Economics Working Paper Series

2014/014

Alternation Bias and Reduction in St. Petersburg Gambles: An Experimental Investigation

Kim Kaivanto and Eike B. Kroll

The Department of Economics
Lancaster University Management School
Lancaster LA1 4YX
UK

LUMS home page: http://www.lums.lancs.ac.uk/

Alternation Bias and Reduction in St. Petersburg Gambles: An Experimental Investigation*

Kim Kaivanto[a†], Eike B. Kroll[b]

[a]Lancaster University, Lancaster LA1 4YX, UK

[b]Otto-von-Guericke University Magdeburg, 39104 Magdeburg, Germany

this version: January 20, 2014

Abstract

Reduction of compound lotteries is implicit both in the statement of the St. Petersburg Paradox and in its resolution by Expected Utility (EU). We report three real-money choice experiments between truncated compound-form St. Petersburg gambles and their reduced-form equivalents. The first tests for differences in elicited Certainty Equivalents. The second develops the distinction between 'weak-form' and 'strong-form' rejection of Reduction, as well as a novel experimental task that verifiably implements Vernon Smith's dominance precept. The third experiment checks for robustness against range and increment manipulation. In all three experiments the null hypothesis of Reduction is rejected, with systematic deprecation of the compound form in favor of the reduced form. This is consistent with the predictions of alternation bias. Together these experiments offer evidence that the Reduction assumption may have limited descriptive validity in modelling St. Petersburg gambles, whether by EU or non-EU theories.

Keywords: St. Petersburg Paradox, reduction axiom, alternation bias, dominance precept, law of small numbers, test of indifference

JEL classification: D81, C91

†Corresponding author: tel +44(0)1524594030; fax +44(0)1524594244; e-mail k.kaivanto@lancaster.ac.uk

1 Introduction

Throughout its 300-year existence,[1] the St. Petersburg Paradox has hinged on an implicit assumption that has become so deeply imbedded in the mathematics and economics professions that it is commonly deployed without any perceived need for separate justification. The assumption is that, for the purpose of modeling choice, compound lotteries may be reduced to their probabilistically equivalent simple lotteries "whose prizes are all the possible prizes of the compound lottery ticket, each evaluated with the compound probabilities that the classical algebra of probability defines" (Samuelson, 1952, p. 671). Within the program to formalize Expected Utility (EU), this assumption was made explicit and instated as an axiom of rational preferences: the Reduction of Compound Lotteries Axiom,[2] by which rational decision makers are required to be indifferent between a multi-stage *compound lottery* and its probabilistically equivalent 'collapsed' single-stage *simple lottery*.

Reduction is implicit both in the statement of the St. Petersburg Paradox as well as in its accepted resolution. The statement of the paradox hinges on powers i in the payoff function, $\$2^i$, precisely off-setting corresponding degrees i of compounding in the probability of Heads, $\left(\frac{1}{2}\right)^i$, for all $i \in \mathbb{Z}_{++}$, yielding an infinite sum. Daniel Bernoulli's resolution, just as the modern resolution by EU, also implements Reduction[3] in specifying the probability of heads on the i^{th} toss as $P_i(H) = \left(\frac{1}{2}\right)^i \ \forall\, i \in \mathbb{Z}_{++}$.

Nevertheless the experimental literature on probability perception suggests that *alternation bias* – a subjective distortion of conditional probability in binary sequences – should be empirically relevant for St. Petersburg Gamble (StPG) coin-toss sequences. Indeed specifically coin-toss sequences have been investigated in many 'perception of randomness' experiments (e.g. Rapoport and Budescu, 1997; Kareev, 1995; Budescu, 1987; for a review see Bar-Hillel and Wagenaar, 1991). Purely from a theoretical standpoint, augmentation of mathematical expectation with alternation bias is sufficient *by itself* to ensure that Willingness To Pay (WTP) for the StPG is finite and within the generally accepted empirical range (Kaivanto, 2008). Moreover, insofar as alternation bias is manifest as the subjective attribution of negative autocorrelation to objectively memoryless and unbiased Bernoulli processes, it suggests that Reduction may not

[1] Nicolaus Bernoulli posed the problem in a letter to Reimond de Montmort on September 9, 1713. Daniel Bernoulli's solution, precursor to the modern economics of decision making under risk, was published in the 1738 *Memoirs of the Imperial Academy of Sciences in St. Petersburg*, for which reason it has come to be known as the St. Petersburg Paradox.

[2] von Neumann and Morgenstern (1947) refer to it as Axiom 3:C:b (p. 26).

[3] implicit in the former, explicit in the latter

hold empirically for StPGs.

The present paper addresses the as-yet untested empirical question, *Is Reduction violated in StPGs?* The experimental design developed for this purpose incorporates several innovations. First, as is pertinent given the emphasis on sequential effects, the lotteries in this experiment are *truncated* real-money StPGs.[4] Second, we introduce reduced-form truncated StPGs, which are implemented with a single draw from an Urn. And third, we introduce a new choice list instrument. By design, this instrument permits simultaneous investigation of (i) weak-form violation of Reduction and (ii) strong-form violation of Reduction, while (iii) providing a gauge of subject-level heterogeneity in task-engagement cost[5] and treatment effect magnitude, and (iv) satisfying the dominance precept of value inducement in the presence of this heterogeneity.

In total we report three real-money experiments in which the standard compound-lottery form of the (truncated) StPG is explicitly juxtaposed with its reduced form. The former is implemented with a coin toss sequence consistent with convention, while the latter is implemented with a single random draw from a probabilistically equivalent urn. In the first experiment, we elicit subjects' certainty (cash) equivalents for each form of the gamble. In the second experiment, subjects face a choice list consisting of 11 choice tasks between the compound and the reduced form of the gamble, where each choice task has a distinct bonus payment added to one of the alternatives. The list starts with €5 added to the reduced form (Urn) and ends with €5 added to the compound form (Coin), changing in increments of €1 between choice tasks. This configuration offers a test of 'strong-form' violation of Reduction, in which subjects reveal with real-money choices whether they violate Reduction, and if so, how much they explicitly forgo in doing so. In the third experiment, we investigate possible range and increment effects with an 11-item price list that ranges from €1 added to the reduced form to €1 added to the compound form, changing in increments of 20 Euro cents between choice tasks.

In the first experiment we find that the distribution of certainty equivalents for the reduced form stochastically dominates the distribution of certainty equivalents for the compound form. Within subjects, the reduced form's certainty equivalent is statistically significantly larger than that of the objectively identical reduced form. Therefore we conclude that Reduction is violated in this 'judged valuation' task, revealing a bias toward the reduced form. This bias is borne out

[4]Previous experiments – e.g. by Cox et al. (2009) – have employed finite but not *truncated* StPGs. However, one of the finite St. Petersburg Gambles in Michael Birnbaum's (1998) yet-to-be published manuscript qualifies as *truncated* in the proper sense.

[5]meaning δ, the monetary equivalent value of the subjective cost of supplying the null hypothesis response instead of the alternative hypothesis response

in the second experiment. When choice is costless – i.e. in a straight choice between the reduced form and the compound form – 90% of the subjects choose the reduced form over the compound form. This constitutes a weak-form violation of Reduction in which the null hypothesis is premised on strict adherence to the axioms of EU and thus the absence of 'secondary criteria' influencing choice. Furthermore, 47.5% of the subjects forgo a sure €1 added to the compound form in order to obtain the reduced form.[6] These 47.5% violate Reduction in the strong-form sense, whereby a distinct preference for the reduced form is expressed through choice which involves giving up a certain €1. In the third experiment, a full 100% choose the reduced form over the compound form when it is objectively costless to do so (weak-form violation of Reduction), and 87.5% choose to forgo 20 Euro cents to obtain the reduced form rather than the objectively equivalent compound form (strong-form violation of Reduction).

In each of these experiments, both the rejection of the EU-based null hypothesis as well as the direction of this departure are consistent with the operation of *alternation bias*. This is a distortion of conditional probability, distinct from distortion of outcomes[7] and distortion of unconditional probabilities.[8] By design, the present experiments preclude the possibility that the observed choice behavior may be due to distortion of outcomes (concavity of utility) or distortion of unconditional probabilities (e.g. probability weighting). Altogether, these experiments provide evidence that the Reduction assumption (Axiom) may have limited descriptive validity in St. Petersburg Gambles. These results carry implications for both the demonstration of the St. Petersburg Paradox as well as for its theoretical resolution, each of which invokes Reduction without separate justification.

The rest of this paper is organized as follows. Reduction is presented in Section 2 both from a theoretical and empirical standpoint. Alternation bias is presented in Section 3. The concepts of weak-form and strong-form violation of Reduction are developed in Section 4. Section 5 on materials and methods presents the new constructs developed for this investigation: truncated StPGs, reduced-form StPGs, and a multiple-price list for implementing weak- and strong-form (dominance satisfying) tests of Reduction. Experiments I, II and III are presented in Sections 6, 7 and 8. Further analysis of alternation bias is presented in Section 9, and Section 10 discusses the results obtained here in light of previous studies. Section 11 concludes.

[6]Conversely, none of the subjects is willing to forgo a sure €1 added to the reduced form in order to obtain the compound form.

[7]Bernoulli's 'moral worth' and concave utility of money

[8]Yaari's dual theory and Prospect Theory's probability weighting

2 Reduction of compound lotteries

EU incorporates Reduction of compound lotteries, sometimes stated as an assumption, sometimes as an axiom. Some formalizations stipulate that simple one-stage lotteries are the basic objects of choice to which theory applies, and that multi-stage compound lotteries are 'reduced' into such simple lotteries through algebra alone (Samuelson, 1952; Hauser, 1978). Other formalizations require the decision maker to be indifferent between a multi-stage compound lottery and its probabilistically equivalent simple one-stage lottery (e.g. Harrison et al., 2012). von Neumann and Morgenstern (1947) denote it as Axiom 3:C:b and describe it as an expression of the 'algebra of combining' (p. 26). In parts of the EU literature the axiom is known by this latter label (Aumann, 1962; Fishburn, 1978). Luce and Raiffa (1957) instead invoke it as an assumption rather than as an axiom (Assumption 2, p. 26).

The Prospect Theory literature has adopted a variety of different approaches for dealing with compound prospects. Prospect Theory as originally introduced simply excluded compound prospects from consideration by restricting the domain of representable preferences to simple prospects (Kahneman and Tversky, 1979). In an exploratory investigation by the same authors, Reduction was found to be violated by the certainty effect and the pseudo-certainty effect, focusing attention on the second-stage prospect (Tversky and Kahneman, 1981). The 'cumulative' variant of Prospect Theory also restricts the domain of representable preferences to simple prospects (Tversky and Kahneman, 1992). There are several approaches to accommodating multi-stage lotteries, each with its own drawbacks (Wakker, 2010, Appendix C).

The normative appeal of Reduction notwithstanding, from a descriptive standpoint Reduction is a very strong assumption. Hauser (1978) expresses the view that it is "perhaps the strongest assumption in the utility axioms." Fishburn (1978) notes that "...many people exhibit systematic and persistent violations of... ...the reduction or invariance principle which says that *preference or choice between acts depends only on their separate probability distributions over outcomes*" (p. 492).[9]

Numerous experimental studies have investigated various aspects of Reduction (Bar-Hillel, 1973; Kahneman and Tversky, 1979; Keller, 1985; Bernasconi and Loomes, 1992; Gneezy, 1996; Budescu and Fischer, 2001; Halevy, 2007). Recently, Kaivanto and Kroll (2012) found *weak-form* violation of Reduction in real-money choices between probabilistically equivalent com-

[9]emphasis added

pound (two-stage) and simple (single-stage) lotteries offering a 1-in-10 chance of €100 ($136). The Kaivanto and Kroll (2012) experiment, which includes control treatments for ratio bias and computational limitations, finds that 80% of subjects choose the reduced form over the compound form, thereby violating the Reduction Axiom in the *weak-form* sense. Nevertheless, according to the most recent study of Reduction by Harrison et al. (2012), such violations of Reduction should be mere artifacts of the Random Lottery Incentive (RLI) scheme, in which a number of questions are asked, but only one randomly selected question is 'played out and paid out' for real money.[10] Harrison et al. (2012) find that violation of Reduction disappears when each question is individually 'played out and paid out' for real money. However, Kaivanto and Kroll (2012) find substantial *weak-form* violation of Reduction in choice between single-stage reduced-form and two-stage compound-form ($\frac{1}{10}$, €100) lotteries, even though they too 'play out and pay out' each question individually.

3 Alternation bias

Alternation bias is a binary sequence manifestation of the local representativeness effect, whereby the population (or infinite limit) properties of a stochastic process are attributed, erroneously, to small finite samples. The local representativeness effect was introduced into the economics literature by Rabin (2002) as a psychological *law of small numbers*, whereby people misjudge and "exaggerate how likely it is that a *small* sample resembles the parent population from which it is drawn" (p. 775).

For finite Bernoulli sequences generated by an objectively fair and memoryless coin, the local representativeness effect leads a subject to expect, within finite sequences, (i) close to a 50%–50% balance between Heads and Tails, and (ii) excessive local irregularity, i.e. too many reversals between Heads and Tails. It is this latter subjective predisposition to expect too many reversals that we call *alternation bias*. This may be understood more formally as a negatively distorted conditional subjective probability belief, or alternatively as an alternation rate that is subjectively upward-distorted $P_S(H|T) = P_S(T|H) > 0.5$.

Alternation bias was first hypothesized by Reichenbach (1934), and it has been amply documented and replicated in the 'perception of randomness' experimental literature (see summary in Bar-Hillel and Wagenaar, 1991). Experimental studies place the magnitude of first-order

[10] (Grether and Plott, 1979; Starmer and Sugden, 1994; Cubitt et al., 1998)

alternation bias at $P_S(H|T) = P_S(T|H) = 0.6$ (Budescu, 1987; Bar-Hillel and Wagenaar, 1991; Kareev, 1995). This forms a lower bound, as alternation bias effects have been estimated up to sixth order (Budescu, 1987). Within economics there are multiple studies, using both observational and laboratory data, that substantiate and replicate local representativeness, alternation bias and their manifestations the Gambler's Fallacy and the Hot Hand effect (Asparouhava et al., 2009; Clotfelter and Cook, 1993; Terrell, 1994, 1998; Croson and Sundali, 2005).

Alternation bias holds clear implications for preferences between compound lotteries and their probabilistically equivalent reduced-form lotteries. *Ceteris paribus*, subjects whose perception of randomness is characterized by alternation bias will not be indifferent between the compound form of a lottery (where alternation bias is operative) and its reduced-form equivalent (where there is no sequential structure to trigger alternation bias). In other words, the *indifference* between reduced- and compound-form lottery variants stipulated by Reduction is predicted to be violated under alternation bias.

Moreover, specifically for StPGs, alternation bias carries implications for mathematical expectation embodying this subjective distortion.[11] Let $\tilde{n} \in \mathbb{Z}_{++}$ be the (random) index of the first toss on which a fair coin first turns up 'Heads'. As $\tilde{n}$ is characterized by the geometric distribution with parameter $p = \frac{1}{2}$, the $n = 1, 2, ...$ stage probabilities are $p_n = \frac{1}{2}(1 - \frac{1}{2})^{n-1} = 2^{-n}$. Based on first-order alternation bias alone, $P_S(H|T) = 0.6$ and $P_S(T|T) = 0.4$, so the subjective (distorted) probability of the coin landing 'Heads' for the first time on toss n becomes

$$p_n^{fo} = \begin{cases} P_S(H) = \frac{1}{2} & \text{for } n = 1 \\ \frac{1}{2} P_S(H|T) P_S(T|T)^{n-2} = 0.3 \cdot 0.4^{n-2} & \text{for } n \geq 2 \end{cases} \tag{3.1}$$

giving a subjectively distorted mathematical expectation of

$$E_S^{fo}(G_{StP}) = \sum_{n=1}^{\infty} p_n^{fo} 2^n = 7.0 \tag{3.2}$$

without any need to invoke risk aversion or unconditional probability weighting.

4 Hypothesis development: weak-form and strong-form violation

For the purpose of formal testing, the Reduction assumption taken in isolation is insufficient for deriving a null hypothesis concerning choice behavior.[12] Reduction may only be tested as part

[11] First pointed out in Kaivanto (2008).

[12] Additional assumptions, such as completeness and monotonicity, are required.

of a joint hypothesis, ideally derived from an axiomatic theory of choice. In the present context, the natural candidate is EU, which as generally understood incorporates Reduction as an axiom or an assumption, and constitutes the modern counterpart to Daniel Bernoulli's 'moral worth' solution of the St. Petersburg Paradox.

The Reduction Axiom of EU stipulates that compound lotteries are evaluated as their equivalent single-stage reduced forms, and thus that indifference holds between compound lotteries and their reduced-form simple-lottery counterparts. However EU does not provide explicit guidance for choice when indifference holds. Any choice from among lotteries judged indifferent is consistent with EU.[13] Where indifference holds between two lotteries, the decision maker loses no utility regardless of which lottery she chooses; choice is 'costless'.

Nevertheless it is not the case that EU places *no* restrictions on choice between lotteries for which the indifference relation holds. EU restricts attention to lottery payoffs and probabilities. Under EU, no other characteristics are admitted as being choice-relevant. Moreover, under the Reduction Axiom, the only attributes to be legitimately (rationally) consulted in making a choice are the probabilities and payoffs of the reduced-form, single-stage, simple lottery. Under EU, preference is blind to all distinctions[14] between the compound form and its probabilistically equivalent reduced form.

This extends even to the difference in 'complexity' between compound and reduced forms.[15] Hence augmentation of EU with a further lexicographic criterion – albeit potentially an intuitively appealing rationalization of any revealed bias toward the reduced form – is in fact formally incompatible with the theoretical formulation of EU. In view of this strict and pure theoretical interpretation, lexicographic biasing of choice toward either the reduced form or the compound form would constitute a violation of EU.[16] Such lexicographic biasing of choice cannot be detected with a single observation per subject, but can in principle be detected with repeated observations or in the sample average of single-choice observations. Given the stringency of the

[13] As an example, note that Nash equilibrium in mixed strategies exploits this property.

[14] including e.g. complexity, the presence or absence of compoundness, the number of levels of compounding, simultaneous vs. sequential resolution of uncertainty, and the uniformity vs. non-uniformity of the time spacing of the resolution of uncertainty

[15] For the effect of complexity, see Sonsino et al. (2002) and Sonsino (2011).

[16] Apparent preference reversals (Schmidt and Hey, 2004; Butler and Loomes, 2007; Blavatskyy, 2009; Bardsley et al., 2010, Ch. 7) and apparent violations of betweenness (Blavatskyy, 2006) result from the differential impact of a zero-mean error term across two choice tasks. Either a constant-variance error term has a bigger impact on one choice task because the alternatives are 'closer together', or for other reasons – differential complexity being one possible example – the variance of the error term is larger in one choice task than in the other. However the test for weak-form violation of Reduction discussed here involves only a single choice task. Hence zero-mean error cannot result in a systematically biased choice pattern.

theoretical assumptions being maintained, however, we refer to this as a *weak-form* violation of Reduction. Null hypotheses for tests of weak-form violation of Reduction stipulate symmetrical empirical choice frequencies for compound-form and reduced-form lotteries.[17]

Following Vernon Smith's precepts for valid microeconomic experiments (induced value theory), we must nevertheless recognize that subjects face numerous costs in supplying the null hypothesis response instead of the alternative hypothesis response (Marschak, 1968; Smith, 1982; Wilcox, 1993; Harrison, 1994). These costs, the aggregate of which we denote with the symbol δ, include the cost of cognitive effort, concentration, fighting distraction or boredom, and the effect of other components of the subject's utility that are higher under the alternative hypothesis than under the null hypothesis. Moreover, specifically when a subject is indifferent between alternatives, "*secondary criteria* may be quite important and apparently small or seemingly irrelevant changes to the framing of decisions (e.g. positioning of words) may have a marked effect" (italics added, Bernasconi and Loomes, 1992). To overcome these costs and secondary criteria, the *dominance precept* requires that "the rewards corresponding to the null hypothesis are perceptively and motivationally greater [by at least $\delta > 0$] than the rewards corresponding to the alternative hypothesis" in order to "overcome any costs (e.g. the psychic cost of effort or of concentration) or components of the subject's utility that might induce a response that is not in accordance with the null hypothesis" (Harrison, 1994).

As observed by Chernoff (1954), Savage (1954), Bernasconi and Loomes (1992), Mandler (2005) and Danan (2008), EU in fact offers a crisp prediction when a sure bonus[18] $x \in \mathbb{R}_{++}$ is added to one of the lotteries. Due to the monotonicity axiom of EU, such a sure bonus x, no matter how small, causes the indifference relation to be replaced by strict preference for the bonus-augmented lottery. Hence adding a small bonus x to either the compound form or the reduced form causes EU decision makers to choose the bonus-augmented option. The formal experimental design property of *dominance* is satisfied when $x > \delta$. But as the value of δ is unknown and subject-specific, a range of values of x may be employed. Where larger values of x lead to lower rates of deviation from the EU prediction – both where x augments the compound-form lottery *and* where x augments the reduced-form lottery – an associated *weak-*

[17] To our knowledge this is the first time that the logic of the '50%–50% randomization' null hypothesis has been spelled out. The assumptions involved – strict adherence to an axiomatic theory that leaves no room for secondary criteria or for $\delta > 0$ – are implausibly strong. Nevertheless this null hypothesis often crops up in the literature.

[18] called a 'small monetary perturbation' by Savage (1954)

form violation of Reduction based on costless choice[19] may be ascribed to the failure to satisfy *dominance*. However, if choice is EU consistent in *all* bonus-augmented reduced-form (Urn+x) questions, an inference concerning δ is possible. Namely, it may be inferred that $\delta < \min x$. This is particularly helpful in interpreting EU *in*consistent bonus-augmented compound-form (Coin+x) questions, as $\delta < \min x$ now ensures that the failure to choose in accordance with the EU-derived null hypothesis is *not* attributable to absence of dominance, but *strong-form* violation of Reduction. With $\delta < \min x$ pinned down, larger values of x in Coin+x no longer offset larger subject-specific values of δ, but instead offset the effects of more extreme degrees of alternation bias.[20]

5 Materials and methods

Truncated StPGs The experiments reported here employ *truncated* StPGs. Cox et al. (2009), for instance, employ finite StPGs that pay nothing in the event of all coin tosses in the sequence landing 'Tails'. Since, in the unrestricted StPG, the player's payoff increases the *longer* the run of Tails, it is important to recognize that if the i^{th} toss lands Tails, extending a run of $i-1$ Tails by one, the player is entitled to a *minimum payout* of €2^{i+1}. We employ StPGs that are truncated at $k = \{2, 6\}$ tosses. In the $k=2$ case a run of 2 Tails pays off €2^3, while in the $k=6$ case a run of 6 Tails pays off €2^7. Thus the StPGs employed here are *truncations* in a formal and proper sense.

Reduced-form StPGs Denote the 'probability of landing Heads for the first time on toss n' in an StPG truncated to k tosses ($n \leq k$) as $p_{n(k)}$. The vector of probabilities of landing Heads for the first time on toss $n = (1, 2, ..., k)$ then becomes $\boldsymbol{p}_{(k)} = (p_{1(k)}, p_{2(k)}, ..., p_{k(k)})$. Thus the probabilities of the $k+1$ possible payoffs $(2^1, 2^2, ..., 2^k, 2^{k+1})$ in the k-truncated StPG(k) may be written as $\boldsymbol{p}_{(k,1)} = (p_{1(k)}, p_{2(k)}, ..., p_{k(k)}, p_{k+1(k)})$. Therefore the $k = 6$ truncated StPG(6) is characterized by the probability vector $\boldsymbol{p}_{(6,1)} = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \frac{1}{64}, \frac{1}{64}\right)$. Here $p_{6(6)} = p_{7(6)} = \frac{1}{2^6} = \frac{1}{64}$, so the $k=6$ reduced-form StPG(6) may be implemented with an urn containing $2^6 = 64$ balls. In general the k-truncated StPG(k) may be implemented with an urn containing no fewer than 2^k balls.

[19] $x = 0 < \delta$

[20] i.e. larger $P_S(H|T)$

Table 1: Example multiple-price list with $\triangle = 1$

Q.	Alternative A	Alternative B
1.	Urn variant + 5 Euro	Coin variant
2.	Urn variant + 4 Euro	Coin variant
3.	Urn variant + 3 Euro	Coin variant
4.	Urn variant + 2 Euro	Coin variant
5.	Urn variant + 1 Euro	Coin variant
6.	Urn variant	Coin variant
7.	Urn variant	Coin variant + 1 Euro
8.	Urn variant	Coin variant + 2 Euro
9.	Urn variant	Coin variant + 3 Euro
10.	Urn variant	Coin variant + 4 Euro
11.	Urn variant	Coin variant + 5 Euro

Multiple-price list The multiple-price list developed here subsumes a 'costless choice' weak-form violation of Reduction test and five incrementally more costly strong-form violation-of-Reduction tests, which do double duty in recording across-subjects heterogeneity in δ and in alternation bias strength corresponding to the magnitude of $P_S(H|T)$.

The range of the multiple-price list is a function of the number of items (questions) and the increment size. The number of items may be further decomposed into $j \in \mathbb{Z}_{++}$ items on either side of the central 'costless choice' item, giving a total of $2j+1$ items. Defining the inter-item increment to be $\triangle \in \mathbb{R}_{++}$, then the total range of the choice list becomes $2j\triangle$.

Across all experiments and treatments reported here, we fix $j = 5$, giving 11 items in total (see Table 1). In Experiment II we employ an increment size of €1, giving a total range of $2 \cdot 5 \cdot 1 =$ €10. In Experiment III we employ $\triangle = 0.20$ Euro, giving a total range of $2 \cdot 5 \cdot 0.20 =$ €2.

Under the null hypothesis of Reduction (embedded in EU) combined with negligible δ, choice in Q1–Q5 and Q7–Q11 is determined entirely by the bonus. If violation of Reduction is merely apparent, due to a failure to satisfy the dominance precept, then a symmetrical pattern of deviation may be expected around the costless choice item Q6. From this symmetrical pattern one may infer the sample heterogeneity in δ. If however violation of Reduction is not merely artifactual, then an asymmetric pattern around the costless choice item Q6 may be expected. From one side of Q6 the upper bound of δ may be inferred, while from the other side of Q6 one may infer sample heterogeneity in treatment effect size.

Table 2: Zero computation cost presentation format of the StPG(6) Alternative A.

Urn variant			Coin variant		
Ball no.	Probability	Payoff	'Heads' on	Probability	Payoff
1 – 32	1/2	2 Euro	1st toss	1/2	2 Euro
33 – 48	1/4	4 Euro	2nd toss	1/4	4 Euro
49 – 56	1/8	8 Euro	3rd toss	1/8	8 Euro
57 – 60	1/16	16 Euro	4th toss	1/16	16 Euro
61 – 62	1/32	32 Euro	5th toss	1/32	32 Euro
63	1/64	64 Euro	6th toss	1/64	64 Euro
64	1/64	128 Euro	All 'Tails'	1/64	128 Euro

6 Experiment I

This experiment is designed to elicit subjects' *judged* valuations of compound- and reduced-form StPGs. We implement standard Certainty Equivalent elicitation for both forms of the StPG. If Reduction holds, the StPGs, whether in compound form or reduced form, should be valued identically.

6.1 subjects and procedures

The experiment was conducted at Karlsruhe Institute of Technology (KIT). We recruited 73 students from the KIT engineering and computer science degree programs and assigned them to sessions of no more than 10 subjects using ORSEE (Greiner, 2004). Each subject had already completed at least one course in mathematics and statistics as part of his/her degree program. subjects ranged in age from 20 to 25, with an average of 22.3 years. 68% of the subjects were male, comparable to the KIT student population average.

Each subject was presented with two banks of 15 choices between Alternative A – the 'coin variant' compound-form StPG(6) in one bank, the 'urn variant' reduced-form StPG(6) in the other bank – and Alternative B (a sure payoff). Within each bank the sure payoff ranged from €1 to €15 in increments of €1 between questions (see Table 3). Each choice (i.e. question) was presented individually. To pre-empt response heterogeneity arising from individual differences in short-term memory and computation costs[21], a complete description of Alternative A – including a complete enumeration of the possible events, their probabilities, and associated payoffs – was provided on each question screen (see Table 2).

[21]resulting in e.g. miscalculation of compound probabilities

Despite being presented individually (and not in list form on a single screen), the €1 to €15 range can conceivably serve to prime or anchor subjects' responses. However as we are interested in the difference between the Certainty Equivalents of the compound and reduced forms – and not the levels *per se* – then even if there were a priming or anchoring effect, our results should remain unbiased because this €1 to €15 range is present in both compound-form and reduced-form 'banks'. These choice tasks were computer-administered with z-Tree (Fischbacher, 2007) as part of a larger experiment (103 questions in total). At the end of the experiment, a single question was randomly selected (out of the 103 questions) for each individual, to be 'played out an payed out' for real money (RLI mechanism).

At the beginning of each session, the subjects received written instructions describing the general types of questions they would face during the experiment as well as a detailed explanation of the RLI mechanism. After all session participants had read the instructions and explanations, the experimenter demonstrated the Urn and Coin Toss randomization devices to be used. In advance of commencing on the two banks of questions in z-Tree, subjects were given the opportunity to seek clarification on any aspect of the experiment. subjects wishing to ask a question were individually led outside of the laboratory, where neither the question nor the answer could be heard by the remaining session subjects. After all subjects completed their choice tasks, the experimenter proceeded to implement the RLI payoff scheme for each subject individually. Standard laboratory protocols to minimize the risk of experimenter demand effects were followed, as were measures to preclude individuals from illusion of control and deriving utility from direct personal involvement in 'playing out' the RLI.

6.2 Results

59 subjects switch exactly once from choosing the lottery to choosing the sure payoff in both 15-question sequences. We refer to the first sure payoff that the subject chooses as the subject's 'Certainty Equivalent' for that lottery.[22] 4 subjects choose the lottery in all 15 tasks. Consequently their Certainty Equivalents can not be determined with finite precision but only as falling 'above €15'. We code these subjects' Certainty Equivalents as €16 and use the median for the purpose of comparing treatments (N= 63).

subjects judge the value of the reduced-form StPG to be higher: the median Certainty

[22]The certain money sum that yields an indifference relation with the lottery will be *less than or equal to* this integer, but strictly larger than the next-smaller integer. Here we are interested in the difference between the reduced-form and the compound-form lotteries, rather than the levels of the Certainty Equivalents themselves.

Table 3: Fifteen choices (presented individually, not all at once) between the lottery (Alternative A: either the reduced-form Urn-implemented StPG(6) or the compound-form Coin-implemented StPG(6)) and a fixed sure amount (Alternative B).

Q.	**Alternative A**	**Alternative B: certain sum**
1.	Lottery	1 Euro
2.	Lottery	2 Euro
3.	Lottery	3 Euro
4.	Lottery	4 Euro
5.	Lottery	5 Euro
6.	Lottery	6 Euro
7.	Lottery	7 Euro
8.	Lottery	8 Euro
9.	Lottery	9 Euro
10.	Lottery	10 Euro
11.	Lottery	11 Euro
12.	Lottery	12 Euro
13.	Lottery	13 Euro
14.	Lottery	14 Euro
15.	Lottery	15 Euro

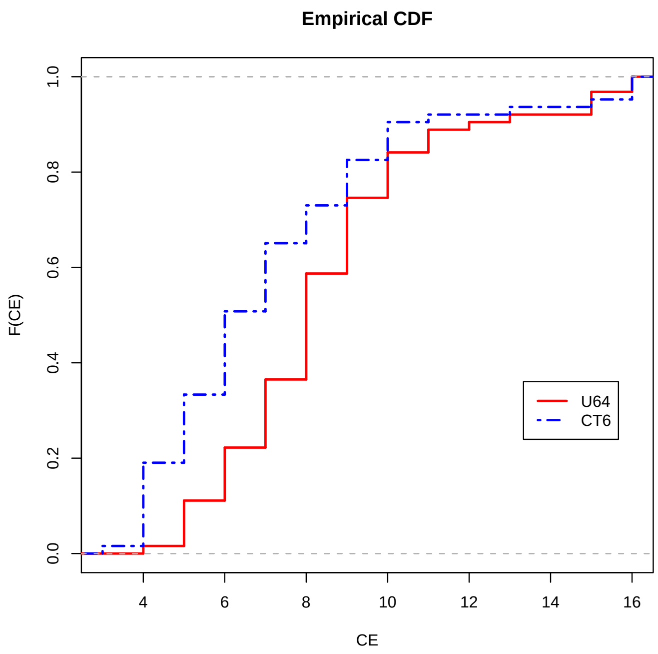
Equivalent for the reduced-form StPG is €8, while it is €6 for the compound-form StPG. Note that the mathematical expectation of StPG(6) is €8, while the first-order alternation bias that rationalizes a subjectively distorted mathematical expectation of €6 is $P_S(H|T) = 0.583$. This estimate is consistent with previous experimental findings on the magnitude of first-order alternation bias (Budescu, 1987; Bar-Hillel and Wagenaar, 1991; Kareev, 1995).

The null hypothesis under Reduction is that there is no difference within subjects between the reduced-form lottery CE and the compound-form lottery CE. The Wilcoxon signed-rank test rejects the null that the median of the paired differences is zero (Asymptotic Wilcoxon-Signed-Rank Test $Z = 4.8017$, p-value $= 1.574 \times 10^{-06}$). The distribution of within-subject differences is negatively skewed.[23] For robustness to skewness we also implement the paired sign test, and find that the null of zero median difference is rejected (Dependent-samples Sign-Test $S = 42$, p-value $= 2.458 \times 10^{-08}$). Furthermore, the comparison of the empirical Cumulative Distribution Functions (CDFs) reveals a stochastic dominance relationship: among those subjects who reported a finite Certainty Equivalent within the task (N= 59), the reduced-form (Urn) First-Order Stochastically Dominates the compound-form (Coin); among all subjects including

[23]D'Agostino skewness test $\gamma_1 = -1.9636$, $z = -3.2857$, p-value $= 0.001017$

those reporting a non-finite Certainty Equivalent in the task (N= 63), the reduced-form (Urn) Second-Order Stochastically Dominates the compound-form (Coin).

Figure 1: Empirical CDFs of the Certainty Equivalents of the reduced-form StPG(6) (U64, red solid line) and the compound-form StPG(6) (CT6, blue dotted line).



7 Experiment II

The purpose of this experiment is to generate data that is suitable for implementing weak-form and strong-form tests of Reduction. In a direct choice task between probabilistically equivalent reduced-form and compound-form StPGs, any revealed bias for the reduced-form StPG only constitutes a weak-form violation, due to the very strong assumptions involved. Therefore we add direct choice tasks in which subjects must forgo a sure payoff in order to violate Reduction. We term the latter *strong-form* violations of Reduction.

7.1 subjects and procedures

The procedural aspects of the experiment, the subject pool, the method of recruitment, the laboratory, and the incentive scheme (RLI) are identical to Experiment I above. A total of 40 subjects participated in Experiment II. All subjects answered the 11 multiple-price list questions

(Table 1: $j = 5$, $\triangle = 1$), presented one at a time on the z-Tree display, for both the $k = 2$ truncated StPG(2) as well as the $k = 6$ truncated StPG(6). To control for 'position effects' (Bernasconi and Loomes, 1992), the Urn option was presented on the left-hand side of the screen for half of the subjects (N=20) and on the right-hand side for the other half (N=20). Within each of these groups, half (N=10) received the StPG(2) questions before the StPG(6) questions, while the remaining half (N=10) received the StPG(6) questions before the StPG(2) questions. At the end of the experiment, a single question was randomly selected (out of the 22 questions) for each individual, to be 'played out and payed out' for real money (RLI mechanism). Unlike the experiments of Bar-Hillel (1973) and Budescu and Fischer (2001), the subject is not personally involved in the selection and implementation of his/her real-money gamble. This is done to ensure that choice is not influenced by illusion of control, fun, excitement or other components of utility directly associated with participation in the process of gambling.

7.2 Results

Table 4: Proportion of subjects (of $N = 40$) choosing the urn with the associated 95% confidence intervals (Jeffreys prior) for the 'max 2 tosses' StPG(2) and 'max 6 tosses' StPG(6) price lists.

	Urn +5	**Urn +4**	**Urn +3**	**Urn +2**	**Urn +1**	**Urn +0**	**Coin +1**	**Coin +2**	**Coin +3**	**Coin +4**	**Coin +5**
StPG(2)	1	1	1	1	1	.95	.25	.10	0	0	0
(95% CI)	(.94,1)	(.94,1)	(.94,1)	(.94,1)	(.94,1)	(.85,.99)	(.14, .40)	(.03, .22)	(0,.06)	(0,.06)	(0,.06)
StPG(6)	1	1	1	1	1	.90	.475	.175	0	0	.025
(95% CI)	(.94,1)	(.94,1)	(.94,1)	(.94,1)	(.94,1)	(.78,.97)	(.33,.63)	(.08,.31)	(0,.06)	(0,.06)	(.003,.11)

The Urn + 0 entries may be tested for weak-form violation of Reduction ($H_0 : \hat{p} = p_0$ and $H_1 : \hat{p} > p_0$) using the Exact Binomial Test with $N = 40$ and $p_0 = 0.5$. On the StPG(2) task, under the null hypothesis, the probability of observing $\hat{p} \geq .95$ is 7.467×10^{-10}. On the StPG(6) task, under the null hypothesis, the probability of observing $\hat{p} \geq .90$ is 9.285×10^{-08}. So under the Binomial test's assumption that the Urn-choice probability is homogeneous within task tested, the null hypothesis of 'symmetrical choice probability' is rejected. Thus the data display weak-form violation of Reduction.

There is strong asymmetry to the left and right of the costless choice item (Urn+0). On the left-hand side of Table 4, note that all bonus sums – including €1 – are recognized as pure windfall when they are combined with the Urn. All subjects elect to collect this windfall by systematically choosing bonus-augmented Urn options over the Coin. We infer that $\delta < 1$ and that dominance is satisfied. Hence entries on the right-hand-side of the table reflect strong-form violation of Reduction. In StPG(6), 47.5% of subjects forgo a sure bonus of €1 in order to obtain Urn (reduced form) instead of the Coin (compound form). In StPG(2), 25% of subjects forgo €1 in order to obtain the Urn. So with the $\triangle = 1$ increment, 25% of subjects display strong-form violation of reduction on StPG(2) and 47.5% of subjects display strong-form violation of Reduction on StPG(6).

8 Experiment III

This experiment is designed to test the robustness of Experiment II's results by changing the increment size to $\triangle = 0.20$ Euro and hence the range of the choice list to $2j\triangle =$ €2.

8.1 subjects and procedures

To ensure comparability, Experiment III is identical to Experiment II in terms of procedures, method of recruitment, and incentive scheme (RLI). However Experiment III was conducted at the Magdeburg Experimental Laboratory (MaXLab) at the Faculty of Economics and Management of Otto-von-Guericke-University Magdeburg. Using ORSEE (Greiner, 2004), we recruited 40 subjects who were enrolled in an engineering or computer science degree program at the time of participating in the study, ensuring comparability of the subjects' training in probability theory with those of Experiments I and II. The subjects ranged in age from 19 to 23, with an average of 22.5 years; 48% were male.

8.2 Results

Here, with restricted increment and range, *all subjects* choose the Urn when choice is costless (i.e. in the Urn+0 column of Table 5). This is more extreme than in Experiment II, and the null hypothesis of symmetrical choice probabilities is rejected (Exact Binomial Test one-sided p-value $= 9.095 \times 10^{-13}$). Hence weak-form rejection of Reduction proves robust to increment and range manipulation.

As in Experiment II, here there is again strong asymmetry to the left and right of the costless choice item (Urn+0) of Table 5. From the left-hand side of the table we infer that $\delta < 0.20$ and

Table 5: Proportion of subjects (of $N=40$) choosing the urn with the associated 95% confidence intervals (Jeffreys prior) for the 'max 2 tosses' StPG(2) and 'max 6 tosses' StPG(6) price lists.

	Urn +1	**Urn +.80**	**Urn +.60**	**Urn +.40**	**Urn +.20**	**Urn +0**	**Coin +.20**	**Coin +.40**	**Coin +.60**	**Coin +.80**	**Coin +1**
StPG(2)	1	1	1	1	1	1	.70	.525	.35	.275	.20
(95% CI)	(.94,1)	(.94,1)	(.94,1)	(.94,1)	(.94,1)	(.94,1)	(.55, .82)	(.37,.67)	(.22,.50)	(.16,.43)	(.10,.34)
StPG(6)	1	1	1	1	1	1	.875	.75	.65	.575	.50
(95% CI)	(.94,1)	(.94,1)	(.94,1)	(.94,1)	(.94,1)	(.94,1)	(.75,.95)	(.60,.86)	(.50,.78)	(.42,.72)	(.35,.65)

that dominance is satisfied. Hence entries on the right-hand-side of the table reflect strong-form violation of Reduction. To be consistent with Reduction, the Urn-choice fractions should be zero on the right-hand side of Table 5. Instead, 70% of subjects in the StPG(2) task and 87.5% of subjects in the StPG(6) task forgo a sure bonus of 20 Euro cents in order to obtain the reduced form instead of the compound form.

9 Further analysis of alternation bias

The above test results are consistent with the predictions of alternation bias. The structure of Experiments II and III allow further direct tests of whether alternation bias is the *sole* driver of the weak- and strong-form violation of Reduction documented here.

Take the Urn+0 and Coin+1 columns of Tables 4 and 5. The StPG(2) Urn-choice fraction in the Urn+0 column under the $\triangle = 1$ increment (Table 4) should – if there are no other drivers of choice – be equal to the Urn-choice fraction in the Urn+0 column under the $\triangle = 0.20$ increment (Table 5). Similarly for StPG(6) in the Urn+0 columns across the two tables as well as for StPG(2) and StPG(6) in the Urn+1 columns across the two tables, giving a total of four subtables.

With the restricted increment and range of Experiment III, all subjects choose the Urn in both the StPG(2) and StPG(6) tasks when choice is costless (in the Urn+0 column of Table 5). The zero count in the Coin cell of the tabular presentation of these data causes problems for standard Chi-squared tests and logistic regression. These problems may be overcome by making recourse to exact tests. Fisher's Exact Test is commonly applied in these circumstances.

Here, the null hypothesis of no association between increment size ($\triangle = 1$ or $\triangle = 0.20$) and Urn-choice propensity fails to be rejected in each of the four subtables when using Fisher's Exact Test.[24] However, these four tests have very low power. In fact Fisher's ET is known to have low power in 2×2 tables (Lydersen et al., 2009). Being a conditional test, this test assumes that both margins of the 2×2 table are fixed. So not only does Fisher's ET have low power, but it is premised on an assumption that is inconsistent with the structure of our experiments, where the number of subjects in each increment-size condition is fixed by the experimenter, but the Urn/Coin choice margin is not fixed. Unconditional tests such as Barnard's ET and Boschloo's ET do not fix both margins, and consequently achieve higher power. Across the present four subtables, Barnard's ET achieves uniformly higher power than Boschloo's ET. We report the former.

Table 6: Barnard's unconditional Exact Test of association between increment size and Urn-choice propensity: p-values (one-sided), associated power, and adjusted p^{BH}-values in which the overall False Discovery Rate is controlled with the Benjamini-Hochberg method.

			Urn count	**Coin** count	**Barnard's ET** p-value	power	**FDR** p^{BH}-value
Urn+0	StPG(2)	$\triangle = 1$	38	2	0.103	0.32	0.206
		$\triangle = .20$	40	0			
	StPG(6)	$\triangle = 1$	36	4	0.021	0.78	0.084
		$\triangle = .20$	40	0			
Coin+1	StPG(2)	$\triangle = 1$	10	30	0.340	0.12	0.453
		$\triangle = .20$	8	32			
	StPG(6)	$\triangle = 1$	19	21	0.456	0.067	0.456
		$\triangle = .20$	20	20			

Even with Barnard's ET, the achieved power is very low in the Coin+1 subtables. It is

[24] Using Fisher's Exact Test, we find no significant association between the Urn-choosing fraction and increment size in comparing (a.i) the Urn-choosing fraction from the StPG(2) row of the Urn+0 column of Table 4 ($\frac{38}{38+2} = 0.95$) with that appearing in the corresponding cell of Table 5 ($\frac{40}{40+0} = 1$) ($p = 0.494$ 2-sided; $p = 0.247$ 1-sided), (a.ii) the Urn-choosing fraction from the StPG(6) row of the Urn+0 column of Table 4 ($\frac{36}{36+4} = 0.90$) with that appearing in the corresponding cell of Table 5 ($\frac{40}{40+0} = 1$) ($p = 0.116$ 2-sided; $p = 0.058$ 1-sided), (b.i) the Urn-choosing fraction from the StPG(2) row of the Coin+1 column of Table 4 ($\frac{10}{10+30} = 0.25$) with that appearing in the corresponding cell of Table 5 ($\frac{8}{8+32} = 0.20$) ($p = 0.790$ 2-sided; $p = 0.397$ 1-sided), and (b.ii) the Urn-choosing fraction from the StPG(6) row of the Coin+1 column of Table 4 ($\frac{19}{19+21} = 0.48$) with that appearing in the corresponding cell of Table 5 ($\frac{20}{20+20} = 0.50$) ($p = 1.000$ 2-sided; $p = 0.500$ 1-sided).

better in the Urn+0 subtables, achieving a level of .78 in the StPG(6) subtable, where the null hypothesis is rejected with a p-value of 0.021. This result is suggestive of there being an increment-size effect on Urn-choice propensity for $\alpha = 0.05$. But adjusting the critical value to control the Family-Wise Error Rate of the four simultaneous tests using Bonferroni's correction $\alpha^{\mathrm{B}} = 0.05/4 = 0.0125$ or the more powerful Dunn-Šidák correction $\alpha^{\mathrm{DS}} = 1-(1-\alpha)^{1/4} = 0.0127$ suggests that the null of 'no association' is not rejected in even one of the tests. Similarly, when employing the Benjamini and Hochberg (1995) procedure for adjusting the p-values to control the False Discovery Rate across the four simultaneous hypothesis tests, the null hypothesis again remains unrejected.

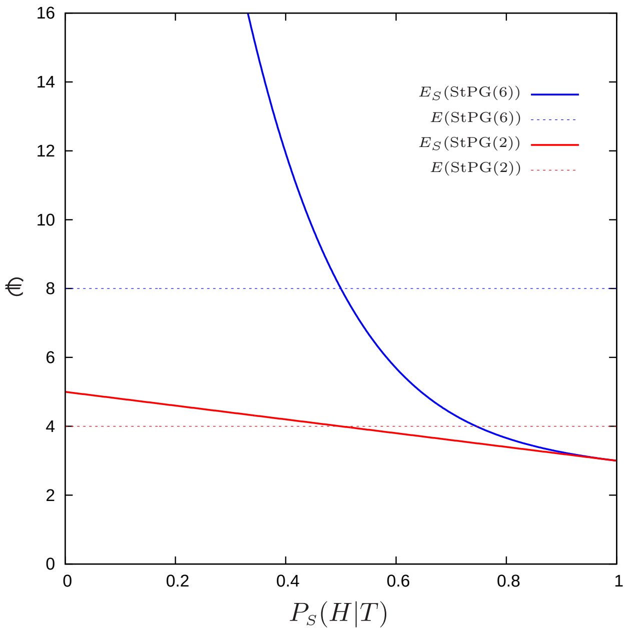
The results of Bayesian analysis are more conclusive. Starting from uniformly distributed priors on Urn-choice propensity $p(\theta_{€1}) = p(\theta_{€.2}) =$ Beta(1,1) and the null hypothesis of no difference in these propensities $H_0 : \theta_{€1} - \theta_{€.2} = 0$, $H_1 : \theta_{€1} - \theta_{€.2} \neq 0$, the Bayes factor $BF_{01} = \frac{P(D|H_0)}{P(D|H_1}$ is given by $BF_{01} = \frac{(n_1+1)(n_2+1)}{n_1+n_2+1}\binom{n_1}{s_1}\binom{n_1}{s_2}\binom{n_1+n_2}{s_1+s_2}^{-1}$. In the Urn+0 StPG(2) subtable, where $n_1 = n_2 = 40$, $s_1 = 38$ and $s_2 = 40$, the Bayes factor is $BF_{01} = 5.1$. In the Urn+0 StPG(6) subtable $BF_{01} = 1.2$, while for the Coin+1 subtables, the Bayes factor is $BF_{01} = 3.8$ for StPG(2) and $BF_{01} = 3.6$ for StPG(6). Three of these Bayes factors fall in the range $3 < BF \leq 10$, indicating that these subtables offer substantial evidence in favor of the null hypothesis of there being no increment-size effect (Jeffreys, 1961).

Table 4 also contains information on the heterogeneity of alternation bias in the sample. We may calculate the alternation bias required to rationalize the entries in Table 4, separately for StPG(2) and StPG(6).

Let us take StPG(6) first. The 17.5% of subjects who forgo €2 in favor of the Urn satisfy the inequality $E_S^{fo}(\text{StPG}(6))+2 \leq E(\text{StPG}(6)) + \delta$. From this and the fact that none of these 17.5% of subjects is willing to forgo €3 in order to obtain the Urn, the first-order alternation bias for this 17.5% of the subjects may be calculated as $P_S(H|T) \in [0.583, 0.646)$ for those subjects with $\delta = 0$ and $P_S(H|T) \in [0.537, 0.583)$ for those subjects with δ arbitrarily close to 1. Similarly, the 47.5% who forgo €1 (but not €2) in favor of the Urn satisfy the inequality $E_S^{fo}(\text{StPG}(6))+1 \leq E(\text{StPG}(6)) + \delta$ and their first-order alternation bias may be calculated as $P_S(H|T) \in [0.537, 0.583)$ for those subjects with $\delta = 0$ and $P_S(H|T) \in [0.5, 0.537)$ for those subjects with δ arbitrarily close to 1. Overall, these alternation bias estimates are consistent with Experiment I as well as with existing literature (Budescu, 1987; Bar-Hillel and Wagenaar, 1991; Kareev, 1995).

However, StPG(2) throws up an inconsistency. Choosing to forgo €1 in order have the Urn StPG(2) instead of the Coin StPG(2) requires an alternation bias of $P_S(H|T) = 1$ for those subjects with $\delta = 0$ and $P_S(H|T) \in [0.5, 1)$ for those subjects with δ arbitrarily close to 1. Note that alternation bias of $P_S(H|T) = 1$ falls outside the empirical range established in the existing literature, where $P_S(H|T)$ has been found to reside in the neighborhood of 0.6. However, choosing to forgo €2 in order to have the Urn StPg(2) instead of the Coin StPG(2) – as is done by 10% of the subjects in Table 4 – cannot be rationalized by any degree of alternation bias $P_S(H|T) \in [0, 1]$ for subjects with $\delta \in [0, 1)$. This is because, even with the maximum alternation bias of $P_S(H|T) = 1$, the subjectively distorted expectation of the Coin-implemented StPG(2) is $E_S^{fo}(\text{StPG}(2)) = 3$ (see Figure 2), from which follows that $E_S^{fo}(\text{StPG}(2)){+}2 > E(\text{StPG}(2)) + \delta$ for $\delta \in [0, 1)$.[25]

Figure 2: Implications of subjective distortion of conditional probability on subjective expected value: $E_S(\text{StPG}(2))$ and $E_S(\text{StPG}(6))$ as functions of $P_S(H|T)$.



For StPG(6), the implied alternation bias ranges are consistent with existing empirical estimates and these ranges successfully rationalize the departures from Reduction in Table 4. However for StPG(2), alternation bias is insufficient, on its own, to rationalize all departures from

[25]where the range for δ is implied by the unanimous choice of the Urn in the Urn+1 items on the left-hand side of Table 4

Reduction in Table 4. Instead, the observed within-subjects pattern across Coin+2 StPG(2) and Coin+2 StPG(6) implies a model of the form

$$E_{S_i}^{fo}(\text{StPG}(2)) + 2 \leq E(\text{StPG}(2)) + \delta_i + \Phi_i \ \ , \tag{9.1}$$

where Φ_i reflects the monetary value of the strength of preference difference between the Urn and the Coin that subject $i \in \{1, 2, ..., N\}$ does not attribute to $P_{S_i}(H|T)$ or δ_i. This can be understood in terms of *source dependence*. Whereas the source dependence studied by Tversky and Wakker (1995) and Abdellaoui et al. (2011) operates via the probability weighting function, here it is clear from the within-subjects contrast between Coin+2 StPG(2) and Coin+2 StPG(6) that the effect driving the Urn-choice fraction in Coin+2 StPG(2) is independent of the probability weighting function, else the Urn-choice fraction in Coin+3 StPG(6) would be biased upward above the level observed in Table 4. On the other hand, the source dependence proposed by Smith (1969) and estimated econometrically by Harrison et al. (2012) operates via the utility function: individuals are permitted to have different utility functions specific to different random processes (randomization devices). This form of source dependence is consistent with the choice pattern in Coin+2 StPG(2), Coin+2 StPG(6) and Coin+3 StPG(6). Disentangling and quantifying the behavioral effects aggregated within Φ_i – including e.g. of preference for simplicity over complexity, of preference for few levels of compounding over many levels of compounding, and of preference for simultaneous resolution of uncertainty over sequential resolution of uncertainty – requires lottery manipulations that shift the focus away from truncated StPGs, and is left for future investigation.

10 Discussion

Recognition that alternation bias is empirically operative in StPGs creates an alternative to the radical reparameterization of CPT suggested by Blavatskyy (2005).[26] Prospect Theory, and by extension CPT, was conceived to be a descriptive workhorse model to account for a central subset of decision-making biases and heuristics, while abstracting from all those remaining.[27] Original Prospect Theory is explicitly limited to 'simple prospects', i.e. reduced-form prospects, and therefore makes no claims to capture the behavioral implications of StPG-defining compoundness. Meanwhile, only one parametrization of one particular CPT probability weighting

[26] Similarly, it suggests that Pfiffelmann's (2011) calibrated restriction of the shape of the probability weighting function may be re-implemented under alternation bias.

[27] "Theories of choice are at best approximate and incomplete" (Tversky and Kahneman, 1992, p. 317).

function satisfies the simplest probabilistic reduction of compound gambles condition.[28] Using this power function specification of the probability weighting function, Blavatskyy (2005) shows that the CPT value of StPGs is infinite. Augmenting CPT to incorporate subjective distortion of conditional probability in binary sequence compound prospects – as the present experiments suggest is required in a descriptive theory – renders the otherwise conventionally parameterized CPT valuation finite, falling within the accepted empirical range (Kaivanto, 2008).

Methodologically, the present study extends and complements existing experimental tests of Reduction, the principal features of which are summarized in Table 7. Although the experimental designs employed in these studies display variety and heterogeneity, they are also characterized by similarities and regularities. For instance, six of the studies employ 2-Alternative Forced Choice (2AFC) tasks, one employs 3-Alternative Forced Choice (3AFC) tasks, and one elicits lottery reservation prices with the Becker-DeGroot-Marschak (BDM) method. None of the eight studies has a design that demonstrably satisfies the dominance precept. Incentives range from nothing (hypothetical choice tasks) to €100 ≈ US$136, some paying out 1-in-1 while others pay out according to an RLI mechanism, but not one of these studies is able to demonstrate that the payoff function being employed compensates for δ_i, the monetary value of the psychological cost of supplying the null hypothesis response in place of the alternative hypothesis response. In part, this is due to designs which do not accommodate and reflect possible heterogeneity in the latent δ_i parameter. The development and in-use demonstration of just such an instrument is one of the key undertakings of the present work.

[28]When Prelec's two-parameter family $W(p) = [-\beta(-\ln p)^\alpha]$ is specialized to $W(p) = p^\beta$ by setting $\alpha = 1$, then $((x, p), q) \sim (x, pq)$, as shown by Luce (2001).

Table 7: Summary of existing experiments' design parameters

Study	**Incentive**	**Task**	**Design**	$\boldsymbol{H_0}$	**Controlling**	**¬ Controlling**
BH73[1]	≈US$3 1-in-4	2AFC×4	repeated measures	$\hat{p} = 0.5$		δ_i, ill of control, fun, excitement, util of gambling
KT79[2]	hypo-thetical	2AFC pairs	repeated measures	across-task consistency $\hat{p}^{T1} = \hat{p}^{T2}$		δ_i
K85[3]	hypo-thetical	2AFC pairs	repeated measures	across-task consistency $\hat{p}^{T1} = \hat{p}^{T2}$	pictoral representations, training	δ_i
BL92[4]	hypo-thetical	3AFC×2, swap option	2× one sample, swap	$\hat{p}^{Ak} = \frac{1}{3}$ $k = 1, 2, 3$	secondary criteria, word placement	δ_i
BF01[5]	NIS2 1-in-1	2AFC	repeated measures	$\hat{p} = 0.5$	gains/losses, complexity, ratio bias, simult/sequent, opaqueness, incr v decr p, temporal intvls	δ_i, hope, fun, impatience
H07[6]	CA$2 1-in-1 $\forall j$	BDM reser-vation prices	repeated measures $j = 1, ..., 4$	$\hat{F}_{Tj} = F$ $\forall j$	stakes ×10, task order, BDM compr, wording, sampl efficiency	δ_i
KK12[7]	€100 ≈US$136 1-in-1	2AFC	3× one sample	$\hat{p} = 0.5$	ratio bias, miscalculation	δ_i
HMS12A[8]	US$0,10, 20,35,70 1-in-1 1-in-40	2AFC triples s-c,s-ae,c-ae[9]	between subjects	across-task consistency $\hat{p}_j^{\frac{\text{c}}{\text{ae}}} = \hat{p}_k^{\frac{\text{c}}{\text{ae}}} = \psi$ $\forall j,k \in \{1,\ldots,10\}, j \neq k$ $\hat{p}_j = 0.5$ $\forall j = 1, ..., 10$ $\hat{p}^{\text{s-c}} = \hat{p}^{\text{s-ae}}$	1-in-1/40	δ_i
HMS12B[8]	US$0,10, 20,35,70 1-in-1 1-in-40	2AFC triples s-c,s-ae,c-ae[9]	source dependent structural model	$\hat{r}\hat{c}=0$, $\hat{\gamma}\hat{c}=0$	1-in-1/40, demographic variables	δ_i

[1] Bar-Hillel (1973)
[2] Kahneman and Tversky (1979)
[3] Keller (1985)
[4] Bernasconi and Loomes (1992)
[5] Budescu and Fischer (2001)
[6] Halevy (2007)
[7] Kaivanto and Kroll (2012)
[8] Harrison et al. (2012)
[9] simple vs. compound, simple vs. actuarially equivalent, compound vs. actuarially equivalent

Bar-Hillel (1973), Budescu and Fischer (2001) and Kaivanto and Kroll (2012) implement incentivized weak-form tests of Reduction ($H_0 : \hat{p} = 0.5$), with the latter supplying a non-trivial monetary incentive both in nominal terms as well as in expectation. Bernasconi and Loomes (1992) implement hypothetical 3AFC tasks, and thus the null hypothesis of their weak-form tests[29] are of the form $H_0 : \hat{p} = \frac{1}{3}$. Altogether half of the existing studies (four in total) implement weak-form tests of Reduction.[30] Bar-Hillel's (1973) rejections of Reduction are psychologically inclusive, in the sense of not controlling for the effects of miscalculation of probabilities, illusion of control, fun, excitement or other components of utility directly associated with participation in the process of gambling (that is, the 'utility of gambling' deriving from, e.g., rolling the dice, picking marbles from an urn, or choosing a card). Budescu and Fischer (2001) ask their subjects to record the reasons for their choices; these reported reasons reflect the psychological inclusivity of the design and implementation. The most frequently given reasons were: sample space size (ratio bias), number of events (fewer being preferred), stress (of sequential compound events), simplicity (being preferred), hope, fun (of there being 'more action'), illusion of control, and time (preferring an immediate answer). Meanwhile, the Kaivanto and Kroll (2012) experiments were implemented with laboratory procedures designed (i) to attenuate and suppress the experience of gambling and the possibility for illusion of control, as well as (ii) to mitigate presentation format regularities that manifest as response bias attributable to 'secondary criteria', while explicitly controlling for ratio bias (sample space effects) and miscalculation-induced response bias.[31]

In the modern experimental economics sense, weak-form violations of Reduction fall short of meaningful and valid rejections of Reduction. The EU values of a compound lottery and its reduced form are identical, and since *any* choice pattern between these two lotteries is consistent with identical EU values, 2AFC tasks cannot offer evidence against Reduction. To resolve this methodological dilemma, experiments II and III follow Chernoff (1954), Savage (1954), Bernasconi and Loomes (1992), Mandler (2005) and Danan (2008) in utilizing a small monetary 'bonus' to deform the indifference relation into a strict preference relation, and furthermore present the subject with a multiple-price list of such bonus-augmented 2AFC tasks such that heterogeneity in δ_i may be revealed and conformity with the dominance precept may be directly

[29]n.b. the 'swap option' that Bernasconi and Loomes (1992) include provides additional information beyond the weak-form test

[30]Harrison et al. (2012) also implement weak-form tests of Reduction using the binomial test on 10 different 2AFC tasks separately.

[31]The laboratory procedures in Experiments I–III are also designed to achieve control in these dimensions.

observed *if* the choice of reduced form over its compound equivalent is not merely a manifestation of the failure to satisfy dominance.

As a conceptual matter however, we may ask whether a set of assumptions exists from which the symmetric choice frequency null hypothesis follows. Modern experimental economics tests theories in conjunction with a set of auxiliary hypotheses first set out by Vernon Smith (1969). In the absence of these auxiliary hypotheses – making recourse solely to the axioms of EU – it is nevertheless not the case that EU places no restrictions whatsoever on what lottery attributes the decision maker may consult in making choices. Rather, as quotations from Paul Samuelson and Peter Fishburn above make plain, the only parameters consulted by EU decision makers are outcomes and reduced-form probabilities. This rules out systematic non-symmetrical response distributions in 2AFC tasks between lotteries with identical EU values such as compound lotteries and their reduced-form equivalents. In the eyes of a strictly EU decision maker, there simply are no distinguishing features between compound- and reduced-form lottery pairs, and this prevents systematically biased choice patterns from being formed. As noted above, this is a very strict theoretical interpretation of EU, and accordingly we suggest that symmetric choice frequency null hypotheses be labeled *weak-form* tests.

Those acquainted with the precepts of modern experimental economics rightly point to 'secondary criteria' and absence of dominance as potential causes of asymmetric response patterns in such weak-form tests. Hence it is a matter for the consumer of the weak-form test result to evaluate (i) whether sufficient precautions against secondary criteria have been taken, and (ii) whether the incentives offered satisfy the dominance precept. The three studies cited above illustrate the difference that (i) and (ii) can make. The Bar-Hillel (1973) experiments, which are implemented with weak incentives and direct subject involvement in 'playing out' the lotteries – hence likely to give rise to the illusion of control, fun, excitement, and utility from the process of gambling itself – find that subjects prefer compound lotteries over reduced-form lotteries. This result, which is known in the literature as an example of the *overweighting of conjunctive probabilities*, is precisely the opposite of what Budescu and Fischer (2001) and Kaivanto and Kroll (2012) find with experimental designs more attuned to (i) controlling secondary criteria and (ii) offering strong incentives.

Three studies – Kahneman and Tversky (1979), Keller (1985) and Harrison et al. (2012) – employ tests that substitute the strict interpretation of EU's axioms with a requirement for across-task consistency (or homogeneity) in choice propensity. This is the configuration employed in the classical preference reversal experiments, which remain agnostic with regard to what constitutes the normative choice propensity in any one task individually. Instead, formal statistical testing is used to determine whether choice propensity *changes* across the 2AFC pair as a result of each having a different presentation format (one including a compound-form lottery, the other including the equivalent reduced-form lottery).

Halevy (2007) employs a valuation-based design, specifically the elicitation of reservation prices through the BDM mechanism. He tests whether the four repeated measures (distributions of reservation prices) are drawn from the same parent distribution, using the Friedman test. As in Experiment I of this paper (Section 6) Halevy (2007) finds that the reduced-form lottery attracts a higher valuation than compound-form lotteries. Moreover, Halevy (2007) shows that this difference in valuation survives a robustness treatment condition in which the monetary stakes are increased by a factor of 10, the task order is varied, subjects' comprehension of the BDM mechanism is reinforced, the question wording is varied, and the efficiency of the sampling method is increased.

Finally, Harrison et al. (2012) estimate structural econometric models in which the value and probability weighting functions are specified with parametric forms that admit dependence on the compound lottery form, i.e. source dependence. Vuong and Clarke tests are used to select the best-fitting model (between EU and RDU), and then within the best-fitting model (RDU) structural tests of Reduction are implemented by testing the $H_0 : \hat{r}\hat{c} = 0$ in $U(x|\text{compound lottery}) = x^{(1-r-rc)}/(1-r-rc)$ and $H_0 : \hat{\gamma}\hat{c} = 0$ in $\omega(p|\text{compound lottery}) = p^{\gamma+\gamma c}$. The authors do not find evidence of source dependence in the 1-in-1 incentive condition, but *do* find evidence of source dependence (Reduction violation) in the 1-in-40 RLI condition. Furthermore, they note that the inference would be reversed – i.e. source dependence (Reduction violation) in the 1-in-1 condition and no source dependence in the 1-in-40 condition – if one were to erroneously assume EU instead of RDU.

Ultimately, it would be desirable to test for alternation bias and Reduction within a structural econometric modeling framework. In its present state of development however, where a parametrically convenient linear specification reflects the state of the art, it is not clear how distortions of conditional probability (i.e. alternation bias) should be incorporated.[32] At the same

[32] As Harrison et al.'s (2012) lotteries are restricted to a maximum of one level of compounding, their estimates of probability weighting function source dependence simultaneously impound both unconditional probability weight-

time, it is not yet clear how to design the tasks required for structural econometric modeling in such a way as to capture or accommodate heterogeneity in δ_i. The trade-off struck in the design of Experiments I–III favors alternation bias based hypotheses and direct revelation of δ_i over parametric statistical modeling. Yet as elaborated in Section 9, results obtained within this framework reinforce Harrison et al.'s (2012) structural econometric finding that the utility function is source dependent in the manner first conjectured by Smith (1969). With careful interpretation, these different methodological approaches prove to be complementary rather than rival.

11 Conclusion

Together Experiments I–III cast doubt on the descriptive validity of Reduction in StPGs. Since Reduction is an axiom that is necessary for EU and especially its application to StPGs, the weak- and strong-form violations uncovered here also impinge upon EU's status as a wholly satisfactory resolution of the St. Petersburg Paradox.

The direction of the present empirical findings is consistent with the operation of alternation bias, which is a subjective distortion of conditional probability, distinct from the major avenues to resolving the St. Petersburg Paradox pursued thus far: outcome distortion (concave utility for money) and probability distortion (unconditional probability weighting). Nevertheless, we also find that alternation bias is not the sole explanator of the preference for reduced-form StPGs over compound-form StPGs. This points to a requirement for experimental designs that can accommodate more elaborate source dependence structures than those contemplated thus far.

The present methodology complements recent advances by Halevy (2007) and Harrison et al. (2012). To our knowledge, the multiple-price list utilized here is the first instrument to facilitate revelation of heterogeneity in δ_i and to enable the bounding of these latent δ_i. Consequently, it allows direct demonstration of conformity with Smith's (1969) dominance precept.

Acknowledgements: We especially wish to thank Peter Wakker, Vernon Smith, Glen Harrison, Yoram Halevy and Andrea Isoni for commenting on this work. Furthermore, this manuscript benefitted greatly from comments by session participants at the 2011 European ESA meeting in Luxembourg, the 2012 ESA International meeting in New York, and the 2012 FUR XV meeting in Atlanta. The usual disclaimer applies.

ing and conditional probability weighting in unknown proportion.

References

Abdellaoui, M., Baillon, A., Placido, L., & Wakker, P. (2011). The rich domain of uncertainty: Source functions and their experimental implementation. *American Economic Review*, 101(2), 695–723.

Asparouhava, E., Hertzel, M., & Lemmon, M. (2009). Inference from streaks in random outcomes: Experimental evidence on beliefs in regime shifting and the law of small numbers. *Management Science*, 55(1), 1766–1782.

Aumann, R.J. (1962). Utility theory without the completeness axiom. *Econometrica*, 30(3), 445-462.

Bardley, N., Cubitt, R., Loomes, G., Moffatt, P., Starmer, C., & Sugden, R. (2010). *Experimental Economics; Rethinking the Rules*. Princeton, NJ: Princeton University Press.

Bar-Hillel, M. (1973). On the subjective probability of compound events. *Organizational Behavior and Human Performance*, 9(3), 396–406.

Bar-Hillel, M., & Wagenaar, W.A. (1991). The perception of randomness. *Advances in Applied Mathematics*, 12(4), 428–454.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, 57(1), 289–300.

Bernasconi, M., & Loomes, G. (1992). Failures of the reduction principle in an Ellsberg-type problem. *Theory and Decision*, 32(1), 77–100.

Birnbaum, M. (1998). Finite versions of the St. Petersburg Paradox. Working paper of the California State University Fullerton Decision Research Center 9-06-98.

Blavatskyy, P.R. (2005). Back to the St. Petersburg Paradox?. *Management Science*, 51(4), 677-678.

Blavatskyy, P.R. (2006). Violations of betweenness or random errors?. *Economics Letters*, 91(1), 34–38.

Blavatskyy, P.R. (2009). Preference reversals and probabilistic decisions. *Journal of Risk and Uncertainty*, 39(3), 237–250.

Budescu, D.V. (1987). A markov model for generation of random binary sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 13(1), 25-39.

Budescu, D.V., & Fischer, I. (2001). The same but different: An empirical investigation of the reducibility principle. *Journal of Behavioral Decision Making*, 14(3), 187–206.

Butler, D.J., & Loomes G.C. (2007). Imprecision as an account of the preference reversal phenomenon. *American Economic Review*, 97(1), 277–297.

Chernoff, H. (1954). Rational selection of decision functions. *Econometrica*, 22(2), 422–443.

Clotfelter, C., & Cook, P. (1993). The 'gambler's fallacy' in lottery play. *Management Science*, 39(12), 1521-1525.

Cox, J.C., Sadiraj, V., & Vogt, B. (2009). On the empirical relevance of St. Petersburg lotteries. *Economics Bulletin*, 29(1), 214–220.

Croson, R., & Sundali, J. (2005). The gambler's fallacy and the hot hand: Empirical data from casinos. *Journal of Risk and Uncertainty*, 30(3), 195–209.

Cubitt, R., Starmer, C., & Sugden, R. (1998). On the validity of the random lottery incentive system. *Experimental Economics*, 1(1), 115–131.

Danan, E. (2008). Revealed preference and indifferent selection. *Mathematical Social Sciences*, 55(1), 24–37.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2), 171–178.

Fishburn, P.C. (1978). Mixture axioms in linear and multilinear utility theories. *Theory and Decision*, 9(2), 161–171.

Fishburn, P.C. (1986). Comment on G. Shafer, "Savage revisited" . *Statistical Science*, 1(4), 492–495.

Gneezy, U. (1996). Probability judgments in multi-stage problems: Experimental evidence of systematic bias. *Acta Psychologica*, 93(1-3), 59–68.

Greiner, B. (2004). The online recruitment system ORSEE 2.0 – A guide for the organization of experiments in economics. University of Cologne Department of Economics working paper.

Grether, D.M., & Plott, C.R. (1979). Economic theory of choice and the preference reversal phenomenon. *American Economic Review*, 69(4), 623–638.

Halevy, Y. (2007). Ellsberg revisited: An experimental study. *Econometrica*, 75(2), 503–536.

Harrison, G.W. (1994). Expected utility theory and the experimentalists. *Experimental Economics*, 19(2), 223–253.

Harrison, G.W., Martinez-Correa, J., & Swarthout, J.T. (2012). Reduction of compound lotteries with objective probabilities: Theory and evidence. Georgia State University Center for the Economic Analysis of Risk working paper WP 2012-05.

Hauser, J.H. (1978). Consumer preference axioms: Behavioral postulates for describing and predicting stochastic choice. *Management Science*, 24(13), 1331–1341.

Jeffreys, H. (1961). *Theory of Probability*. Oxford: Oxford University Press.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291.

Kaivanto, K. (2008). Alternation bias and the parameterization of cumulative prospect theory. In: Abdellaoui, M., Hey, J.D. (Eds.), *Advances in Decision Making Under Risk and Uncertainty*. Heidelberg: Springer-Verlag, pp. 91–107.

Kaivanto, K., & Kroll, E.B. (2012). Negative recency, randomization device choice, and reduction of compound lotteries. *Economics Letters*, 115(2), 263–267.

Kareev, Y. (1995). Positive bias in the perception of covariation. *Psychological Review*, 102(3), 490–502.

Keller, L.R. (1985). Testing of the 'reduction of compound alternatives' principle. *Omega*, 13(4), 349–358.

Luce, R.D. (2001). Reduction invariance and Prelec's weighting functions. *Journal of Mathematical Pscychology*, 45(1), 167–179.

Luce, R.D., & Raiffa, H. (1957). *Games and Decisions*. New York, NY: John Wiley & Sons.

Lydersen, S, Fagerland, M.W., & Laake, P. (2009). Recommended tests for association in 2×2 tables. *Statistics in Medicine*, 28, 1159–1175.

Mandler, M. (2005). Incomplete preferences and rational transitivity of choice. *Games and Economic Behavior*, 50(2), 255–277.

Marschak, J. (1968). Economics of inquiring, communicating, deciding. *American Economic Review*, 58(2), 1–18.

von Neumann, J., & Morgenstern, O. (1947). *Theory of Games and Economic Behavior*, 2nd edition. Princeton, NJ: Princeton University Press.

Pfiffelmann, M. (2011). Solving the St. Petersburg Paradox in cumulative prospect theory: The right amount of probability weighting. *Theory and Decision*, 71(3), 325–341.

Prelec, D. (1998). The probability weighting function. *Econometrica*, 66(3), 497–452.

Prelec, D. (2000). Compound invariant weighting functions in prospect theory. In: Kahneman, D., Tversky, A. (Eds.), *Choices, values and frames*. Cambridge, MA: Cambridge University Press, pp. 67–92.

Rabin, M. (2002). Inference by believers in the law of small numbers. *Quarterly Journal of Economics*, 117(3), 775–816.

Rapoport, A., & Budescu, D.V. (1997). Randomization in individual choice behavior. *Psychological Review*, 104(3), 603-617.

Reichenbach, H. (1934). *Wahrscheinlichkeitslehre: Eine Untersuchung uber die logischen und mathematischen Grundlagen der Wahrscheinlichkeitsrechnung*. Translated as Reichenbach, H. (1949). *The theory of probability; An inquiry into the logical and mathematical foundations of the calculus of probability*. Berkeley, CA: University of California Press.

Rieger, M.O., & Wang, M. (2006). Cumulative prospect theory and the St. Petersburg Paradox. *Economic Theory*, 28(3), 24–55.

Samuelson, P.A. (1952). Probability, utility, and the independence axiom. *Econometrica*, 20(4), 670–678.

Savage, L.J. (1954). *The Foundations of Statistics*. New York, NY: John Wiley & Sons.

Schmidt, U., & Hey, J.D. (2004). Are preference reversals errors? An experimental investigation. *Journal of Risk and Uncertainty*, 29(3), 207–218.

Smith, V.L. (1969). Measuring nonmonetary utilities in uncertain choices: the Ellsberg urn. *Quarterly Journal of Economics*, 83(2), 324–329.

Smith, V.L. (1982). Microeconomic systems as an experimental science. *American Economic Review*, 72(5), 923–955.

Sonsino, D., Benzion, U., & Mador, G. (2002). The complexity effects on choice with uncertanty – experimental evidence. *Economic Journal*, 112(482), 936–965.

Sonsino, D. (2011). A note on negativity bias and framing response asymmetry. *Theory and Decision*, 71(2), 235–250.

Starmer, C., & Sugden, R. (1994). Does the random-lottery incentive system elicit true preferences? An experimental investigation. *American Economic Review*, 81(4), 971–978.

Terrell, D. (1994). A test of the gambler's fallacy: Evidence from pari-mutuel games. *Journal of Risk and Uncertainty*, 8(3), 309-317.

Terrell, D. (1998). Biases in assessments of probabilities: New evidence from greyhound races. *Journal of Risk and Uncertainty*, 17(2), 151–166.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481), 453–458.

Tversky, A., & Kahneman, D. (1992). Advances in Prospect Theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297–323.

Tversky, A., & Wakker, P. (1995). Risk attitudes and decision weights. *Econometrica*, 63(1), 1255–1280.

Wakker, P. (2010). *Prospect Theory for Risk and Ambiguity*. Cambridge, MA: Cambridge University Press.

Wilcox, N.T. (1993). Lottery choice: Incentives, complexity and decision time. *Economic Journal*, 103(421), 1397–1417.