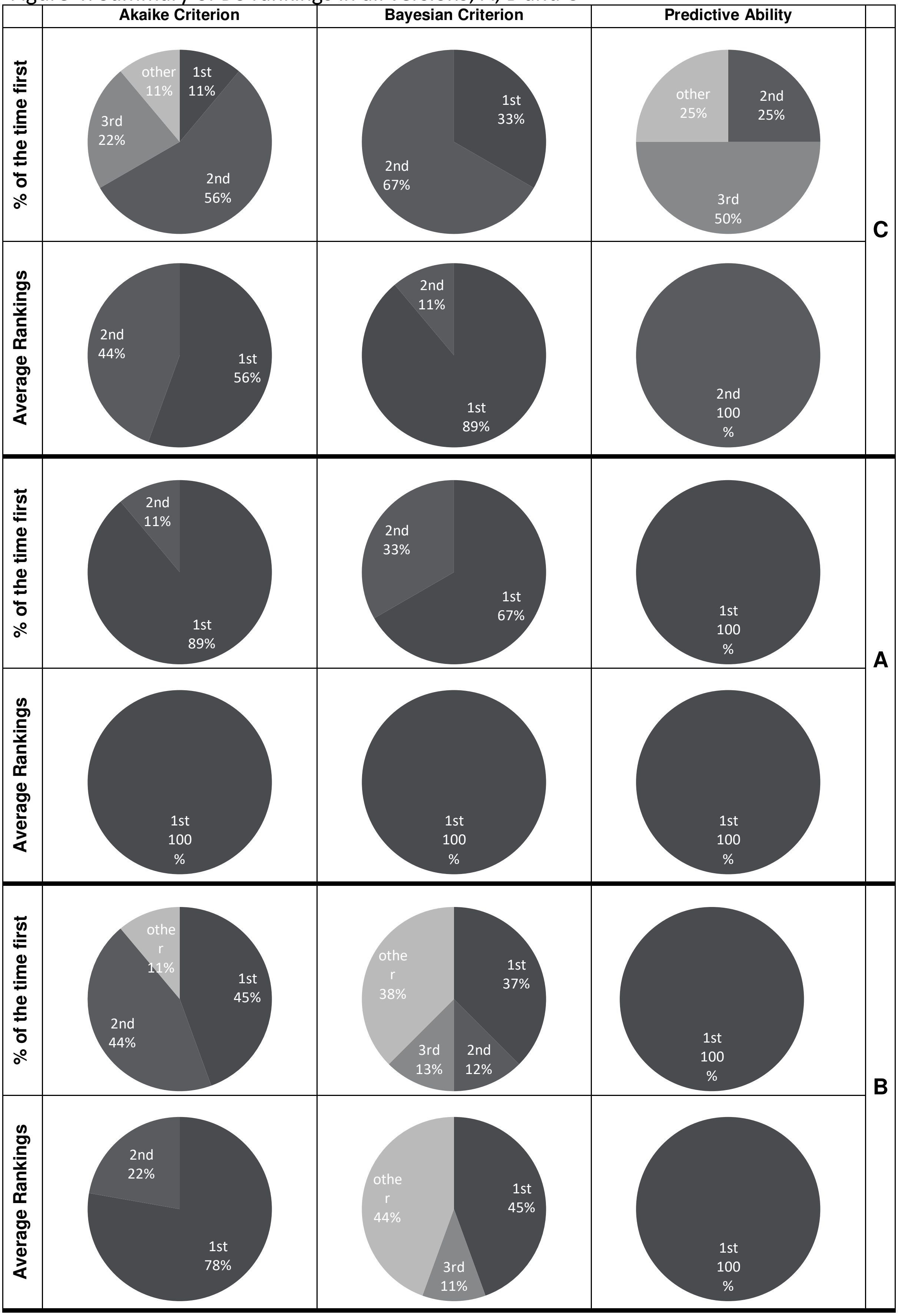
Table 9: Average Rankings; Luce Model

	EU	**DA**	**DS**	**PR**	**RL**	**ST**	**WU**
Akaike Criterion							
1st 100	4.2	4.7	**2.6**	3.8	3.5	4.9	4.4
2nd_100	4.0	4.5	**2.6**	4.1	3.2	5.4	4.0
3rd 100	4.1	4.5	**2.9**	3.8	3.3	5.0	4.2
4th 100	4.1	4.7	**3.0**	3.7	3.1	5.5	3.8
5th 100	4.0	4.7	3.1	3.7	**2.8**	5.5	4.1
all 500	5.1	4.8	2.8	4.2	**2.4**	5.4	3.3
1st 400	5.1	4.8	2.8	3.9	**2.5**	5.4	3.5
1st 300	4.9	4.7	2.8	4.1	**2.5**	5.3	3.6
1st 200	4.6	4.5	**2.6**	4.0	2.9	5.2	4.2
Bayesian Criterion							
1st 100	3.1	4.5	**2.5**	3.6	4.3	4.7	5.2
2nd_100	2.6	4.2	**2.4**	3.8	4.5	5.2	5.2
3rd 100	2.8	4.2	**2.8**	3.8	4.4	4.9	5.0
4th 100	3.0	4.4	**2.7**	3.5	4.1	5.4	4.8
5th 100	**2.8**	4.5	2.9	3.5	4.0	5.3	5.0
all 500	4.1	4.7	**2.3**	3.9	3.2	5.3	4.4
1st 400	3.9	4.7	**2.4**	3.7	3.3	5.3	4.7
1st 300	3.8	4.6	**2.4**	3.8	3.6	5.1	4.8
1st 200	3.2	4.4	**2.4**	3.8	4.1	4.9	5.2
Predictive Ability							
1st 400	4.8	4.2	3.4	3.9	**2.7**	5.3	3.6
1st 300	5.0	4.1	3.2	3.8	**2.8**	5.2	3.7
1st 200	4.5	3.9	3.6	4.0	**3.5**	4.8	3.7
1st 100	4.2	3.7	3.6	3.7	**3.5**	4.7	4.5

Notes: The table reports the number of times a model comes first in descriptive ability (using the Akaike and Bayesian Information Criteria) and predictive ability. The explanation for row labels can be found in Section 4. Model abbreviations from left to right: EU: Expected Utility; DA: Disappointment Aversion DS: Dispersion Skewness; PR: Prospective Reference; RL: Rank Dependent Utility with Prelec Function; ST; Salience; WU: Weighted Utility.

Figure 4 below presents a visual representation of the results of these rankings, including also the results for the other two stochastic specifications – Versions A and B. The figure focuses on the performance of DS relative to the other models. The key result is that DS is mostly the 1st or 2nd best model, with its closest rival being RDUT; but we note that DS can explain behavioural patterns that RDUT cannot explain. We also note that DS does relatively better on the average rankings, rather than the rankings based on the ‘percentage of times first’; from this, one may conclude that DS is more ‘robust’ than the others.

Figure 4: Summary of DS rankings in all versions; A, B and C



7. Discussion and Conclusion

We present a new model of decision-making under risk, which incorporates the dispersion, and skewness of the utilities of the outcomes of a lottery. We test this model against six others standard in the literature and show that it outperforms most, particularly in explanatory ability. As we noted above, there is no objectively 'best' way of ranking models. We presented two ways of comparison: (1) rankings based on percentage of times a model comes first and (2) average rankings. Overall, the simplest version of DS presented in the main body of the paper performs better under (2). We could possibly conclude that one can judge its goodness of fit as being between the average ranking and those based on the percentage of times a model comes first.

We also estimated two other stochastic specifications (versions A and B), both of which incorporate a simple editing phase (details can be found in Appendix D). In this editing phase, in order to decrease the cognitive cost the DM looks for dominance between lotteries. If one lottery first-order dominates the other, DM chooses the dominant one without mentally calculating the DS value of both. We incorporate this property, assuming the DM implements it with a tremble, by including a tremble probability in our estimations. In version A, the tremble probability is exogenously given; we see that DS outperforms all models. One might argue that the tremble element gives an unfair advantage to DS. Therefore, in version B, instead of an exogenously given tremble probability, we estimate it and penalize DS for this extra parameter. We see that DS maintains its outperforming position[13].

We would also like to attract attention to the simplification that we made when passing from (4) to (5): in (4) it was assumed that the pessimism $\alpha_{S(\tilde{z})}$ is a function of skewness; in (5) we assumed a simple functional form: $\alpha_{S(\tilde{z})} = 1 - S(\tilde{z})/2$. Future work on developing different functional forms for $\alpha_{S(\tilde{z})}$, as well as incorporating other factors, such as the past experience of the DM, might improve the fit of DS and its predictive ability.

We have shown (in Section 3.1) that our new theory can explain prominent standard 'anomalies': the common consequence and common ratio effects, valuation gaps and

[13] We agree that we could also incorporate an editing phase and a tremble in the other models, yet these models – except for the original version of RDUT (Prospect Theory) – did not include these things when they were first introduced.

preference reversals. More importantly, we have shown (in Section 3.2) that the theory can explain behavioural patterns observed in experiments, that the closest rival regarding empirical fit (RDUT) cannot accommodate, such as reverse common ratio effects, violations of ordinal independence/upper tail independence, violations of stochastic dominance and empirically observed compliance with lower distribution independence (See Section 3.2). We acknowledge the fact that RDUT is the closest rival of DS regarding the empirical fit especially in Version A, but we would like to emphasize the advantage of DS in explaining behavioural patterns that RDUT cannot accommodate.

A clear advantage of the new model is its parsimony – having only one parameter more[14] than EU. It is also behaviourally plausible, incorporating the fact that DMs take into account not only the expected utility of a lottery but also its dispersion and skewness; the vital elements of the distribution of the utility of outcomes is important to the DM.

[14] Though this depends upon the Version. See Appendix D.

REFERENCES

Bayrak, O.K., and B. Kriström (2016). Is there a valuation gap? The case of interval valuations, *Economics Bulletin*, 36(1), 218–236.

Bayrak, O.K. and Hey, J.D. (2020). Understanding Preference Imprecision. *Journal of Economic Surveys*, 34: 154-174. doi:10.1111/joes.12343

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

Bateman, I., Munro, A., Rhodes, B., Starmer, C., and Sugden, R. (1997). A test of the theory of reference-dependent preferences. *The Quarterly Journal of Economics*, 112(2), 479-505.

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

Birnbaum, Michael. (2005). Three new tests of independence that differentiate models of risky decision making. *Management Science,* 51(9), 1346–1358.

Blavatskyy, P. R. (2010). Reverse common ratio effect. *Journal of Risk and Uncertainty*, 40(3), 219-241.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2012). Salience theory of choice under risk. *The Quarterly Journal of Economics*, 127(3), 1243-1285.

Busemeyer, J. R., And Wang, Y. M. (2000). Model comparisons and model selections based on generalization criterion methodology. *Journal of Mathematical Psychology*, 44(1), 171-189.

Butler, D., and G. Loomes (1988). Decision difficulty and imprecise preferences, *Acta Psychologica*, 68(1–3), 183–196. http://doi.org/10.1016/0001-6918(88)90054-6.

Carbone, E., and Hey, J. (1995). A Comparison of the Estimates of EU and non EU Preference Functionals. *Geneva Papers on Risk and Insurance-Issues and Practice*, 20, 111-133.

Chew, S. H. (1983). A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox. *Econometrica*, 51, 1065–1092.

Cubitt, R. P., D. Navarro-Martinez, and C. Starmer (2015). On preference imprecision, *Journal of Risk and Uncertainty*, 50(1), 1–34. http://doi.org/10.1007/s11166-015-9207-6.

Dekel, E. (1986). An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom. *Journal of Economic Theory*, 40(2), 304-318.

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

Dubourg, W. R., M. W. Jones-Lee and G. Loomes (1997). Imprecise preferences and survey design in contingent valuation, *Economica*, 64, 681–702.

Ebert, S. (2015). On skewed risks in economic models and experiments, *Journal of Economic Behavior & Organization*, 112, 85-97.

Forrest, D., R. Simmons, and N. Chesters (2002). Buying a dream: Alternative models of the demand for lotto, *Economic Inquiry*, 40, 485–496.

Garrett, T., and R. Sobel (1999). Gamblers favor skewness not risk: Further evidence from United States lottery games, *Economics Letters*, 63(1), 85–90.

Golec, J., and M. Tamarkin (1998). Betters love skewness, not risk, at the horse track, *Journal of Political Economy*, 106, 205–225.

Green, J. R., and JULLIEN, B. (1988). Ordinal independence in nonlinear utility theory. *Journal of Risk and Uncertainty*, 1(4), 355-387.

Gul, F. (1991). A theory of disappointment aversion. *Econometrica: Journal of the Econometric Society*, 667-686.

Grossman, P. J., and Eckel, C. C. (2015). Loving the long shot: Risk taking with skewed lotteries, *Journal of Risk and Uncertainty*, 51(3), 195-217.

Hagen, O. (1991). Decisions under risk: A descriptive model and a technique for decision making, *European Journal of Political Economy*, 7(3), 381-405.

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

Hull, C. L. (1943). Principles of behavior: An introduction to behavior theory, Oxford, England: Appleton-Century.

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.

Loomes, Graham, and Sugden, Robert. (1998). Testing different stochastic specifications of risky choice. *Economica*, 65, 581–598.

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

Plott, C. R., and Zeiler, K. (2005). The willingness to pay-willingness to accept gap, the" endowment effect," subject misconceptions, and experimental procedures for eliciting valuations. *American Economic Review*, 95(3), 530-545.

Prelec, D. (1998). The Probability Weighting Function. *Econometrica*, 66(3), 497. doi:10.2307/2998573

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization*, 3(4), 323-343.

Tversky, A., and Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business*, S251-S278.

Teitelbaum, J. C., and Zeiler, K. (Eds.). (2018). Research handbook on behavioral law and economics. Edward Elgar Publishing.

Tversky, A., and D. Kahneman (1992). Advances in prospect theory: Cumulative representation of uncertainty, *Journal of Risk and Uncertainty*, 5(4), 297-323.

Viscusi, W. K. (1989). Prospective reference theory: Toward an explanation of the paradoxes. *Journal of Risk and Uncertainty*, 2(3), 235-263.

Wilcox, N. T. (2008). Stochastic models for binary discrete choice under risk: A critical primer and econometric comparison, in James C. Cox, Glenn W. Harrison (ed.) *Risk Aversion in Experiments (Research in Experimental Economics, Volume 12),* Emerald Group Publishing Limited, 197 - 292.

Wu, G. (1994). An empirical test of ordinal independence. *Journal of Risk and Uncertainty*, 9(1), 39-60.

Wu, C. C., Bossaerts, P., and Knutson, B. (2011). The affective impact of financial skewness on neural activity and choice, *PloS one*, 6(2), e16838.

Appendix A: Estimated Preference Functionals

1)EU: Expected Utility—subjects choose on the basis of expected utility:

$$V(p) = p_2u_2 + p_3u_3 + p_4u_4 \tag{17}$$

Parameters estimated: u_2, u_3 and u_4.

2)DA: Disappointment Aversion—subjects choose on the basis of expected (modified) utility—where utility is modified ex post to take account of any disappointment or delight experienced:

$$V(p) = \min(W_1, W_2, W_3) \tag{18}$$

$$W_1 = (1+\beta)p_2u_2 + (1+\beta)p_3u_3 + p_4u_4 + \beta p_1 + \beta p_2 + \beta p_3 \tag{19}$$

$$W_2 = (1+\beta)p_2u_2 + p_3u_3 + p_4u_4 + \beta p_1 + \beta p_2 \tag{20}$$

$$W_3 = p_2u_2 + p_3u_3 + p_4u_4 + \beta p_1 \tag{21}$$

Parameters estimated u_2, u_3, u_4 and β.

3)DS: See Section 2.

Parameters estimated u_2, u_3, u_4 and k.

4)PR: Prospective Reference—subjects choose on the basis of a weighted average of the expected utility calculated using the correct probabilities and the expected utility calculated using equal probabilities for all the non-null outcomes:

$$V(p) = \lambda(p_2u_2 + p_3u_3 + p_4u_4) + (1-\lambda)(a_2u_2 + a_3u_3 + a_4u_4) \tag{22}$$

$a_i = 1/n(p)$ and $n(p)$ is the number of non-zero elements in p.

Parameters estimated: u_2, u_3, u_4 and λ.

5)RL: Rank dependent with Prelec weighting function—subjects choose on the basis of expected utility where the (cumulative) probabilities are distorted by a weighting function which takes the power function form:

$$V(p) = w(p_2 + p_3 + p_4)u_2 + w(p_3 + p_4)(u_3 - u_2) + w(p_4)(u_4 - u_3) \tag{23}$$

where $w(.)$ is the Prelec function $w(p) = \exp\left(-\beta \cdot (-\ln p)^{\alpha}\right)$.

Parameters estimated u_2, u_3, u_4, α and β.

6)ST: Salience Theory—When choosing between *(x,p)* and *(y,q)* (each of which has at most four states) subjects consider the 16 possible states: with probability p_iq_j *(i=1..4, j=1..4)* get x_i

if *(x,p)* chosen and get y_j if *(y,q)* chosen. Then decide on the basis of whether V is positive (choose *(x,p)*) or negative (choose *(y,q)*) where *V* is given by:

$$V = \sum_{i=1}^{4}\sum_{j=1}^{4} \mu(x_i, y_j)[u(x_i) - u(y_i)]p_i q_j \tag{24}$$

where $u(x_1) = 0$, $u(x_2) = u_2$, $u(x_3) = u_3$, and $u(x_4) = u_4$, and $\mu(x,y) = |x - y|^{\beta}$.

This is the Salience Function.

Parameters estimated: *u2, u3, u4* and β.

7)WU: Weighted Utility—subjects choose on the basis of expected weighted utility:

$$V(p) = \frac{w_2 p_2 u_2 + w_3 p_3 u_3 + p_4 u_4}{p_1 + w_2 p_2 + w_3 p_3 + p_4} \tag{25}$$

Parameters estimated: *u2, u3, u4,* w_2 and w_3.

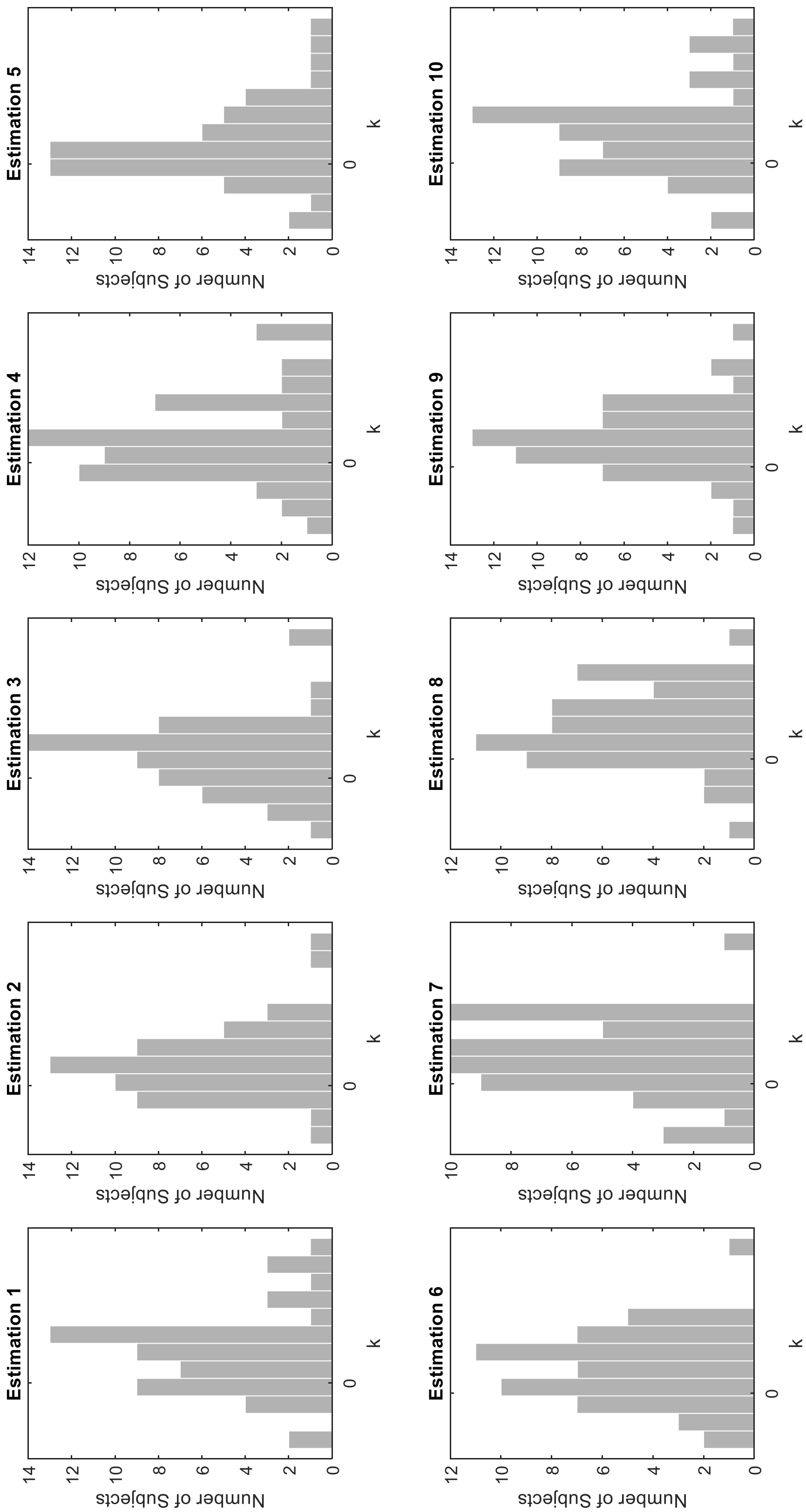
Appendix B: Lotteries in the experiment

L	p1	p2	p3	p4	q1	q2	q3	q4
1	0	0	0.875	0.125	0	0.125	0	0.875
2	0	0	0.875	0.125	0	0.125	0	0.875
3	0	0	0.875	0.125	0	0.125	0.5	0.375
4	0	0	0.875	0.125	0	0.375	0	0.625
5	0	0	0.875	0.125	0	0.375	0.125	0.5
6	0	0	0.875	0.125	0	0.375	0.25	0.375
7	0	0	0.875	0.125	0	0.625	0	0.375
8	0	0.125	0.5	0.375	0	0.375	0	0.625
9	0	0.125	0.5	0.375	0	0.375	0.125	0.5
10	0	0.125	0.875	0	0	0.375	0	0.625
11	0	0.125	0.875	0	0	0.375	0.125	0.5
12	0	0.125	0.875	0	0	0.375	0.25	0.375
13	0	0.125	0.875	0	0	0.375	0.5	0.125
14	0	0.125	0.875	0	0	0.625	0	0.375
15	0	0.125	0.875	0	0	0.875	0	0.125
16	0	0.25	0.75	0	0	0.375	0	0.625
17	0	0.25	0.75	0	0	0.375	0.125	0.5
18	0	0.25	0.75	0	0	0.375	0.25	0.375
19	0	0.25	0.75	0	0	0.375	0.5	0.125
20	0	0.25	0.75	0	0	0.375	0.5	0.125
21	0	0.25	0.75	0	0	0.625	0	0.375
22	0	0.25	0.75	0	0	0.875	0	0.125
23	0	0.375	0.5	0.125	0	0.625	0	0.375
24	0	0.125	0.875	0	0	0.25	0.75	0
25	0	0.375	0.125	0.5	0	0.375	0.25	0.375
26	0	0	0.5	0.5	0.125	0	0.25	0.625
27	0	0	0.5	0.5	0.125	0	0.25	0.625
28	0	0	0.875	0.125	0.125	0	0.25	0.625
29	0	0	0.875	0.125	0.125	0	0.625	0.25
30	0	0	0.875	0.125	0.375	0	0.375	0.25
31	0	0	0.875	0.125	0.5	0	0	0.5
32	0	0	0.875	0.125	0.75	0	0	0.25
33	0	0	1	0	0.125	0	0.25	0.625
34	0	0	1	0	0.125	0	0.625	0.25
35	0	0	1	0	0.375	0	0.375	0.25
36	0	0	1	0	0.5	0	0	0.5
37	0	0	1	0	0.75	0	0	0.25
38	0	0	1	0	0.75	0	0	0.25
39	0	0	1	0	0.75	0	0.125	0.125
40	0.125	0	0.625	0.25	0.5	0	0	0.5
41	0.25	0	0.75	0	0.375	0	0.375	0.25
42	0.25	0	0.75	0	0.5	0	0	0.5
43	0.25	0	0.75	0	0.75	0	0	0.25
44	0.25	0	0.75	0	0.75	0	0.125	0.125
45	0.375	0	0.375	0.25	0.5	0	0	0.5
46	0.375	0	0.625	0	0.5	0	0	0.5
47	0.375	0	0.625	0	0.75	0	0	0.25
48	0.375	0	0.625	0	0.75	0	0.125	0.125
49	0.25	0	0.75	0	0.375	0	0.625	0
50	0.75	0	0	0.25	0.75	0	0.125	0.125
51	0	0.75	0	0.25	0.25	0.375	0	0.375
52	0	0.75	0	0.25	0.375	0.125	0	0.5
53	0	0.75	0	0.25	0.625	0	0	0.375
54	0	0.875	0	0.125	0.25	0.375	0	0.375
55	0	0.875	0	0.125	0.375	0.125	0	0.5
56	0	0.875	0	0.125	0.5	0.25	0	0.25
57	0	0.875	0	0.125	0.625	0	0	0.375
58	0	0.875	0	0.125	0.625	0.125	0	0.25
59	0.125	0.75	0	0.125	0.25	0.375	0	0.375
60	0.125	0.75	0	0.125	0.375	0.125	0	0.5

61	0.125	0.75	0	0.125	0.5	0.25	0	0.25
62	0.125	0.75	0	0.125	0.625	0	0	0.375
63	0.125	0.75	0	0.125	0.625	0.125	0	0.25
64	0.125	0.875	0	0	0.25	0.375	0	0.375
65	0.125	0.875	0	0	0.375	0.125	0	0.5
66	0.125	0.875	0	0	0.5	0.25	0	0.25
67	0.125	0.875	0	0	0.625	0	0	0.375
68	0.125	0.875	0	0	0.625	0.125	0	0.25
69	0.125	0.875	0	0	0.75	0.125	0	0.125
70	0.125	0.875	0	0	0.875	0	0	0.125
71	0.125	0.875	0	0	0.875	0	0	0.125
72	0.25	0.375	0	0.375	0.375	0.125	0	0.5
73	0.5	0.25	0	0.25	0.625	0	0	0.375
74	0.5	0.25	0	0.25	0.625	0	0	0.375
75	0	0.75	0	0.25	0.125	0.75	0	0.125
76	0	0.75	0.25	0	0.125	0	0.875	0
77	0	0.75	0.25	0	0.125	0.375	0.5	0
78	0	0.75	0.25	0	0.375	0.125	0.5	0
79	0	0.75	0.25	0	0.375	0.25	0.375	0
80	0	0.75	0.25	0	0.5	0	0.5	0
81	0	0.75	0.25	0	0.5	0.125	0.375	0
82	0	1	0	0	0.125	0	0.875	0
83	0	1	0	0	0.125	0.375	0.5	0
84	0	1	0	0	0.25	0.625	0.125	0
85	0	1	0	0	0.375	0.125	0.5	0
86	0	1	0	0	0.375	0.25	0.375	0
87	0	1	0	0	0.5	0	0.5	0
88	0	1	0	0	0.5	0	0.5	0
89	0	1	0	0	0.5	0.125	0.375	0
90	0	1	0	0	0.75	0.125	0.125	0
91	0.25	0.625	0.125	0	0.375	0.125	0.5	0
92	0.25	0.625	0.125	0	0.375	0.25	0.375	0
93	0.25	0.625	0.125	0	0.5	0	0.5	0
94	0.25	0.625	0.125	0	0.5	0.125	0.375	0
95	0.375	0.25	0.375	0	0.5	0	0.5	0
96	0.375	0.25	0.375	0	0.5	0	0.5	0
97	0.375	0.625	0	0	0.5	0	0.5	0
98	0.375	0.625	0	0	0.5	0.125	0.375	0
99	0.375	0.625	0	0	0.75	0.125	0.125	0
100	0.375	0.125	0.5	0	0.5	0.125	0.375	0

Appendix C

Figure C1: Histogram of k values (Version C: No editing, Luce Error)



Appendix D: A Possible Editing Phase

Ideally, we would prefer to include an editing phase conducted by the DM before the evaluation phase of the theory: in this editing phase, the DM looks to see if one lottery first-order dominates the other. In such a case, the DM chooses the dominant option without proceeding with the evaluation phase of the theory. In the other problems the DM proceeds as in the theory.

While we think that this editing phase is an important feature of our theory, we want to be fair to the other theories in our ‘horse-race’. It seems that the only other theory which also has an explicit editing phase is Rank Dependent Expected Utility, at least in its earliest incarnation – Prospect Theory. However, recent descriptions of both it and the other theories seem to omit descriptions of such an editing phase – perhaps regarding it as implicit.

So, to be fair to the other theories, we omit any editing phase in any of the models in all the econometric analyses reported in the main body of the paper. For those readers interested in the effects of the incorporation of an editing phase in DS we include Appendices E and F which report the results. There are two Versions reported there: Version A in which the DM trembles in the dominating problems with an exogenous probability; and Version B in which the DM trembles in the dominating problems with an endogenous probability. Not surprisingly DS does much better with this editing phase included. These are in addition to the model without editing, which is referred to as Version C.

Appendix E: Estimations of versions which has editing phase: A and B with Luce model

Table E1a: % of the times that each model comes first; Version A; Luce Model

	EU	**DA**	**DS**	**PR**	**RL**	**ST**	**WU**
Akaike Criterion							
1st 100	8	9	**38**	8	17	11	9
2nd_100	4	6	**51**	4	23	2	13
3rd 100	9	9	**40**	4	23	6	9
4th 100	8	8	**36**	11	21	0	17
5th 100	9	6	26	9	**30**	4	15
all 500	0	4	**72**	2	17	2	4
1st 400	0	2	**74**	0	15	2	8
1st 300	0	6	**70**	0	13	4	8
1st 200	0	2	**60**	2	25	4	8
Bayesian Criterion							
1st 100	**30**	9	28	8	8	11	6
2nd_100	**40**	4	36	4	13	4	2
3rd 100	32	9	**34**	2	9	6	8
4th 100	30	8	**36**	11	8	2	6
5th 100	**40**	4	26	9	9	4	8
all 500	2	2	**77**	2	11	2	4
1st 400	2	2	**79**	0	11	2	4
1st 300	4	4	**74**	2	9	4	4
1st 200	17	2	**58**	4	11	6	2
Predictive Ability							
1st 400	4	9	**45**	6	19	6	11
1st 300	0	9	**49**	4	17	8	15
1st 200	2	9	**55**	4	13	11	6
1st 100	9	9	**43**	4	13	13	8

Table E1b: Average Rankings; Version A; Luce Model

	EU	DA	DS	PR	RL	ST	WU
Akaike Criterion							
1st 100	4.4	4.8	**1.8**	4.0	3.5	5.0	4.5
2nd_100	4.3	4.6	**1.8**	4.2	3.4	5.5	4.2
3rd 100	4.3	4.7	**2.1**	4.0	3.4	5.0	4.3
4th 100	4.4	4.8	**2.2**	3.9	3.2	5.6	3.9
5th 100	4.2	4.8	**2.4**	3.9	3.0	5.6	4.1
all 500	5.3	5.0	**1.4**	4.3	2.8	5.5	3.6
1st 400	5.4	4.9	**1.4**	4.2	2.9	5.5	3.7
1st 300	5.2	4.9	**1.5**	4.3	2.9	5.3	3.9
1st 200	4.9	4.8	**1.5**	4.2	3.1	5.2	4.3
Bayesian Criterion							
1st 100	3.1	4.6	**2.0**	3.8	4.4	4.8	5.2
2nd_100	2.7	4.3	**1.9**	3.9	4.5	5.2	5.2
3rd 100	2.9	4.4	**2.3**	4.0	4.5	4.9	5.1
4th 100	3.2	4.5	**2.0**	3.6	4.2	5.5	4.9
5th 100	2.9	4.6	**2.3**	3.7	4.1	5.4	5.0
all 500	4.4	4.9	**1.3**	4.0	3.5	5.4	4.5
1st 400	4.2	4.8	**1.3**	3.9	3.5	5.4	4.8
1st 300	4.2	4.8	**1.4**	4.0	3.7	5.1	4.8
1st 200	3.5	4.6	**1.5**	3.9	4.2	5.0	5.2
Predictive Ability							
1st 400	5.0	4.4	**2.0**	4.2	3.1	5.4	3.8
1st 300	5.2	4.4	**2.0**	4.1	3.1	5.3	3.9
1st 200	4.7	4.2	**2.0**	4.3	3.8	4.9	4.0
1st 100	4.3	3.9	**2.3**	4.0	3.9	4.8	4.6

Table E2a: % of the times that each model comes first; Version B; Luce Model

	EU	DA	DS	PR	RL	ST	WU
Akaike Criterion							
1st 100	15	9	17	8	**23**	15	13
2nd_100	15	9	21	8	**30**	4	17
3rd 100	17	9	19	9	**28**	9	9
4th 100	15	9	9	13	**26**	6	21
5th 100	11	6	17	11	**38**	4	15
all 500	0	4	**68**	2	17	2	8
1st 400	0	2	**66**	2	21	2	8
1st 300	0	8	**57**	2	19	8	8
1st 200	2	6	**45**	6	28	4	9
Bayesian Criterion							
1st 100	**40**	9	6	8	15	17	6
2nd_100	**55**	8	8	8	13	6	6
3rd 100	**32**	11	9	15	15	9	9
4th 100	**38**	9	8	15	15	4	11
5th 100	**42**	4	11	11	21	6	8
all 500	11	6	**53**	4	17	4	6
1st 400	11	6	**45**	6	17	8	8
1st 300	17	9	**32**	8	17	9	8
1st 200	**34**	8	15	13	15	11	4
Predictive Ability							
1st 400	4	9	**47**	6	19	6	11
1st 300	0	6	**55**	4	17	8	15
1st 200	2	9	**55**	4	13	11	6
1st 100	8	9	**49**	4	11	11	8

Table E2b: Average Rankings; Version B; Luce Model

	EU	DA	DS	PR	RL	ST	WU
Akaike Criterion							
1st 100	4.1	4.7	**2.9**	3.7	3.3	4.9	4.3
2nd_100	4.0	4.5	**2.8**	4.0	3.2	5.4	3.9
3rd 100	4.1	4.5	**3.1**	3.8	3.2	4.9	4.2
4th 100	4.1	4.7	3.1	3.7	**3.0**	5.5	3.8
5th 100	4.0	4.7	3.3	3.7	**2.7**	5.5	4.0
all 500	5.3	5.0	**1.5**	4.3	2.8	5.5	3.6
1st 400	5.4	4.9	**1.6**	4.1	2.8	5.5	3.7
1st 300	5.2	4.8	**1.7**	4.3	2.8	5.3	3.8
1st 200	4.8	4.7	**1.9**	4.1	3.0	5.2	4.2
Bayesian Criterion							
1st 100	**2.9**	4.2	4.0	3.2	4.2	4.6	5.0
2nd_100	**2.5**	3.8	4.1	3.3	4.2	5.0	4.9
3rd 100	**2.8**	3.9	4.1	3.5	4.1	4.7	4.8
4th 100	**3.0**	4.2	4.1	3.0	3.8	5.3	4.6
5th 100	**2.7**	4.2	4.2	3.2	3.6	5.2	4.8
all 500	4.3	4.8	**1.7**	4.0	3.3	5.4	4.4
1st 400	4.1	4.7	**1.9**	3.8	3.3	5.3	4.8
1st 300	3.9	4.7	**2.3**	3.8	3.6	5.1	4.7
1st 200	3.2	4.3	**3.1**	3.5	4.0	4.8	5.1
Predictive Ability							
1st 400	5.0	4.4	**1.9**	4.2	3.2	5.4	3.8
1st 300	5.2	4.4	**1.8**	4.1	3.2	5.3	3.9
1st 200	4.8	4.2	**1.9**	4.2	3.9	4.9	4.0
1st 100	4.3	3.9	**2.2**	4.0	3.9	4.9	4.6

Appendix F: Results with white noise/Fechner Error Specification: Version A, B and C

Table F1a: % of the times that each model comes first; Version A with White Noise

	EU	DA	DS	PR	RL	ST	WU
Akaike Criterion							
1st 100	9	9	**34**	6	17	15	9
2nd_100	4	6	**49**	6	19	6	13
3rd 100	9	9	**40**	4	23	6	9
4th 100	9	9	**38**	11	19	0	13
5th 100	9	4	28	8	**30**	8	13
all 500	0	4	**70**	0	21	0	6
1st 400	0	4	**68**	0	19	0	9
1st 300	0	6	**66**	0	17	4	8
1st 200	0	2	**51**	4	26	8	9
Bayesian Criterion							
1st 100	**26**	9	**26**	6	11	15	6
2nd_100	**38**	4	**38**	4	11	6	2
3rd 100	30	11	**34**	4	9	6	6
4th 100	28	11	**32**	11	9	2	6
5th 100	**38**	2	28	9	8	8	8
all 500	2	2	**77**	0	15	0	4
1st 400	2	4	**79**	0	11	2	2
1st 300	2	6	**74**	2	8	6	4
1st 200	15	4	**55**	4	13	8	2
Predictive Ability							
1st 400	6	9	**49**	4	17	9	9
1st 300	0	9	**51**	2	13	9	15
1st 200	4	8	**47**	6	13	15	8
1st 100	6	9	**45**	6	13	13	8

Table F1b: Average Rankings; Version A with White Noise

	EU	DA	DS	PR	RL	ST	WU
Akaike Criterion							
1st 100	4.4	4.7	**1.9**	4.1	3.5	4.8	4.5
2nd_100	4.4	4.5	**1.9**	4.2	3.4	5.4	4.1
3rd 100	4.5	4.4	**2.2**	4.0	3.4	5.1	4.4
4th 100	4.4	4.6	**2.2**	3.9	3.2	5.6	4.0
5th 100	4.2	4.8	**2.2**	4.0	3.0	5.5	4.2
all 500	5.3	4.8	**1.5**	4.3	2.9	5.3	3.8
1st 400	5.5	4.7	**1.6**	4.2	2.8	5.3	3.8
1st 300	5.3	4.9	**1.6**	4.3	2.8	5.2	3.8
1st 200	5.1	4.7	**1.9**	4.1	3.0	5.2	4.0
Bayesian Criterion							
1st 100	3.3	4.6	**2.0**	4.0	4.3	4.6	5.2
2nd_100	2.9	4.2	**2.0**	3.9	4.5	5.2	5.2
3rd 100	3.2	4.2	**2.2**	3.9	4.3	5.0	5.2
4th 100	3.3	4.3	**2.2**	3.7	4.0	5.6	4.8
5th 100	3.1	4.5	**2.2**	3.8	4.1	5.3	4.9
all 500	4.7	4.8	**1.4**	4.1	3.4	5.2	4.5
1st 400	4.6	4.6	**1.4**	4.0	3.4	5.2	4.7
1st 300	4.4	4.7	**1.5**	4.1	3.6	4.9	4.8
1st 200	3.7	4.5	**1.8**	4.0	4.0	5.0	5.0
Predictive Ability							
1st 400	5.2	4.2	**2.0**	4.5	3.0	5.2	3.8
1st 300	5.2	4.5	**1.9**	4.1	3.2	5.1	3.9
1st 200	4.5	4.2	**2.2**	4.2	3.7	4.6	4.1
1st 100	4.5	4.1	**2.2**	3.7	3.8	5.0	4.4

Table F2a: % of the times that each model comes first; Version B with White Noise

	EU	DA	DS	PR	RL	ST	WU
Akaike Criterion							
1st 100	13	11	15	8	**21**	19	13
2nd_100	13	11	17	9	**28**	6	19
3rd 100	15	9	**26**	9	25	8	9
4th 100	17	11	11	13	**25**	4	19
5th 100	13	4	19	9	**36**	8	13
all 500	0	8	**60**	0	23	2	8
1st 400	0	4	**64**	2	21	0	9
1st 300	0	6	**55**	4	23	6	8
1st 200	2	2	**43**	6	30	8	9
Bayesian Criterion							
1st 100	**28**	11	9	6	17	23	6
2nd_100	**51**	11	6	13	11	6	4
3rd 100	**32**	13	13	13	13	9	8
4th 100	**38**	13	8	15	15	2	9
5th 100	**38**	6	13	11	15	9	8
all 500	11	8	**49**	0	23	2	8
1st 400	11	9	**43**	2	19	8	8
1st 300	13	9	**32**	8	19	11	8
1st 200	**28**	11	15	9	19	13	4
Predictive Ability							
1st 400	6	9	**51**	4	17	9	9
1st 300	0	9	**51**	2	13	9	15
1st 200	4	8	**47**	6	13	15	8
1st 100	6	11	**43**	6	13	13	8

Table F2b: Average Rankings; Version B with White Noise

	EU	DA	DS	PR	RL	ST	WU
Akaike Criterion							
1st 100	4.2	4.6	**2.9**	3.8	3.4	4.8	4.3
2nd_100	4.2	4.3	**2.7**	4.0	3.2	5.4	3.9
3rd 100	4.3	4.3	**2.9**	3.8	3.2	5.1	4.2
4th 100	4.2	4.5	3.2	3.7	**3.0**	5.6	3.8
5th 100	4.0	4.6	3.2	3.8	**2.8**	5.5	4.0
all 500	5.3	4.8	**1.7**	4.3	2.8	5.3	3.7
1st 400	5.5	4.7	**1.7**	4.2	2.8	5.3	3.8
1st 300	5.3	4.9	**1.8**	4.2	2.7	5.2	3.8
1st 200	5.1	4.7	**2.1**	4.1	2.9	5.1	3.9
Bayesian Criterion							
1st 100	**3.0**	4.1	3.9	3.3	4.1	4.5	5.0
2nd_100	**2.7**	3.7	4.1	3.3	4.2	5.0	4.9
3rd 100	**3.0**	3.8	3.9	3.4	4.1	4.8	4.9
4th 100	3.2	3.9	4.2	**3.1**	3.6	5.3	4.5
5th 100	**2.8**	4.1	4.0	3.4	3.7	5.2	4.7
all 500	4.5	4.7	**1.9**	4.1	3.2	5.1	4.4
1st 400	4.4	4.6	**2.0**	4.0	3.2	5.1	4.6
1st 300	4.1	4.6	**2.4**	3.9	3.4	4.9	4.7
1st 200	3.5	4.2	**3.2**	3.6	3.8	4.9	4.8
Predictive Ability							
1st 400	5.2	4.2	**1.8**	4.5	3.0	5.3	3.9
1st 300	5.2	4.5	**1.9**	4.1	3.2	5.1	3.9
1st 200	4.5	4.2	**2.2**	4.2	3.7	4.6	4.1
1st 100	4.5	4.1	**2.2**	3.8	3.8	5.0	4.4

Table F3a: % of the times that each model comes first; Version C with White Noise

	EU	DA	DS	PR	RL	ST	WU
Akaike Criterion							
1st 100	13	11	17	8	**21**	19	13
2nd_100	13	9	23	8	**25**	8	19
3rd 100	17	11	17	9	**26**	8	11
4th 100	19	11	9	13	**25**	4	19
5th 100	15	4	17	11	**34**	8	13
all 500	28	28	25	25	**57**	21	28
1st 400	9	13	9	9	**45**	4	11
1st 300	9	11	13	9	**40**	8	11
1st 200	8	13	19	9	**36**	8	9
Bayesian Criterion							
1st 100	**26**	11	19	6	15	19	6
2nd_100	**45**	6	28	4	11	8	2
3rd 100	**34**	13	21	8	11	8	6
4th 100	**38**	13	17	11	11	2	8
5th 100	**40**	4	21	9	11	9	8
all 500	34	26	32	26	**47**	21	25
1st 400	17	13	**23**	9	**23**	8	9
1st 300	15	15	**25**	11	17	11	8
1st 200	**30**	13	23	8	17	9	2
Predictive Ability							
1st 400	9	13	15	8	**40**	9	13
1st 300	0	8	9	15	**34**	11	23
1st 200	9	17	11	15	**28**	13	8
1st 100	15	13	11	11	**19**	**19**	11

Table F3b: Average Rankings; Version C with White Noise

	EU	DA	DS	PR	RL	ST	WU
Akaike Criterion							
1st 100	4.2	4.5	**2.8**	3.9	3.4	4.8	4.4
2nd_100	4.2	4.5	**2.6**	4.2	3.1	5.2	4.0
3rd 100	4.2	4.4	**2.9**	3.8	3.2	5.1	4.3
4th 100	4.2	4.4	3.2	3.7	**3.2**	5.5	3.8
5th 100	3.9	4.5	3.2	3.7	**2.9**	5.4	4.1
all 500	4.4	3.9	2.6	3.5	**2.2**	4.6	3.1
1st 400	5.3	4.4	2.9	4.0	**2.4**	5.4	3.6
1st 300	5.1	4.7	2.9	3.9	**2.5**	5.2	3.6
1st 200	4.8	4.4	**2.8**	3.9	2.9	5.2	3.9
Bayesian Criterion							
1st 100	3.1	4.4	**2.7**	3.7	4.2	4.6	5.2
2nd_100	3.0	4.2	**2.3**	3.9	4.4	5.0	5.1
3rd 100	2.9	4.2	**2.7**	3.7	4.3	4.9	5.1
4th 100	3.3	4.1	**2.9**	3.5	4.0	5.5	4.8
5th 100	**2.9**	4.2	3.0	3.5	3.9	5.3	4.9
all 500	3.8	3.8	**2.2**	3.4	2.6	4.5	3.8
1st 400	4.2	4.3	**2.5**	3.8	3.2	5.2	4.7
1st 300	4.2	4.5	**2.5**	3.8	3.5	4.9	4.7
1st 200	3.5	4.2	**2.6**	3.7	3.9	5.0	5.0
Predictive Ability							
1st 400	5.1	3.9	3.4	4.1	**2.5**	5.1	3.7
1st 300	5.1	4.3	3.2	3.7	**2.8**	5.1	3.6
1st 200	4.2	3.8	3.5	3.7	**3.5**	4.7	4.0
1st 100	4.1	4.0	**3.5**	**3.5**	**3.5**	4.8	4.2