# Electronic Supplementary 

## Material

## for

## "Prince: An Improved Method for

## Measuring Incentivized Preferences"

Cathleen Johnson, Aurélien Baillon, Han Bleichrodt, Zhihua Li, Dennie van Dolder, \& Peter P. Wakker

December, 2020

This Electronic Supplementary Material consists of two parts. Part 1, comprising §§ES.1-ES.8, contains regular appendixes, and has a regular page numbering [E.1], E.2, E.3, .... E.45. Part 2, comprising §§ES.9-ES.15, reproduces stimuli of Experiment 4 and exactly reproduces the page numberings of those stimuli.

In our statistical analyses, we throughout first calculate an index per subject, and then use between-subject variations in our tests. We throughout report t-tests. Wilcoxon tests never gave essential differences.

We end here with an afterthought. After having finished our experiment, we came to prefer a more symmetric reformulation of the implementation for matching questions of Prince, asking for an indifference value rather than a threshold, and then writing, for instance:

## E. 2

If $x$ in my envelope falls below (is worse than) my indifference value, then give me Option 1 (mug).
If $x$ in my envelope equals my indifference value, then I am indifferent.**
If $x$ in my envelope exceeds (is better than) my indifference value, then give me
Option 2 ( $€ \mathrm{x}$ ).
:
** Then a coin toss will decide whether Option 1 or 2 is given.

In this reformulation, upper and lower cases are treated symmetrically, reducing the risk of biases in either direction. For example, Figure 1 now becomes:

Figure 1'. Instructions for WTA with matching and endowment (Question 1)

In each of the 50 envelopes, one option is to keep your mug, and the other option is to give up your mug for a money amount. The note in each envelope is as follows.

## Option 1: Keep your mug

Option 2: Give up your mug for $€ \mathrm{X}$

The money amount $x$ varies between $€ 0$ and $€ 10$ in different envelopes.*
Please give us instructions which option from your envelope to give to you. Do so by specifying an indifference value (in cents).

## My indifference value is $€$

If $x$ in my envelope falls below (is worse than) my indifference value, then give me Option 1
(mug).
If $x$ in my envelope equals my indifference value, then I am indifferent.**
If $x$ in my envelope exceeds (is better than) my indifference value, then give me Option 2 ( $€ x$ x).

* Five envelopes contain a randomly generated amount between $€ 0$ and $€ 1$, five contain a randomly generated amount between $€ 1$ and $€ 2, \ldots$, and five contain a randomly generated amount between $€ 9$ and $€ 10$. Thus the amount in your envelope can be any amount, in cents, between $€ 0$ and $€ 10$.
** Then a coin toss will decide whether Option 1 or 2 is given.


## E. 3

The main text mentioned many anomalies in $\S 3.5$ when no range was specified (such as values $\mathrm{x}>4$ in Figure ES.2.2 below, for type $\lambda$ envelopes). We expect that the above reformulation will reduce such anomalies and, in general, is clearer.
E. 4

## Part 1. Experiments 1-3

## §ES. 1 of Electronic Supplementary Material: Questions 2 and 3 of Experiment 2

This section presents Questions 2 and 3 of Experiment 2, with matching and a choice list for chocolate (instead of the mug of Experiment 1).

# Instructions for envelopes of type $\boldsymbol{\tau}$ (Question 2 of <br> <br> Experiment 2) 

 <br> <br> Experiment 2)}

In each of the 10 envelopes of type $\tau$, one option is the chocolate you just saw, and the other option is a money amount. The note in each envelope of type $\tau$ is as follows.

## Type $\tau$

## Option 1: the chocolate

## Option 2: €x

The money amount $\mathbf{x}$ varies between $€ \mathbf{0}$ and $€ 10$ in different envelopes. One of the envelopes contains a randomly generated amount between $€ 0$ and $€ 1$, one envelope contains a randomly generated amount between $€ 1$ and $€ 2$, one contains a randomly generated amount between $€ 2$ and $€ 3$, and so on, with finally one envelope containing a randomly generated amount between $€ 9$ and $€ 10$. Thus the amount in your envelope can be any amount, in cents, between $€ 0$ and $€ 10$.

Please give us instructions, for each possible envelope of type $\tau$ that your envelope may be, whether we should give you the money amount or the chocolate. Do so by specifying a threshold (in cents).

## My threshold is

If the money amount x in my envelope is equal to or above the threshold, then give me that money amount.
If the money amount $x$ in my envelope is below the threshold then give me the chocolate.

## E. 7

## Instructions for envelopes of type $\psi$ (Question 3 of Experiment 2)

In each of the 10 envelopes of type $\psi$, one option is the chocolate you just saw, and the other option is a money amount. The money amount x varies between $€ 0.50$ and $€ 9.50$ in different envelopes (see below).

The note in each envelope of type $\psi$ is as follows.

## Type $\psi$

## Option 1: the chocolate

Option 2: € x

In the following list, each line describes the content of one envelope of type $\psi$. On each line, cross out the square before the option that we should give you if that line describes the two options in your envelope.

| 1. $\square \mathrm{CHOCOLATE}$ | $\square € 0.50$ |
| :--- | ---: |
| 2. $\square \mathrm{CHOCOLATE}$ | $\square € 1.50$ |
| 3. $\square \mathrm{CHOCOLATE}$ | $\square € 2.50$ |
| 4. $\square \mathrm{CHOCOLATE}$ | $\square € 3.50$ |
| 5. $\square \mathrm{CHOCOLATE}$ | $\square € 4.50$ |
| 6. $\square \mathrm{CHOCOLATE}$ | $\square € 5.50$ |
| 7. $\square \mathrm{CHOCOLATE}$ | $\square € 6.50$ |
| 8. $\square \mathrm{CHOCOLATE}$ | $\square € 7.50$ |
| 9. $\square \mathrm{CHOCOLATE}$ | $\square € 8.50$ |
| 10. $\square \mathrm{CHOCOLATE}$ | $\square € 9.50$ |

# §ES. 2 of Electronic Supplementary Material: Questions 4, 5, and 6 of Experiment 2: Testing preference reversals 

Questions 4, 5, and 6 of Experiment 2 investigate classical preference reversals for probability distributions over money. Objective probabilities p were generated using two ten-sided dice. Figure ES.2.1 displays Question 4 (choice question), which concerns a choice between $4_{0.970}$ (receiving $€ 4$ with probability 0.97 and $€ 0$ otherwise) and $16_{0.3} 10$. Question 5 (Figure ES.2.2) is called the CE-P-bet question. It measures the certainty equivalent ( $C E$ ), i.e. the equally preferred sure amount, of 40.970 . The latter bet is called the $P$-bet because it has a high probability of winning. Question 6 (Figure ES.2.3) is called the CE-\$-bet question. It measures the CE of 160.310 , called the $\$$-bet because it involves the highest money amount (in dollars when receiving its name; Lichtenstein \& Slovic 1971).

This time we specified no range or further information on the values x in Questions 5 and 6, leaving its choice entirely to the subjects. This neither affects incentive compatibility nor transparency. It still is perfectly clear to subjects that providing wrong instructions can only harm themselves, possibly leading to less preferred options.

Figure ES.2.1. Instructions for choice between lotteries (Question 4 of Experiment 2), the choice question)

## Instructions for envelopes of type $\boldsymbol{\theta}$

In the 10 envelopes of type $\theta$, both options are probability-contingent money. The note in each envelope of type $\theta$ is as follows.

## Type $\theta$

## Option 1: €4 with a probability of $97 \%$ <br> Option 2: €16 with a probability of $31 \%$

Please give us instructions whether we should give you option 1 or option 2 if your envelope is of type $\theta$.

## Give me option

Figure ES.2.2. Instructions for certainty equivalent of P-bet (Question 5 of Experiment 2, the CE-P-bet)

## Instructions for envelopes of type $\boldsymbol{\lambda}$

In the 10 envelopes of type $\lambda$, one option is probability-contingent money, and the other option is a sure money amount. The note in each envelope of type $\lambda$ is as follows.

## Type $\lambda$

## Option 1: €4 with a probability of $97 \%$

## Option 2: €x

The money amount $\mathbf{x}$ varies between the envelopes. Please give us instructions for each possible envelope of type $\lambda$ that your envelope may be, whether we should give you the sure money amount or the probability-contingent money. Do so by specifying a threshold (in cents):

## My threshold is $€$

If the money amount $x$ in my envelope is equal to or above the threshold, then give me that money amount.
If the money amount $x$ in my envelope is below the threshold then give me the probability-contingent money.

Figure ES.2.3. Instructions for certainty equivalent of \$-bet (Question 6 of Experiment 2, CE-\$-bet)

## Instructions for envelopes of type $\rho$

In the 10 envelopes of type $\rho$, one option is probability-contingent money, and the other option is a sure money amount. The note in each envelope of type $\rho$ is as follows.

## Type $\rho$

## Option 1: €16 with a probability of $31 \%$

## Option 2: €x

The money amount $\mathbf{x}$ varies between the envelopes. Please give us instructions for each possible envelope of type $\rho$ that your envelope may be, whether we should give you the sure money amount or the probability-contingent money. Do so by specifying a threshold (in cents):

## My threshold is $€$

If the money amount $x$ in my envelope is equal to or above the threshold, then give me that money amount.
If the money amount $x$ in my envelope is below the threshold then give me the probability-contingent money.

Analysis. In Question 5 of Experiment 2, the CE-P-bet, 15 subjects violated stochastic dominance with CEs exceeding 5 . We remove these subjects from the analyses.

Violations of stochastic dominance up to 1 can be explained by choice errors, rounding, and an indifference-bias-upwards. ${ }^{1}$ So as to treat upward and downward

[^0]biases symmetrically and reduce the biasing effects of removing only high values, we also removed subjects with CE-P values below 3; this concerned only 3 subjects. Finally, one subject did not provide an indifference value for the CE-P-bet, and is also removed from the analyses. This leaves 61 subjects for our analyses. Incorporating the removed subjects would not affect any conclusion. The average CE of the CE-P-bet is 3.96. The average CE of the CE-\$-bet is larger: 5.74 , which is in agreement with common findings ( $t_{60}=5.31, \mathrm{p}<0.001$ ). However, 79 percent of the subject preferred to $\$$-bet in direct choice, and only 21 percent preferred the P-bet ( $t_{60}=5.43$, $\mathrm{p}<0.001$ ), deviating from common findings. Our group averages suggest that, unlike common findings, we have no preference reversals. This conclusion is confirmed by an analysis at the individual level: Normal preference reversals (higher CE of the \$ bet but choosing the P bet) occurred for $11 \%$ of the subjects, and the opposite preference reversal (higher CE of the P bet but choosing the $\$$ bet) happened for 7\% of the subjects. These percentages are not significantly different ( $p=0.37$ ), suggesting again noise in the data but no systematic direction.

Discussion of preference reversals. There are only few preference reversals, and they are random, resulting from random choice inconsistencies commonly found (Schmidt \& Hey 2004). Our findings deviate from other studies into preference reversals, where the normal preference reversals are found in great majorities (surveyed by Seidl 2002). Our findings suggest that preference reversals reflect errors in the measurement of preferences (procedural variance) rather than genuine properties of preferences such as intransitivities (Tversky, Slovic, \& Kahneman 1990). Prince has restored consistency between choice and matching, thus resolving preference reversals.

Discussion of no provision of range in matching. In the CE-P questions, we found many violations of stochastic dominance. For the preference reversals we used the classical and often used stimuli of Lichtenstein \& Slovic (1971). These stimuli for preference reversals involve small outcomes and weak strengths of preferences, generating extra noise. Further noise is generated because we specified no range for matching, unlike in Question 3. In the CE-P questions, 24 CEs thus exceeded 4. It is plausible that the true indifference value, very close to $€ 4$, is rounded to $€ 4$, and then

## E. 13

because of the indifference-bias-upwards a value exceeding $€ 4$ may be chosen.
However, 15 CEs even exceeded 5 (and these were removed for being erratic). This result suggests that specifying a range as in $\S 2.2$ is preferable, supporting Birnbaum’s (1992) views. The indifference-bias-upwards, which also played a role here, can be avoided by implementing random choice (rather than always the sure amount) at the stated switching value, as in Figure 1'.

## §ES. 3 of Electronic Supplementary Material:

 Questions 7, 8, and 9 of Experiment 2:Measuring subjective probabilities and ambiguity attitudes

Questions 7 (Figure ES.3.1), 8, and 9 of Experiment 2 show how Prince can be used to measure subjective probabilities and ambiguity attitudes. They were taken from Baillon \& Bleichrodt (2015 Study 1). Question 7 measures the probability p such that

$$
\begin{equation*}
10_{\mathrm{A}} 0 \sim 10_{\mathrm{p}} 0, \tag{ES.3.1}
\end{equation*}
$$

using matching. Here A denotes an uncertain event that may or may not happen. In Figure ES.3.1, A means that the Dutch AEX stock index increases or decreases by no more than $0.5 \%$ during the experiment (which we immediately verified online). $10_{\mathrm{A}} 0$ means that the subject receives $€ 10$ if A happens, and nothing otherwise. $10_{\mathrm{p}} 0$ similarly means that the subject received $€ 10$ with objective probability p. The objective probability p in Eq. ES.3.1 is called the matching probability of event A, denoted $m(A)$. Matching probabilities were widely used in classical Bayesian models. They have recently been found to serve the analysis of ambiguity attitudes well (Wakker 2010 p. 321).

Question 8 (Figure ES.3.2) is like Question 7, but with event A replaced by event B: the AEX stock index increases by more than $0.5 \%$ during the experiment. Question
9 (Figure ES.3.3) concerns event $\mathrm{A} \cup \mathrm{B}$, meaning the AEX stock index does not decrease by more than $0.5 \%$ during the experiment. We will compare our values $m(A), m(B), m(A \cup B)$, and the nonadditivity index $m(A \cup B)-m(A)-m(B)$, with those obtained by Baillon \& Bleichrodt (2015). These authors used classical choice lists with the classical RIS instead of our matching with Prince. They found a positive nonadditivity index.

Figure ES.3.1. Instructions for matching probability (Question 7 of Experiment 2)

## Instructions for envelopes of type $\varphi$

In each of the 10 envelopes of type $\varphi$, one option is AEX-contingent money, and the other option is probability-contingent money. The note in each envelope of type $\varphi$ is as follows.

## Type $\varphi$

Option 1: $€ 10$ if the AEX (Dutch stock index) increases by more than $-0.5 \%$ and by less than $0.5 \%$ during the experiment


Option 2: $\quad € 10$ with a probability of $\mathrm{p} \%$

The probability $\mathbf{p}$ varies between $\mathbf{0 \%}$ and $\mathbf{1 0 0 \%}$ in different envelopes. One of the envelopes contains a randomly selected probability between 0 and $10 \%$, one envelope contains a randomly selected probability between $10 \%$ and $20 \%$, one contains a randomly selected probability between $20 \%$ and $30 \%$, and so on, with finally one containing a randomly selected probability between $90 \%$ and $100 \%$. Thus the probability in your envelope can be any percentage between $0 \%$ and $100 \%$.

Please give us instructions for each possible envelope of type $\varphi$ that your envelope may be, whether we should give you the AEX-contingent money or the probabilitycontingent money. Do so by specifying a threshold (in percentage):

## My threshold is \%

If the probability $p$ in my envelope is equal to the threshold or above, then give me the probability-contingent option 2.
If the probability p in my envelope is below the threshold then give me the AEXcontingent option 1.

FIGURE ES.3.2. Instructions for matching probability (Question 8 of Experiment 2)

## Instructions for envelopes of type $\omega$

In each of the 10 envelopes of type $\omega$, one option is AEX-contingent money, and the other option is probability-contingent money. The note in each envelope of type $\omega$ is as follows.

## Type $\omega$

Option 1: $€ 10$ if the AEX (Dutch stock index) increases by more than $0.5 \%$ during the experiment


AEX
Option 2: €10 with a probability of $\mathrm{p} \%$

The probability $\mathbf{p}$ varies between $\mathbf{0 \%}$ and $\mathbf{1 0 0 \%}$ in different envelopes. One of the envelopes contains a randomly selected probability between 0 and $10 \%$, one envelope contains a randomly selected probability between $10 \%$ and $20 \%$, one contains a randomly selected probability between $20 \%$ and $30 \%$, and so on, with finally one containing a randomly selected probability between $90 \%$ and $100 \%$. Thus the probability in your envelope can be any percentage between $0 \%$ and $100 \%$.

Please give us instructions for each possible envelope of type $\omega$ that your envelope may be, whether we should give you the AEX-contingent money or the probabilitycontingent money. Do so by specifying a threshold (in percentage):

My threshold is
If the probability $p$ in my envelope is equal to the threshold or above, then give me the probability-contingent option 2.
If the probability p in my envelope is below the threshold then give me the AEXcontingent option 1.

Figure ES.3.3. Instructions for matching probability (Question 9 of Experiment 2)

## Instructions for envelopes of type $\chi$

In each of the 10 envelopes of type $\chi$, one option is AEX-contingent money, and the other option is probability-contingent money. The note in each envelope of type $\chi$ is as follows.

## Type $\chi$

Option 1: $€ 10$ if the AEX (Dutch stock index) increases by more than $-0.5 \%$ during the experiment


## AEX

Option 2: €10 with a probability of p\%

The probability $\mathbf{p}$ varies between $\mathbf{0 \%}$ and $\mathbf{1 0 0 \%}$ in different envelopes. One of the envelopes contains a randomly selected probability between 0 and $10 \%$, one envelope contains a randomly selected probability between $10 \%$ and $20 \%$, one contains a randomly selected probability between $20 \%$ and $30 \%$, and so on, with finally one containing a randomly selected probability between $90 \%$ and $100 \%$. Thus the probability in your envelope can be any percentage between $0 \%$ and $100 \%$.

Please give us instructions for each possible envelope of type $\chi$ that your envelope may be, whether we should give you the AEX-contingent money or the probabilitycontingent money. Do so by specifying a threshold (in percentage):

## My threshold is <br> \%

If the probability p in my envelope is equal to the threshold or above, then give me the probability-contingent option 2.
If the probability p in my envelope is below the threshold then give me the AEXcontingent option 1.

Results. Figure ES.3.4 gives results.
Figure ES.3.4. Mean of the matching probabilities of variables


Under expected utility we should have

$$
\begin{equation*}
\mathrm{m}(\mathrm{~A})+\mathrm{m}(\mathrm{~B})=\mathrm{m}(\mathrm{~A} \cup \mathrm{~B}) \tag{ES.3.2}
\end{equation*}
$$

but this equality is rejected (Table ES.3.1).

TABLE ES.3.1. $t$-test of additivity of matching probabilities (Eq. ES.3.2)

|  |  | Mean | Std. Deviation | N | $t$-test for equality of means |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  | $t$ |  | $\mathrm{p}(2$-tailed) |  |  |
| $\mathrm{m}(\mathrm{A})+\mathrm{m}(\mathrm{B})$ | 0.94 | 0.39 | 80 | 6.67 | 79 | .000 |  |
| $\mathrm{~m}(\mathrm{~A} \cup \mathrm{~B})$ | 0.66 | 0.28 | 80 |  |  |  |  |

None of our values $m(A), m(B), m(A \cup B)$, and the nonadditivity index $m(A \cup B)$ $-m(A)-m(B)$, were significantly different from those of Baillon \& Bleichrodt (2015). It confirms once again that matching under Prince has the same validity as choice, and that our measurement of subjective probabilities is more efficient without being less valid.

Discussion. The inequality found $(m(A)+m(B)-m(A \cup B)>0)$ is often referred to as subadditivity (Tversky \& Fox 1995). We briefly mention some implications for ambiguity theories. The inequality confirms a(mbiguity-generated likelihood)insensitivity (Abdellaoui et al. 2011). It means, roughly, that the inverse-S shaped processing of likelihoods, moving subjective likelihoods towards fifty-fifty, is more pronounced for ambiguity than for risk. Insensitivity reflects a lack of understanding and discriminatory power, and for our stimuli it entails a violation of Schmeidler's (1989) ambiguity aversion. Yet it is the common empirical finding (Einhorn \& Hogarth 1985; l’Haridon et al. 2018; Riege \& Teigen 2013; Trautmann \& van de Kuilen 2015; Wakker 2010 end of §10.4.2).

Lower subadditivity and a-insensitivity are genuine properties of preference and not artefacts of measurement. Hence Prince will not remove them. Here, as throughout, the advantage of Prince is that we obtained our data more quickly and precisely than preceding papers did without losing validity.

TABLE ES.3.2. A $t$-test between Baillon \& Bleichrodt (2015, Study 1) and our (Prince) Questions 7,8,9 of Experiment 2

|  |  | $t$-test for Equality of Means |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | T | df | $\begin{aligned} & \text { p (2- } \\ & \text { tailed) } \end{aligned}$ | Mean Difference | Std. Error <br> Difference |
| m(B) | Equal variances assumed** | -0.75 | 114.00 | 0.45 | -0.04 | 0.05 |
|  | Equal variances not assumed | -0.83 | 89.90 | 0.41 | -0.04 | 0.04 |
| m(A) | Equal variances assumed*** | 1.36 | 114.00 | 0.18 | 0.07 | 0.05 |
|  | Equal variances not assumed | 1.53 | 94.89 | 0.13 | 0.07 | 0.05 |
| m(AUB) | Equal variances assumed** | 0.74 | 114.00 | 0.46 | 0.04 | 0.06 |
|  | Equal variances not assumed | 0.83 | 95.01 | 0.41 | 0.04 | 0.05 |
| $\begin{aligned} & \hline \mathrm{m}(\mathrm{AUB})- \\ & \mathrm{m}(\mathrm{~A})- \\ & \mathrm{m}(\mathrm{~B}) \end{aligned}$ | Equal variances assumed | -0.06 | 114.00 | 0.96 | 0.00 | 0.07 |
|  | Equal variances not assumed | -0.06 | 72.71 | 0.96 | 0.00 | 0.07 |

** the $p$ value of the test of equal variance is significant at $5 \%$ level.
*** the $p$ value of the test of equal variance is significant at $1 \%$ level.
nonadditivity index: $m(A)+m(B)-m(A \cup B)$

# §ES. 4 of Electronic Supplementary Material: Further details of Experiment 3 (adaptive measurement of utility) 

## ES.4.1 Preceding studies and real incentives

Wakker \& Deneffe (1996) introduced the tradeoff (TO) method. If real incentives are implemented using the traditional BDM method and RIS, then subjects can answer strategically and may rationally overstate the values of $r^{1}, r^{2}$, and $r^{3}$. Hence, most implementations were hypothetical. Exceptions, using real incentives, include Abdellaoui (2000), Bleichrodt, Cillo, \& Diecidue (2010), Schunk \& Betsch (2006), and van de Kuilen \& Wakker (2011). Subjects never noticed the possibility to answer strategically (van de Kuilen \& Wakker 2011, end of §4).

We use Prince to implement real incentives in the TO method. Then the RCS has been randomly selected for each subject before any decision making. Subjects possess the RCS but with details remaining unknown. Subjects then obviously do not have any opportunity to influence the RCS and, furthermore, this is perfectly clear to them. Thus, there is no possibility to answer strategically. Providing the RCS prior to the experiment also rules out any illusion of strategic answering in nonadaptive experiments (Wang, Venkatesh, \& Chatterjee 2007 p. 203).

## ES.4.2 Construction and use of envelopes for real incentives

For the 100 envelopes constructed before the experiment, we chose 12 types of envelopes, one for each question from TO1.1 - TO3.4. The envelope of type TOi.j ( $\mathrm{i}=$ $1-3 ; j=1-4)$ contains the two lotteries of Question TOi,j with values substituted as indicated in Table ES.4.1. The values $\overline{\mathrm{x}^{0}}=\mathrm{x}^{0}=18, \overline{\mathrm{y}^{0}}=\mathrm{y}^{0}=25$, and $\overline{\mathrm{z}^{0}}=\mathrm{z}^{0}=210$ are as in Figure 6. For example, the envelopes of type TO1.3 results from Figure 7b and contains the lotteries $322_{2} 0$ and $27_{2 / 3} 9$. The only exception is type TO3.4, which
contains $\overline{z^{3^{\prime}}}=3050$ instead of $\mathrm{z}^{3}=342$ for $\mathrm{r}^{3 .}$ The $\#$ numbers in Table ES.4. 1 indicate how many of the 100 envelopes were of the particular type. For example, there were three envelopes of type TO1.2.

Table ES.4.1. The pre-set TO values

|  | $\mathrm{j}=1$ | $\mathrm{j}=2$ | $\mathrm{j}=3$ | $\mathrm{j}=4$ |
| :--- | :--- | :--- | :--- | :--- |
| TO1 | $\overline{\mathrm{x}^{1}}=23$ for $\mathrm{x}^{1}(\# 33)$ | $\overline{\mathrm{x}^{2}}=27$ for $\mathrm{x}^{2}(\# 3)$ | $\overline{\mathrm{x}^{3}}=32$ for $\mathrm{x}^{3}(\# 3)$ | $\overline{\mathrm{x}^{4}}=36$ for $\mathrm{x}^{4}(\# 3)$ |
| TO2 | $\overline{\mathrm{y}^{1}}=46$ for $\mathrm{y}^{1}(\# 33)$ | $\overline{\mathrm{y}^{2}}=69$ for $\mathrm{y}^{2}(\# 3)$ | $\overline{\mathrm{y}^{3}}=92$ for $\mathrm{y}^{3}(\# 3)$ | $\overline{\mathrm{y}^{4}}=116$ for $\mathrm{y}^{4}(\# 3)$ |
| TO3 | $\overline{z^{1}}=253$ for $\mathrm{z}^{1}(\# 9)$ | $\overline{\mathrm{z}^{2}}=297$ for $\mathrm{z}^{2}(\# 3)$ | $\overline{z^{3}}=342$ for $\mathrm{z}^{3}(\# 3)$ | $\overline{z^{4}=3100 \text { for } \mathrm{z}^{4 *}(\# 1)}$ |

* In this choice question (Figure 6, TO3, $\mathrm{j}=4$ ), we used $\overline{\mathrm{z}^{3}}=3050$ (instead of $\overline{\mathrm{z}^{3}}=$ 342) for $z^{3}$.

The stimuli of TO1 had been used by Abdellaoui (2000) and those of TO2 had been used by Booij, van Praag, \& van de Kuilen (2010), both scaled up. Those of TO3 were new. They serve to study high outcomes.

Subjects were informed that: (a) at least one of every ten of them would play for real; (b) if playing for real, the average gain under random ("blind") choosing is 53.27; (c) at least one prize exceeded 3000. They were told that the types of envelope did not occur equally often, which should be obvious with 100 envelopes of 12 types.

In other respects the organization was as with the Prince experiments reported before. For example, the questionnaire asked subjects to give us instructions about which option from their own envelope to give them at the end of the experiment. One more difference was as follows. In the previous experiments in this paper we demonstrated how Prince can be used with a regular performance-contingent real payment for every subject. We now implement Prince with performance-contingent real payments only for some subjects, to facilitate larger payments. Abdellaoui et al. (2011, Online Appendix) suggested that such payment schemes work best to motivate subjects. For this purpose, at the end we collected the three front pages numbered TOj. 0 of each subject that contained their answers, whereas the subjects kept the rest
of their questionnaires that also contained their answers. ${ }^{2}$ The front pages of each subject were folded together so as to be unrecognizable, and were put in an opaque case and shuffled. Then we let one subject draw some (three in the first session and six in the second) triples of questionnaires from this case, after which the corresponding subjects played for real.

## ES.4.3 Maximizing overlap with experimental questions

The numbers $x^{j}$ for TO1 in Table ES.4.1 result from expected value maximization, and those for TO2 and TO3 (except TO3.4) result from Tversky \& Kahneman's (1992) prospect theory and their parameters. We rounded the $x^{j}$ s. We constructed the 100 envelopes, in particular the frequency of each type of envelope among the 100 envelopes, so as to generate an appropriate overall expected value of the game for the subjects.

We chose most envelopes of type TOj. 1 because here three rather than two outcomes were known beforehand. Hence, each answer $x^{1}$ by the subject implied a choice from the corresponding envelope, depending on whether $x^{1}$ exceeded the corresponding value from the envelope or not. ${ }^{3}$ The same observations hold for $\mathrm{y}^{1}$ and $\mathrm{z}^{1 .}$ We thus obtain a high overlap between experimental answers and RCSs.

In the explanations given at the beginning of the experiment, an example was used that occurred in a pilot: In TO1.1 a subject in a pilot had answered $\mathrm{x}^{1}=50$, implying an indifference $10_{1 / 2} 8 \sim 50_{1 / 2}$. Then, in case the choice in the envelope had been between the options $10_{1 / 2} 8$ and $32_{1 / 2} 1$, the former, $10_{1 / 2} 8$, would be given to the subject. We explained that the subject's preference would then only be reinforced if outcome 10 in his preferred lottery had been increased to 24 , and that he would want the former even more from the pair $\left\{24_{1 / 2} 8,32_{1 / 2} 1\right\}$. This pair was actually contained in the envelope of this subject, who indeed received $24_{1 / 2} 8$. In general, if the value $\bar{r}^{j}$ in the envelope exceeded the answer $r^{j}$ given by the subject and the value $\overline{r^{i-1}}$ in the

[^1]envelope was below $r^{i-1}$ (reinforcing the preference), then we would give the lottery with outcome $\overline{r^{j}}$ to the subject. Conversely, if the value $\overline{r^{j}}$ in the envelope was below the answer $\mathrm{r}^{\mathrm{j}}$ given by the subject and the value $\overline{r^{i-1}}$ in the envelope exceeded $r^{i-1}$ (reinforcing the preference), then we would give the lottery with outcome $\overline{r^{i-1}}$ to the subject. We, thus, used transitivity and dominance to derive choices, so that the answers given during the experiment pertained to many possible envelopes. This was explained to the subjects.

## ES.4.4 Verifiable absence of deception

As in the other experiments in this paper, subjects could verify that there was no deception. They again verified, at the beginning, through sets of 10 numbered envelopes, that all numbered envelopes were present. Subjects collected, shuffled, and selected envelopes from bags themselves. At the end, when subjects who had been selected for payment came to the front of the room, a list describing the content of all 100 envelopes was handed out to the other subjects, with calculations showing that our information about expected value under random play and maximal amounts was correct. These other subjects were asked to open their envelope and verify that the list correctly described its content. For the subjects who were in front of the class, everyone in the class could see that the description of their envelopes was correct. Similar procedures were followed in all experiments described in this paper. ${ }^{4}$

One of the experimenters carried over $€ 3000$ in cash with him, showing to the subjects that this amount could and would be paid on the spot if the lucky case arose. The random selection of at least one per 10 subjects to play for real (three out of 25 in the first session and 6 out of 55 in the second) was done by letting the subjects put their questionnaires in an opaque bag, after which the questionnaires to be implemented for real were randomly selected by a subject. Lotteries were generated

[^2]using a six-sided die, thrown by a subject selected for this role by the subject who was playing for real.

## ES.4.5 Parametric analysis of Experiment 3 on utility measurement

We used Eq. 1 in the main text with the two most common parametric utility families, CARA (constant absolute risk aversion, or linear-exponential) and CRRA (constant relative risk aversion, or log-power) utility. They are defined as follows.

CARA utility:

$$
\begin{align*}
& \text { for } \alpha>0, U(r)=1-e^{-\alpha r} ;  \tag{ES.4.1}\\
& \text { for } \alpha=0, U(r)=r ;  \tag{ES.4.2}\\
& \text { for } \alpha<0, U(r)=e^{-\alpha r}-1 . \tag{ES.4.3}
\end{align*}
$$

CRRA utility ${ }^{5}$ :

$$
\begin{align*}
& \text { for } \rho>0, \mathrm{U}(\mathrm{r})=\mathrm{r}^{\rho} \text {; }  \tag{ES.4.4}\\
& \text { for } \rho=0, \mathrm{U}(\mathrm{r})=\ln (\mathrm{r}) \text {; }  \tag{ES.4.5}\\
& \text { for } \rho<0, \mathrm{U}(\mathrm{r})=-\mathrm{r}^{\rho} . \tag{ES.4.6}
\end{align*}
$$

We used Eq. 2 in the main text as the basis of our parametric analysis, rewriting it as

$$
\begin{equation*}
\mathrm{r}^{\mathrm{j}}=\mathrm{U}^{-1}\left(2 \mathrm{U}\left(\mathrm{r}^{\mathrm{j}-1}\right)-\mathrm{U}\left(\mathrm{r}^{\mathrm{j}-2}\right)\right) \text { for } \mathrm{j}=2,3,4 . \tag{ES.4.7}
\end{equation*}
$$

We assumed a Fechnerian error model with the error term directly imposed on answers $\mathrm{r}^{j}$ that subjects produced:

$$
\begin{equation*}
\mathrm{r}^{\mathrm{j}}=\mathrm{U}^{-1}\left(2 \mathrm{U}\left(\mathrm{r}^{\mathrm{j}-1}\right)-\mathrm{U}\left(\mathrm{r}^{\mathrm{j}-2}\right)\right)+\varepsilon_{R R} \text { for } \mathrm{j}=2,3,4 . \tag{ES.4.8}
\end{equation*}
$$

We chose to impose error terms upon the direct measured values, being $r^{j}$ (Eq. ES.4.7), rather than upon utilities as often done in the literature. Our error model is similar to Wilcox's (2011) contextual approach. The error term $\varepsilon_{R R}$ has expectation 0

[^3]and standard deviaton $\sigma>0$ per $€$. That is, $\sigma$ is to be interpreted as the standard deviation when differences $\mathrm{r}^{\mathrm{j}}-\mathrm{r}^{\mathrm{j}-1}$ are in the order of magnitude of $€ 1$. Our error model is thus close to the thinking process of subjects, because the TO method enhances the direct comparison of $\mathrm{r}^{\mathrm{j}}$ to $\mathrm{r}^{\mathrm{j}-1}$ during the experiment. Given that $\mathrm{x}^{\mathrm{j}}-\mathrm{x}^{\mathrm{j}-1}=4.5$ under expected value maximization, we chose the standard deviation $\varepsilon_{\mathrm{x}}$ for $\mathrm{r}^{\mathrm{j}}=\mathrm{x}^{\mathrm{j}}$ equal to $4.5 \sigma$. We similarly used expected value approximations to set $\varepsilon_{y}=13 \sigma$ and $\varepsilon_{z}=40 \sigma$. We estimate $\sigma$ and the utility parameter so as to maximize likelihood.

## ES.4.6 Results of parametric analysis of Experiment 3 on utility measurement

Our main analysis is at the individual level, finding the best utility parameter and error variance for all observations $\mathrm{x}^{\mathrm{j}}, \mathrm{y}^{\mathrm{j}}$, and $\mathrm{z}^{\mathrm{j}}$ simultaneously. Normal distributions of the parameters are rejected. We, hence, report medians, Wilcoxon signed rank, Mann-Whitney, and Kruskal-Wallis tests.

The median CARA parameter $\alpha$ is 0.0001 (more precisely, $\alpha=0.000099$ ), yielding risk tolerance $1 / \alpha=€ 10000$. It might suggest concave utility, but the value is not significantly different from linearity ( $\alpha=0$ ). The median power $\rho$ is 0.96 , again suggesting weak concavity but again not deviating significantly from linearity ( $\rho=1$ ).

The standard deviation $\sigma$ is larger for CARA utility than for CRRA utility ( $\mathrm{p}=$ $0.001)^{6}$. Thus CRRA utility fits the data better than CARA utility. We also fitted parameters to the x stimuli, y stimuli, and z stimuli separately, and compared them. Kruskal-Wallis tests reveal that CARA's $\alpha$ marginally depends on $x, y$, and $z(p=$ 0.09 ), being higher for x than for y stimuli $(\mathrm{p}=0.03$ ), marginally higher for x than for $z$ stimuli ( $p=0.09$ ), and not different between $y$ and $z$ stimuli. These results confirm the commonly assumed decreasing absolute risk aversion (references in Wakker 2010 p. 83). A reason that we find no differences for the z stimuli may be because the z stimuli, while at higher levels of wealth, also involve bigger differences, which in itself will enhance risk aversion. CRRA's $\rho$ does depend on $x, y$, and $z(p=0.01)$,

[^4]
## E. 27

being smaller for the x than for the y and z stimuli ( $\mathrm{p} \leq 0.01$ ), and not different between the $y$ and $z$ stimuli. This suggests decreasing relative risk aversion. Although most authors have conjectured increasing, rather than decreasing relative risk aversion, several studies found the opposite (Gollier 2001 §4; Ogaki \& Zhang 2001) and no consensus has emerged on this point.

# §ES. 5 of Electronic Supplementary Material: Demographic questions 

The following questions were asked to our subjects at the end of Experiment 2.

Now please open your envelope.
Your envelope is of type .... Your instruction for type ... is .... Now it is clear that you get .... Do you agree?

At the very end, we would like to ask some feedback questions. These questions have no influence on what your get from your envelope but will help us in our investigation.

Finally, please indicate your age, gender and nationality:

## Age

Gender : male O female O

Nationality :

# §ES. 6 of Electronic Supplementary Material: Experiment 4: A comparison between Prince and standard RIS 

We next present details of Experiment 4, comparing Prince and a standard implementation of RIS, in $\S 5$ in the main text. A well-known difficulty when evaluating methods of preference measurement is that we usually have no gold standard for true preference (Infante, Lecouteux, \& Sugden 2016; Janis \& Mann 1977 p. 11; Pedroni et al. 2017 2nd para; Thaler \& Sunstein 2008; Tversky \& Kahneman $1981,2^{\text {nd }}$ to last para). A second, related, difficulty is that there is no clearly leading method of preference measurement to be compared with at present. Whichever method and implementation we choose for comparison, many readers will prefer a different one. For example, the winner of the horse race by Cox, Sadiraj, \& Schmidt (2015; CSS henceforth), discussed below, was newly introduced by them and has not yet been widely used, which is why we chose a more common method.

A horse race in predictive power cannot be well implemented because different measurement methods need different stimuli as inputs. Therefore, prediction tests commonly concern different theories, and not different measurement methods (Erev et al. 2010). For the aforementioned reasons, papers introducing new preference measurement methods have typically refrained from comparative tests, leaving those to other researchers and the test of time. ${ }^{7}$

We nevertheless carried out a comparative experiment, exploratory whenever a gold standard for true preference was absent. As the method competing with Prince we chose a standard RIS. RIS may be the most popular implementation of real incentives for individual preference measurements today. We replicated a number of remarkable findings in the literature on choices under risk. We first discuss Prince's

[^5]secondary effect, being that it improves the quality of decisions somewhat. Given the moderate stakes typically involved in experiments, it may be argued that expected value maximization is rational there (de Finetti 1937; Kahneman \& Lovallo 1993; l'Haridon \& Vieider 2019 p. 189; Luce 2000 p. 86; Samuelson 1959 p. 35; Savage 1971 p. 786). We will, therefore, report the degree of deviations from expected value in the choices considered below. It will be accepted as a criterion for decision quality only by those readers who accept expected value maximization as optimal for our stimuli.

As for the primary purpose of Prince, of better revealing true preference irrespective of its rationality, we gave supportive theoretical arguments in the main text. The two cases where we could specify true preferences have been discussed in §5 in the main text. We here present the other tests. We will signal differences between Prince and RIS for these stimuli, leaving interpretations regarding true preferences to the readers.

Many choice questions were repeated on a different page, so that subjects had probably forgotten and answered independently. These provide a test of consistency. Many other choice questions were repeated on the same page, with the first times this explicitly pointed out, so that the repetition was salient. They provide a test of deliberate randomization of choice by subjects.

## ES.6.1. Subjects, incentives, and stimuli

$\mathrm{N}=51$ subjects, recruited from undergraduates of Erasmus School of Economics in Rotterdam, were randomly divided into a group for Prince ( $\mathrm{n}=26$ ) and one for the standard RIS ( $\mathrm{n}=25$ ). They received $€ 5$ show-up fee and could gain additional money from a lottery choice, $€ 11.50$ on average if they chose randomly. We chose 60 choice situations, displayed in §ES.10, and discussed further later. In what follows, Fn.m refers to the choice situation/pair in Figure n.m in §ES.10. Throughout, Fn.m = 1 indicates a left choice, and Fn.m $=0$ indicates a right choice.

## ES.6.2. Procedure for Prince

Each of the 60 choice situations (each describing a choice between two lotteries) was written on a slip of paper and then randomly inserted in one of 60 numbered envelopes, which were then closed. The envelopes were displayed at the entrance of the lab on a table, with numbers consecutively and visibly ordered in 6 rows of 10 colums. Subjects entered the laboratory one by one. Each was guided to a cubicle by one of three experimenters, after having observed and confirmed that there were exactly 60 numbered envelopes with each number appearing once. In the cubicle were written explanations to be read (§ES.9), and a questionnaire to be filled out (§ES.10). After all subjects had been seated, they were asked to come forward again and choose one of the 60 envelopes on the table, which was not allowed to be opened. They then returned to their cubicle and filled out their questionnaire.

After 25 minutes, they were told that, when they were done, they could raise their finger. One of the experimenters then went to their cubicle, opened their envelope to see what their real choice situation was, indicated it by a circle and her paraph on the questionnaire with the subject confirming it, and guided the subject to the front table. There another experimenter implemented the lottery chosen in the real choice pair, using the pages of §ES. 11 with their green text to specify which die numbers generated the winning probabilities, and then had the subject throw the two dies as described in the explanations to the subjects (§ES. 9 p. 3) to determine their gain. The subject received the money gained from the lottery plus the show-up fee and signed a receipt. The subject then received a detailed list describing the content of every numbered envelope (§ES.12), and was asked to verify that her envelope contained what was described there. ${ }^{8}$ The list further gave the calculations of the promised overall expected value of $€ 11.50$ under random choice, so that subjects could also verify this.

To the best of our knowledge, there is no way in which the experimenters could have deceived the subjects, and this was completely verifiable for every subject. Although this complete verifiability is not part of Prince's s definition, it is a desirable feature that this is possible.

[^6]
## ES.6.3. Procedure for RIS

Many details were the same as with Prince, and we only describe differences. There were no 60 envelopes at the end, but subjects were immediately guided to a cubicle with written explanations (§ES.13) and a questionnaire (§ES.14) to be filled out. They concerned the same 60 choice situations as with Prince.

25 minutes later subjects were told that, when they were done, they could raise their finger. One of the experimenters then guided the subject to the front table. There another experimenter first let the subject throw two dies to determine which of the 60 choice situations was implemented for real, as explained in §ES. 13 p .2 , and then things proceeded as with Prince. The list with the 60 choice situations given at the end of course did not refer to envelopes, but simply listed the 60 choice situations on the questionnaire with calculations of expected values.

## ES.6.4. Stimuli continued, and tests

As for the competing method, RIS, there are many ways to implement it and, whichever way we chose, many readers would have preferred several details differently. One aim was to have as many details of Prince and RIS the same as possible, so that only the four essential differences reflected in the acronym Prince (see the main text) are different. We therefore organized the pages of the stimuli as we did (§ES. 10 and §ES.14). For both methods, this organization with figures is very clear to subjects.

In what follows, ( $\mathrm{p}_{1}: \mathrm{x}_{1}, \ldots, \mathrm{p}_{\mathrm{n}}: \mathrm{x}_{\mathrm{n}}$ ) denotes a lottery (probability distribution over money) yielding $€ \mathrm{x}_{\mathrm{j}}$ with probability $\mathrm{p}_{\mathrm{j}}, \mathrm{j}=1, \ldots, \mathrm{n}$, where probabilities are nonnegative and sum to $1 . \alpha_{p} \beta$ denotes ( $\mathrm{p}: \alpha, 1-\mathrm{p}: \beta$ ), as in the main text.

In our statistical tests, we throughout predict that Prince is closer to the (arguably rational) expected value maximization and consistency than RIS, leading to one-sided t-tests, which will be the default. Only in cases where there is no clearly predicted direction of effects or deviation from expected value, do we use two-sided t-tests, which then is always stated explicitly.

## Getting acQuainted with the stimuli

The first pairs, F1.1-F1.5 present trivial choices to get subjects acquainted with the stimuli. We used potential losses because they have been found to increase the attention of subjects (Yechiam, Retzer, Telpaz, \& Hochman 2015).

## CONSISTENCY

Several pairs were repeated on different pages, so that subjects had forgotten and chose independently. These allow for verification of consistency in our experiment. It concerns the pairs F1.6 \& F10.4, F2.2 \& F9.3, F2.3 \& F5.4, F2.4 \& F6.1, F3.1 \& F10.5, F3.4 \& F9.4, F4.1 \& F10.6, F4.4 \& F9.5, F5.1 \& F6.6. We do not include pairs F2.1 \& F9.2 and F2.5 \& F10.2 here because they test spillover effects, and pairs F7.2 \& F10.3 because they test insurance framings. Tested variable: (ABS(F1.6 - F10.4) + ABS(F3.1 - F10.5) + ABS(F3.4 - F9.4) + ABS(F4.1 - F10.6) + ABS(F4.4 - F9.5) + ABS(F5.1 - F6.6) $+\operatorname{ABS}(F 5.4-\mathrm{F} 2.3)+\mathrm{ABS}(\mathrm{F} 6.1-\mathrm{F} 2.4)+\mathrm{ABS}(\mathrm{F} 2.2-\mathrm{F} 9.3)) / 9$. ABS denotes absolute value.

Violations of stochastic dominance (Tversky \& Kahneman 1986; §5 in our main text)

Tversky \& Kahneman (1986) presented a choice pair that led a majority of subjects to violate stochastic dominance. Birnbaum \& Navarrete (1998) presented a general recipe for developing such stimuli. Our pair F2.2 (Figure 9 in the main text), repeated in pair F9.3, is Problem 7 in Tversky \& Kahneman (1986, p. 178). Tested variable: F2.2 + F9.3-1.

Horse race between incentive systems (Cox, Sadiraj, \& Schmidt 2015; see §5 in the main text)

Cox, Sadiraj, \& Schmidt (2015), CSS15 henceforth, tested several incentive systems in a horse race. They used a design that, further, tested several other effects found in the literature. We replicated their choices:

Pair F6.1 ( $3_{0.25} 0$ vs $5_{0.20}$ ): CSS15 Table 1 row 1.
Pair F6.2 ( $10_{0.80} 0$ vs 6 ): CSS15 Table 1 row 2.
Pair F6.3 ( 60.250 vs $10_{0.20} 0$ ): CSS15 Table 1 row 3.
Pair F6.4 (( $0.75: 12,0.20: 10,0.05: 0)$ vs $\left.12_{0.756}\right)$ : CSS15 Table 1 row 4.

Pair F6.5 ( $22_{0.80} 12$ vs 18 ): CSS15 Table 1 row 5.

The tests comparing different incentive systems are discussed in $\S 5$ in the main text, and will not be repeated here. We focus here on the other effects tested by these stimuli.

Pairs F6.1 and F6.3 provide a test of constant relative risk aversion. Multiplying all outcomes in pair F6.1 by 2 yields pair F6.3. Tested variable: F6.3-F6.1.

Pairs F6.2 and F6.5 provide a test of constant absolute risk aversion. Adding 12 to all outcomes in pair F6.2 yields pair F6.5. Tested variable: F6.5-F6.2.

Pairs F6.2 and F6.3 provide a test of the common ratio effect. Dividing all probabilities of nonzero outcomes in pair F6.2 by 4 yields pair F6.3. Tested variable: 1-F6.3-F6.2.

Pairs F6.3 and F6.4 provide a test of the common consequence effect (although there is no riskless lottery here). Replacing the common outcome 0 received with probability 0.75 in pair F6.3 by common outcome 12 yields pair F6.4. Tested variable: 1-F6.4-F6.3.

## Certainty Effect

Several pairs contained one sure (riskless) option, directly testing risk aversion or, more specifically, the certainty effect. This concerned pairs F1.1, F1.2, F1.3, F1.4, F1.5, F1.6, F2.1, F2.3, F2.5, F2.6, F5.4, F5.5, F5.6, F6.2, F6.5, F7.2, F7.4, F7.6, F8.1, F8.2, F8.3, F8.4, F8.5, F8.6, F9.1, F9.2, F9.6, F10.1, F10.2, F10.3, and F10.4. Tested variable: 1 - (F1.1 + F1.2 + F1.3 + F1.4 + F1.5 + F1.6 + F2.1 + F2.3 + F2.5 + F2. $6+$ F5.4 + F5.5 + F5.6 + F6.2 + F6.5 + (1-F7.2) + (1-F7.4) + (1-F7.6) + F8.1 + F8.2 + F8.3 + F8.4 + F8.5 + F8.6 + F9.1 + F9.2 + F9.6 + F10.1 + F10.2 + F10.3 + F10.4)/31.

SPILLOVER EFFECTS (Cox, Sadiraj, \& Schmidt 2014)
Stewart, Reimers, \& Harris (2015) and Cox, Sadiraj, \& Schmidt (2014), CSS14 henceforth, tested spillover effects, where the precedence of one choice pair can impact later choices. We replicated the two tests of CSS14 that gave significant results in their paper.

Pair F2.1 \& pair F9.2 ( $10_{0.50} 0$ vs 4): CSS14 Experiment 1, Option B vs. Option A in Groups 1 and Group 2.1. Pair F2.1 was in a neutral context (as with CSS14's

Group 1), but pair F9.2 was preceded by pair F9.1 ( $12_{0.50} 0$ vs 3 ; CSS14 Option D vs. C) that could induce a spillover effect, as with CSS14's Group 2.1, also presented on one page.

Pair F2.5 \& pair F10.2 ( $10_{0.50} 0$ vs 5) : CSS14 Experiment 2, Option B vs. Option A in Groups 1 and 3.1. Pair F2.5 was in a neutral context as with CSS14's Group 1, but pair F10.2 was preceded by pair F10.1 ( $8_{0.50} 0$ vs 6 ; CSS Option F vs. E) that could induce a spillover effect as with CSS14’s Group 3.1, also presented on one page. Tested variable: (F2.1-F9.2 + F10.2-F2.5)/2.

Deliberate randomization in hard choices (Agranov \& Ortoleva 2017) Agranov \& Ortoleva (2017), A\&O henceforth, showed that subjects deliberately randomize their choices. A\&O presented the same choice pairs three times to subjects, explicitly pointing this out. For their four hard choice pairs (Hard1-Hard4 in their Table 1), subjects deliberately did not make the same choice in such repetitions. We replicate those four choices. ${ }^{9}$ As did A\&O, we repeated them three times on the same page, pointing this out explicitly, but repeated them one more time on a different page, where subjects would have forgotten. We used:

Pairs F3.1-F3.3 (3.850.251.90 vs $4.70_{0.50} 0.80$ ): A\&O Hard1. Repeated in pair F10.5. Pairs F3.4-F3.6 ((0.25:2.80, 0.50:2.25, $0.25: 1.60)$ vs. $\left.4.50_{0.50} 0.50\right)$ : A\&O Hard2. Repeated in pair F9.4.

Pairs F4.1-F4.3 ((0.25:10, 0.25:5.25, 0.25:4.20, $0.25: 0.30)$ vs. ( $0.25: 6.75,0.25: 5.85$, 0.25:3, 0.25:2.70)): A\&O Hard3. Repeated in pair F10.6.

Pairs F4.4-F4.6 ((0.25:4.05, $0.25: 2.55,0.25: 1.50,0.25: 0.65)$ vs. (0.25:4.30, 0.25:1.90, 0.25:1.60, 0.25:0.95)): A\&O Hard4. Repeated in pair F9.5.

[^7]
## E. 36

Deliberate randomization in Common consequence (Starmer \& Sugden 1991)
Starmer \& Sugden (1991), S\&S henceforth, tested Allais’ (1953) common consequence effect. We replicate their stimuli. We present the same choice pair three times on the same page, as did S\&S. For uniformity of stimuli throughout our experiment, though, we allowed our subjects to choose three times independently, allowing these choices to be different. S\&S, to the contrary, imposed the same choice for all three situations. They did so to test reduction of compound lotteries, and then isolation in RIS. We obtain, instead, another test of deliberate randomization for simpler stimuli than considered before, and we obtain a test of the common consequence effect for moderate outcomes.

Pairs F5.1-F5.3 ( $10_{0.20} 0$ vs 70.250 ): S\&S Question 22.
Pairs F5.4-F5.6 ((0.20:10, 0.75:7, 0.05:0) vs. 7): S\&S Question 21.

## Deliberate Randomization in Insurance

It is well-known that formulating risky decisions as insurance decisions increases risk aversion (Hershey, Kunreuther, \& Schoemaker 1982 p. 949/950). We tested this. Subjects considered

Pair F7.2: prior endowment of $€ 15$; then -5 versus $(-15)_{0.32} 0$, formulated as insurance decision versus pair F10.3: prior endowment of $€ 15$; then -5 versus ( -15$)_{0.32} 0$, formulated as neutral choice.

Pair F7.1 provided a trivial choice to get subjects accustomed to the insurance framing. Pairs F7.1/F7.2 were repeated on the same page in pairs F7.3/F7.4 and F7.5/F7.6, to provide another test of deliberate randomization.

## E. 37

The tested variable for deliberate randomization was calculated as follows: x31 $=$ F3.1 + F3.2 + F3.3 (nr. of left choices, signaling consistency if 0 or 3 and inconsistency if 1 or 2 )
x34 = F3.4 + F3.5 + F3.6
$\mathrm{x} 41=\mathrm{F} 4.1+\mathrm{F} 4.2+\mathrm{F} 4.3$
$\mathrm{x} 44=\mathrm{F} 4.4+\mathrm{F} 4.5+\mathrm{F} 4.6$
x51 = F5.1 + F5.2 + F5.3
x54 = F5.4 + F5.5 + F5.6
$\mathrm{x} 71=\mathrm{F} 7.1+\mathrm{F} 7.3+$ F7.5
$\mathrm{x} 72=\mathrm{F} 7.2+$ F7.4 + F7. 6
$\mathrm{y} 31=\mathrm{ABS}(\mathrm{x} 31-1.5)-0.5$ (is 0 if inconsistency with x 31 is 1 or 2 ;
is 1 with consistency if x 31 is 0 or 3 )
y34 $=\operatorname{ABS}(x 34-1.5)-0.5$
y41 $=$ ABS (x41-1.5) -0.5
$\mathrm{y} 44=\operatorname{ABS}(\mathrm{x} 44-1.5)-0.5$
$\mathrm{y} 51=\operatorname{ABS}(\mathrm{x} 51-1.5)-0.5$
y52 $=$ ABS(x52-1.5) -0.5
y71 $=$ ABS (x71-1.5) -0.5
y72 $=$ ABS (x72-1.5) -0.5
Tested variable $=(\mathrm{y} 31+\mathrm{y} 34+\mathrm{y} 41+\mathrm{y} 44+\mathrm{y} 51+\mathrm{y} 52+\mathrm{y} 71+\mathrm{y} 72) / 8$ is an index of inconsistency of a subject, with values between 0 and 1 .

## ES.6.5. Analyses and results

CONSISTENCY
Prince had inconsistency 0.24 and RIS had 0.25 , in agreement with other findings in the literature (Agranov \& Ortoleva 2017; Hey \& Orme 1994; Stott 2006). The difference between Prince and RIS is not significant.

## VIoLations of stochastic dominance

We calculated for each subject the number of wrong (left-sided) choices in pairs F2.2 and F9.3, i.e. the number of violations of stochastic dominance. Prince had fewer ( $\mathrm{p}=$ 0.04 ). We also tested the related question for pair F2.2 alone ( $p=0.04$ ) and pair F9.3
alone ( $p=0.16$ ). For RIS, the group average violated stochastic dominance ( $p=0.02$, two-sided), but for Prince there was exact indifference. We conclude that Prince has fewer violations of stochastic dominance than RIS, and is closer to true preference.

## Horse race between incentive systems, and Starmer \& Sugden’s (1991) COMMON CONSEQUENCE EFFECT

The results regarding Cox, Sadiraj, \& Schmidt's (2015) horse race between different incentive systems are in $\S 5$ in the main text. We give here the results regarding the other phenomena that Cox, Sadiraj, \& Schmidt tested.

The differences between Prince and RIS were not significant regarding violations of constant relative risk aversion (Prince: 8\%; RIS: 12\%), violations of constant absolute risk aversion (Prince: 38\%; RIS: 52\%), or common ratio violations of expected utility in the common direction ${ }^{10}$ (Prince: 0.15 ; RIS ${ }^{11}: 0.32$ ). The common consequence violations of expected utility (tested in pair F6.3 versus F6.4) were stronger under RIS ( 0.52 ) than under Prince ( 0.12 ) ( $\mathrm{p}=0.01$ ). We found similar results in our replication of the common consequence test of Starmer \& Sugden (1991). Here we took for each subject the average number of safe choices in F5.4-F5.6 ((1-(F 5.4 + F5.5 + F5.6)/3)) minus the average number of safe choices in F5.1-F5.3 ((1-(F 5.1 + F5.2 + F5.3)/3)), as test statistic (Prince: 0.19; RIS: 0.48), and their difference. The violations were stronger under RIS than under Prince ( $p=0.047$ ).

## Certainty Effect

Prince tended to exhibit risk aversion through the certainty effect less strongly than RIS: subjects in the Prince treatment chose the certain option 65\% of the time, while those in the RIS treatment chose the certain option $71 \%$ of the time. This difference,

[^8]however, only reaches marginal significance overall ( $p=0.055$ ). The same is true if we focus on insurance-framed choices (tested variable: 3-F7.2-F7.4-F7.6): subjects in the Prince treatment opt for insurance $62 \%$ of the time, while subjects in the RIS treatment do so $80 \%$ of the time ( $\mathrm{p}=0.06$ ).

## SPILLOVER EFFECTS

Regarding the spillover effect, for each subject we took the average inconsistency over pairs F2.1 vs. F9.2 and then F2.5 vs. F10.2. It was, joining the two tests, $23 \%$ for Prince and $26 \%$ for RIS, and this difference is not significant.

## Deliberate randomization

We tested deliberate randomization (not choosing the same prospect in questions that repeated on a page) in eight pairs. Overall, for Prince it was $11 \%$, and for RIS $13 \%$. Their difference is not significant. Prince and RIS were neither significantly different for separate pairs or subgroups of pairs.

## An additional test

We found nonsignificant differences for six cases: inconsistency, nonconstant relative risk aversion, common ratio violations of expected utility, risk aversion, spillover effects, and deliberate randomization. In each case, Prince was closer to consistency or expected value than RIS. The null hypothesis that this is coincidence can be rejected ( $p=0.02$ ), showing once more that Prince is closer to consistency and expected value than RIS.

# §ES. 7 of Electronic Supplementary Material: Further stimuli for Experiments 1-3 

Further stimuli and background material for Experiments 1-3 is here:
http://personal.eur.nl/wakker/data/21.1.prince/links.htm

# §ES. 8 of Electronic Supplementary Material: References for Electronic Supplementary Material 

Abdellaoui, Mohammed (2000) "Parameter-Free Elicitation of Utility and Probability Weighting Functions," Management Science 46, 1497-1512.
Abdellaoui, Mohammed, Aurélien Baillon, Laetitia Placido, \& Peter P. Wakker (2011) "The Rich Domain of Uncertainty: Source Functions and Their Experimental Implementation," American Economic Review 101, 695-723. Agranov, Marina \& Pietro Ortoleva (2017) "Stochastic Choice and Preferences for Randomization," Journal of Political Economy 125, 40-68.

Allais, Maurice (1953) "Fondements d’une Théorie Positive des Choix Comportant un Risque et Critique des Postulats et Axiomes de l'Ecole Américaine," Colloques Internationaux du Centre National de la Recherche Scientifique (Econométrie) 40, 257-332. Paris: Centre National de la Recherche Scientifique. Translated into English, with additions, as "The Foundations of a Positive Theory of Choice Involving Risk and a Criticism of the Postulates and Axioms of the American School." In Maurice Allais \& Ole Hagen (1979, eds.) Expected Utility Hypotheses and the Allais Paradox, 27-145, Reidel, Dordrecht.

Andreoni, James, Michael A. Kuhn, \& Charles Sprenger (2015) "On Measuring Time Preferences," Journal of Economic Behavior and Organization 16, 451-456.

Andreoni, James \& Charles Sprenger (2012) "Estimating Time Preference from Convex Budgets," American Economic Review 102, 3333-3356.

Attema, Arthur E., Han Bleichrodt, Yu Gao, Zhenxing Huang, \& Peter P. Wakker (2016) "Measuring Discounting without Measuring Utility," American Economic Review 106, 1476-1494.

Baillon, Aurélien \& Han Bleichrodt (2015) "Testing Ambiguity Models through the Measurement of Probabilities for Gains and Losses," American Economic Journal: Microeconomics 7, 77-100.

Birnbaum, Michael H. (1992) "Should Contextual Effects in Human Judgment Be Avoided?," Book Review of: E. Christopher Poulton (1989) "Bias in Quantifying Judgments," Erlbaum, Hillsdale NJ; Contemporary Psychology 37, 21-23.

Birnbaum, Michael H. \& Juan B. Navarrete (1998) "Testing Descriptive Utility Theories: Violations of Stochastic Dominance and Cumulative Independence," Journal of Risk and Uncertainty 17, 49-78.
Bleichrodt, Han, Alessandra Cillo, \& Enrico Diecidue (2010) "A Quantitative Measurement of Regret Theory," Management Science 56, 161-175.

Booij, Adam S., Bernard M.S. Van Praag, \& Gijs van de Kuilen (2010) "A Parametric Analysis of Prospect Theory's Functionals," Theory and Decision 68, 115-148.

Choi, Syngjoo, Raymond Fishman, Douglas Gale, \& Shachar Kariv (2007) "Consistency and Heterogeneity of Individual Behavior under Uncertainty," American Economic Review 97, 1921-1938.

Cox, James C., Vjollca Sadiraj, \& Ulrich Schmidt (2014) "Asymmetrically Dominated Choice Problems, the Isolation Hypothesis and Random Incentive Mechanisms," PLoS ONE 9, e90742.

Cox, James C., Vjollca Sadiraj, \& Ulrich Schmidt (2015) "Paradoxes and Mechanisms for Choice under Risk," Experimental Economics 18, 215-250. de Finetti, Bruno (1937) "La Prévision: Ses Lois Logiques, ses Sources Subjectives," Annales de l'Institut Henri Poincaré 7, 1-68. Translated into English by Henry E. Kyburg Jr., "Foresight: Its Logical Laws, its Subjective Sources," in Henry E. Kyburg Jr. \& Howard E. Smokler (1964, eds.) Studies in Subjective Probability, 93-158, Wiley, New York; 2nd edn. 1980, 53-118, Krieger, New York.

Einhorn, Hillel J. \& Robin M. Hogarth (1985) "Ambiguity and Uncertainty in Probabilistic Inference," Psychological Review 92, 433-461.

Epper, Thomas, Helga Fehr-Duda, \& Adrian Bruhin (2011) "Viewing the Future through a Warped Lens: Why Uncertainty Generates Hyperbolic Discounting," Journal of Risk and Uncertainty 43, 163-203.

Erev, Ido, Eyal Ert, Alvin E. Roth, Ernan Haruvy, Stefan M. Herzog, Robin Hau, Ralph Hertwig, Terrence Stewart, Robert West, \& Christian Lebiere (2010) "A Choice Prediction Competition: Choices from Experience and from Description," Journal of Behavioral Decision Making 23, 15-47.
Gollier, Christian (2001) "Wealth Inequality and Asset Pricing," Review of Economic Studies 68, 181-203.

Hershey, John C., Howard C. Kunreuther, \& Paul J.H. Schoemaker (1982) "Sources of Bias in Assessment Procedures for Utility Functions," Management Science 28, 936-953.

Hey, John D. \& Chris Orme (1994) "Investigating Generalizations of Expected Utility Theory Using Experimental Data," Econometrica 62, 1291-1326.

Holt, Charles A. \& Susan K. Laury (2002) "Risk Aversion and Incentive Effects," American Economic Review 92, 1644-1655.

Infante, Gerardo, Guilhem Lecouteux, \& Robert Sugden (2016) "Preference Purification and the Inner Rational Agent: A Critique of the Conventional Wisdom of Behavioural Welfare Economics," Journal of Economic Methodology 23, 1-25.

Janis, Irving L. \& Leon Mann (1977) "Decision Making; A Study of Conflict, Choice and Commitment." The Free Press, New York.

Kahneman, Daniel \& Dan Lovallo (1993) "Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking," Management Science 39, 17-31.
l'Haridon, Olivier \& Ferdinand Vieider (2019) "All over the Map: A Worldwide Comparison of Risk Preferences," Quantitative Economics 10, 185-215.
l’Haridon, Olivier, Ferdinand Vieider, Diego Aycinena, Augustinus Bandur, Alexis Belianin, Lubomir Cingl, Amit Kothiyal, \& Peter Martinsson (2018) "Off the Charts: Massive Unexplained Heterogeneity in a Global Study of Ambiguity Attitudes," Review of Economics and Statistics 100, 664-677.

Lichtenstein, Sarah \& Paul Slovic (1971) "Reversals of Preference between Bids and Choices in Gambling Decisions," Journal of Experimental Psychology 89, 46-55.

Luce, R. Duncan (2000) "Utility of Gains and Losses: Measurement-Theoretical and Experimental Approaches." Lawrence Erlbaum Publishers, London.

Ogaki, Masao \& Qiang Zhang (2001) "Decreasing Relative Risk Aversion and Tests of Risk Sharing," Econometrica 69, 515-526.
Pedroni, Andreas, Renato Frey, Adrian Bruhin, Gilles Dutilh, Ralph Hertwig, \& Jörg Rieskamp (2017) "The Risk Elicitation Puzzle," Nature Human Behaviour 1, forthcoming.

Riege, Anine H. \& Karl Halvor Teigen (2013) "Additivity Neglect in Probability Estimates: Effects of Numeracy and Response Format," Organizational Behavior and Human Decision Processes 121, 41-52.

Samuelson, Paul A. (1959) "The St. Petersburg Paradox as a Divergent Double Limit," International Economic Review 1, 31-37.

Savage, Leonard J. (1971) "Elicitation of Personal Probabilities and Expectations," Journal of the American Statistical Association 66, 783-801.

Schmeidler, David (1989) "Subjective Probability and Expected Utility without Additivity," Econometrica 57, 571-587.
Schmidt, Ulrich \& John D. Hey (2004) "Are Preference Reversals Errors? An Experimental Investigation," Journal of Risk and Uncertainty 29, 207-218.

Schunk, Daniel \& Cornelia Betsch (2006) "Explaining Heterogeneity in Utility Functions by Individual Differences in Decision Modes," Journal of Economic Psychology 27, 386-401.
Seidl, Christian (2002) "Preference Reversal," Journal of Economic Surveys 16, 621655.

Starmer, Chris \& Robert Sugden (1991) "Does the Random-Lottery Incentive System Elicit True Preferences? An Experimental Investigation," American Economic Review 81, 971-978.

Stewart, Neil, Stian Reimers, \& Adam J.L. Harris (2015) "On the Origin of Utility, Weighting, and Discounting Functions: How They Get Their Shapes and how to Change Their Shapes," Management Science 61, 687-705.

Stott, Henry P. (2006) "Cumulative Prospect Theory’s Functional Menagerie," Journal of Risk and Uncertainty 32, 101-130.
Thaler, Richard H. \& Cass R. Sunstein (2008) "Nudge: Improving Decisions about Health, Wealth, and Happiness." Yale University Press, New Haven.

Trautmann, Stefan T. \& Gijs van de Kuilen (2015) "Ambiguity Attitudes." In Gideon Keren \& George Wu (eds.), The Wiley Blackwell Handbook of Judgment and Decision Making (Ch. 3), 89-116, Blackwell, Oxford, UK.

Tversky, Amos \& Craig R. Fox (1995) "Weighing Risk and Uncertainty," Psychological Review 102, 269-283.
Tversky, Amos \& Daniel Kahneman (1981) "The Framing of Decisions and the Psychology of Choice," Science 211, 453-458.

Tversky, Amos \& Daniel Kahneman (1986) "Rational Choice and the Framing of Decisions," Journal of Business 59, S251-S278.

Tversky, Amos \& Daniel Kahneman (1992) "Advances in Prospect Theory: Cumulative Representation of Uncertainty," Journal of Risk and Uncertainty 5, 297-323.

Tversky, Amos, Paul Slovic, \& Daniel Kahneman (1990) "The Causes of Preference Reversal," American Economic Review 80, 204-217.

## E. 45

van de Kuilen, Gijs \& Peter P. Wakker (2011) "The Midweight Method to Measure Attitudes toward Risk and Ambiguity," Management Science 57, 582-598.
Wakker, Peter P. (2010) "Prospect Theory: For Risk and Ambiguity." Cambridge University Press, Cambridge, UK.

Wakker, Peter P. \& Daniel Deneffe (1996) "Eliciting von Neumann-Morgenstern Utilities when Probabilities Are Distorted or Unknown," Management Science 42, 1131-1150.

Wang, Tuo, Ramaswamy Venkatesh, \& Rabikar Chatterjee (2007) "Reservation Price as a Range: An Incentive-Compatible Measurement Approach," Journal of Marketing Research 64, 200-213.

Wilcox Nathaniel T. (2011) " ‘Stochastically More Risk Averse:’ A Contextual Theory of Stochastic Discrete Choice under Risk," Journal of Econometrics 162, 89-104.

Yechiam, Eldad, Matan Retzer, Ariel Telpaz, \& Guy Hochman (2015) "Losses as Ecological Guides: Minor Losses Lead to Maximization and not to Avoidance," Cognition 139, 10-17.

# Part 2. The stimuli used in 

## Experiment 4

## §ES. 9 of Electronic Supplementary Material:

## Prince explanations for subjects in Experiment 4

The following 7 pages present the explanations given to subjects for the Prince treatment of Experiment 4. The page numbers [1,] 2-7 of those explanations have been kept.

## Explanation of the Experiment

## Welcome!

You will receive:
€5 for participation
$+$
an additional prize.

## Do not open your envelope.

The additional prize comes from your envelope.
Your envelope contains a pair of prospects
(explained later). You will get one of these two prospects.
During the experiment, you write instructions to us about which of the two prospects to give to you at the end.
Your goal: get most-wanted prospect from your envelope.

Average additional prize if you give instructions randomly: $€ 11.50$. But your instructions will be better than random!

# There are some negative prizes (losses) that would be subtracted from the participation fee, but you can always instruct us to avoid those. 

Please do not communicate with other participants.

There are no right or wrong answers.
Give us instructions according to what you want most.
That (what you want) is also what we are interested in, and want to investigate.

## Explaining prospects

We have two 10 -sided dice.
One can take the values: $\quad 0,1,2,3, \ldots, 9$
The other can take the values: $00,10,20,30, \ldots, 90$

You will throw both dice and their numbers are added, leading to a range of 100 values, from 0 up to 99 .
Each value is equally likely.

## Example of a prospect:


$€ 16$ with a probability of 0.31 :

- If value of dice < 31, then you win $€ 16$
- If value of dice $\geq 31$, then you win nothing

The probability of the prize indeed is 0.31 .
If you get this prospect, we will use the figure

to specify in green which values of the throws of dice give which outcomes.

## How you give instructions to us

There is a questionnaire containing 10 pages numbered $1 \ldots$ 10 , each with 6 pairs of prospects, numbered 1.1 ... 1.6, ..., $10.1, \ldots$ 10.6. These 60 pairs of prospects were randomly inserted into the 60 envelopes. So, your envelope contains one of these pairs.

For each possible pair of prospects in your envelope, you will write us instructions about which one of that pair we should give to you.

## Of course:

If you write what you want, then you get what you want!

## Verification

At end of experiment:
you receive a list describing the pairs in all 60 envelopes.
You will then check that the description of your numbered envelope is correct, i.e., as described in the list. Can then also check that the average additional prize indeed is €11.50.

Questions at any time: please raise your hand. Experimenters will come.

## Remember

There are no right or wrong answers; it is only about what you want.

Instructions:
If you write what you want, then you get what you want.

Do not communicate with other participants.

Do not open your envelope.

## Now the experiment can begin.

For each pair of prospects in the questionnaire, instruct us which one to give to you by marking the box above that prospect with an x , such as in the example below where you would prefer the prospect to the right.


## §ES. 10 of Electronic Supplementary Material:

## Prince stimuli for Experiment 4

The following 11 pages present the Prince stimuli of Experiment 4. The page numbers [0], 1-10 of those stimuli have been kept.

Overview of the 10 pages (reduced) of the questionnaire containing 6 prospect pairs each


The 10 pages follow next, nonreduced.



Please give us instructions on which prospect
to give to you, by marking its box with an x.

Below, questions will be repeated. That is, you will see the same pair three times, one after the other.


Below, questions will be repeated. That is, you will see the same pair three times, one after the other.


Please give us instructions on which prospect to give to you, by marking its box with an $x$. _ -

Below, questions will be repeated. That is, you will see the same pair three times, one after the other. 5







## §ES. 11 of Electronic Supplementary Material:

## Implementation of lotteries for Experiment 4

The following 11 pages present the pages used by an experimenter when implementing lotteries of Experiment 4, both for Prince and for RIS. The page numbers [0], 1-10 of those pages have been kept.

For experimenter: die values added in green to implement probabilities



Below, questions will be repeated. That is, you will see the same pair three times, one after the other.









# §ES. 12 of Electronic Supplementary Material: Verification of lotteries and envelopes for Prince for Experiment 4 

The following 13 pages present the pages given to subjects at the end of the Prince treatment of Experiment 4. They made it possible for subjects to verify that the experiment had been done honestly, with the envelopes having the proper content. The page numbers [1], 2-13 of those pages have been kept.

## Verification document for Prince experiment

This document, first, provides the expected value (EV) calculations of the additional prize (showup fee not included), which is $€ 11.50$ or, more precisely, $€ 11.57$. Next it indicates for each choice pair in which envelope it was contained. This allows you to verify that the experiment was conducted honestly, in particular, that your randomly chosen envelope had the proper content -as it did with the other subjects.

## 1. Expected value calculation

prospect1.1 left has EV -1.15; prospect1.1 right has EV 0.00; pair 1.1 has EV -0.58 . prospect1.2 left has EV -1.15; prospect1.2 right has EV 0.00; pair 1.2 has EV -0.35. prospect1.3 left has EV -0.60; prospect1.3 right has EV 0.00; pair 1.3 has EV -0.30 . prospect1.4 left has EV -0.40; prospect1.4 right has EV 0.00; pair 1.4 has EV -0.20 . prospect1.5 left has EV -0.20; prospect1.5 right has EV 0.00; pair 1.5 has EV -0.10 . prospect1.6 left has EV 5.00; prospect1.6 right has EV 5.00; pair 1.6 has EV 5.00. EV page $1=(-0.58-0.35-0.30-0.20-0.10+5.00) / 6=0.58$
prospect2.1 left has EV 5.00; prospect2.1 right has EV 4.00; pair 2.1 has EV 4.50. prospect2.2 left has EV 3.45; prospect2.2 right has EV 3.55; pair 2.2 has EV 3.50. prospect2.3 left has EV 7.25; prospect2.3 right has EV 7.00; pair 2.3 has EV 7.13. prospect2.4 left has EV 0.75 ; prospect2.4 right has EV 1.00; pair 2.4 has EV 0.88 . prospect2.5 left has EV 5.00; prospect2.5 right has EV 5.00; pair 2.5 has EV 5.00. prospect2.6 left has EV 45.00; prospect2.6 right has EV 45.00; pair 2.6 has EV 45.00. EV page $2=(4.50+3.50+7.13+0.88+5.00+45.00) / 6=11.00$
prospect3.1 left has EV 2.39; prospect3.1 right has EV 2.75; pair 3.1 has EV 2.57. prospect3.2 left has EV 2.39; prospect3.2 right has EV 2.75; pair 3.2 has EV 2.57. prospect3.3 left has EV 2.39; prospect3.3 right has EV 2.75; pair 3.3 has EV 2.57. prospect3.4 left has EV 41.83; prospect3.4 right has EV 2.50; pair 3.4 has EV 22.16. prospect3.5 left has EV 41.83; prospect3.5 right has EV 2.50; pair 3.5 has EV 22.16. prospect3.6 left has EV 2.22; prospect3.6 right has EV 2.50; pair 3.6 has EV 2.36. EV page $3=(2.57+2.57+22.57+22.16+22.16+2.36) / 6=9.07$
prospect4.1 left has EV 4.94; prospect4.1 right has EV 4.57; pair 4.1 has EV 4.76. prospect4.2 left has EV 4.94; prospect4.2 right has EV 4.57; pair 4.2 has EV 4.76. prospect4.3 left has EV 4.94; prospect4.3 right has EV 4.56; pair 4.3 has EV 4.75.
prospect4.4 left has EV 2.19; prospect4.4 right has EV 2.19; pair 4.4 has EV 2.19. prospect4.5 left has EV 2.19; prospect4.5 right has EV 2.19; pair 4.5 has EV 2.19. prospect4.6 left has EV 2.19; prospect4.6 right has EV 2.19; pair 4.6 has EV 2.19. EV page $4=(4.76+4.76+4.75+2.19+2.19+2.19) / 6=3.47$
prospect5.1 left has EV 2.00; prospect5.1 right has EV 1.75; pair 5.1 has EV 1.88. prospect5.2 left has EV 2.00; prospect5.2 right has EV 1.75; pair 5.2 has EV 1.88. prospect5.3 left has EV 2.00; prospect5.3 right has EV 1.75; pair 5.3 has EV 1.88. prospect5.4 left has EV 7.25; prospect5.4 right has EV 7.00; pair 5.4 has EV 7.13. prospect5.5 left has EV 7.25; prospect5.5 right has EV 7.00; pair 5.5 has EV 7.13. prospect5.6 left has EV 7.25; prospect5.6 right has EV 7.00; pair 5.6 has EV 7.13. EV page $5=(1.88+1.88+1.88+7.13+7.13+7.13) / 6=4.50$
prospect6.1 left has EV 0.75; prospect6.1 right has EV 1.00; pair 6.1 has EV 0.88 . prospect6.2 left has EV 8.00; prospect6.2 right has EV 6.00; pair 6.2 has EV 7.00. prospect6.3 left has EV 1.50; prospect6.3 right has EV 2.00; pair 6.3 has EV 1.75. prospect6.4 left has EV 11.00; prospect6.4 right has EV 10.50; pair 6.4 has EV 10.75. prospect6.5 left has EV 20.00; prospect6.5 right has EV 18.00; pair 6.5 has EV 19.00. prospect6.6 left has EV 2.00; prospect6.6 right has EV 1.75; pair 6.6 has EV 1.88. EV page $6=(0.88+7.00+1.75+10.75+19.00+1.88) / 6=6.88$
prospect7.1 left has EV 3.40; prospect7.1 right has EV 10.20; pair 7.1 has EV 6.80. prospect7.2 left has EV 10.00; prospect7.2 right has EV 10.20; pair 7.2 has EV 10.10. prospect7.3 left has EV 3.40; prospect7.3 right has EV 10.20; pair 7.3 has EV 6.80. prospect 7.4 left has EV 10.00; prospect7.4 right has EV 10.20; pair 7.4 has EV 10.10. prospect7.5 left has EV 3.40; prospect7.5 right has EV 10.20; pair 7.5 has EV 6.80 prospect7.6 left has EV 10.00; prospect7.6 right has EV 10.20; pair 7.6 has EV 10.10. EV page $7=(6.80+10.10+6.80+10.10+6.80+10.10) / 6=8.45$
prospect8.1 left has EV -3.70; prospect8.1 right has EV 0.00; pair 8.1 has EV -1.85. prospect8.2 left has EV -3.50; prospect8.2 right has EV 0.00; pair 8.2 has EV -1.75. prospect8.3 left has EV -3.00; prospect8.3 right has EV 0.00; pair 8.3 has EV -1.50. prospect8.4 left has EV -2.00; prospect8.4 right has EV 0.00; pair 8.4 has EV -1.00. prospect8.5 left has EV -1.00; prospect8.5 right has EV 0.00; pair 8.5 has EV -0.50. prospect8.6 left has EV 320.00; prospect8.6 right has EV 300.00; pair 8.6 has EV 310.00 .
prospect9.1 left has EV 6.00; prospect9.1 right has EV 3.00; pair 9.1 has EV 4.50. prospect9.2 left has EV 5.00; prospect9.2 right has EV 4.00; pair 9.2 has EV 4.50. prospect9.3 left has EV 3.45; prospect9.3 right has EV 3.55; pair 9.3 has EV 3.50. prospect9.4 left has EV 2.22; prospect9.4 right has EV 2.50; pair 9.4 has EV 2.36. prospect9.5 left has EV 2.19; prospect9.5 right has EV 2.19; pair 9.5 has EV 2.19. prospect9.6 left has EV 80.40; prospect9.6 right has EV 75.00; pair 9.6 has EV 77.70. EV page $9=(4.50+4.50+3.50+2.36+2.19+77.70) / 6=15.79$
prospect10.1 left has EV 4.00; prospect10.1 right has EV 6.00; pair 10.1 has EV 5.00. prospect10.2 left has EV 5.00; prospect10.2 right has EV 5.00; pair 10.2 has EV 5.00. prospect10.3 left has EV 10.20; prospect10.3 right has EV 10.00; pair 10.3 has EV 10.10 .
prospect10.4 left has EV 5.00; prospect10.4 right has EV 5.00; pair 10.4 has EV 5.00. prospect 10.5 left has EV 2.39; prospect 10.5 right has EV 2.75; pair 10.5 has EV 2.57. prospect 10.6 left has EV 4.94; prospect 10.6 right has EV 4.57; pair 10.6 has EV 4.76. EV page $10=(5.00+5.00+10.10+5.00+2.57+4.76) / 6=5.40$

Overall EV $=(0.58+11.00+9.07+3.47+4.50+6.88+8.45+50.57+15.79+$ 5.40)/10 =

### 11.57

2. Which envelope contained which choice pair



Implementation with values of dice



Implementation with values of dice


Implementation with values of dice



Implementation with values of dice


Implementation with values of dice




Pair 7.1: env. 16
First, you receive $€ 15$. But, next, you cannot avoid and have to choose a risk of a loss.


Implementation with values of dice
Pair 7.2: env. 24
First, you receive $€ 15$. But, next, you risk a loss. You can avoid it by taking insurance.


Implementation with values of dice
Pair 7.3: env. 6
First, you receive $€ 15$. But, next, you cannot avoid and have to choose a risk of a loss.


Pair 7.4: env. 18
First, you receive $€ 15$. But, next, you risk a loss. You can avoid it by taking insurance.


Implementation with values of dice

Pair 7.5: env. 38
First, you receive $€ 15$. But, next, you cannot avoid and have to choose a risk of a loss.


Pair 7.6: env. 34
First, you receive $€ 15$. But, next, you risk a loss. You can avoid it by taking insurance.


Implementation with values of dice


Implementation with values of dice


Implementation with values of dice


Implementation with values of dice



## §ES. 13 of Electronic Supplementary Material:

## RIS explanations for subjects in Experiment 4

The following 6 pages present the explanations given to subjects for the RIS treatment of Experiment 4. The page numbers [1], 2-6 of those explanations have been kept.

## Explanation of Experiment

Welcome!

You will receive:
€5 for participation
+
an additional prize.

The additional prize comes from a prospect (explained later).
You will be asked to choose between prospects.

The average additional prize if you choose randomly: $€ 11.50$. But you will choose better than randomly!

There are some negative prizes (losses) that would be subtracted from the participation fee, but you can always choose to avoid those.

Please do not communicate with other participants.

There are no right or wrong answers.
Always choose what you want most. That (what you want) is also what we are interested in, and want to investigate.

## Explaining prospects

We have two 10-sided dice.
One can take the values: $\quad 0,1,2,3, \ldots, 9$
The other can take the values: $00,10,20,30, \ldots, 90$

You will throw both dice and their numbers will be added, leading to a range of 100 values,
from 0 up to 99.
Each value is equally likely.

Example of a prospect:

$€ 16$ with a probability of 0.31 :

- If value of dice $<31$, then you win $€ 16$
- If value of dice $\geq 31$, then you win nothing

The probability of the prize indeed is 0.31 .
If you get this prospect, we will use the figure

to specify in green which values of the throws of dice give which outcomes.

## Procedure

There is a questionnaire containing 10 pages numbered 1 .. 10 , each with 6 pairs of prospects, numbered $1.1 \ldots 1.6, \ldots$, 10.1 ... 10.6.

During the experiment, for each pair, you will indicate which prospect you prefer.

At the end of the experiment:

- First you throw a 10 -sided die to select a page.
- Next, given this page, you throw a 6-sided die to select a prospect pair from that page.
- From this pair, the prospect is chosen for which you indicated your preference.
- Finally, your prospect is played out as explained before.


## Recommendation

It is in your best interest to indicate all preferences truthfully. Then, from the prospect pair played for real, you get the prospect you want.

## Verification

At end of experiment:
you receive a list describing all 60 prospect pairs. You will then check that your prospect pair played for real is correct, i.e., as described in the list. You can then also check that the average additional prize indeed is $€ 11.50$.

Questions at any time: please raise your hand. Experimenters will come.

## Remember

There are no right or wrong answers; it is only about what you want.

If you reply truthfully,
you get the prospect you want.

Do not communicate with other participants.

## Now the experiment can begin.

For each pair of prospects in the questionnaire, mark the box above the prospect you prefer with an $x$, such as in the example below where you would prefer the prospect to the right.


## §ES. 14 of Electronic Supplementary Material:

## RIS stimuli for Experiment 4

The following 11 pages present the RIS stimuli of Experiment 4. The page numbers [0], 1-10 of those stimuli have been kept.

Overview of the 10 pages (reduced) of the questionnaire containing 6 prospect pairs each


The 10 pages follow next, nonreduced.



Below, questions will be repeated. That is, you will see the same pair three times, one after the other.


Below, questions will be repeated. That is, you will see the same pair three times, one after the other.








## §ES. 15 of Electronic Supplementary Material: Verification of lotteries and envelopes for RIS

## for Experiment 4

The following 3 pages present the pages given to subjects at the end of the RIS treatment of Experiment 4. They made it possible for subjects to verify that the experiment had been done honestly, with the envelopes having the proper content. The page numbers [1], 2, 3 of those pages have been kept.

## Verification document for RIS experiment

This document provides the expected value (EV) calculations of the additional prize (showup fee not included), which is $€ 11.50$ or, more precisely, $€ 11.57$.

## 1. Expected value calculation

prospect1.1 left has EV -1.15; prospect1.1 right has EV 0.00; pair 1.1 has EV -0.58. prospect1.2 left has EV -1.15; prospect1.2 right has EV 0.00; pair 1.2 has EV -0.35 . prospect1.3 left has EV -0.60; prospect1.3 right has EV 0.00; pair 1.3 has EV -0.30. prospect1.4 left has EV -0.40; prospect1.4 right has EV 0.00; pair 1.4 has EV -0.20. prospect1.5 left has EV -0.20; prospect1.5 right has EV 0.00; pair 1.5 has EV -0.10 . prospect1.6 left has EV 5.00; prospect1.6 right has EV 5.00; pair 1.6 has EV 5.00. EV page $1=(-0.58-0.35-0.30-0.20-0.10+5.00) / 6=0.58$
prospect2.1 left has EV 5.00; prospect2.1 right has EV 4.00; pair 2.1 has EV 4.50 . prospect2.2 left has EV 3.45; prospect2.2 right has EV 3.55; pair 2.2 has EV 3.50. prospect2.3 left has EV 7.25; prospect2.3 right has EV 7.00; pair 2.3 has EV 7.13. prospect2.4 left has EV 0.75; prospect2.4 right has EV 1.00; pair 2.4 has EV 0.88 . prospect2.5 left has EV 5.00; prospect2.5 right has EV 5.00; pair 2.5 has EV 5.00. prospect2.6 left has EV 45.00; prospect2.6 right has EV 45.00; pair 2.6 has EV 45.00. EV page $2=(4.50+3.50+7.13+0.88+5.00+45.00) / 6=11.00$
prospect3.1 left has EV 2.39; prospect3.1 right has EV 2.75; pair 3.1 has EV 2.57. prospect3.2 left has EV 2.39; prospect3.2 right has EV 2.75; pair 3.2 has EV 2.57. prospect3.3 left has EV 2.39; prospect3.3 right has EV 2.75; pair 3.3 has EV 2.57. prospect3.4 left has EV 41.83; prospect3.4 right has EV 2.50; pair 3.4 has EV 22.16. prospect3.5 left has EV 41.83; prospect3.5 right has EV 2.50; pair 3.5 has EV 22.16. prospect3.6 left has EV 2.22; prospect3.6 right has EV 2.50; pair 3.6 has EV 2.36. EV page $3=(2.57+2.57+22.57+22.16+22.16+2.36) / 6=9.07$
prospect4.1 left has EV 4.94; prospect4.1 right has EV 4.57; pair 4.1 has EV 4.76. prospect4.2 left has EV 4.94; prospect4.2 right has EV 4.57; pair 4.2 has EV 4.76. prospect4.3 left has EV 4.94; prospect4.3 right has EV 4.56; pair 4.3 has EV 4.75. prospect4.4 left has EV 2.19; prospect4.4 right has EV 2.19; pair 4.4 has EV 2.19. prospect4.5 left has EV 2.19; prospect4.5 right has EV 2.19; pair 4.5 has EV 2.19. prospect4.6 left has EV 2.19; prospect4.6 right has EV 2.19; pair 4.6 has EV 2.19. EV page $4=(4.76+4.76+4.75+2.19+2.19+2.19) / 6=3.47$
prospect5.1 left has EV 2.00; prospect5.1 right has EV 1.75; pair 5.1 has EV 1.88. prospect5.2 left has EV 2.00; prospect5.2 right has EV 1.75; pair 5.2 has EV 1.88. prospect5.3 left has EV 2.00; prospect5.3 right has EV 1.75; pair 5.3 has EV 1.88. prospect5.4 left has EV 7.25; prospect5.4 right has EV 7.00; pair 5.4 has EV 7.13. prospect5.5 left has EV 7.25; prospect5.5 right has EV 7.00; pair 5.5 has EV 7.13. prospect5.6 left has EV 7.25; prospect5.6 right has EV 7.00; pair 5.6 has EV 7.13. EV page $5=(1.88+1.88+1.88+7.13+7.13+7.13) / 6=4.50$
prospect6.1 left has EV 0.75; prospect6.1 right has EV 1.00; pair 6.1 has EV 0.88 . prospect6.2 left has EV 8.00; prospect6.2 right has EV 6.00; pair 6.2 has EV 7.00. prospect6.3 left has EV 1.50; prospect6.3 right has EV 2.00; pair 6.3 has EV 1.75. prospect6.4 left has EV 11.00; prospect6.4 right has EV 10.50; pair 6.4 has EV 10.75. prospect6.5 left has EV 20.00; prospect6.5 right has EV 18.00; pair 6.5 has EV 19.00. prospect6.6 left has EV 2.00; prospect6.6 right has EV 1.75; pair 6.6 has EV 1.88. EV page $6=(0.88+7.00+1.75+10.75+19.00+1.88) / 6=6.88$
prospect7.1 left has EV 3.40; prospect7.1 right has EV 10.20; pair 7.1 has EV 6.80. prospect7.2 left has EV 10.00; prospect7.2 right has EV 10.20; pair 7.2 has EV 10.10. prospect7.3 left has EV 3.40; prospect7.3 right has EV 10.20; pair 7.3 has EV 6.80. prospect 7.4 left has EV 10.00; prospect7.4 right has EV 10.20; pair 7.4 has EV 10.10. prospect7.5 left has EV 3.40; prospect7.5 right has EV 10.20; pair 7.5 has EV 6.80. prospect7.6 left has EV 10.00; prospect7.6 right has EV 10.20; pair 7.6 has EV 10.10. EV page $7=(6.80+10.10+6.80+10.10+6.80+10.10) / 6=8.45$
prospect8.1 left has EV -3.70; prospect8.1 right has EV 0.00; pair 8.1 has EV -1.85. prospect8.2 left has EV -3.50; prospect8.2 right has EV 0.00; pair 8.2 has EV -1.75. prospect8.3 left has EV -3.00; prospect8.3 right has EV 0.00; pair 8.3 has EV -1.50. prospect8.4 left has EV -2.00; prospect8.4 right has EV 0.00; pair 8.4 has EV -1.00. prospect8.5 left has EV -1.00; prospect8.5 right has EV 0.00; pair 8.5 has EV -0.50 . prospect8.6 left has EV 320.00; prospect8.6 right has EV 300.00; pair 8.6 has EV 310.00 .

EV page $8=(-1.85-1.75-1.50-1.00-0.50+310.00) / 6=50.57$
prospect9.1 left has EV 6.00; prospect9.1 right has EV 3.00; pair 9.1 has EV 4.50. prospect9.2 left has EV 5.00; prospect9.2 right has EV 4.00; pair 9.2 has EV 4.50. prospect9.3 left has EV 3.45; prospect9.3 right has EV 3.55; pair 9.3 has EV 3.50.
prospect9.4 left has EV 2.22; prospect9.4 right has EV 2.50; pair 9.4 has EV 2.36. prospect9.5 left has EV 2.19; prospect9.5 right has EV 2.19; pair 9.5 has EV 2.19. prospect9.6 left has EV 80.40; prospect9.6 right has EV 75.00; pair 9.6 has EV 77.70. EV page $9=(4.50+4.50+3.50+2.36+2.19+77.70) / 6=15.79$
prospect10.1 left has EV 4.00; prospect10.1 right has EV 6.00; pair 10.1 has EV 5.00. prospect 10.2 left has EV 5.00; prospect 10.2 right has EV 5.00; pair 10.2 has EV 5.00. prospect10.3 left has EV 10.20; prospect10.3 right has EV 10.00; pair 10.3 has EV 10.10 .
prospect10.4 left has EV 5.00; prospect10.4 right has EV 5.00; pair 10.4 has EV 5.00. prospect10.5 left has EV 2.39; prospect10.5 right has EV 2.75; pair 10.5 has EV 2.57. prospect10.6 left has EV 4.94; prospect10.6 right has EV 4.57; pair 10.6 has EV 4.76. EV page $10=(5.00+5.00+10.10+5.00+2.57+4.76) / 6=5.40$

Overall EV $=(0.58+11.00+9.07+3.47+4.50+6.88+8.45+50.57+15.79+$
5.40)/10 =

### 11.57


[^0]:    ${ }^{1}$ The formulation about the implementation in Figure 1 may have encouraged subjects to report the next unit (subjectively chosen by themselves) above their indifference value, because at their indifference point the case is immaterial to them. Hence, it could have led to an upward bias due to indifference. The reformulation in Figure 1' above avoids this bias:
    "If $x$ in my envelope falls below (is worse than) my indifference value, then give me Option 1 (mug). If $x$ in my envelope exceeds (is better than) my indifference value, then give me Option 2 ( $€ x$ ). If $x$ in my envelope equals my indifference value, then toss a coin to decide whether I am given Option 1 (mug) or 2 ( $€ \mathrm{x}$ ).

[^1]:    ${ }^{2}$ The subjects filled out all answers twice: once on the page containing the question, and once on the front page numbered TOj.0.
    ${ }^{3}$ In the case of indifference, subjects could choose on the spot, as always if the experimental answers did not specify a choice.

[^2]:    ${ }^{4}$ For experiments in labs, each subject, upon receiving payment at the end, received a list describing the content of all envelopes, and was asked to verify correctness of content of the own envelope. The list also contained maximal and minimal amounts and calculations giving expected values under random play.

[^3]:    ${ }^{5}$ The x stimuli contain an outcome 0 , at which utility is $-\infty$ for $\rho \leq 0$. Because this outcome cancels from the equations, we still allow for power $\rho \leq 0$ in our data fitting. Only five subjects had $\rho \leq 0$ (minimum -0.75), and they do not affect any result.

[^4]:    ${ }^{6}$ Their medians happen to be the same (0.88), due to different skewnesses. Note that the standard deviations have the same unit (1 per $€$ ) and, hence, can be compared.

[^5]:    ${ }^{7}$ We know only one introduction of a new measurement method in decision theory that immediately included a comparison with another method (Attema et al. 2016), because the other method (Epper, Fehr-Duda, \& Bruhin 2011) could be taken as best at that moment. Andreoni \& Sprenger (2012), Choi et al. (2007), and Holt \& Laury (2002) did not provide a comparative test. Andreoni, Kuhn, \& Sprenger (2015) dedicated a full separate paper to comparing Andreoni \& Sprenger’s (2012) new measurement method for intertemporal choice with an alternative. Their comparison is of limited interest to readers who would have preferred another competing alternative.

[^6]:    ${ }^{8}$ They could also do that later at ease at their home. Especially if the experimenters would had frauded for an envelope with a big prize, subjects could then be expected to return to the issue and protest.

[^7]:    ${ }^{9}$ Agranov \& Ortoleva paid in points where 1 point is $1 \$ / 20$. We immediately pay in money amounts ( $€$ instead of \$) to maintain uniformity of our stimuli. The numbers get more complex this way, making choices even harder, which can be expected to enhance randomization. The italicized text on top of Pages 3 and 4 in our stimuli is taken from their online appendix (Appendix B Part 3), adapted to our context.

[^8]:    ${ }^{10}$ In these tests we follow Cox, Sadiraj, \& Schmidt (2015) and Starmer \& Sugden (1991) in directionally testing for the common directions of violations of expected uitility. That is, we took the average of the number of safe choices in pair F6.2 minus the number of safe choices in pair F6.3 as test statistic for the common ratio effect. For the common consequence effect, we used a similar statistic. Had we looked at violations of expected utility in both directions, then the difference between Prince and RIS would not have been significant in both the common ratio and the common consequence tests that we did.
    ${ }^{11}$ For RIS, the violation of expected utility is significant ( $p=0.004$ ). For Prince, it is not $(p=0.37)$.

