

Consistent inconsistencies? Evidence from decision under risk

Guillaume Hollard¹ · Hela Maafi² ·
Jean-Christophe Vergnaud³

Published online: 30 October 2015
© Springer Science+Business Media New York 2015

Abstract Conventional economic theory assumes that agents should be consistent across decisions. However, it is often observed that experimental subjects fail to report consistent preferences. So far, these inconsistencies are almost always examined singly. We thus wonder whether the more inconsistent individuals in one task are also more inconsistent in other tasks. We propose an experiment in which subjects are asked to report their preferences over risky bets so as to obtain, for each subject, three measures of inconsistencies: classical preference reversals, framing effects and preference instability. In line with previous experimental findings, subjects are largely inconsistent according to each of these three measures and there are considerable individual differences. The main result is that we find no correlation among these three measures of inconsistency.

Keywords Heterogeneity · Consistency · Preference reversals · Framing effects

Electronic supplementary material The online version of this article (doi:[10.1007/s11238-015-9518-8](https://doi.org/10.1007/s11238-015-9518-8)) contains supplementary material, which is available to authorized users.

✉ Hela Maafi
hmaafi@univ-paris8.fr

Guillaume Hollard
guillaume.hollard@polytechnique.edu

Jean-Christophe Vergnaud
vergnaud@univ-paris1.fr

¹ Département d'économie de l'école polytechnique, CNRS, 91128 Palaiseau Cedex, France

² LED, Université Paris 8 (Saint-Denis), 2 Rue de la Liberté, 93200 Saint-Denis, France

³ Centre d'Économie de la Sorbonne, CNRS, 106/112 Boulevard de l'Hôpital, 75647 Paris Cedex 13, France

JEL Classification D81 · D03 · C91

1 Introduction

Conventional economic theory is based on the assumption that agents should exhibit consistency across choices. However, it is often observed that experimental subjects are prone to preference inconsistency: elicited preferences are not independent from the elicitation procedure. Yet, in a typical experiment, a substantial fraction of the subjects appear to be rather consistent and exhibit only low rates of preference inconsistency. Should we thus generalize a bit and consider the possibility that some individuals are immune to inconsistency in general (i.e. not just in one particular situation)? It is, for instance, suggested that individuals who have a higher taste for abstract reasoning, greater cognitive capacity or longer market experience may be more consistent (Ameriks et al. 2003; Burks et al. 2009; Frederick 2005; List 2003). In other words, are individuals consistently consistent or consistently inconsistent? To the best of our knowledge this is an open question since experimental studies usually deal with only one specific type of inconsistency at a time. We here propose to test whether different causes of inconsistency can have similar effects on the same individuals.

Our focus is on decisions made under risk. We here consider specific types of inconsistency that are considered as violations of the invariance principle. The invariance principle states that subjects should reveal identical preferences in two theoretically equivalent decision problems. As such, two theoretically equivalent versions of a decision problem should elicit the same preferences. Empirically, violations are frequent when two equivalent decision problems are presented in a different way or at a different point of time.

We here consider two experimental tasks known to generate substantial, systematic and replicable violations of the invariance principle. The first task is based on Lichtenstein and Slovic's (1971) original preference-reversal experiment: preferences between a safe lottery and a risky lottery are elicited using two different mechanisms, namely binary choice and valuation (or pricing). The second task is a "framing effect" task inspired by the Asian disease problem (Tversky and Kahneman 1981): choices are presented either in a loss frame or a gain frame. Subjects have to choose between a risky lottery and a sure amount of money. Following De Martino et al. (2006), the sure amount is described as an amount "to keep" in the gain frame, while it is described as an amount "to lose" in the loss frame.

Subjects perform the two experimental tasks successively. The first task provides a measure of the propensity to display classical preference reversals. Subjects are inconsistent when the preferences inferred from binary choice are different from those inferred from valuation. The second task measures the individual's sensitivity to framing effect. In the general case, subjects are sensitive to the frame when they prefer the sure amount in the gain frame but prefer the risky lottery in the loss frame. The frequency of inconsistent choices between the loss and gain frame is our second measure of inconsistency. In the second task, some choices are performed twice within a few minutes' interval. The frequency of preference instability is our third measure of inconsistency.

For each individual, our protocol thus elicits three distinct measures of inconsistency: classical preference reversals, framing effects and preference instability. An individual who displays consistent consistency or consistent inconsistency will get a similar score according to each measure. To illustrate, consider first expected-utility subjects (EU-subject) who are asked to exhibit their preferences over options of a set of paired-lotteries. Note that the options of a given paired-lottery in our protocol have similar expected value. An extreme risk-averse (seeking) EU-subject will prefer the safest (riskiest) option in all tasks. She will appear to be consistent in all three tasks and get a very low score (around 0 %) on each of the three measures of inconsistency. A risk-neutral expected-utility subject will be indifferent between the two proposed options of a given paired-lottery. She will consistently report a high score (around 50 %) for each measure of inconsistency. Moderate risk aversion attitude will lie between these two extreme. By computing pairwise correlation for our three measures of inconsistency, we are thus able to detect the presence of consistently consistent expected-utility individuals and consistently inconsistent expected-utility individuals.

For non expected-utility individuals, we can also expect some correlation between our three measures of inconsistency since they all represent a failure of the same principle of rationality, namely the invariance principle. However, for non-EU subjects, there is no unified model that allows us to predict the behavior simultaneously in the three tasks. For instance, prospect theory, in its third generation (Schmidt et al. 2008), does propose both an explanation for classical preference reversals and traditional framing effect. But third-generation prospect theory remains silent regarding preference instability as well as the counter-framing effect. The main issue with non-expected utility models is that they are related to some specific tasks (eg. prospect theory to loss aversion and behavior in gain vs losses, scale compatibility on valuations, error models on preference instability). An important step towards a unified model that explains sensitivity to the context would be to know whether the different types of inconsistency studied so far are related or not.

We find (1) average amounts of inconsistency which are in line with past experimental findings, (2) a large heterogeneity between subjects regarding each measure of inconsistency and (3) no within-subject correlation between our three measures of inconsistency: classical preference reversals, framing effects and preference instability. Our results suggest that there is nothing like a single characteristic measuring individual sensitivity to preference inconsistency.

The remainder of the paper is organized as follows: Section 2 presents the three measures of inconsistency and Sect. 3 presents the experimental design. The results are presented in Sects. 4 and 5. Section 6 discusses our results and concludes.

2 Our three measures of inconsistencies

In this paper we focus on three types of inconsistencies that were extensively examined and replicated in the literature: framing effect, classical preference reversals and preference instability. We describe in the following these inconsistencies and present some experimental evidence that will serve as a benchmark for our analysis.

2.1 Framing effect

The seminal paper by [Tversky and Kahneman \(1981\)](#) demonstrated that framing one question in terms of a gain or a loss results in different attitudes toward risk. Decision makers generally exhibit risk aversion when a question is framed in terms of a gain but risk seeking when this same question is framed in terms of a loss. For example, a subject which is endowed with €100, prefers the safe option of “keeping €50 of the €100” to the risky option of “gambling the €100 with a 50 % chance of winning” (positive frame). On the contrary, the same subject endowed with €100, prefers the risky option of “gambling the €100 with a 50 % chance of winning” to the safe option of “losing €50 of the €100” (negative frame). Using variations of this example, [De Martino et al. \(2006\)](#) observed that subjects were risk averse in the gain frame, tending to choose the sure option over the gamble option (the proportion of gambling in the gain frame was 42.9 %), and were risk seeking in the loss frame, tending to choose the gamble option over the sure option (the proportion of gambling in the loss frame was 61.6 %). Thus, the proportion of gambling increases by about 20 % points when we move from a gain frame to a loss frame, which shows that elicited preferences are significantly affected by the framing manipulation. The term “framing effect” was then extended to encompass sensitivity to the description of choice problems that are otherwise logically equivalent. [Levin et al. \(1998\)](#) review a long list of related experiments.

2.2 Classical preference reversals

[Lichtenstein and Slovic \(1971\)](#) show that preferences cannot be defined independently of the elicitation mechanism used. A typical preference-reversal experiment involves two risky lotteries: one lottery featuring a high probability of winning a small amount of money, called the probability bet or “P-bet”; and another lottery featuring a low probability of winning a large amount of money, called the dollar bet or “\$-bet”.¹ The two bets have roughly the same expected value. Subjects are asked to reveal their preferences over the two bets using two different, but mathematically-equivalent and incentive-compatible, elicitation mechanisms: *binary choice* and *pricing*. Preference reversals occur when choice is inconsistent with announced prices. Many pieces of experimental evidence (see [Seidl 2002](#); [Berg et al. 2010](#)) have confirmed that preferences elicited using different mechanisms do not always coincide. Preference reversals were also found to be systematic: subjects generally prefer the P-bet in binary choice, but reverse their preference and prefer the \$-bet in valuation. This is known as “standard preference reversals”. Subjects less frequently prefer the \$-bet to the P-bet in the binary-choice task and report a higher price for the P-bet than for the \$-bet; this is known as “non-standard preference reversals”. [Grether and Plott \(1979\)](#) observed that

¹ To illustrate, consider the following bets (source [Lichtenstein and Slovic 1971](#), Table 3): \$-bet = (\$16, 11/36; -\$1.50, 25/36) and P-bet = (\$4, 35/36; -\$1, 1/36). Here, the \$-bet offers an 11/36 chance of winning \$16 and a 25/36 chance of losing \$1.50, while the P-bet offers a 35/36 chance of winning \$4 and a 1/36 chance of losing \$1. Both bets have an expected value of approximately \$3.85.

subjects reversed their preferences 33 % of the time. The rate of standard preference reversals (69 %) outnumbers that of non-standard preference reversals (13 %).²

2.3 Preference instability

Experimental evidence shows that decision makers often fail to report the same preferences when they have the opportunity to perform exactly the same task more than once. In a typical experiment, subjects face the same pairwise choice problem twice. This repetition is separated by a short period of time during which the subjects carry out some other choice problems. In these experiments, the rate of unstable choices lies between 15 and 30 %. For instance, in [Hey and Orme \(1994\)](#), subjects perform four sets of the same 25 pairs of questions. They found an average rate of instability of around 25 %. [Camerer \(1989\)](#) observed that 31.6 % of subjects reverse their preferences when they answer the same binary choice question twice. [Starmer and Sugden \(1989\)](#) found a proportion of unstable subjects of between 25.8 and 28.3 % according to the questions involved. In [Ballinger and Wilcox \(1997\)](#), the median estimate of the rate of instability was around 20.8 %. These empirical findings are often interpreted as evidence that preferences are noisy or entail some random component. For instance, we may consider that preferences changed because of some unobserved changes in the context, such as the emotional state of the decision maker.

3 Experimental design

The experiment was divided into two parts. In the first part, subjects responded to preference-reversals task. In the second part, they responded to framing-effect task.

3.1 Participants

We recruited 41 (25 males and 16 females) subjects at the University of Paris 1 (France). Volunteers ranged in age between 18 and 56 years with a mean of age of 23. The majority of participants (85 %) were students, 37 % of whom were majoring in Economics. We ran three sessions, and no subject participated in more than one session.

3.2 Organization

As the subjects arrived, they were asked to sit at separate individual computers. At the start of the experiment, subjects received written instructions that were read aloud by the experimenter. Subjects were told that they will make a number of decision problems and that they will be given €5 for their participation. Afterwards, the experimenter described in detail the nature of gambles considered in the first part and the

² Cf experiment 1 (with incentives) where 91 choices out of 273 were inconsistent with the announced selling prices (overall preference reversals), 69 out of the 99 choices of the P-bet were inconsistent with the announced selling prices (standard preference reversals) and 22 out of the 174 choices of the \$-bet were inconsistent with the selling prices (non-standard preference reversals).

Table 1 Organization of the experiment

| | Organization of the experiment | Nature of the task |
|--------------|---|-------------------------------|
| Part 1 | Instruction for preference reversals task | |
| | Preference reversals task | Binary choices and valuations |
| Part 2 | Instruction for framing effect task | |
| | Framing effect task | Binary choices |
| Payment part | Play for real one question of Part 1 | Binary choice or valuation |
| | Play for real one question of Part 2 | Binary choice |

way participants were to be paid. Participants were told that they will perform some supplementary questions (Part 2) once they finish this part and that the corresponding instructions would be described afterwards. They were also told that once the experiment was started, they were not allowed to communicate with each other, but they were allowed to ask the experimenter questions in private. The experiment began when subjects had no more questions. Once all participants had finished the first part, the experimenter distributed the written instructions for the second part and described out loud the gambles involved and the payment mechanism. After responding to all questions individually, the experimenter announced the beginning of second part of the experiment. Table 1 summarizes the different phases of the experiment. The experiment was computerized using software developed under REGATE (Zeiliger 2000).

3.3 The preference-reversal task

In the first part of the experiment, we examined the rate of preference reversals using six paired-lotteries that are similar to those used in Grether and Plott (1979). The gambles used here (see Table 2) do not involve losses, and probabilities were described to subjects via an urn containing 100 balls. Urns contain winning and losing balls, the proportions of which are objectively known. The winning (losing) balls were red (black) colored and refer to the positive non-zero (zero) outcome. To illustrate, consider pair I in Table 2. Here, the P-bet offers €5 if a winning ball is drawn from an urn containing 80 winning balls and 20 losing balls, and the \$-bet offers €20 if a winning ball is drawn from an urn containing 20 winning balls and 80 losing balls. The paired gambles have similar expected value.


In this part, subjects were first asked to specify their minimum willingness to accept (WTA) for the 12 bets (six P-bets and six \$-bets) using the Becker–DeGroot–Marschak (1964) (BDM) mechanism (see Fig. 1 for an illustration). This mechanism is widely used in the preference-reversal literature. Subjects were then asked to choose between the \$-bet and its corresponding P-bet for the six pairs of lotteries in Table 2 (see Fig. 2 for an illustration). Note that we did not control for the order of tasks here as the elicitation of prices before or after binary choice does not seem to affect the pattern of reversals. Grether and Plott (1979) have explicitly tested this issue, and note that choice patterns and reversal rates appear to be the same for choices made before and after the elicitation of selling prices (p. 632).

Table 2 Preference-reversal paired-lotteries

| Pairs | Type | Probability of winning | Amount to be won (€) | EV |
|-------|------|------------------------|----------------------|------|
| I | \$ | 20/100 | 20 | 4 |
| | P | 80/100 | 5 | 4 |
| II | \$ | 30/100 | 16 | 4.8 |
| | P | 90/100 | 5 | 4.5 |
| III | \$ | 20/100 | 10 | 2 |
| | P | 90/100 | 4 | 3.6 |
| IV | \$ | 40/100 | 6 | 2.4 |
| | P | 90/100 | 3 | 2.7 |
| V | \$ | 50/100 | 12 | 6 |
| | P | 90/100 | 6 | 5.4 |
| VI | \$ | 25/100 | 27 | 6.75 |
| | P | 70/100 | 8 | 5.6 |

Give your minimum selling price for a ticket that offers you **20 Euros** if you draw a **winning** ball from this urn.

Your Price: Euros



20 Winning
80 Losing

100 balls


- The buying price will be picked randomly between €0,1 and €20
 - Your price > buying price: you don't sell (you play the ticket)
 - Your price ≤ buying price: you receive the buying price

OK

Fig. 1 Illustration of the valuation task (the \$-bet of pair I in Table 2)

Choose between ticket A and ticket B


Ticket A: 20 Euros if you draw a winning ball from Urn A below



20 Winning
80 Losing

100 balls

Ticket B: 5 Euros if you draw a winning ball from Urn B below



80 Winning
20 Losing

100 balls

OK

Fig. 2 Illustration of the binary-choice task (paired-lottery I in Table 2)

At the very end of the experiment (i.e. after part 2), the computer randomly chose one task from this part (WTA or binary choice) and one question, picked randomly from that task, was played for real. For example, if round 1 was chosen, then the subject plays the BDM mechanism. The computer then randomly chose one question, say the \$-bet of pair I. Afterwards, an offer between €0.1 and €20 was chosen randomly. If the random offer exceeded the expressed WTA, the participant received the random offer. If the random offer was below the expressed WTA, the subject played the ticket. In the latter case, the computer drew a ball from the urn, and the subject won €20 if the drawn ball was a winning ball and nothing otherwise. In the BDM mechanism, the maximum offer for a given bet $L = (\text{€}x, p)$, offering € x with probability p , was € x .

3.4 The framing effect task

In the second part of the experiment, we examined sensitivity to framing effects by replicating De Martino et al.'s (2006) design. We chose this protocol because it nicely captures a typical framing effect by simply changing the formulation of the safe option. This part of the experiment involves three rounds composed of 32 binary-choice questions each: 16 in the loss frame and 16 in the gain frame. Subjects were not aware that they were moving from one round to the other. Round 1 and round 3 use the same set of questions (see Table 3).

In each of the questions in rounds 1 and 3, the expected values of the sure and gamble options are always equivalent and are mathematically equivalent between frames (see set 1 in Table 4). The expected values were, however, unbalanced in round 2 (see set 2 in Table 4), in order to (1) check that subjects do not respond to questions randomly, and (2) separate rounds 1 and 3 which are similar. Subjects were not told that rounds 1 and 3 were the same, but were simply informed that they were to make 96 binary choices. In each round, the 32 corresponding questions were presented to subjects randomly. Subjects were given 5 s to respond to each question. Subjects were told that if they fail to answer within the allocated time (5 s), they will receive nothing if the unanswered question were to be selected for payment.

The framing task was constructed as follows: At the start of each question, subjects were shown a message with their initial endowment, "You receive € X ". A new screen then appeared and subjects were told that they have to choose between a "sure option" and a "gamble offering € X with probability p ". The sure option was framed as the "amount to keep" in the gain frame and as the "amount to lose" in the loss frame. As in De Martino et al. (2006), in each round we used four different initial endowments: (€10, €20, €30 and €40), and four probabilities of winning: 20, 40, 60 and 80 %. The resulting lotteries are listed in Table 4.

Table 3 Experiment: description of Part 2 of the experiment

| Round | Task |
|-------|---------------------------|
| 1 | 32 binary choices (set 1) |
| 2 | 32 binary choices (set 2) |
| 3 | 32 binary choices (set 1) |

Table 4 Framing-effect gambles

| | X | p (%) | Set 1 | | Set 2 | |
|-------------------------------------|-----|---------|-------|-------|-------|-------|
| | | | S_G | S_L | S_G | S_L |
| | 10 | 20 | 2 | 8 | 1 | 7 |
| | 10 | 40 | 4 | 6 | 6 | 8 |
| | 10 | 60 | 6 | 4 | 8 | 6 |
| | 10 | 80 | 8 | 2 | 7 | 1 |
| | 20 | 20 | 4 | 16 | 6 | 18 |
| | 20 | 40 | 8 | 12 | 7 | 11 |
| | 20 | 60 | 12 | 8 | 11 | 7 |
| | 20 | 80 | 16 | 4 | 18 | 6 |
| | 30 | 20 | 6 | 24 | 4 | 20 |
| | 30 | 40 | 12 | 18 | 16 | 22 |
| | 30 | 60 | 18 | 12 | 16 | 10 |
| | 30 | 80 | 24 | 6 | 28 | 10 |
| | 40 | 20 | 8 | 32 | 12 | 36 |
| | 40 | 40 | 16 | 24 | 14 | 22 |
| X initial endowment, p the | 40 | 60 | 24 | 16 | 28 | 20 |
| winning probability, S_G the sure | 40 | 80 | 32 | 8 | 30 | 6 |
| gain and S_L the sure loss | | | | | | |

To illustrate, consider the second question of set 1 in Table 4. Here, subjects were endowed with €10 and were asked to choose between: (i) “keeping €4” or “gambling the €10 with a 40 % chance of winning” in the gain frame, and (ii) “losing €6” or “gambling the €10 with a 40 % chance of winning” in the loss frame (see Fig. 3). Note that options (i) and (ii) are identical: for both frames, the sure amount to be won is the same (only the frame changes), and the gamble option is identical. The gamble options were represented by a pie-chart depicting the winning probability.

At the end of the experiment, after playing one question for real from part 1, the computer randomly chose one question from part 2 to be played for real. If the subject had chosen the sure option, then she won “the amount to keep” in the gain frame, or “the initial endowment – the amount to lose” in the loss frame. If the subject had chosen the gamble option, then a spinning black disk appears and she was asked to click on it to determine her gain. The subject won the initial endowment if she clicked on the green zone and nothing if she clicked on the red zone. Subjects earned an average of €18.78 (median gain = €18).

4 Descriptive results

The purpose of this section is twofold. First, we present and construct our three measures of inconsistency. Second, we check whether our results are in line with previous evidence. We observe that the pattern of inconsistencies, their rates and the individual heterogeneity for each measure are in line with previous findings.

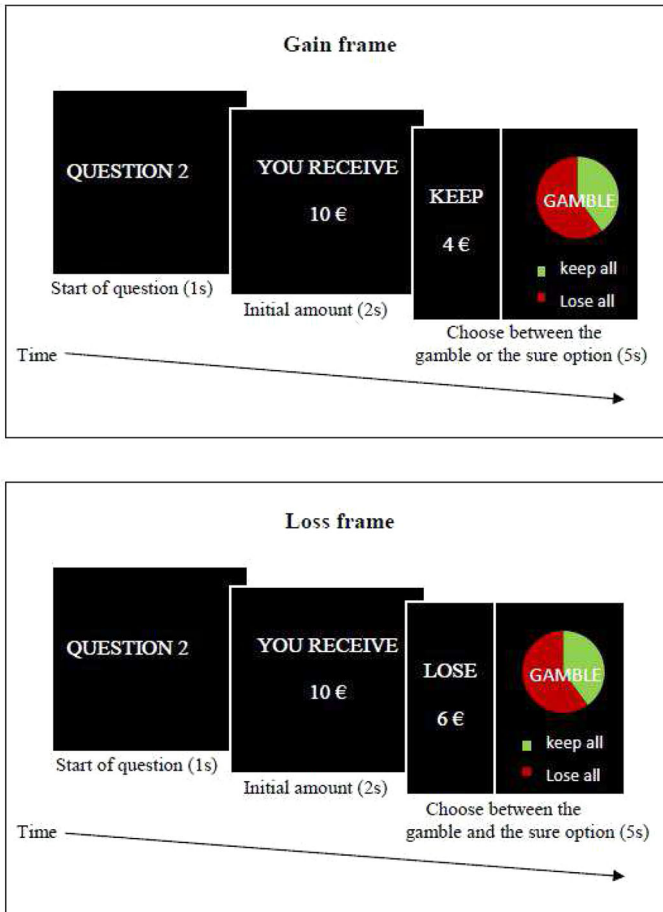


Fig. 3 Illustration of the framing effects tasks (question 2 of set 1 in Table 4)

4.1 Classical preference reversals

For a given paired lottery of type $L^k : P^k = (x_P^k, p_P^k)$ and $\$^k = (x_S^k, p_S^k)$,³ a decision maker exhibits preference reversals if her choice is inconsistent with her announced prices (this is called “overall preference reversals”). Decision makers reverse their preferences in a systematic way: standard preference reversals (i.e., subjects prefer the P-bet in the binary-choice task and report a higher price for the $\$$ -bet in the valuation task, $P \succ \$$, and $V(\$) > V(P)$) outnumber non-standard preference reversals (i.e. subjects prefer the $\$$ -bet to the P-bet in the binary-choice task and report a higher price for the P-bet than for the $\$$ -bet, $\$ \succ P$ and $V(P) > V(\$)$).

³ Lottery $P^k = (x_P^k, p_P^k)$ offers $\text{€}x_P^k$ with probability p_P^k , and lottery $\$^k = (x_S^k, p_S^k)$ offers $\text{€}x_S^k$ with probability p_S^k .

Table 5 Frequencies of classical preference reversals

| Bet | Choices | Selling prices | | |
|--------------------|---------|----------------|--------------|-------|
| | | Consistent | Inconsistent | Equal |
| P | 153 | 48 | 94 | 11 |
| \$ | 92 | 75 | 13 | 4 |
| Total ($N = 41$) | 245 | 123 | 107 | 15 |

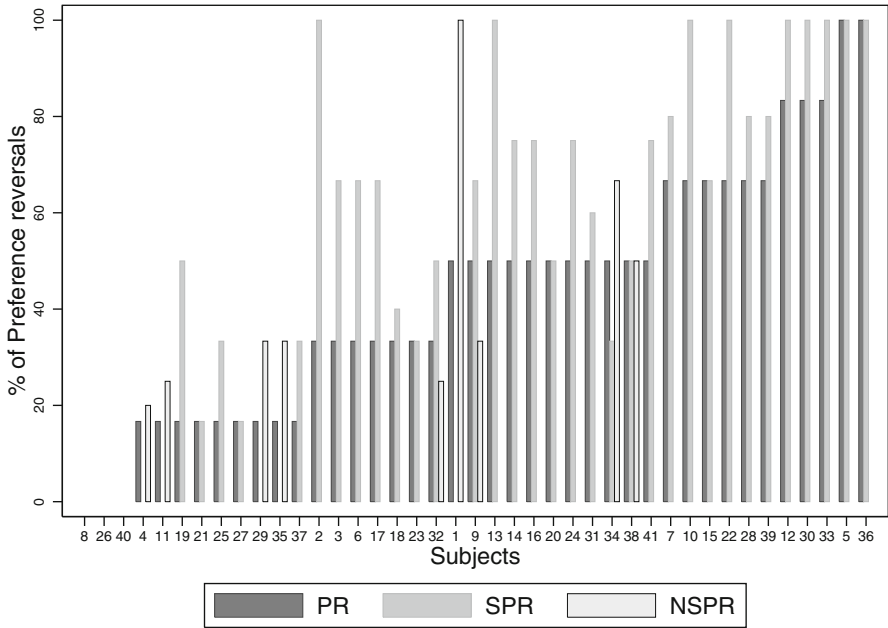


Fig. 4 Individual proportions of preference reversals

Table 5 shows the results for choice and valuation in the six paired-lotteries. The overall rate of preference reversals is 43 % (107 choices out of 245 were inconsistent with the announced selling prices). The rate of standard preference reversals is 61 % (of the 153 choices of P-bets, 94 were inconsistent with the announced selling price) and that of non-standard preference reversals is 14 % (13 of the 92 choices of \$-bets were inconsistent with the announced selling prices). These rates are consistent with those in previous work (see Seidl 2002; Berg et al. 2010).

For each subject, we computed the overall proportion of reversals (overall PR), the proportion of standard preference reversals (SPR) and the proportion of non-standard preference reversals (NSPR). Almost all participants exhibited sensitivity to the elicitation mechanism. Figure 4 shows that subjects vary in their degree of preference reversals. The proportions of overall preference reversals, standard preference reversals and non-standard preference reversals range between 0 and 100 %. There is notable heterogeneity between subjects in the extent of preference reversals, which shows that subjects are different: there are those who react strongly to the use of different elicitation mechanisms, while others do not.

Table 6 Testing risk aversion in the gain frame and risk seeking in the loss frame: H_0 : % age of gambling = 50 % (binomial test)

| | No. of obs | Gain | | Loss | |
|----------------|------------|-------------|----------------|-------------|-----------------|
| | | % of gamble | <i>p</i> value | % of gamble | <i>p</i> -value |
| Round 1 | 642 | 45.48 | 0.0244 | 54.82 | 0.016 |
| Round 3 | 652 | 39.97 | 0.0000 | 44.32 | 0.0042 |
| Rounds 1 and 3 | 1294 | 42.58 | 0.0000 | 49.53 | 0.7598 |

The number of observations is different across sets because subjects did not report their preferences in some cases as they have 5 s to respond to each question. During the instruction phase, they were told that if the question picked randomly for payment was not performed, they gain nothing

Table 7 Comparing the % age of gambling in the gain frame and in the loss frame (McNemar's Chi-square test)

| | No. of observations | Gain | Loss | Exact McNemar <i>p</i> value |
|----------------|---------------------|-------|-------|------------------------------|
| Round 1 | 642 | 45.48 | 54.82 | 0.0000 |
| Round 3 | 652 | 39.97 | 44.32 | 0.0154 |
| Rounds 1 and 3 | 1294 | 42.58 | 49.53 | 0.0000 |

4.2 Framing effects

In what follows, we focus on preferences in round 1 and round 3, where the sure amount is equal to the expected value of the corresponding gamble.⁴ For a given question in Table 4, we constructed four dummies reflecting the subjects' decisions: (1) G_{sure} = choosing the sure option in the gain frame; (2) G_{gamble} = choosing the gamble option in the gain frame; (3) L_{sure} = choosing the sure option in the loss frame; and (4) L_{gamble} = choosing the gamble option in the loss frame. We define risk aversion and risk seeking with respect to risk neutrality where subjects choose the gamble in 50 % of cases. We find that the frame of the questions affects attitude towards risk. The observed behavior is partially consistent with prospect theory. In the gain frame, subjects were risk averse, tending to choose the sure option over the gamble option. Table 6 shows that the proportion of gambling in the gain frame is significantly less than 50 %. Nevertheless, subjects were risk-seeking in the loss frame only in round 1, where the proportion of gambling is significantly greater than 50 %.⁵ Besides, we observe from Table 7 that the proportion of gambling in the loss frame is significantly higher than that in the gain frame.

For each subject, we computed a "framing effect measure", henceforward FE, that captures her sensitivity to the frame effect. This measure is the difference between the

⁴ Following De Martino et al. (2006), we use the questions of set 2 to ensure that subjects remain engaged in the tasks and do not respond randomly throughout this part and also to separate rounds 1 and 3. Some descriptive results for set 2 are presented in the "Appendix".

⁵ Consistent with De Martino et al's (2006) results, a framing effect was also found across the four probabilities and the four starting amounts used here (see Figs. 15, 16, 17, 18, 19, 20 in the "Appendix").

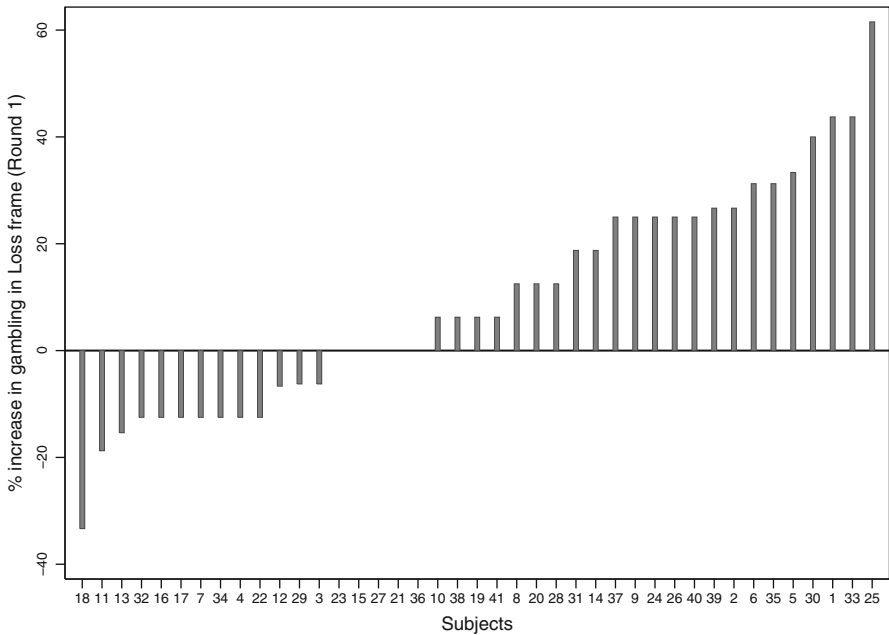


Fig. 5 Individual proportions of framing effects (round 1)

proportion of questions in which the subject chose the gamble option in the loss frame as compared to the gain frame. For each subject in a given round, we compared 16 choices in the gain frame to the corresponding 16 choices in the loss frame. Formally, for subject i in round j :

$$FE_j^i = \text{Mean}(L_{\text{gamble}} - G_{\text{gamble}}) \tag{1}$$

For subject i , FE_j^i is based on the number of paired-choices in the considered round j (16 when we consider rounds 1 and 3 separately and 32 when we aggregate rounds 1 and 3).

Figure 5 shows the distribution of the FE across individuals in round 1 (see Figs. 11 and 12 for round 3 and aggregated results). We first observe that subjects differ in their sensitivity to the frame manipulation. As for preference reversals, there is notable heterogeneity between subjects in the degree of the framing effect. We second see that some subjects exhibit an overall counter-framing effect.

In line with the prediction of prospect theory, and with De Martino et al.’s results, around 70 % of subjects gamble more in the loss frame than in the gain frame, thereby exhibiting a traditional framing effect. However, the fact that the remaining 30 % of subjects exhibit a counter-framing effect simply reminds us of the importance of subject heterogeneity in these sorts of data.⁶

⁶ We find that 13 subjects out of 41 exhibit a counter-framing effect in round 1, 12 out of 41 in round 3, and 14 out of 41 when we consider all of the observations in set 1 (rounds 1 and 3 together).

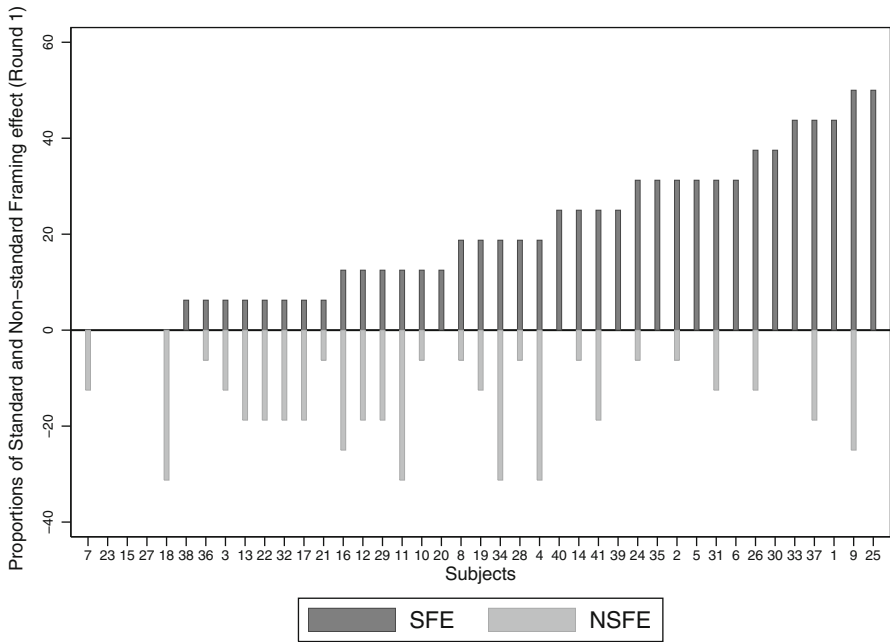


Fig. 6 Individual proportions of standard and non-standard framing effects (round 1)

For each subject, we also computed the proportion of framing effect as predicted by prospect theory, which we call “standard framing effect” (SFE), and the proportion of counter-framing effect, called “non-standard framing effect” (NSFE). Standard framing effect measures the frequency of decisions in accordance with the traditional framing effect. Specifically, for subject i in round j , SFE_j^i is the proportion of pairs for which the gamble option was chosen in the loss frame and the safe option in the gain frame. It is computed as follows:

$$SFE_j^i = \text{Mean}[I(L_{\text{gamble}} > G_{\text{gamble}})] \tag{2}$$

with $I(\cdot)$ is the indicator function such that $I(x > y) = 1$ if $x > y$ and $I(x > y) = 0$ otherwise.

Non-standard framing effect measures the frequency of decisions that are counter to the traditional framing effect. For subject i in round j , $NSFE_j^i$ is the proportion of pairs for which the safe option was chosen in the loss frame and the gamble option in the gain frame⁷:

$$NSFE_j^i = \text{Mean}[I(L_{\text{gamble}} < G_{\text{gamble}})] \tag{3}$$

Figure 6 shows the individual proportions of standard and non-standard framing effects in round 1 (see Figs. 13 and 14 for round 3 and aggregated results). We observe that subjects differ markedly in their level of standard and non-standard framing effects:

⁷ For subject i , SFE_j^i and $NSFE_j^i$ are based on the number of paired-choices in round j (16 when we consider rounds 1 and 3 separately and 32 when we aggregate rounds 1 and 3).

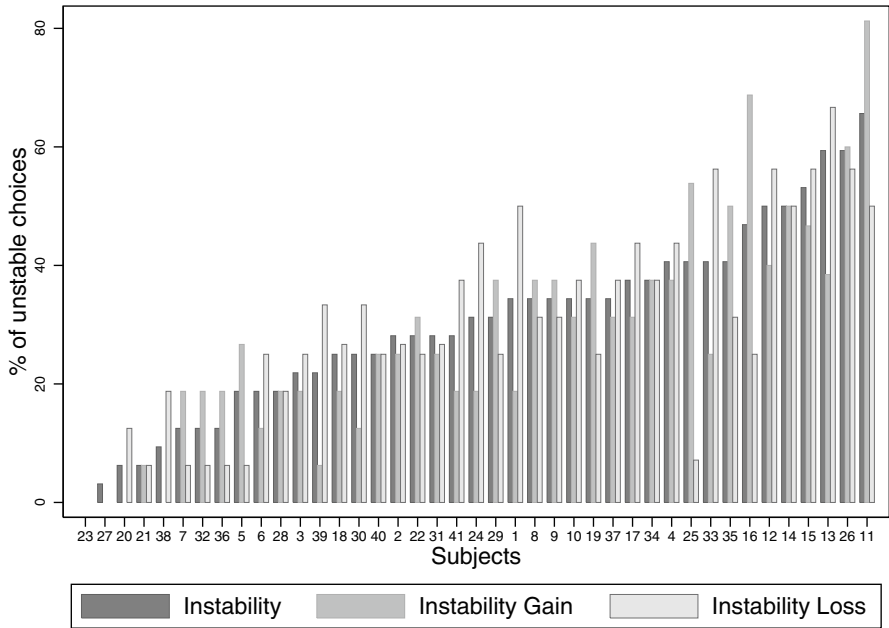


Fig. 7 Individual proportions of instability

the proportion of standard framing effect ranges between 0 and 50 % and that of non-standard framing effect between 0 and 31.25 %.⁸

4.3 Preference instability

In Part 2 of our experiment, subjects perform the binary-choice task with the 32 paired-options of set 1 twice, with the set of questions being repeated under exactly the same conditions. By this repetition, we can examine the stability of preferences over time, using the 16 paired-options in the gain frame and the 16 paired-options in the loss frame.

Subjects are observed to reverse their preferences in 30 % of cases. This result is consistent with previous findings (Camerer 1989; Starmer and Sugden 1989; Hey and Orme 1994; Ballinger and Wilcox 1997; Loomes et al. 2002). Specifically, the rate of instability is 28.75 % in the gain frame and 29.92 % in the loss frame. The rate of instability in the gain frame is not significantly different from that in the loss frame (*t*-test, degree of freedom = 1291, *t* = -0.5574). For each subject, we computed the proportion of unstable choices for the 32 paired-options taken together: for the 16 paired-options in the gain frame, and for the 16 paired options in the loss frame. Figure 7 shows that subjects differ notably in their level of instability.

⁸ Although 21 % of subjects are not consistent with the predictions of prospect theory, the proportion of standard framing effect is significantly higher than that of non-standard framing effect at better than the 1 % level in set 1. We obtain the same result if we consider the observations in rounds 1 and 3 as separate or not.

Table 8 Correlation between observed inconsistencies (Spearman correlations)

| | Instability | FE _{rounds1&3} | FE _{round1} | FE _{round3} |
|----------------------|----------------|-----------------------------|----------------------|----------------------|
| FE ₁₃ | −0.0434 | | | |
| FE ₁ | −0.0458 | 0.8513*** | | |
| FE _{round3} | −0.0334 | 0.7641*** | 0.3559** | |
| PR | −0.1096 | 0.0375 | 0.0573 | − 0.0006 |

Bold values indicate the absence of correlation discussed in the paper

(i) FE₁₃ = FE in rounds 1 and 3, FE₁ = FE in round 1 and FE₃ = FE in round 3

(ii) Levels of significance * 10 %, ** 5 %, *** 1 %

This heterogeneity between subjects shows that some subjects are more unstable than others in revealing their preferences. Figure 7 shows also the considerable correlation between the rates of instability in the gain frame and in the loss frame (Spearman's $\rho = 0.8142$, $p < 0.01$).

5 Consistency across tasks: results and interpretations

We examined the correlation between our three measures of inconsistency: the proportions of preference reversals (PR), framing effects (FE) and instability (see “Appendix 1”). Table 8 shows that there is no significant correlation between these three measures. Subjects who perform well in preference-reversal tasks are not significantly less sensitive to the framing effect and are not more stable in revealing their preferences. The absence of correlation between the three measures of inconsistency suggests a heterogeneity of behavior *within* subjects.

We can be a little more precise regarding the relation between SPR and SFE. As it often happens in such experiment, some subjects exhibit non-standard preference reversals (NSPR) in task 1 or non-standard framing effects (NSFE) in task 2. Since the measure of overall preference reversals and that of the overall framing effect encompass the non-standard inconsistencies, we may attribute the absence of correlation to the difference in the measurement of overall inconsistencies. Consequently, examining the correlation between standard (non-standard) inconsistencies is more compelling. We thus examined whether there is a positive correlation between standard inconsistencies (standard preference reversals and standard framing effect) on the one hand, and non-standard inconsistencies on the other hand (non-standard preference reversals and non-standard framing effect). Table 9 lists the Spearman correlations between standard and non-standard inconsistencies. We find no significant correlation between standard preference reversals and standard framing effect, nor between non-standard preference reversals and non-standard framing effect. However, we do see that instability is positively and significantly correlated with standard framing effect and non-standard framing effect. We also note that measures in the table that are obtained by use of *valuation* (SPR and NSPR) are not correlated with measures based on binary *choices* (e.g. all framing effects and instability measures).

Table 9 Correlation between standard inconsistencies, non-standard inconsistencies and instability (Spearman correlations)

| | Instability | SFE1 | NSFE1 | SFE13 | NSFE13 | SFE3 | NSFE3 | SPR |
|--------|---------------|---------------|---------------|---------------|---------------|--------------|--------------|-----------|
| SFE1 | 0.319* | | | | | | | |
| NSFE1 | 0.396** | -0.190 | | | | | | |
| SFE13 | 0.379** | 0.899*** | -0.083 | | | | | |
| NSFE13 | 0.571*** | -0.109 | 0.859*** | -0.043 | | | | |
| SFE3 | 0.347** | 0.543*** | 0.084 | 0.836*** | 0.072 | | | |
| NSFE3 | 0.547*** | 0.021 | 0.198 | -0.006 | 0.638*** | -0.028 | | |
| SPR | -0.155 | -0.105 | -0.153 | -0.046 | -0.106 | 0.052 | 0.027 | |
| NSPR | 0.141 | 0.091 | 0.229 | 0.102 | 0.204 | 0.120 | 0.089 | -0.545*** |

Bold values indicate the absence of correlation discussed in the paper

(i) SFE_{13} = SFE in rounds 1 and 3, SFE_1 = SFE in round 1, SFE_3 = SFE in round 3, $NSFE_{13}$ = NSFE in rounds 1 and 3, $NSFE_1$ = NSFE in round 1 and $NSFE_3$ = NSFE in round 3

(ii) Levels of significance * 10 %, ** 5 %, *** 1 %

It is important to note that the correlations which have been found to be significant in other experiments, like those between different measures of framing, are still significant. This is important because we may otherwise have thought that the absence of correlation was due to a lack of statistical power. In the same vein, we may wonder whether the absence of correlation between our measures of consistency is due to some non-linear relation between the different measures. However, the Spearman correlation coefficient allows us to reject the possibility of any monotonically relationship.

6 Conclusion

This paper has explored whether different types of inconsistency are related. We define and measure three types of inconsistency for each decision-maker. The degree of each type of inconsistency varies greatly across individuals. But, we find no significant correlation between our three measures. This rules out the possibility of some individuals being more or less consistent than others in absolute terms. If something like market experience or cognitive skills were to partly explain inconsistency, we would observe some correlation across our three inconsistency measures. This is not the case. The hope of capturing the propensity to exhibit preference inconsistencies, using a simple individual characteristic, falls short of empirical support. This empirical finding also has some theoretical implications. Different models have been built to explain different types of preference inconsistencies. For instance, prospect theory relates to framing effects, while models with an error term are more suitable to explain preference instability. In the light of the presented results, building a unified model of preference inconsistencies appears even more challenging than initially thought. Absent such a model, having different explanations for different types of inconsistencies is almost unavoidable.

Appendix 1: Correlation: supporting figures

See Figs. 8, 9 and 10.

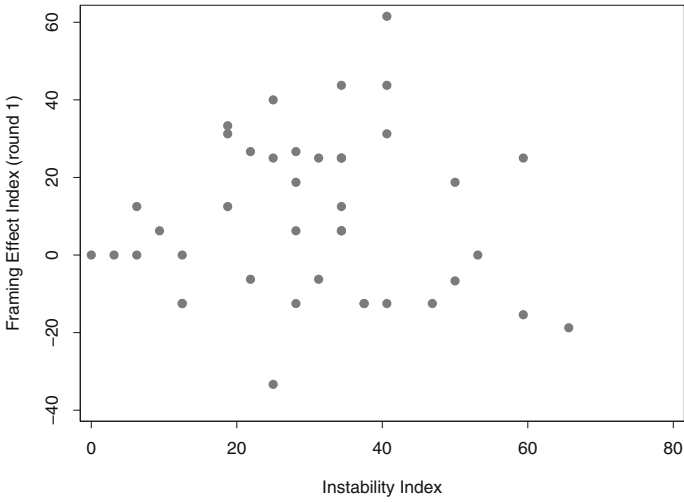


Fig. 8 FE₁ vs instability (individual data)

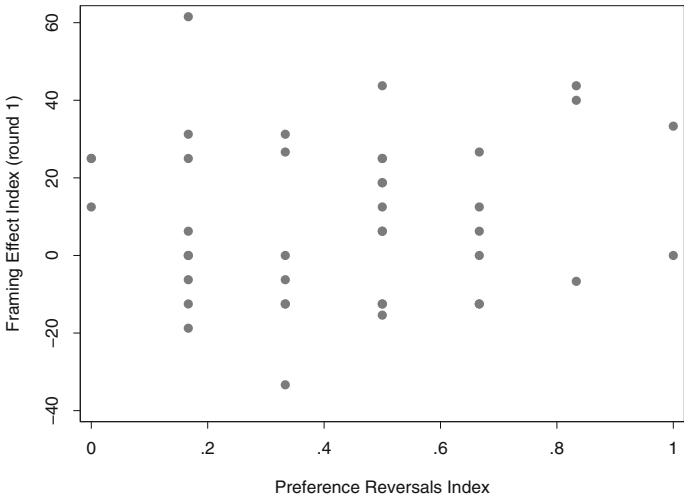


Fig. 9 FE₁ vs PR (individual data)

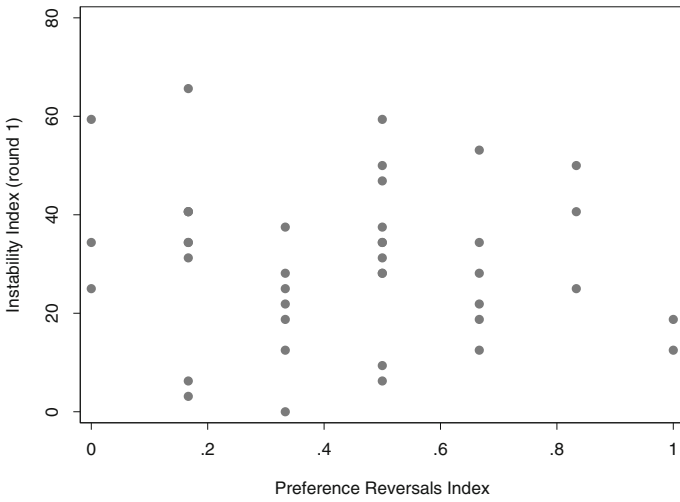


Fig. 10 Instability vs PR (individual data)

Appendix 2: Framing effect: supporting figures

See Figs. 11, 12, 13 and 14.

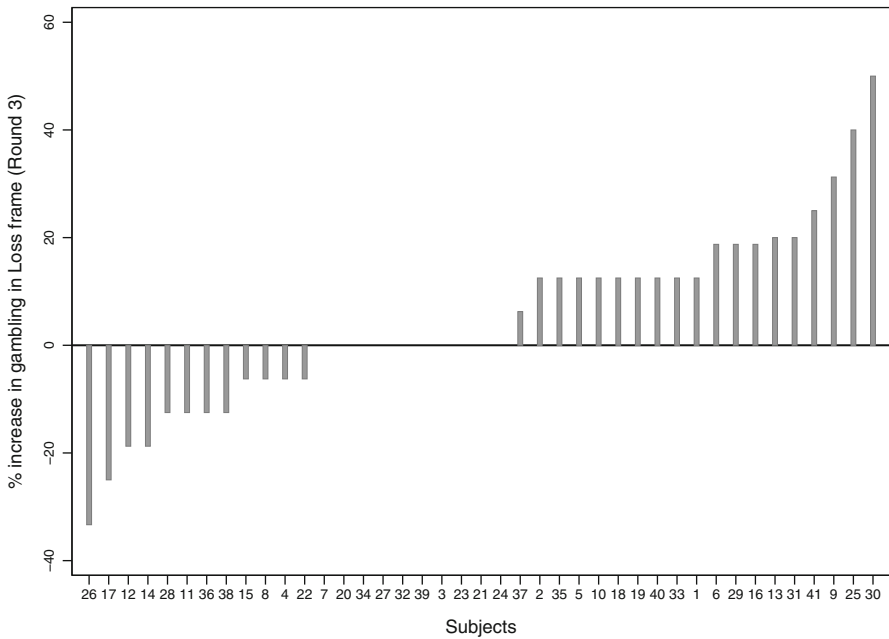


Fig. 11 Individual proportions of framing effects (round 3)

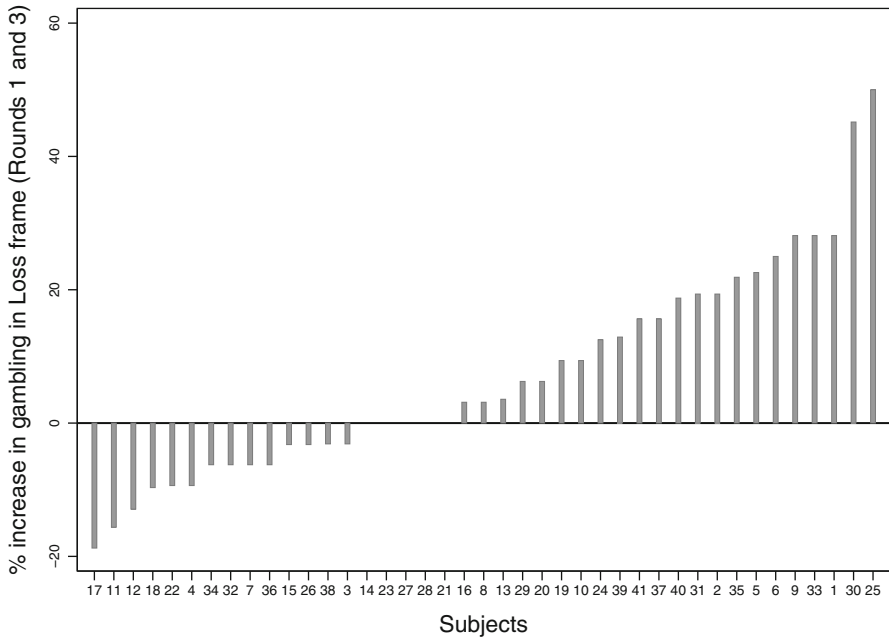


Fig. 12 Individual proportions of framing effects (rounds 1 and 3)

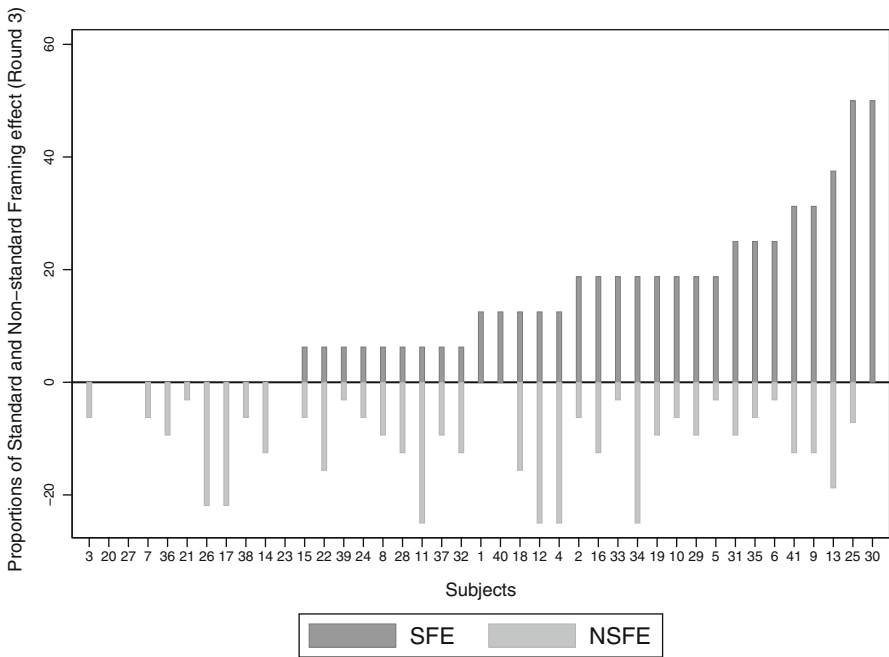


Fig. 13 Individual proportions of standard and non-standard framing effects (round 3)

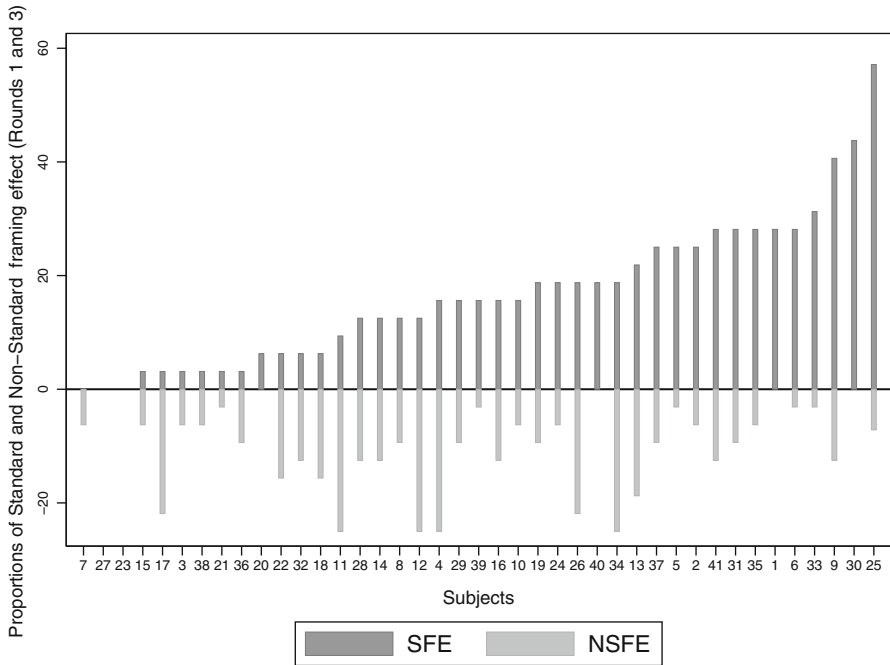


Fig. 14 Individual proportions of standard and non-standard framing effects (rounds 1 and 3)

Appendix 3: Framing effect: varying probability of winning and initial endowment

- Figures 15, 16 and 17 show decisions across varying probabilities: the bars show the percentages (%) of trials in which subjects chose the gamble option in the gain frame (dark bar) and in the loss frame (lighter bar), for the four winning probabilities in the gamble option (20, 40, 60 and 80 %). Consistent with De Martino et al. (2006), we observe that subjects are more risk averse in the gain as compared to the loss frame for the four probabilities of winning.
- Figures 18, 19 and 20 show the decisions by the amount at stake: the bars show the percentages (%) of trials in which subjects chose the gamble option in the gain frame (dark bar) and in the loss frame (lighter bar), for the four amounts at stake in the gamble option (€10, €20, €30 and €40). As in De Martino et al. (2006), risk attitude is affected by the framing of questions for the different endowments, especially for high payoffs.

Appendix 4: Framing effect: results in round 2

In round 2 (33 % of all the trials), the expected outcomes of the sure and the gamble option were unbalanced. We used these “catch trials” (1) to ensure that subjects remain

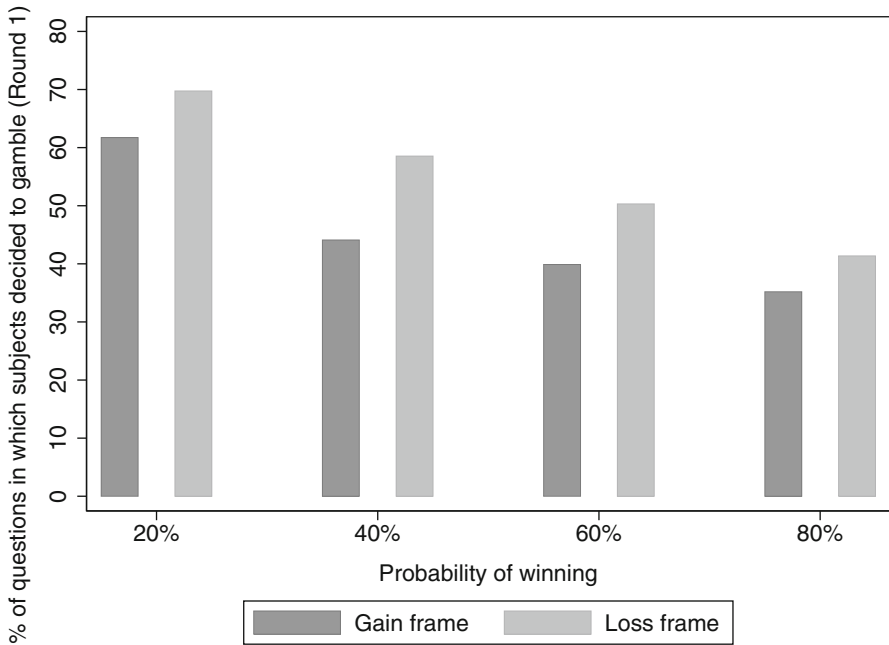


Fig. 15 The proportion gambling by winning probability (round 1)

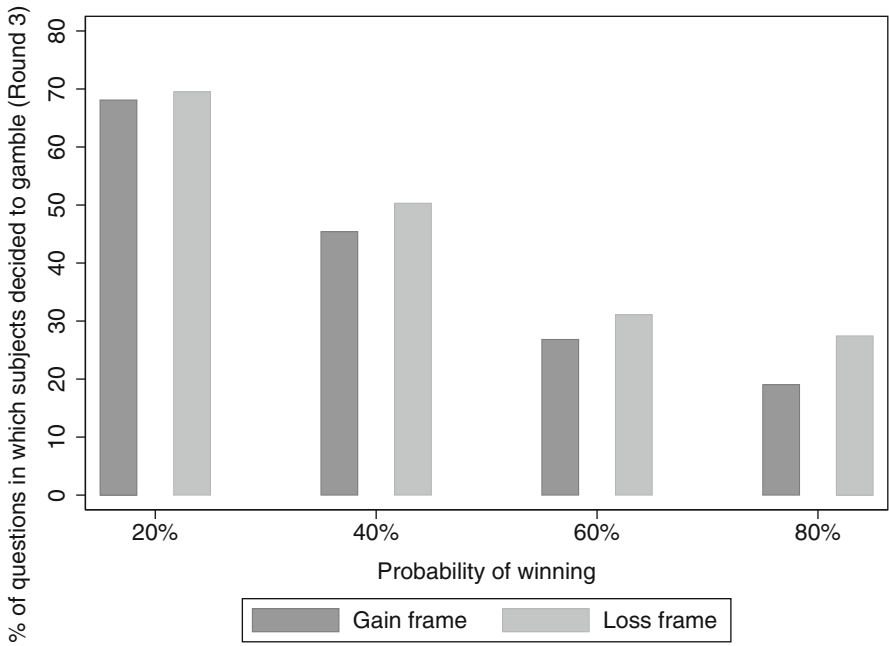


Fig. 16 The proportion gambling by winning probability (round 3)

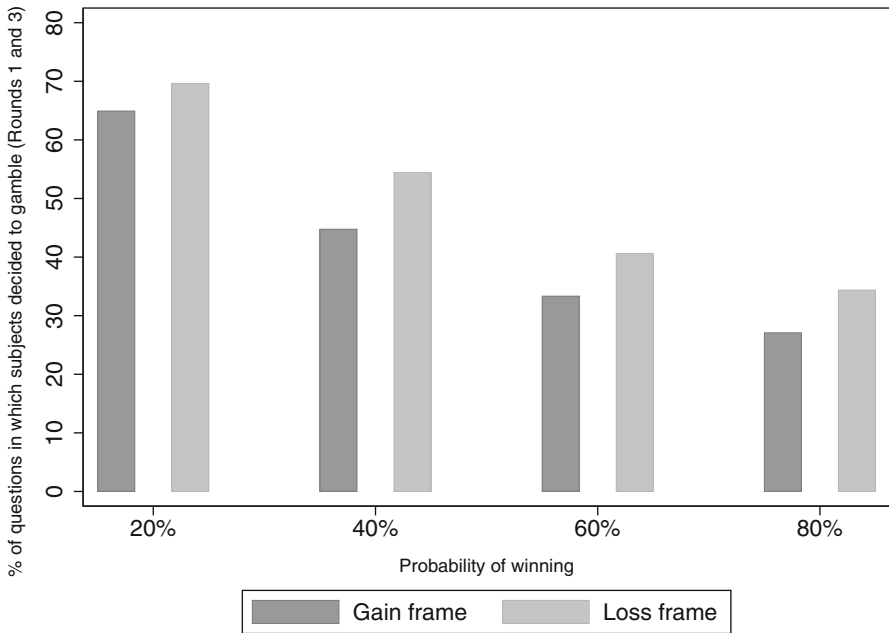


Fig. 17 The proportion gambling by winning probability (rounds 1 and 3)

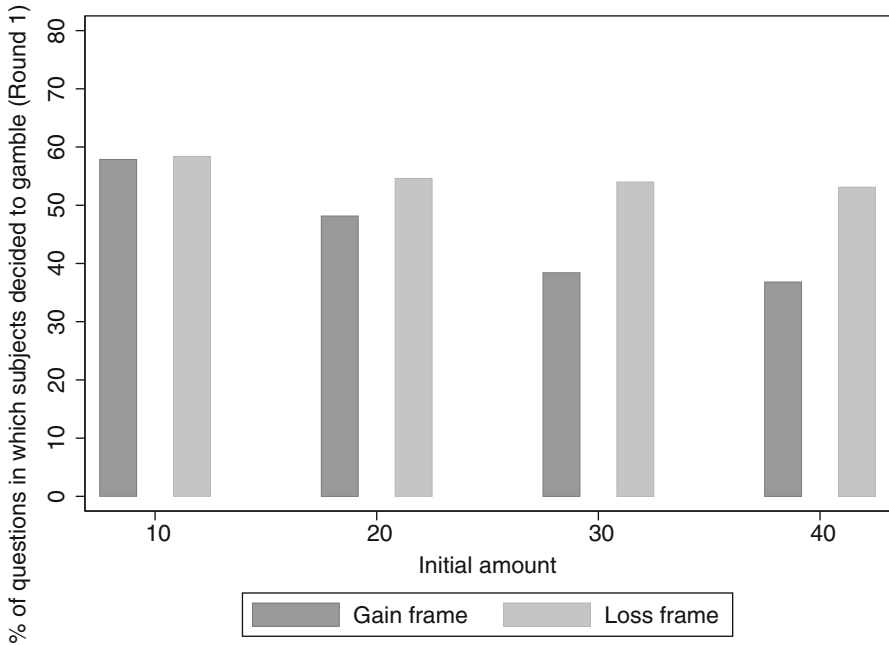


Fig. 18 The proportion gambling by the initial endowment (round 1)

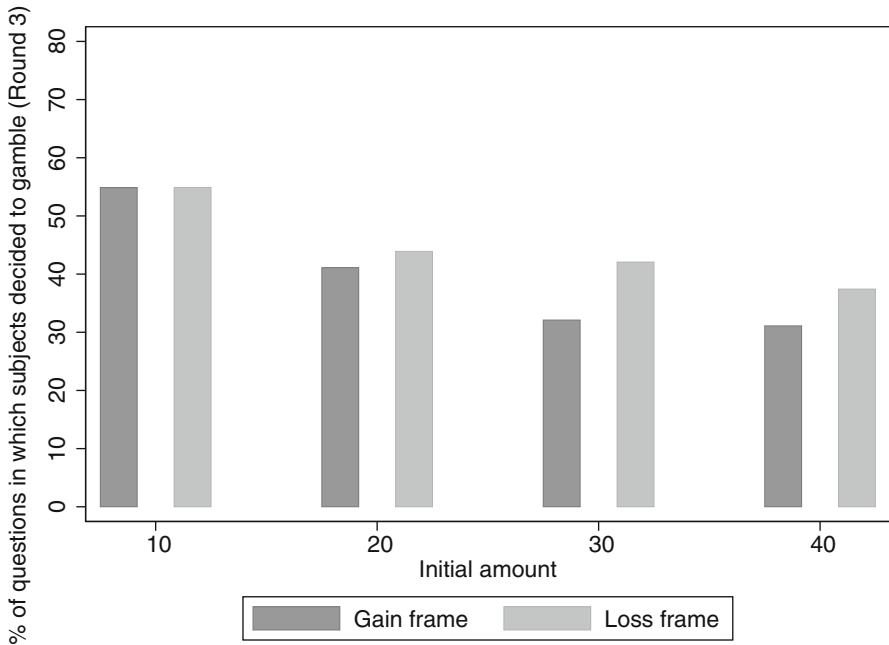


Fig. 19 The proportion gambling by the initial endowment (round 3)

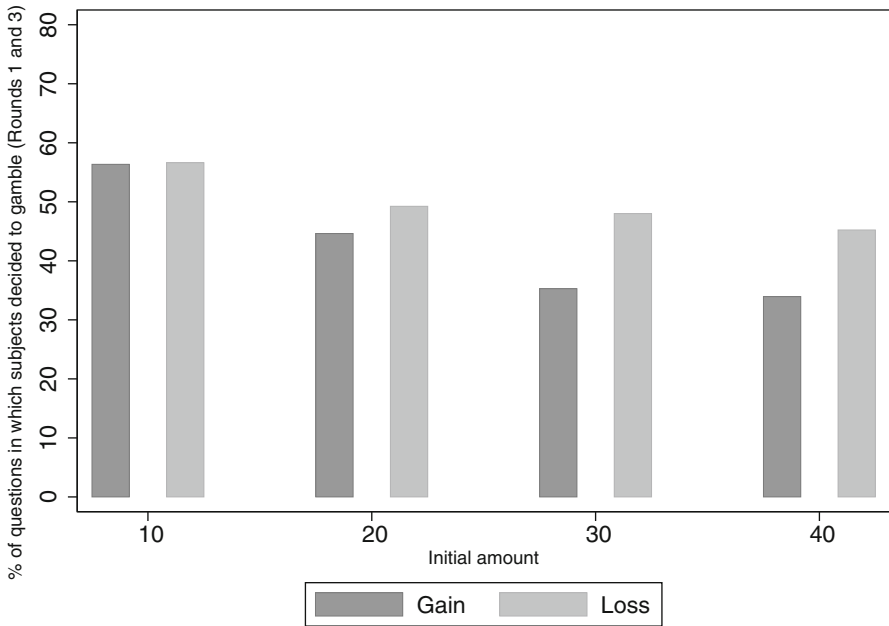


Fig. 20 The proportion gambling by the initial endowment (rounds 1 and 3)

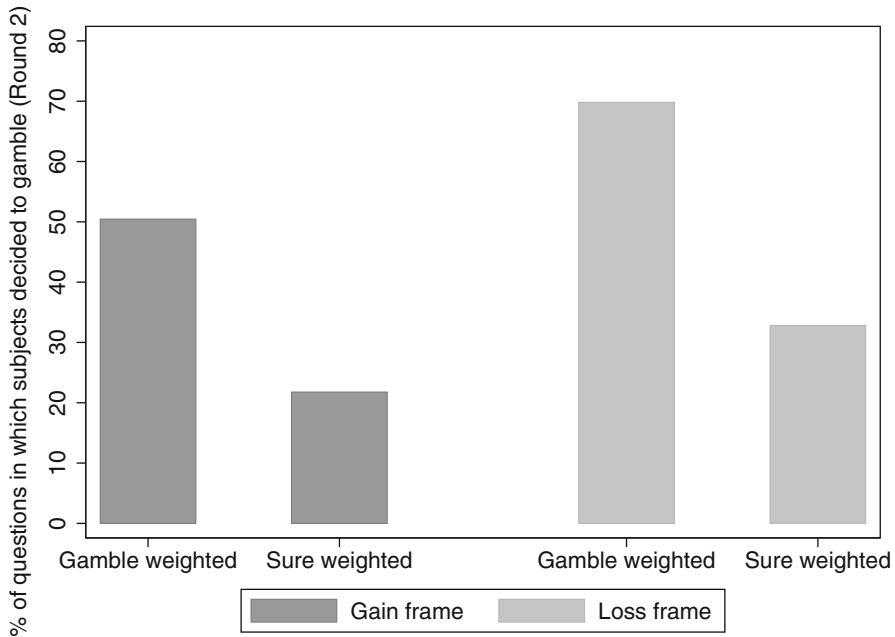


Fig. 21 The proportion gambling by the frame and the attractiveness of the sure and gamble options

engaged in the experiment (as in De Martino et al. 2006), and (2) to separate the two sets of questions where the sure and the gamble options had the same expected outcome. For both frames (gain and loss), we used two types of catch trials: the “gamble weighting” where the gamble option is preferable to the sure option, and the “sure weighting” where the sure option is preferable to the gamble option. We varied the attractiveness of the sure and gamble options in both frames (see Table 4) to examine the accuracy of optimal decisions according to the attractiveness of the options.

Figure 21 shows the proportion of gambling in the gain frame (dark bar) and in the loss frame (lighter bar). Subjects were accurate in making optimal choices, by generally gambling more (less) when the gamble option was more (less) favorable than the sure option. More precisely, we constructed an “attractiveness index” as the difference between the expected outcomes of the gamble and the sure option and found that the proportion of gambling is positively and significantly correlated with the attractiveness of the gamble option over the sure option (gain frame: $\rho = 0.2799$, $p < 0.01$. Loss frame: $\rho = 0.3324$, $p < 0.01$).

References

- Ameriks, J., Caplin, A., & Leahy, J. (2003). Wealth accumulation and the propensity to plan. *The Quarterly Journal of Economics*, 118(3), 1007–1047.
- Ballinger, T. P., & Wilcox, N. T. (1997). Decisions, error and heterogeneity. *Economic Journal*, 107(443), 1090–1105.

- Becker, G. M., Degroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, *9*(3), 226–232.
- Berg, J. E., Dickhaut, J. W., & Rietz, T. A. (2010). Preference reversals: The impact of truth-revealing monetary incentives. *Games and Economic Behavior*, *68*(2), 443–468.
- Burks, S. V., Carpenter, J. P., Goette, L., & Rustichini, A. (2009). Cognitive skills explain economic preferences, strategic behavior, and job attachment. *PNAS*, *106*(19), 7745–7750.
- Camerer, C. F. (1989). An experimental test of several generalized utility theories. *Journal of Risk and Uncertainty*, *2*(1), 61–104.
- De Martino, B., Kumaran, D., Seymour, B., & Dolan Raymond, J. (2006). Frames, biases, and rational decision-making in the human brain. *Science*, *313*(5787), 684–687.
- Fredrick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42.
- Grether, D. M., & Plott, C. R. (1979). Economic theory of choice and the preference reversal phenomenon. *American Economic Review*, *69*(2), 623–638.
- Hey, J., & Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica*, *62*(6), 1291–1326.
- Levin, I. P., Schneider, S. L., & Gaeth, G. J. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes*, *76*(2), 149–188.
- Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology*, *89*(1), 46–55.
- List, J. A. (2003). Does market experience eliminate market anomalies? *The Quarterly Journal of Economics*, *118*(1), 41–71.
- Loomes, G., Moffatt, P. G., & Sugden, R. (2002). A microeconomic test of alternative stochastic theories of risky choice. *Journal of Risk and Uncertainty*, *24*(2), 103–130.
- Schmidt, U., Starmer, C., & Sugden, R. (2008). Third-generation prospect theory. *Journal of Risk and Uncertainty*, *36*(3), 203–223.
- Seidl, C. (2002). Preference reversal. *Journal of Economic Surveys*, *16*(5), 621–655.
- Starmer, C., & Sugden, R. (1989). Probability and juxtaposition effects: An experimental investigation of the common ratio effect. *Journal of Risk and Uncertainty*, *2*(2), 159–178.
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458.
- Zeiliger, R. (2000). A presentation of regate, internet based software for experimental economics. GATE, Lyon. <http://www.gate.cnrs.fr/~zeiliger/regate/RegateIntro.ppt>