Decision Analysis

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

The Impact of Learning by Thought on Violations of Independence and Coalescing

Michael H. Birnbaum, Ulrich Schmidt

DECISION ANALYSIS
Vol. 12, No. 3, September 2015, pp. 144–152
ISSN 1545-8490 (print) | ISSN 1545-8504 (online)
http://dx.doi.org/10.1287/deca.2015.0316

The Impact of Learning by Thought on Violations of Independence and Coalescing

Michael H. Birnbaum
Department of Psychology, California State University, Fullerton, Fullerton, California 92831,
mbirnbaum@exchange.fullerton.edu

Ulrich Schmidt
Department of Economics, University of Kiel, Olshausenstr. 40, 24098 Kiel, Germany; Kiel Institute for the World Economy, Kiel, Germany; and Department of Economics and Econometrics, University of Johannesburg, South Africa,
us@bwl.uni-kiel.de

This paper reports results from a repeated experiment on decision making under risk where subjects must address the same choice problems in several rounds. We investigate how behavior changes in the course of the experiment. The design focuses on choice problems allowing for direct tests of independence and coalescing. We show that inconsistencies in responses as well as violations of independence and coalescing decrease from earlier to later rounds. Our results provide evidence in favor of expected utility in conjunction with the discovered preference hypothesis.

Keywords: independence axiom; splitting effects; coalescing; errors; experiment
History: Received on October 10, 2014. Accepted by Editor-in-Chief Rakesh K. Sarin on March 16, 2015, after 1 revision. Published online in *Articles in Advance* July 10, 2015.

1. Introduction

An important issue in the behavioral economics literature is the question of whether observed behavioral biases and market anomalies are persistent or whether they tend to disappear if subjects have the possibility to learn. To analyze this question, many studies run repeated experiments and compare behavior in later rounds with that in earlier rounds (e.g., Cox and Grether 1996, Loomes et al. 2003, Dufwenberg et al. 2005). This paper also presents a repeated experimental study and investigates individual decision making under risk. In this context, the well known discovered preference hypothesis (Plott 1996; see also Smith 1989, Harrison 1994, Binmore 1999) which proposes that "individuals have a consistent set of preferences over states, but such preferences only become known to the individual with thought and experience" (Myagkov and Plott 1997, p. 821) implies that choice behavior in later rounds may reveal different preferences than in earlier rounds and, in particular, that observed irrationalities should decrease. By contrast, several random preference theories of choice assume that choice responses satisfy assumptions of independence and identical distributions (iid). According to these theories choices should not change from trial to trial, but instead show stationarity (Birnbaum 2012, 2013).

In this paper we collect repeated binary choice data in an experiment with four repetitions. This allows us to analyze the question of whether typical violations of expected utility (EU) decrease with experience, as implied in the discovered preference hypothesis. In our study we focus on two important failures of EU, violations of independence and coalescing. The common consequence and common ratio effect of Allais are well known experimental designs where substantial violations of independence have been observed. These violations motivated the development of alternative theories such as rank-dependent utility (RDU) (Quiggin 1981, 1982; Luce 1991; Luce and Fishburn 1991), cumulative prospect theory (CPT) (Tversky and Kahneman 1992, Wakker and Tversky 1993), and configural weight models such as TAX and RAM (Birnbaum and McIntosh 1996), which rely on weaker independence conditions. Apart from common consequence and common ratio effects, our experimental design also tests for these weaker conditions. Coalescing demands that if two branches in a gamble lead to the same consequence, they can be combined by adding their probabilities without altering the utility of the gamble. Violations of coalescing (also called splitting effects) have been observed in a number of studies (Starmer and Sugden 1993; Humphrey 1995, 2001). They are particularly troublesome as they can be used to generate violations of first-order stochastic dominance (Birnbaum and Navarette 1998). Therefore, it would be a welcome result if these violations decrease with experience.

To compare our study with the related literature, one should distinguish, (as in Myagkov and Plott 1997), between learning by thought and learning by experience. In our study subjects have to address each choice problem four times without any feedback between choices. Thus we can only observe learning by thought, as subjects do not experience the consequences of their choices. Note also that our study considers only decisions based on description and not cases wherein the probabilities of consequences are learned via experience, as in the literature reviewed by Hertwig and Erev (2009).

Learning by thought and learning by experience have been analyzed in a study by van de Kuilen and Wakker (2006) using a common ratio design and the random lottery incentive system. Subjects make choices in 15 common ratio problems and after each choice the preferred lottery is played out. At the end of the experiment one of the choices is randomly selected and the previously determined payoff is paid for real. There is also a control group that did not receive any feedback after the single choices. Thus control group subjects could only learn by thought. According to the results of van de Kuilen and Wakker (2006) learning by thought and experience leads to a significant increase in consistency with EU (the violation rate of independence decreases from 46.15% in the first round to 23.08% in the fifteenth round) while learning by thought only leads to no such increase. In a related study, van de Kuilen (2009) finds the same effect for nonadditive probability weighting, i.e., learning by experience reduces probability weighting while learning by thought only does not. By contrast, Nicholls et al. (2015) report that learning by thought reduces violations of the sure-thing-principle, which is the analogue of independence in choice under uncertainty.

Also, Hey (2001) considers repeated binary choice problems in five repetitions without feedback (i.e., learning by thought only). When analyzing the data, he considers each subject and each of the five repetitions separately. More precisely, for each subject he fits the parameters of a number of different preference functional (including EU and alternative theories such as RDU) repetition by repetition and compares their goodness of fit. He finds some (limited) evidence that the majority of subjects converge to EU as best fitting functional as required by the discovered preference hypothesis. As only learning by thought is involved, this conclusion is in contrast to the results from van de Kuilen and Wakker (2006) and van de Kuilen (2009).

Our study is a synthesis of the studies of Hey (2001) and van de Kuilen and Wakker (2006). While our experimental set-up is similar to Hey (2001), we analyze data as van de Kuilen and Wakker (2006), i.e., choice problem by choice problem and not subject by subject. Note that in the van de Kuilen and Wakker (2006) set-up, subjects address 15 different common ratio problems only once whereas in Hey (2001) (and in our design) they must repeatedly respond to a larger number of identical choice problems. It may be that learning by thought requires subjects to consider the same choice problems several times, which does not happen in the van de Kuilen and Wakker (2006) study. Analyzing the data choice problem by choice problem instead of subject by subject may have some disadvantages (i.e., not analyzing differences between individuals, in particular the different degree of variability in their responses). Yet it does not incur the potential problems of fitting preference functional such as pre-specifying functional forms. By focusing on choice problems instead of subjects we can also analyze whether violation rates for some independence conditions decrease more than for others. Moreover, our analysis of violations of coalescing also requires distinguishing between choice problems instead of subjects.

Rank-dependent utility and cumulative prospect theory explain violations of independence by probability weighting. According to van de Kuilen (2009), learning by thought only does not reduce probability weighting and should, therefore, have no influence on violations of independence. By contrast, configural weight models explain violations of independence by violations of coalescing. There is some evidence that common ratio and common consequence effects are caused by violations of coalescing (Birnbaum 2004, Schmidt and Seidl 2014). As the impact of learning by thought on violations of coalescing has not yet been analyzed, it is an open question whether learning by thought can reduce violations of independence caused by failures of coalescing, at least in a design where subjects repeatedly address identical choice problems.

For decision analysis, particularly decision support, it is an important question whether learning by experience only or also learning by thought can help to decrease biases and inconsistent choice behavior. While in experiments with relatively low payoffs learning by experience can be easily implemented, for complex managerial decisions with high economic impact this is not the case. Consider, for instance, the decision of a firm whether to make a costly investment in a new production technology. Learning by experience would require making this decision repeatedly while experiencing real consequences in each case. This is impossible as the decision can only be made once. Obviously, before making the real decision the firm can make a series of hypothetical decisions, thereby allowing for learning by thought. If learning by thought decreases biases and inconsistencies, it can lead to an improved elicitation of utilities and consequently to better decisions.

This paper seeks to analyze the impact of learning by thought on violations of independence and coalescing. The next section presents our experimental design. §3 presents our results and §4 provides some concluding observations.

2. Experimental Design

The experiment was conducted at the University of Kiel with 54 subjects, primarily economics and business administration students (all undergraduates). There were six sessions, approximately 90 minutes each. Nine subjects participated in each session. Subjects were free to set their own pace. Each subject received a €5 show-up fee and had to respond to 176 pairwise choice questions, arranged in four booklets of 44 choices each. After a subject finished all four booklets one of her choices was randomly chosen and played out for real. The average payment was €19.14 for 90 minutes, i.e., €12.76 per hour, which considerably exceeds the usual student wage of about €8 per hour.

Figure 1 Presentation of Lotteries

A:	50% *to win* €20	B:	33% *to win* €10
	30% *to win* €30		34% *to win* €15
	20% *to win* €40		33% *to win* €60

Lotteries were presented as shown in Figure 1 and subjects had to circle their choice. Prizes were ordered from lowest to highest. Explanation and playing out of lotteries involved a container of tickets numbered 1 to 100. For example, suppose a subject could play out lottery A in Figure 1. Then she would win €20 when drawing a ticket from 1 to 50, €30 for a ticket between 51 and 80, and €40 for a ticket between 81 and 100. All this was explained in instructions, which were given to the students in printed form and read aloud. After the instructions were received, subjects had to answer four transparent dominance questions that were controlled by the experimenter before proceeding.

Lotteries in the booklets were presented in a pseudo-random order. The ordering of lotteries was different in each booklet; no choice problem was followed by another testing the same independence property. Only after finishing one booklet did a subject received the next one. For half the subjects each booklet contained only coalesced or only split choice problems. For the other half split and coalesced choice problems were intermixed in each booklet. Our stimuli involved 11 tests of independence conditions, nine of which were investigated in coalesced and split forms. All 20 tests were replicated four times with counterbalanced left-right positioning. Additionally, to test the attentiveness of the subjects, each booklet included two transparent stochastic dominance questions, one based on outcome monotonicity and one on event monotonicity.

Our tests of independence conditions and the involved lottery pairs are presented in Table 1. Each lottery pair consists of a safe lottery S (in which you can win prize s_i with probability p_i), and a risky lottery R for which possible prizes and probabilities are

Table 1 The Lottery Pairs

Property	Lottery No.	p_1 / s_1	p_2 / s_2	p_3 / s_3	q_1 / r_1	q_2 / r_2	q_3 / r_3	Average choice % R	Violations of independence: Rep 1 and 2 (split)	Violations of independence: Rep 3 and 4 (split)
CCE1	1	*0.80*	*0.20*		*0.90*	*0.10*		96	17	17
		0	19		0	44				
	2	*0.40*	*0.20*	*0.40*	*0.50*	*0.50*		83	(17)	(7)
		0	19	44	0	44				
CCE2	5	*0.89*	*0.11*		*0.90*	*0.10*		56	39	33
		0	16		0	32				
	6	*1.00*			*0.01*	*0.89*	*0.10*	58	(34)	(33)
		16			0	16	32			
CCE3	9	*0.80*	*0.20*		*0.90*	*0.10*		37	38	26
		0	19		0	44				
	10	*1.00*			*0.10*	*0.80*	*0.10*	11	(25)	(10)
		19			0	19	44			
CCE4	5	*0.70*	*0.30*		*0.80*	*0.10*	*0.10*	56	38	32
		0	21		0	21	42			
	13	*0.70*	*0.20*	*0.10*	*0.80*	*0.20*		30	(36)	(21)
		0	21	42	0	42				
CRE1	15	*0.98*	*0.02*		*0.99*	*0.01*		71	55	57
		0	23		0	46				
	16	*1.00*			*0.50*	*0.50*		15	(44)	(48)
		23			0	46				
CRE2	20	*0.80*	*0.20*		*0.86*	*0.14*		50	35	27
		0	28		0	44				
	19	*0.40*	*0.60*		*0.58*	*0.42*		29	(23)	(15)
		0	28		0	44				
UTI	29	*0.73*	*0.02*	*0.25*	*0.74*	*0.01*	*0.25*	82	49	45
		0	15	60	0	33	60			
	30	*0.73*	*0.02*	*0.25*	*0.74*	*0.26*		44	(29)	(18)
		0	15	33	0	33				
LTI	33	*0.75*	*0.23*	*0.02*	*0.75*	*0.24*	*0.01*	91	25	20
		1	34	36	1	33	60			
	34	*0.75*	*0.23*	*0.02*	*0.99*	*0.01*		77	(14)	(9)
		33	34	36	33	60				
UCI	37	*0.20*	*0.20*	*0.60*	*0.20*	*0.20*	*0.60*	74	41	33
		9	10	24	3	21	24			
	38	*0.40*	*0.60*		*0.20*	*0.80*		62	(22)	(25)
		9	21		3	21				
LDI	23	*0.60*	*0.20*	*0.20*	*0.60*	*0.20*	*0.20*	6	7	5
		1	18	19	1	2	32			
	24	*0.10*	*0.45*	*0.45*	*0.10*	*0.45*	*0.45*	9	(—)	(—)
		1	18	19	1	2	32			
UDI	25	*0.20*	*0.20*	*0.60*	*0.20*	*0.20*	*0.60*	80	14	14
		6	7	20	1	19	20			
	26	*0.45*	*0.45*	*0.10*	*0.45*	*0.45*	*0.10*	78	(—)	(—)
		6	7	20	1	19	20			

Note. The first lottery pair of a choice problem always characterizes the lotteries S and R, and the second one the lotteries S' and R'.

denoted by r_i and q_i, respectively. We took the lotteries from previous studies that reported high violation rates, but adjusted outcomes to obtain an average expected value of about €12. Table 1 shows only the coalesced forms of the lottery pairs. For the tests of independence conditions in split variants we used the canonical split form of these pairs. In the canonical split form of a pairwise choice, both lotteries are split so that there are equal probabilities on corresponding ranked branches and the number of branches is equal in both gambles and minimal. Appendix A shows the lottery pairs used in the split tests. Note that each pairwise choice problem presented in Table 1 has a unique canonical split form. Presenting subjects the lottery pairs in both coalesced and split form allows us to test whether coalescing is satisfied, i.e., whether choices in the split form do not differ systematically from those in the coalesced form.

The first six independence tests in Table 1 are four common consequence effects (CCE1-4) and two common ratio effects (CRE1 and 2). Such tests have been widely used to find the independence axiom of EU; the paradoxes of Allais are special variants of a CCE and a CRE. CCEs can be formally described by $S = (x, p_1; s_2, p_2; s_3, p_3)$, $R = (x, q_1; r_2, q_2; r_3, q_3)$, $S' = (x, p_1 - \alpha; s_2, p_2; s_3, p_3; x', \alpha)$, and $R' = (x, q_1 - \alpha; r_2, q_2; r_3, q_3; x', \alpha)$, i.e., S' and R' are constructed from S and R by shifting probability mass (α) from the common consequence x to a different common consequence x'. Consequently, an EU maximizer will prefer S over R if and only if she will prefer S' over R'. Note that in Table 1 the first row of a choice problem always characterizes the lotteries S and R and the second row the lotteries S' and R'. For CCE1 we have, for instance, $x = 0$, $p_1 = 0.8$, $p_2 = 0.2$, $s_2 = 19$, $p_3 = 0$ for S, $q_1 = 0.90$, $q_2 = 0.10$, $r_2 = 44$, $q_3 = 0$ for R and S' and R' are constructed by setting $\alpha = 0.4$ and $x' = 44$. The lotteries in the four CCEs are taken from Starmer (1992) who observed high violation rates for these lotteries. The typical pattern of violations in CCE1-4 is that subjects prefer R over S but S' over R'. The same is true for the two CREs (i.e., CRE1 and 2) presented in Table 1. A CRE can be formally described by $S = (x, 1-\beta(1-p_1); s_2, \beta p_2)$, $R = (x, 1-\beta(1-q_1); r_2, \beta q_2)$, $S' = (x, p_1; s_2, p_2)$, and $R' = (x, q_1; r_2, q_2)$, i.e., S and R are constructed from S' and R' by multiplying all probabilities by β and assigning the remaining probability $1-\beta$ to the common consequence x. Again, EU implies that people choose the risky or the safe lottery in both choice problems. In CRE1 (taken from Birnbaum 2001) and CRE2 (taken from Starmer and Sugden 1989), however, substantial violations of EU have been observed with many people choosing R and S'.

The remaining five independence properties in Table 1 are weakened variants of the independence axiom of EU that were used to derive alternative theories. We focus on variants that are implied by RDU, CPT, and configural weight models. A central property in this context is tail independence (TI), which was introduced by Green and Jullien (1988) using the term ordinal independence. Formally, TI demands that $S = (x_1, p_1; \ldots; x_i, p_i; x_{i+1}, p_{i+1}; \ldots; x_n, p_n) \succeq R = (x_1, p_1; \ldots; x_i, p_i; x_{i+1}, q_{i+1}; \ldots; x_n, q_n)$ if and only if $S' = (x_1, q_1; \ldots; x_i, q_i; x_{i+1}, p_{i+1}; \ldots; x_n, p_n) \succeq R' = (x_1, q_1; \ldots; x_i, q_i; x_{i+1}, q_{i+1}; \ldots; x_n, q_n)$ where $x_1 \geq x_2 \geq \cdots \geq x_n$. This means that if two lotteries share a common tail (i.e., identical probabilities of receiving any outcome better than x_{i+1}), then the preference between these lotteries must not change if this tail is replaced by a different common tail. Note that in the definition above the upper tail is the common tail and thus the condition is called upper tail independence (UTI). TI, however, also demands that preferences not change if lower common tails are exchanged; this will be called lower tail independence (LTI). TI is a very general property that is implied by many models including all variants of RDU as well as CPT. Therefore, rejecting TI would provide serious evidence against all these models. In his experiments, Wu (1994) observed UPI violation rates of up to 50%. Similar evidence has been reported by Birnbaum (2001). Our study tries to discover whether the reported TI violations may be due to splitting effects and/or errors. The lotteries we use for the UTI test are taken from Wu (1994). To our knowledge, LTI has not been tested before. Our construction of lotteries in the LTI test is similar to that used in the UTI test.

Another property implied by CPT and the common versions of RDU is upper cumulative independence (UCI), which demands that decision weights depend only on cumulative probabilities. Formally, UCI demands that if $S = (s_1, p_1; s_2, p_2; \alpha, p_3) \prec R = (r_1, p_1; \gamma, p_2; \alpha, p_3)$ then $S' = (s_1, p_1 + p_2; \gamma, p_3) \prec R' = (r_1, p_1; \gamma, p_2 + p_3)$, where $\alpha > \gamma > s_2 > s_1 > r_1$. Birnbaum and Navarette (1998) and Birnbaum et al. (1999) reported substantial violations of UCI. Our lottery pairs are taken from the latter paper, which observed violation rates of 40.1% for these pairs, where the typical violation pattern is RS'.

The final property we test is distribution independence (DI). Whereas configural weight models and original prospect theory imply that DI holds, it should be violated according to RDU and CPT, at least if the weighting function is inverse-S shaped as commonly suggested by empirical research. For three outcome lotteries, DI demands that $S = (s_1, \beta; s_2, \beta; \alpha, 1-2\beta) \succeq R = (r_1, \beta; r_2, \beta; \alpha, 1-2\beta)$ if and only if $S' = (s_1, \delta; s_2, \delta; \alpha, 1-2\delta) \succeq R' = (r_1, \delta; r_2, \delta; \alpha, 1-2\delta)$ where α is either the highest or the lowest outcome in both lotteries. If α is the highest outcome, the condition is called upper distribution independence (UDI), otherwise lower distribution independence (LDI). The lotteries used in our tests of UDI and LDI are taken from Birnbaum (2005). The evidence reported in that paper and in Birnbaum and Chavez (1997) indicates that one should observe either no violations or violations contrary to CPT with inverse-*S* weighting function. The data obtained from the present experiment have been used by Birnbaum et al. (2009) to analyze a different question, i.e., whether violations of independence can be attributed to errors. The issue of learning is not addressed in that paper.

3. Results

(i) *Inconsistencies*. We define that an inconsistency occurs if a subject in a given choice problem chooses the risky lottery in one repetition and the safe lottery in another repetition. According to the discovered preference hypotheses, inconsistencies should decrease in the course of the experiment. Comparing behavior in repetitions 1 and 2, we observe inconsistencies in 21.67% of all choices. This number decreases to 18.71% for repetitions 2 and 3 and further to 13.62% for repetitions 3 and 4. According to Wilcoxon tests both decreases are significant at the 1%-level ($z = -2.57$ for comparing 1 and 2 with 2 and 3 and $z = -4.26$ for comparing 2 and 3 with 3 and 4, two-sided). As inconsistencies are likely caused by erroneous responses, this evidence indicates that errors are decreasing through learning by thought. Hey (2001) reported similar evidence.

Violations of transparent stochastic dominance may be regarded as a further manifestation of inconsistencies. Recall that in each repetition we had two choice problems that tested consistency with transparent stochastic dominance. With 54 subjects, there could be a maximum of 108 violations per repetition. Observed violation rates are, however, very low. We observed one violation in repetition 1, three in repetition 2, two in repetition 3, and 1 in repetition 4. These low numbers indicate that subjects were rather attentive and that attentiveness did not decrease during the experiment, which could be the case in view of the relatively large number of choice problems subjects had to address.

Table 2 **Violations of Independence (in Percent)**

	Rep 1–2	Rep 3–4	*z*-value
Mean_C	33	29	−3.73***
Mean_S	28	22	−3.25***
EU	33	27	−3.52***
EU_C	37	32	−2.50**
EU_S	30	23	−2.51**
NEU	28	24	−3.42***
NEU_C	29	25	−2.80***
NEU_S	26	21	−2.10**
Mean	31	26	−4.91***

***Indicates a two-sided significance level of 1%.
**Two-sided significance level of 5%.

(ii) *Violations of independence*. Violations of independence are analyzed in the last columns of Table 1, which reports for each independence test the observed violation rate both in the coalesced and split version averaged over repetitions 1 and 2 and over repetitions 3 and 4. A summary of results is presented in Table 2 where the subscript *C* (*S*) denotes the coalesced (split) variant. The last row "mean" reports the average violation rates and the last column gives results of Wilcoxon tests. Overall, violation rates are only slightly decreasing by 5% but, as indicated by the last column of the table, the decrease is significant at the 1% level. The first two rows distinguish between violations in coalesced and split form, indicating that the latter occur less frequently. Subsequent rows distinguish between independence conditions, which are only implied by EU, and the weaker conditions, which are also implied by nonexpected utility (NEU) models. This analysis shows that the weaker independence conditions are less frequently violated, though this result is not significant. The lowest violation rates occur for weaker independence conditions presented in split form.

(iii) *Violations of coalescing*. Violation rates of coalescing are reported in Table 3, again averaged over repetitions 1 and 2, and over repetitions 3 and 4. The table shows that violations are decreasing in 12 cases and remain unchanged in 2. The last row reports that average violation rates are substantially decreasing

Table 3 Violations of Coalescing (in Percent)

Problem	Property	Rep 1–2	Rep 3–4
1–3	CCE1	15	8
5–7	CCE2	46	27
9–11	CCE3	39	12
10–12	CCE3	26	17
13–14	CCE4	42	24
15–17	CRE1	30	19
19–21	CRE2	31	17
20–22	CRE2	39	29
29–31	UTI	31	23
30–32	UTI	42	42
33–35	LTI	15	10
34–36	LTI	28	23
38–39	UCI	39	25
41–42	—	44	44
Mean		33	23

by 10%. This decrease is significant at the 1% level ($z = -3.06$, $p < 0.01$, two-sided).

4. Conclusions

This paper presented a repeated experiment on decision making under risk where people address the same choice problems on four occasions. As lotteries are only played out at the very end, any differences in behavior in the course the experiment can only be attributed to learning by thought, but not to learning by experience.

Our results show that learning by thought has a strong impact on risk preferences as inconsistencies, violations of independence, and violations of coalescing are significantly decreasing from earlier to later rounds. Our data are also consistent with the interpretation of an NEU model in which parameters are affected by stimuli presented in the lab (Birnbaum 2013). We cannot rule out that while the rates of violation of expected utility for certain choice problems decrease, they increase for other choice problems not tested.

In view of the contradictory results on the impact of learning by thought in the studies of Hey (2001) and van de Kuilen and Wakker (2006), our study supports the results of Hey (2001). This seems to indicate that learning by thought is only effective if identical choice problems are repeatedly addressed, several times, which was not the case in the van de Kuilen and Wakker (2006) design. Decision support and assessment of utility functionals should benefit from such repetitions. Altogether, as Hey (2001), we provide evidence supporting the view that EU in conjunction with the discovered preference hypothesis could be a reasonable characterization of individual behavior towards risk.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/deca.2015.0316.

Acknowledgments

The authors thank the associate editor and two anonymous referees for helpful comments on an earlier version of this paper.

Appendix

Table A.1 Split Variant of Independence Tests

Property	Lottery No.	p_1 / s_1	p_2 / s_2	p_3 / s_3	p_4 / s_4	q_1 / r_1	q_2 / r_2	q_3 / r_3	q_4 / r_4	Average choice % R
$CCE1_S$	3	0.89	0.01	0.10		0.89	0.01	0.10		88
		0	16	16		0	0	32		
	4	0.01	0.89	0.10		0.01	0.89	0.10		85
		16	16	16		0	16	32		
$CCE2_S$	7	0.80	0.10	0.10		0.80	0.10	0.10		30
		0	19	19		0	0	44		
	8	0.10	0.80	0.10		0.10	0.80	0.10		53
		19	19	19		0	19	44		
$CCE3_S$	11	0.70	0.10	0.10	0.10	0.70	0.10	0.10	0.10	24
		0	21	21	21	0	0	21	42	
	12	0.70	0.10	0.10	0.10	0.70	0.10	0.10	0.10	21
		0	21	21	42	0	0	42	42	

Table A.1 (Continued)

Property	Lottery No.	p_1 / s_1	p_2 / s_2	p_3 / s_3	p_4 / s_4	q_1 / r_1	q_2 / r_2	q_3 / r_3	q_4 / r_4	Average choice % R
$CCE4_S$	7	0.80	0.10	0.10		0.80	0.10	0.10		30
		0	19	19		0	0	44		
	14	0.40	0.10	0.10	0.40	0.40	0.10	0.10	0.40	50
		0	19	19	44	0	0	44	44	
$CRE1_S$	17	0.98	0.01	0.01		0.98	0.01	0.01		57
		0	23	23		0	0	46		
	18	0.50	0.50			0.50	0.50			12
		23	23			0	46			
$CRE2_S$	21	0.80	0.06	0.14		0.80	0.06	0.14		21
		0	28	28		0	0	45		
	22	0.40	0.18	0.42		0.40	0.18	0.42		27
		0	28	28		0	0	45		
UTI_S	31	0.73	0.01	0.01	0.25	0.73	0.01	0.01	0.25	72
		0	15	15	60	0	0	33	60	
	32	0.73	0.01	0.01	0.25	0.73	0.01	0.01	0.25	72
		0	15	15	33	0	0	33	33	
LTI_S	35	0.75	0.23	0.01	0.01	0.75	0.23	0.01	0.01	91
		1	34	36	36	1	33	33	60	
	36	0.75	0.23	0.01	0.01	0.75	0.23	0.01	0.01	88
		33	34	36	36	33	33	33	60	
UCI_S	37	0.20	0.20	0.60		0.20	0.20	0.60		74
		9	10	24		3	21	24		
	39	0.20	0.20	0.60		0.20	0.20	0.60		81
		9	9	21		3	21	21		

B. Instructions (Translated from German to English)

Welcome to this session in which you have to make decisions between gambles. There are no right or wrong decisions, we just want to learn about your individual preferences. You will receive sequentially four booklets. In each booklet you have to make 44 choices between two gambles. An example for such a choice appears below.

See Figure 1 in the main text.

If you would choose gamble A in this example you win €20 with a probability of 50%, €30 with a probability of 30%, and €40 with a probability of 20%. If you choose gamble B, you win €10 with a probability of 33%, €15 with a probability of 34% and €60 with a probability of 33%. To determine the amount to win we use this black urn which contains 100 tickets, numbered from 1 to 100. Note that in all gambles we ordered monetary prized from lowest to highest and the higher the number on the ticket drawn from the urn, the higher will be the prize you win. Suppose you have chosen gamble A. Then for tickets from 1 to 50 (i.e., with 50% probability) you will win €20, for tickets 51–80 (i.e., with 30% probability) you will win €30 and for tickets 81–100 (i.e., with 20%) probability you will win €40. If you have chosen gamble B you win €10 for tickets 1–33, €15 for tickets 34–67, and €60 for tickets 68–100.

We will pay out only one of your choices for real cash at the end of the session. To determine which choice this is, we have two additional urns. We will first draw a ticket from the red urn which contains tickets numbered from 1 to 4. The drawn ticket determines which of the four booklets is relevant for payout. Then we will draw a ticket from the blue urn which contains tickets numbered from 1 to 44. This ticket determines which of the 44 choices from the booklet is relevant for payout. In this way we determine the pair of gambles which is relevant for your payout and you will play the gamble you have chosen. Therefore, you should always choose the gamble you will like most. Now you have to answer four test questions before we start with the session. If you have any questions now or during the session, please raise your hand and one of the experimenters will come to you.

References

Binmore KG (1999) Why experiments in economics? *Econom. J.* 109:F16–F24.

Birnbaum MH (2001) Decision and choice: Paradoxes of choice. Smelser NJ, Baltes PB, eds. *International Encyclopedia of the Social and Behavioral Sciences* (Elsevier, Oxford, UK), 3286–3291.

Birnbaum MH (2004) Causes of allais common consequence paradoxes: An experimental dissection. *J. Math. Psych.* 48:87–106.

Birnbaum MH (2005) Three new tests of independence that differentiate models of risky decision making. *Management Sci.* 51(9):1346–1358.

Birnbaum MH (2012) A statistical test of the assumption that repeated choices are independently and identically distributed. *Judgment and Decision Making* 7:97–109.

Birnbaum MH (2013) True-and-error models violate independence and yet they are testable. *Judgment Decision Making* 8: 717–737.

Birnbaum MH, Chavez A (1997) Tests of theories of decision making: Violations of branch independence and distribution independence. *Organ. Behav. Human Decision Processes* 71:161–194.

Birnbaum MH, McIntosh WR (1996) Violations of branch independence in choices between gambles. *Organ. Behav. Human Decision Processes* 67:91–110.

Birnbaum MH, Navarrete JB (1998) Testing descriptive utility theories: Violations of stochastic dominance and cumulative independence. *J. Risk Uncertainty* 17:17–49.

Birnbaum MH, Patton JN, Lott MK (1999) Evidence against rank-dependent utility theories: Tests of cumulative independence, interval independence, stochastic dominance, and transitivity. *Organ. Behav. Human Decision Processes* 77:44–83.

Birnbaum MH, Schmidt U, Schneider MD (2009) Testing independence conditions in the presence of errors and splitting effects. Working paper.

Cox JC, Grether DM (1996) The preference reversal phenomenon: Response mode, markets and incentives. *Econom. Theory* 7: 381–405.

Dufwenberg M, Lindqvist T, Moore E (2005) Bubbles and experience: An experiment. *Amer. Econom. Rev.* 95:1731–1737.

Green JR, Jullien B (1988) Ordinal independence in nonlinear utility theory. *J. Risk Uncertainty* 1:355–387.

Harrison GW (1994) Expected utility theory and the experimentalists. *Empirical Econom.* 19(2):223–253.

Hey JD (2001) Does repetition improve consistency? *Experiment. Econom.* 4(1):5–54.

Hertwig R, Erev I (2009) The description-experience gap in risky choice. *Trends Cognitive Sci.* 13:517–523.

Humphrey SJ (1995) Regret aversion or event-splitting effects? More evidence under risk and uncertainty. *J. Risk Uncertainty* 11:263–274.

Humphrey SJ (2001) Non-transitive choice: Event-splitting effects or framing effects? *Economica* 68:77–96.

Loomes G, Starmer C, Sugden R (2003) Do anomalies disappear in repeated markets? *Econom. J.* 113:C153–C166.

Luce RD (1991) Rank- and sign-dependent linear utility models for binary gambles. *J. Econom. Theory* 53:75–100.

Luce RD, Fishburn PC (1991) Rank- and sign-dependent linear utility models for finite first-order gambles. *J. Risk Uncertainty* 4:29–59.

Myagkov M, Plott CR (1997) Exchange economies and loss exposure: Experiments exploring prospect theory and competitive equilibria in market environments. *Amer. Econom. Rev.* 87(5): 801–828.

Nicholls N, Romm AT, Zimper A (2015) The impact of statistical learning on violations of the sure-thing principle. *J. Risk and Uncertainty*. Forthcoming.

Plott CR (1996) Rational individual behaviour in markets and social choice processes: The discovered preference hypothesis'. Arrow KJ, Colombatto E, Perlman M, Schmidt C, eds. *The Rational Foundations of Economic Behaviour* (Macmillan, Basingstoke, UK), 225–250.

Quiggin J (1981) Risk perception and risk aversion among Australian farmers. *Australian J. Agricultural Econom.* 25:160–169.

Quiggin J (1982) A theory of anticipated utility. *J. Econom. Behav. Organ.* 3:323–343.

Schmidt U, Seidl C (2014) Reconsidering the common ratio effect: The roles of compound independence, reduction, and coalescing. *Theory and Decision* 77(3):323–339.

Smith VL (1989) Theory, experiment and economics. *J. Econom. Perspect.* 3(1):151–169.

Starmer C (1992) Testing new theories of choice under uncertainty using the common consequence effect. *Rev. Econom. Stud.* 59:813–830.

Starmer C, Sugden R (1989) Probability and juxtaposition effects: An experimental investigation of the common ratio effect. *J. Risk Uncertainty* 2:159–178.

Starmer C, Sugden R (1993) Testing for juxtaposition and event-splitting effects. *J. Risk Uncertainty* 6:235–254.

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

van de Kuilen G (2009) Subjective probability weighting and the discovered preference hypothesis. *Theory Decision* 67:1–22.

van de Kuilen G, Wakker PP (2006) Learning in the allais paradox. *J. Risk Uncertainty* 33:155–164.

Wakker PP, Tversky A (1993) An axiomatization of cumulative prospect theory. *J. Risk Uncertainty* 7:147–176.

Wu G (1994) An empirical test of ordinal independence. *J. Risk Uncertainty* 9:39–60.

Michael H. Birnbaum is a professor of psychology at California State University, Fullerton and Director of the Decision Research Center; in 1991 he was named the university's Outstanding Professor. He conducts research in the fields of judgment and decision making and psychophysical and social judgment. He has been the president of the Society for Mathematical Psychology, Society for Judgment and Decision Making, and Society for Computers in Psychology. In 2014, he won the L. Donald Shields Award.

Ulrich Schmidt is a professor of economics at the University of Kiel and research professor at the Kiel Institute for the World Economy. Additionally, he is a distinguished visiting professor at the University of Johannesburg. Ulrich Schmidt holds Ph.D. degrees in Economics and in Psychology. He worked intensively in the field of (non)expected utility. He is currently on the editorial board of *Journal of Risk and Uncertainty*, *Theory and Decision*, and *Geneva Risk and Insurance Review*.