Title: Exploring the consistency of higher-order risk preferences[*]

Authors:

- Alexander Haering, RWI – Leibniz-Institute for Economic Research, Germany, and Fresenius University of Applied Sciences, Germany
- Timo Heinrich, Durham University Business School, UK
- Thomas Mayrhofer, Harvard Medical School, Harvard University, U.S.A., and School of Business Studies, Stralsund University of Applied Sciences, Germany[1]

[*] Manuscript received June 2018; revised March 2019.

[1] We thank the Co-Editor Masaki Aoyagi and two anonymous referees for very helpful comments. We are also grateful to Han Bleichrodt, Jeannette Brosig-Koch, Louis Eeckhoudt, Glenn Harrison, Heike Hennig-Schmidt, Zwetelina Iliewa, Johanna Kokot, Gijs van de Kuilen, Isabel Marcin, Peter Martinsson, Sonja Opper, Richard Peter, Hendrik Schmitz, Jan Siebert, Stefan Traub, Stefan Trautmann, and Toshio Yamagishi for their valuable feedback. We have also benefited from the comments of seminar participants at the Ruhr University Bochum, the Helmut Schmidt University Hamburg, IE Business School Madrid, and the University of Paderborn, as well as those of conference participants at the GfeW Annual Meeting 2015 in Hamburg, the CEAR/MRIC Behavioral Insurance Workshop 2015 in Atlanta, the Experimental Finance conference 2016 in Mannheim, the FUR conference 2016 at the University of Warwick, the INFORMS Annual Meeting 2016 in Nashville, the 6th ACCER International Workshop 2017 in Duisburg, the Tagung des Sozialwissenschaftlichen Ausschusses 2018 in Essen and the International ESA Meeting 2018 in Berlin. Furthermore, we thank Christine E. Looser, Haiying Qin, Weijing Le, Chaoliang Yang and Guanzhong Yang for their support in organizing the experiments. Constantin Carstens, Johanna Rosebrock, Franziska Then and Stephanie Zimmer provided excellent research assistance. Financial support from the BMBF through the IN-EAST School of Advanced Studies (Grant No. 01UC1304) is gratefully acknowledged.

Please address correspondence to: Timo Heinrich, Durham University Business School, Mill Hill Lane, Durham, DH1 3LB, UK; email: timo.heinrich@uni-due.de.

Abstract: This study measures higher-order risk preferences and their consistency. We explore the role of country differences, the variation of stakes, and the framing of lotteries. We observe a robust dichotomous pattern of choice behavior in China, in the USA and in Germany. The majority of choices are consistent with mixed risk aversion or mixed risk-loving behavior. We also find this pattern after a tenfold increase in the stakes. Finally, our results reveal that this pattern is strengthened if the lotteries are displayed in compound rather than reduced form. In a follow-up study we explore potential explanations for this framing effect.

JEL Classification: C91, D81

Keywords: Mixed risk aversion, prudence, temperance, higher-order risk preferences

Shortened title*:* Higher-order risk preferences

1 Introduction

Within the expected utility framework, most of the commonly used utility functions (e.g. ln(x) and $x^{0.5}$) imply "mixed risk aversion," which means that the derivatives of the utility functions exhibit alternating signs (see Brockett and Golden 1987, Caballé and Pomansky 1996). Therefore, these utility functions assume second-order risk aversion ($U^{\mathrm{II}} < 0$), as well as higher-order risk preferences, such as prudence ($U^{\mathrm{III}} > 0$) – also called third-order risk aversion – and temperance ($U^{\mathrm{IV}} < 0$) – also called fourth-order risk aversion. More recently, higher-order risk preferences have also been defined as preferences over binary lotteries by Eeckhoudt and Schlesinger (2006). These model-free definitions do not require assumptions as far reaching as expected utility theory and lend themselves to experimental investigation. Based on the binary lotteries by Eeckhoudt and Schlesinger (2006), Eeckhoudt, Schlesinger and Tsetlin (2009) define mixed risk aversion as a preference for combining "good" outcomes with "bad" ones. Crainich, Eeckhoudt and Trannoy (2013) then introduced the concept of "mixed risk-loving" behavior, which they define as a preference for combining "good" outcomes with "good" ones. In an expected utility framework, this would imply a utility function for which all the derivatives are strictly positive.[1] While mixed risk averters are second-order risk-averse, prudent and temperate, mixed risk-loving individuals are second-order risk-loving, prudent and intemperate. To put it more generally, mixed risk averters coincide in their choices with mixed risk lovers in the odd orders (e.g., in prudence) but differ in the even orders (e.g., in temperance).

Recently, Deck and Schlesinger (2014) used an economic laboratory experiment to study mixed risk-averse and mixed risk-loving behavior. They made two major observations: First, in their data a non-negligible minority of individuals make consistently second-order risk-loving choices. Second, in line with the theoretical prediction, they observed a consistent pattern of mixed risk-averse and mixed risk-loving behavior. In this paper, we study whether this dichotomy can be regarded as a widespread

[1] As Ebert (2013) points out, neither property follows from risk-loving preferences per se.

pattern explaining the heterogeneity of choices under risk. We conduct a large-scale experiment with a total of 605 participants and add to the literature by measuring higher-order risk preferences (up to order 6) across different countries by employing distinct subject pools in China, in the USA and in Germany. Furthermore, we contribute to previous findings by studying the effect of a tenfold increase in the stakes and the effect of a straightforward change in the framing of the lotteries. In a follow-up study with an additional 224 participants, we explore the influence of framing further.[2]

Previous experimental studies by Deck and Schlesinger (2010), Ebert and Wiesen (2011, 2014), Maier and Rüger (2012) and Noussair, Trautmann, and van de Kuilen (2014) suggest that a majority of aggregate choices are in line with prudence and – with the exception of the studies by Deck and Schlesinger (2010) and Baillon, Schlesinger and van de Kuilen (2017) – in line with temperance (see

[2] Deck and Schlesinger (2014) were not only the first to study mixed risk-averse and mixed risk-loving behavior experimentally, but they were also the first to assess risk preference of orders greater than 4. Risk aversion of order 5 (called "edginess" by Lajeri-Chaherli 2004) or even 6 (named "bentness" by Miles S. Kimball at a conference to honor Louis Eeckhoudt in 2012) have so far rarely been studied. However, utility functions typically imply assumptions across *all* orders of risk aversion, and there is no compelling reason why these assumptions should not be subject to empirical scrutiny. In addition, in an intertemporal consumption problem, an increase in the *n*-th order risk of future income yields an increase in savings if and only if *n*+1-th order risk aversion is present (as shown by Eeckhoudt and Schlesinger 2008, in an expected utility theory framework). In other words, anyone who thinks that *n*-th order risk matters to decision makers will care about their *n*-th *and* (*n*+1)-th order risk attitudes in an intertemporal setting. Also note that in Deck and Schlesinger's (2014) design, the elicitation of higher-order risk attitudes requires rather complex lotteries. We believe that, because of this complexity, assessing behavior in the respective lotteries with fifth and sixth order variations of risk is quite useful because it provides an even tougher test for the theoretical predictions. Very recently, Ebert, Nocetti and Schlesinger (2017) proposed an alternative method to elicit higher-order risk preferences. Their theory is based on greater mutual aggravation and does not require complex doubly-compounded lotteries.

Appendix A1 for a more detailed comparison).[3] In addition, based on representative data from the Netherlands, Noussair, Trautmann and van de Kuilen (2014) find that lottery choices are correlated with behavior in the field. The effects they observe are in line with theoretical predictions. Prudent lottery choices are associated with greater wealth, a greater likelihood of having a savings account and a lower likelihood of credit card debt. Temperate lottery choices are associated with less risky investment portfolios.

Ideally, considering individual differences in higher-order risk preferences will help to build more realistic economic models. In lifecycle savings models, for example, prudence and temperance determine how current savings are influenced by the riskiness of future income (Kimball 1990, 1992). Other areas in which higher-order risk preferences have been theoretically shown to impact behavior include auctions (Esö and White 2004), bargaining games (White 2008), research and development expenditures (Nocetti 2015), prevention (Eeckhoudt and Gollier 2005, Courbage and Rey 2006, 2016, and Peter 2017) and medical decision making (Eeckhoudt 2002, Felder and Mayrhofer 2014, 2017). However, what still needs to be established is the degree to which the previous findings on individual differences in higher-order risk preferences are robust and sufficiently general in different contexts.

During recent decades, it has become evident that many behavioral patterns identified in Western subject populations are by no means universal human traits. For example, economists discovered that human behavior in strategic interaction varies widely across societies, with aggregate behavior covering virtually the complete strategy space (see, e.g., Roth et al. 1991, Oosterbeek, Sloof and van de

---

[3] Baillon, Schlesinger and van de Kuilen (2017) also measure ambiguity prudence and ambiguity temperance based on the preference conditions by Baillon (2017). Other experimental studies have observed that higher-order risks influence precautionary savings (Bostian and Heinzel 2012), as well as behavior in auctions (Kocher, Pahlke and Trautmann 2015) and medical treatment and prevention decisions (Krieger and Mayrhofer 2012, 2017). Moreover, higher-order risk preferences have also been studied experimentally in social settings (Heinrich and Mayrhofer 2018), across multiple domains (Ebert and van de Kuilen 2015, Deck and Schlesinger 2017) and with children (Heinrich and Shachat, 2018). For a detailed review of the recent experimental literature on higher-order risk preferences, please see Trautmann and van de Kuilen (2018).

Kuilen 2004 and Herrmann, Thöni and Gächter 2008). With regard to second-order risk aversion, research in economics and psychology has provided evidence of differences in risk attitudes across countries (see, e.g., Weber and Hsee 1998, Vieider, Chmura et al. 2015 and Vieider, Lefebvre et al. 2015) and across stake sizes (see, e.g., Binswanger 1980, 1981 and Kachelmeier and Shehata 1992b). Measuring risk preferences in different international subject pools provides a tougher test of the generalizability of a theory. Furthermore, experimental research in economics has often been criticized for using small samples and small stakes (e.g., Levitt and List 2007). Conducting the experiment in three countries provides us with a larger aggregate sample. It also allows us to exploit differences in purchasing power to conduct high stakes experiments for (relatively) low cost.

Additionally, it has been observed that displaying lotteries in a reduced rather than compound form may impact the degree of second-order risk aversion (see, e.g., Abdellaoui, Klibanoff and Placido 2015 and Harrison, Martínez-Correa and Swarthout 2015). With respect to mixed risk-averse and mixed risk-loving behavior, Deck and Schlesinger (2014) conjecture that the compound presentation "admittedly also facilitates viewing the problem as 'combining good with bad' or 'combining good with good,' rather than presenting the lotteries in a reduced form, which might obfuscate this interpretation" (Deck and Schlesinger 2014, 1921ff).

Even though subject pool differences, stake size and the framing of lotteries have been studied with respect to second-order risk aversion, there has been very little work on higher-order risk preferences. Since second-order risk aversion and higher-order risk preferences are related theoretically, as well as empirically, we expect that these factors also influence higher-order risk preferences.

The results we report in this paper suggest that a majority of people, across national contexts are second-order risk-averse. Moreover, we confirm the main observations by Deck and Schlesinger (2014): A considerable proportion of people are second-order risk-loving and choices can be explained rather well by a dichotomy of preference types. In total, up to 62% of the participants can be classified as mixed risk-averse and up to 14% as mixed risk-loving. We present the first comparison of higher-order risk preferences across countries (i.e., China, USA and Germany) and under high stakes. Our study reveals that mixed risk aversion is somewhat more prevalent among Germans than among Chinese and that the dichotomy of preference types persists when stakes increase. We also discover that the dichotomy can be strengthened through the framing of the lottery.[4] When we display the lotteries in compound rather than reduced form, we observe significantly more prudent and temperate behavior *within* the same subjects. A follow-up study reveals that the justifications for specific lottery choices differ significantly between both types of framings.

2 Theoretical background and hypotheses

A Theoretical background

In this section we present the theoretical background to our experiment, which follows Deck and Schlesinger (2014). In their experiment, they use a variety of lotteries, which are based on the theoretical work by Eeckhoudt and Schlesinger (2006), Eeckhoudt, Schlesinger and Tsetlin (2009) and Crainich, Eeckhoudt and Trannoy (2013). The lotteries are binary with equal probabilities, i.e. $[x, y]$ denotes a lottery with a 50-50 chance of receiving either outcome x or outcome y. However, x and y might themselves be lotteries.

In the following we will refer to risk aversion of n-th degree as "n-RA" and to risk loving behavior of n-th degree as "n-RL". Figure 1 shows lotteries for eliciting risk preferences up to order 4. Let us assume that W is the initial wealth of an individual, with $W > 0$, and that k_1 and k_2 are sure losses, with $k_1 > 0$ and $k_2 > 0$. Furthermore, ε and δ are independent zero-mean background risks, i.e., lotteries with an expected value of zero.

The first row in Figure 1 illustrates a second-order risk aversion, i.e. 2-RA, task in which risk aversion is a preference for disaggregating harms, i.e., the sure losses k_1 and k_2. Disaggregating these two

[4] A very recent study by Deck and Schlesinger (2017) that was conducted in parallel with ours makes a similar observation with respect to the framing of lotteries (see Sections 2.B and V). Furthermore, they replicate the observations made in Deck and Schlesinger (2014) and also consider choices when the payoffs are non-monetary.

“bad” payoffs reduces the spread between the two possible outcomes. This corresponds to a lower variance, which is a necessary assumption for lower second-order risk. Lottery *A2* has a greater spread (and thus variance) than lottery *B2*. A risk-averse individual would choose lottery *B2* over lottery *A2* and, vice versa, a risk-loving individual would choose lottery *A2* over lottery *B2*. In other words, while both regard a sure loss as “bad,” second-order risk is only “bad” for the risk-averse individual but “good” for the risk-loving individual.

The second row in Figure 1 shows a 3-RA task. In this case the sure loss k_1 is replaced by a zero-mean background risk ε. Eeckhoudt and Schlesinger (2006) define 3-RA as a preference for disaggregating a sure loss and an additional zero-mean background risk. Therefore, a 3-RA individual would prefer lottery *B3* over lottery *A3*, while a 3-RL individual would prefer *A3*. Note that in this case a risk-averse and a risk-loving individual would agree that avoiding a sure loss is “good,” but both differ in their judgment of the zero-mean background risk. The upper arm of *B3* yields a combination of “good” with “bad” for risk-averse individuals and a combination of “good” with “good” for risk-loving individuals.

The third row in Figure 1 exemplifies a 4-RA task. Now the second loss k_2 is also replaced by a second zero-mean background risk δ (which is independent of ε). Eeckhoudt and Schlesinger (2006) define 4-RA as a preference for disaggregating two independent zero-mean background risks. Thus, a 4-RA individual would prefer lottery *B4* over lottery *A4*, while an 4-RL individual would prefer lottery *A4* over lottery *B4*. The lower arm of *A4* yields a combination of “bad” with “bad” for risk-averse individuals. However, for risk-loving individuals this represents a combination of “good” with “good.”

Deck and Schlesinger (2014) now define mixed risk aversion as a preference for combining “good” with “bad”, and mixed risk-loving behavior as a preference for combining “good” with “good”. This yields a pattern in which both types coincide in their lottery choices of odd orders but differ in even orders.

For orders higher than four, Deck and Schlesinger (2014) use a more general approach that is based on the theoretical work by Eeckhoudt, Schlesinger and Tsetlin (2009) and is illustrated in Figure 2.

Following Deck and Schlesinger (2014) we consider a pair of random variables [X_1, Y_1], where Y_1 has more n-th order risk than X_1. According to Ekern (1980), Y_1 has more n-th order risk than X_1 if X_1 is n-th order stochastic dominant compared to Y_1 and the two random variables have the same n – 1 moments (for $n > 1$). Moreover, let us consider a second pair of random variables [X_2, Y_2] where Y_2 has more m-th order risk than X_2. All random variables are statistically independent of each other. Eeckhoudt, Schlesinger, and Tsetlin (2009) show that, for this setting, the 50-50 lottery [$W + X_1 + X_2$, $W + Y_1 + Y_2$] has more ($m + n$)-th order risk than the 50-50 lottery [$W + X_1 + Y_2$, $W + Y_1 + X_2$]. An individual who prefers lotteries with lower ($m + n$)-th order risk is "($m + n$)-th order risk-averse." An individual who is ($m + n$)-th order risk-averse would choose lottery *B* over lottery *A* in Figure 2. This approach is more general, and can be used for all orders.

Moreover, from the viewpoint of a mixed risk-averse individual both random variables X_i can be considered as relatively "good," and both random variables Y_i as relatively "bad." Lottery *A* in Figure 2 shows a 50-50 chance of receiving either "good" with "good" (upper lottery arm) or "bad" with "bad" (lower lottery arm), while lottery *B* shows a combination of "good" with "bad" in both lottery arms. Lottery *B* therefore apportions the good and bad outcomes. A mixed risk-averse individual dislikes risk of any order and therefore always prefers lottery *B*. However, a mixed risk-loving individual only dislikes risk of odd orders and therefore only prefers lottery *B* if $m + n$ is odd (and lottery *A* otherwise).

<<Figure 1 here>>

<<Figure 2 here>>

B Hypotheses

Cross-country differences.—In selecting China, the USA, and Germany, we aim to strike a balance between the economic relevance of the subject pools and their heterogeneity. On the one hand, these countries are those with the highest population on their respective continents, as well as the largest economies in terms of total GDP. On the other hand, these countries differ culturally. In one of the first studies analyzing second-order risk preferences in an international comparison, Hsee and Weber (1999) found that Chinese people are more likely to take risks than Americans with respect to hypothetical payoffs (see also Weber and Hsee 1998 and Statman 2008). They explain their findings based on the cultural trait of individualism as introduced by Hofstede (1980). According to the cushion hypothesis, people from China are less individualistic and thus less likely than Americans to deal on their own with the consequences of risky decisions. In fact, the most recent data on the cultural dimensions of 69 countries by Hofstede, Hofstede and Minkov (2010) also reveals that Germany, the USA and China differ widely with respect to individualism: the USA ranks 1st, Germany 17th and China 52nd out of these 69 countries.

Hofstede (1980) originally identified power distance, masculinity and uncertainty avoidance next to individualism as dimensions that characterize a culture. Of these four dimensions, uncertainty avoidance has been found to be associated with risk preferences: In a comprehensive survey covering 53 countries, Rieger, Wang and Hens (2015) observe that uncertainty avoidance is associated with higher second-order risk aversion. The data on uncertainty avoidance suggests smaller disparities between the countries: Germany ranks 40th, the USA 57th and China 63rd among the 69 countries. Thus, based on individualism and uncertainty avoidance, we would expect Chinese people to be the least risk-averse.[5]

---

[5] Of course, China, the USA and Germany also differ in economic, social and political measures that may correlate with risk preferences. The existing evidence is broadly consistent, with Chinese people being the least

While there are many international comparison studies on second-order risk aversion (see Haering and Heinrich 2017 for an overview), nothing is known about differences in higher-order risk preferences. We follow Deck and Schlesinger's (2014) argument and assume that human behavior is driven by a basic tendency to combine either "good" with "bad" or "good" with "good." Under this assumption, one may assume that the observed differences in second-order risk aversion indicate differences in the distribution of the two types between subject pools. Accordingly, based on the evidence on second-order risk aversion we expect less mixed risk-averse and more mixed risk-loving behavior in China:

Hypothesis 1: *Chinese people make fewer mixed risk-averse and more mixed risk-loving choices than Americans and Germans.*

Differences in stake sizes.—Markowitz (1952) was among the first who argued that second-order risk preferences could change with increasing wealth. He suggested that the utility function, for levels of wealth above present wealth, is first convex and then concave. Therefore, Markowitz assumed that

---

second-order risk-averse. Falk et al. (2015) conducted the first representative survey comparing economic preferences around the globe. The authors correlate the average risk attitude in 76 countries with other country characteristics. Based on the five (weakly) significant correlations they observe, second-order risk aversion should be greatest in Germany. With regard to three measures (degree of redistribution, life expectancy and degree of inequality) we would expect Chinese people to be the least risk-averse. With regard to two other measures (rigidity of employment laws and number of homicides per capita), we would expect Americans to be the least risk-averse. Rieger, Wang and Hens (2015) find a positive correlation between (log) GDP per capita and second-order risk aversion in the gain domain. A similar observation is made in the large experimental study by Vieider, Lefebvre et al. (2015), using monetary incentives. These authors elicit the risk preferences of students in 30 countries and observe a positive correlation between (log) GDP per capita and second-order risk aversion. Based on these correlations, we would expect Chinese people to be the least risk-averse compared to people from Germany and the USA. Further experimental comparisons between China and Western countries have been conducted by Kachelmeier and Shehata (1992a) and Ehmke, Lusk and Tyner (2010).

individuals are second-order risk-loving when the stakes are small and second-order risk-averse when the stakes are high. Pratt (1964) and Arrow (1965), who introduced—independently of each other—the now famous Arrow–Pratt coefficients as measurements for absolute and relative risk aversion, also assumed increasing risk aversion with increasing wealth. Similar assumptions were made by Eeckhoudt and Kimball (1992) and Kimball (1992) regarding 3-RA and by Eeckhoudt, Gollier, and Schlesinger (1996) and Gollier and Pratt (1996) regarding 4-RA.

Empirically testing these theoretical assumptions has been a challenge, since it requires a considerable variation of the stake size. The most common approach to this is to conduct experiments in developing countries, where large monetary incentives can be provided at lower cost than in developed countries. These studies typically observe more second-order risk-averse choices with higher stakes when eliciting risk preferences using binary gambles (Binswanger 1980, 1981, Grisley and Kellog 1987, Wik et al. 2004) or tasks based on eliciting certainty equivalents (Kachelmeier and Shehata 1992a, Fehr-Duda et al. 2010). However, similar observations have been made in developed countries. Holt and Laury (2002, 2005) elicit second-order risk aversion (using a price list format) in the USA. They find that (relative) risk aversion increases with real stakes but not with hypothetical stakes.

There are only two experimental papers that consider the relationship between stake size and higher-order risk preferences. Deck and Schlesinger (2010) confront subjects with ten choices between lottery pairs. These lotteries have an overall expected payoff of $25.80. The comparison of two choices allows them to study the influence of a fivefold increase in the stake on 3-RA; two more comparisons of lottery choices allow them to study the influence on 4-RA. Deck and Schlesinger (2010) find weak support for the hypothesis that 3-RA preferences are more pronounced when stake sizes are higher (approximately one third of their subjects changed their behavior when the stake size increased, and 70% of them made more 3-RA choices). Although they find mostly 4-RL behavior in their subject population, 4-RL behavior is less common when the stakes are higher.

Noussair, Trautmann and van de Kuilen (2014) study the prevalence of 3-RA and 4-RA in a laboratory experiment, as well as in a large representative sample of the Dutch population. In expectation, participants in their real payoff treatments earn €7 (because the lotteries have an expected value of €70 but only one in ten participants is paid). Noussair, Trautmann and van de Kuilen (2014) find that 2-RA and 4-RA increase when the hypothetical stakes are increased (from €70 to €10,500). They find no significant difference between the real monetary payoff treatments and a treatment with hypothetical payoffs (in which lotteries have an expected value of €70, but no one is paid). They also do not find any stake size effect for 3-RA. However, in a direct test of higher-order risk preferences and their relationship to an endowment to risk ratio, Noussair, Trautmann and van de Kuilen (2014) find decreasing absolute 3-RA and decreasing absolute 4-RA. Moreover, their estimated parameters regarding their expo-power utility functions show increasing relative 3-RA and increasing relative 4-RA.

While the theory suggesting two simple types of preferences (for either (*i*) combining "good" with "bad" or (*ii*) combining "good" with "good") does not predict a change of type if the stake size changes, the empirical evidence indicates that (relative) 2-RA increases with higher stakes. In addition, there is limited evidence indicating an increase in (relative) 3-RA and (relative) 4-RA. We formulate our second hypothesis accordingly:

Hypothesis 2: *The number of mixed risk-averse choices increases and the number of mixed risk-loving choices decreases when the stake size increases.*

Differences through displaying reduced rather than compound lotteries.— According to most theories of decision making, displaying actuarially equivalent lotteries as compound or reduced lotteries does not influence choices.

With respect to second-order risk aversion, it has been known for a while that reduced lotteries are often valued differently than compound lotteries. Early experiments in psychology report, for example, that people overestimate the joint probabilities in compound lotteries (see, e.g., Slovic 1969, Bar-Hillel 1973 and the overview in Budescu and Fischer 2001). In economics, the observation that reduced lotteries are valued differently has been used to explain the pattern of preference reversals as observed by Lichtenstein and Slovic (1971), Lindman (1971) and Grether and Plott (1979), as well as ambiguity aversion as identified by Ellsberg (1961). For example, Segal (1988) shows that a violation of the reduction of compound lotteries axiom (ROCL) can generate preference reversals even if the independence axiom holds.[6] Segal (1987, 1990) also shows that ambiguity aversion can be explained by relaxing the ROCL and applying Quiggin's (1982) rank-dependent utility model.

Higher-order risk preferences are typically elicited using compound lotteries. To our knowledge, only Maier and Rüger (2012), Deck and Schlesinger (2017) and Baillon, Schlesinger and van de Kuilen (2017) have used reduced lotteries to elicit higher-order risk preferences. In the gain domain, Maier and Rüger (2012) observe 55% of choices to be 2-RA, 60% to be 3-RA and 58% to be 4-RA.[7] These percentages are at the lower end of the range of observed frequencies in other studies (see also Table A1 in the Appendix) and thus support the conjecture by Deck and Schlesinger (2014), that the reduced form may obfuscate the "good" with "bad" or "good" with "good" interpretation. In a study conducted in parallel to ours, Deck and Schlesinger (2017) find a significant framing effect between the compound and the reduced presentation of the lotteries. They observe less 4-RA and 5-RA in re-

---

[6] Consider a two-stage lottery and a one-stage lottery yielding the same prizes as the two-stage lottery with the probabilities multiplied out. The ROCL then states that a decision maker is indifferent between these two lotteries (see Samuelson 1952). Note that these theoretical results are directly related to the so-called "random lottery incentive mechanism," i.e., the random selection of one of several lotteries for paying subjects in experiments, while treating choices within lotteries as if made in isolation. This is done to elicit preferences across multiple lotteries in an incentive-compatible way while keeping wealth constant. This procedure has become the norm in experimental economics (Baltussen et al. 2012). It is typically justified with reports of small or unsystematic differences between behaviors under different payment protocols (see, e.g., Starmer and Sugden 1991, Beattie and Loomes 1997, Cubitt, Starmer and Sugden 1998). Recently, however, the random lottery incentive mechanism has been criticized by Harrison and Swarthout (2014), Harrison, Martínez-Correa and Swarthout (2015), and Cox, Sadiraj and Schmidt (2015). They mainly point out the logical inconsistency in assuming the independence axiom holds when paying based on the random lottery mechanism, while taking violations of the independence axiom across lottery choices at face value.

[7] Loss and gain framings have been compared previously when eliciting higher-order risk preferences, with little or no difference being reported (Deck and Schlesinger 2010, Maier and Rüger 2012).

duced lotteries but no difference in the frequency of 3-RA choices when lotteries are displayed in a reduced rather than a compound form. Furthermore, in a recent study Baillon, Schlesinger and van de Kuilen (2017) also elicit preferences using reduced lotteries and find little 4-RA: In their study, 84% of choices are 2-RA, 71% are 3-RA and 43% are 4-RA.

In summary, the presentation of lotteries in a reduced rather than a compound form might influence choices, if decision makers violate the independence axiom or the ROCL. Based on the prior empirical results and the conjecture by Deck and Schlesinger (2014), we expect less mixed risk-averse and less mixed risk-loving choices when using a reduced framing:

Hypothesis 3: *The number of mixed risk-averse choices and the number of mixed risk-loving choices decrease when lotteries are displayed in a reduced rather than a compound framing.*

3 Experimental design

A Elicitation method

Our elicitation method follows Deck and Schlesinger (2014) that comprises 38 tasks (see Deck and Schlesinger 2014, 1922ff and Online Appendix O1). Each of these tasks involves choosing between Option A and Option B. Examples of the 2-RA, 3-RA, and 4-RA lotteries (i.e., lotteries of orders 2, 3 and 4) as presented to the participants in compound form are shown on the left-hand side of Figure 3. Each option involves different amounts of money, and each 50-50 lottery is represented as a circle with a line through the middle.

For example, Option A in the 2-RA task (order 2, task 6) involves a 50-50 chance of winning either 10 ECU or 20 ECU (where ECU stands for experimental-currency-units; see the next section and Table 1 for the exchange rate of ECU to the local currency). Following Deck and Schlesinger (2014), all outcomes are shown in the domain of gains. Let us assume that $W = 20$ and $k_1 = k_2 = 5$, where W denotes wealth and k_1 and k_2 (certain) losses that are subtracted from wealth (see Figure 1 in Section

2.A). In Option A, 10 ECU represents $W - k_1 - k_2 = 10$ while (15 + 5) ECU represents the initial wealth $W = 20$. Option B represents the lottery where the harms are disaggregated, i.e. $[W - k_1, W - k_2]$. In this example, this corresponds to a sure outcome of 15 ECU. Both lotteries have the same expected value of 15 ECU. However, Option A is risky while Option B is not. Thus, a 2-RA individual should choose the certain option over the risky one when the expected values are the same.

In the 3-RA task (order 3, task 11), the outcomes of a 50-50 lottery may contain another lottery. For example, Option A involves a second lottery with a 50-50 chance of winning either -2 or 2 ECU. Thus, the participant has a 25% probability of winning 5 – 2 = 3 ECU, a 25% probability of winning 5 + 2 = 7 ECU, and a 50% probability of winning 10 ECU. Since $[X, Y]$ denotes a lottery where there is a 50-50 chance of receiving X and a 50-50 chance of receiving Y, then Option A can also be written as [5 + [-2, 2], 10]. Let us assume that $W = 10$ and $k_2 = 5$. Moreover, the sure loss k_1 is replaced by a zero-mean background risk ε which itself is a lottery (here [-2, 2]). Then Option A corresponds to $[W - k_2 + \varepsilon, W]$ and Option B to $[W - k_2, W + \varepsilon]$. Eeckhoudt and Schlesinger (2006) define 3-RA as a preference for disaggregating a sure loss and an additional zero-mean background risk. Therefore, a 3-RA individual should prefer Option B over Option A.

In the 4-RA task lottery (order 4, task 21), the outcomes in a 50-50 lottery may contain not just one but two other lotteries. The example shown is a composition of (2+2)-th-order risk, since Option A can be written as [[1, 16] + [1,16], [5,12] + [5,12]] and Option B as [[5,12] + [1, 16], [1, 16] + [5,12]]. As the lottery [1, 16] has more 2-nd-order risk than the lottery [5, 12], the different compositions of both lotteries differ in their 4-th-order risk. An individual who prefers lotteries with lower 4-th-order risk is 4-RA and would choose Option B over Option A. Because the outcomes of option B are composed of the more risky and the less risky lottery, it generates less 4-th-order risk than option A (cf. Section 2.A).

The right-hand side of Figure 3 also shows the corresponding reduced lottery pair for each compound lottery pair. The reduced lotteries can be derived by multiplying out the probabilities of the potential outcomes. The resulting lottery is actuarially equivalent to the compound lottery – that is, it yields the same probability distribution over outcomes.

<<Figure 3 here>>

B Experimental treatments

We initially conducted an economic laboratory experiment with sessions in China, the USA and Germany. Subjects faced 38 tasks in randomized order, and one of the tasks was randomly selected for payment. In each task, the position (left or right) of the two lotteries was determined randomly. Subjects had to choose between them, revealing their risk preference (the instructions are shown in Appendix A2).[8]

[8] Paying one randomly determined task adds another layer of compounding to the lotteries. Following Deck and Schlesinger (2014), we nevertheless used the random payment technique, because the subjects' wealth is not influenced during the course of preference elicitation and because this allows for straightforward comparison with previous studies on higher-order risk preferences, all of which use this method (see Appendix A1 for more details). Furthermore, collecting multiple lottery choices from each subject is essential for answering our research questions. Eliciting only one decision per subject would not allow us to identify mixed risk-averse or mixed risk-loving types. Lastly, as Azrieli, Chambers and Healy (2018, 1) point out, the random payment technique "is essentially the only incentive compatible mechanism." In all sessions, the elicitation of lottery preferences was preceded by four control questions. These control questions were also used by Deck and Schlesinger (2014). The subjects were asked to state the potential payoffs in two lotteries, as well as the maximum and minimum payoffs of a compound lottery, as in Deck and Schlesinger (2014) and as shown in Online Appendix O2. All subjects were able to answer these four questions correctly (see Online Appendices O4, O6 and O8 for more details). The elicitation of lottery preferences was followed by the administration of a questionnaire containing basic demographic questions and questions to determine whether the participant had migrated to the current country (these questions were similar to those used by Sutter et al. 2013).

The treatments, the orders of the lotteries and the number of subjects are shown in Table 1. Lottery pairs that were displayed in the original compound framing are identified by the suffix "C." Reduced lottery pairs are identified by the suffix "R." In addition, Table 1 includes the country-specific exchange rate regarding the experimental-currency-units (ECU) (as explained in the following section) and average payoffs in the local currency.

In order to investigate the effects of the stake size, we increased the payoff tenfold for 48 additional Chinese subjects. The participants in the CHN 10x treatment participated in the same sessions as the Chinese subjects with regular payment. This allowed us to randomize the assignment of Chinese subjects to treatments.

In order to investigate the effect on choices of compound and reduced lotteries, we ran additional sessions in Germany (Compound & Reduced). All of the 143 participants faced the original choices in order 1 and order 2 (and with the exception of one lottery in order 1, none of these were displayed in compound framing). Each subject faced lotteries of two additional orders in the original compound framing, as well as in the reduced framing. This allows us to compare the differences in behavior towards compound and reduced lotteries of orders 3 to 6 *within* subjects. All six combinations were run in each session and subjects were randomly assigned to orders.

<<Table 1 here>>

To shed some more light on the effects of framing, we conducted a Follow-up Experiment in Germany with 224 subjects divided randomly into four different conditions in a 2x2 between-subjects design. We confronted participants with an incentivized 3-RA lottery choice or an incentivized 4-RA lottery choice. In both conditions approximately half of the participants saw the respective lottery pair in the compound framing (58 in 3-RA and 54 in 4-RA), while the others saw it in the reduced framing

(56 in 3-RA and 56 in 4-RA). Participants in all four conditions were matched into groups of two and had to make a choice between two lotteries they were facing.[9]

Adapting an experimental design by Burchardi and Penczynski (2014) to lottery choices, our Follow-up Experiment incentivizes decision makers to reveal the reasoning behind their choice. More specifically, participants were asked to send one written free-form message, together with their preferred choice, to the other participant in their group. Subjects knew that their choices could be revised after both pieces of information were exchanged. The instructions stated "*Before you enter your final decision, you have the opportunity to influence the final choice of your partner: Before the decisions are entered, you will send a preferred choice together with a text message to your partner.*" Subjects also knew that the final choice of one member of the group would be randomly selected after both members had entered their final decision. The final lottery choice of the selected member would determine the payoffs of both. Thus, the message was the only way to influence the other group member, who decided on the payoff-relevant lottery with a probability of one half.

We developed a classification scheme to analyze the content of these messages. This classification scheme was based on prior considerations and a non-incentivized survey that we reported in our working paper Haering et al. (2017). Two research assistants who were unaware of the research questions and not involved in any other experimental studies first coded the old survey data, based on the existing classification scheme. Then any discrepancies in their classification were discussed with one of the authors to clarify misunderstandings. In addition, examples for each content category were selected. These examples served the coders as a reference during the classification of the 224 messages from the Follow-up Experiment, which they coded independently.

---

[9] We selected the 3-RA lottery, in which we observed the largest share of preference reversals within-subjects in the experiments described above (subject to having three different potential outcomes). In addition, we selected the 4-RA lottery, in which we observed the largest share of preferences reversals. These are tasks 11 and 21 shown in Figure 3 in Section 3.A. For both lottery pairs, we randomly varied the position of the more 3-RA or more 4-RA option (left or right).

In all experiments, we also implemented the Cognitive Reflection Test (CRT), as well as the Berlin Numeracy Test (BNT). The CRT consists of three questions and was developed by Frederick (2005) to assess the ability to resist reporting the response that first comes to mind. He finds that this measure is correlated with different measures of cognitive ability and varies widely between American universities. In his study, those who answer more questions correctly are also less second-order risk-averse in the gain domain. It has been reported that higher cognitive ability is associated with lower second-order risk aversion (e.g., by Burks et al. 2009, Dohmen et al. 2010). Therefore, we concluded that this test may capture differences in risk taking that are due to differences in cognitive ability between our subject pools. However, note that Noussair, Trautmann and van de Kuilen (2014) find no such relationship, although in their student sample those who score more highly on the CRT are significantly more 3-RA. The BNT was developed by Cokely et al. (2012) and consists of four questions that aim to assess statistical numeracy and risk literacy. Cokely et al. report that the BNT successfully discriminates between participants on the basis of their numeracy in 15 countries, including China, the USA and Germany. Furthermore, they find that the BNT is highly predictive of the ability to make a correct assessment of the everyday risks associated with consumption, health or medical choices.

The sessions of our experiments were conducted at the experimental lab at Nankai University in Tianjin (China), at CLER at Harvard Business School in Boston (USA) and at the elfe laboratory at the University of Duisburg-Essen in Essen (Germany). No subject participated in more than one session. The experiment was computerized and programmed using zTree (Fischbacher 2007). Screenshots are provided in Online Appendix O2.

C Experimental conditions across countries

In order to create similar conditions in the CHN, USA and GER treatments, we followed best practice as described below. To minimize currency effects, the payoffs in ECU were the same in all sessions, but the exchange rate for one ECU was different in every location (see Bohnet et al. 2008, Herrmann, Thöni and Gächter 2008, and Özer, Zheng and Ren 2014 for similar approaches). We selected exchange rates by putting equal weight on the UBS Prices & Earnings survey (UBS 2014) data (this measure is also used by Özer, Zheng, and Ren 2014), and the country-level purchasing power parity provided by the OECD (2015) (this measure is also used by Roth et al. 1991, Buchan and Croson 2004, and Ehmke, Lusk and Tyner 2010). This procedure led to payments that were inside the feasible bandwidth for subject payments in Tianjin and Boston but were somewhat higher than the usual average payoff in Tianjin and somewhat lower than the usual average payoff in Boston. Therefore we adjusted the payments by 5% in the direction of the usual average payoff.[10]

In order to minimize potential experimenter effects, all experimenters followed the same detailed protocol in all countries (see, e.g., Roth et al. 1991, Buchan and Croson 2004, and Herrmann, Thöni and Gächter 2008 for similar approaches). The experiments in China and in the USA were conducted by local experimenters who also spoke German. The two local experimenters also conducted one session each in Germany, which allowed us to control for idiosyncratic experimenter effects (Bohnet et al. 2008, Özer, Zheng and Ren 2014). These measures have also been advocated by Roth et al. (1991). As an additional measure of control, one lead experimenter from Germany was present (but not visible to subjects) to oversee the procedures in China and in the USA (see Buchan and Croson 2004, and Herrmann, Thöni and Gächter 2008 for a similar approach). To ensure that the instructions were similar, we only used written instructions. These, along with all the computer pages, were translated using the back translation procedure (Brislin 1970). This procedure is now commonly applied in cross-cultural research in economics (see, e.g., Buchan and Croson 2004, Bohnet et al. 2008, Herrmann, Thöni and Gächter 2008, Ehmke, Lusk and Tyner 2010 and Özer, Zheng and Ren 2014).

---

[10] There was no reliable data on purchasing power available for Tianjin, Boston and Essen. Thus, we based our calculations on the UBS data for Beijing, New York City and Berlin. This includes country-level data adjusts for the fact that some students commute into the metropolitan areas and many spend a significant amount of time in more rural areas. The rules of the laboratory in Essen, Germany require experimenters to base expected payments on an hourly student wage of €12.50. Using this anchor, we calculated payments in China and the USA. Note that Vieider (2012) finds no influence of small variations in payoffs (+/- 20%) on second-order risk aversion.

We attempted to conduct our study with subject pools that were as similar as possible, despite their different locations. Therefore, we only used student subjects, since they have a similar educational level and are of a similar age. In all three countries the subjects were recruited from a subject database.[11] We were able to recruit samples that were similar in their gender composition in all three countries. However, the databases were either not large enough or did not contain enough information to allow us to recruit samples that were similar for additional demographic characteristics. Therefore, we used the additional information on the participants we collected using a post-experimental questionnaire to control for differences between the subject pools in our analysis. Also note that Vieider, Lefebvre et al. (2015) and Ehmke, Lusk and Tyner (2010) find little difference in experimentally elicited risk preferences between student subject pools at different locations within the same country.[12]

[11] In the USA and Germany, we relied on existing databases. In both countries, this procedure was handled via ORSEE (Greiner 2015). In China, however, we had to build a database from scratch. Recruitment for this database was comparable to the procedures employed in the USA and in Germany. Two student assistants advertised participation by distributing flyers on campus and giving presentations in lectures. The advertisement promised the opportunity to earn a monetary reward for participation in an economic experiment. Potential participants could register via e-mail or text message.

[12] We had to make two adjustments because of American regulations, for which we control in regression analyses. First, in the USA it was necessary to inform subjects about the expected payoff and the nature of our experiment in the recruitment email. Second, it was necessary to present subjects with an IRB consent form in the laboratory prior to the experiments. The IRB form contained additional information regarding the experimental procedure, a short description of the task, and the expected payoffs. Neither of these two measures was required in Germany and China. Therefore, the Chinese participants received neither prior information in the recruitment e-mail nor an IRB consent form. To control for this difference, we used the American procedures in half of the sessions conducted in Germany. In other words, in these sessions German subjects were recruited via a German version of the American e-mail invitation and received a translation of the IRB consent form prior to the experiment.

4 Results

A Summary statistics

Table 2 summarizes the characteristics of all participants. In the sessions for country comparison, slightly more women than men participated in all three countries (CHN, USA and GER). Because we were not able to recruit subjects based on their age and gender in China and the USA, we conducted the sessions in Germany last. In Germany, we aimed to stratify our sample on the basis of the composition of the subjects recruited in the other two countries. However, we were not able to match the previous samples fully because the age structure of student populations differs across countries. Thus, the age distribution of the German participants differs significantly from the joint distribution of the Chinese and American subjects ($p = 0.010$, two-sided Mann-Whitney *U* test). The proportion of female subjects does not differ significantly between the German and the joint subject pools ($p = 0.883$, Fisher's exact test).

With respect to the BNT, on average the Chinese participants were able to solve 2.879 of the 4 questions correctly, which is higher than the 2.047 correct answers in the USA ($p < 0.001$, two-sided Mann-Whitney *U* test). With 1.393 correct answers on average, the German subjects provided even fewer correct answers than the Americans ($p < 0.001$). With respect to the CRT, participants in China and the USA did not differ significantly ($p = 0.568$). They were able to solve a little more than half of the three questions correctly. In Germany the rate was lower, with 1.290 correct answers ($p \leq 0.007$).[13]

Table 2 also summarizes the characteristics of the Chinese subjects, who participated in the high stakes treatment (CHN 10x). There are no significant differences between these participants and the

[13] Frederick (2005) observed that students at Princeton University answered 1.63 questions correctly on average ($N = 121$), while students at the University of Michigan (Ann Arbor) answered 1.18 questions correctly ($N = 1{,}267$). Brañas-Garza, Kujal and Lenkei (2015) provide a meta-study.

subjects facing regular stakes (CHN) ($p \geq 0.318$, two-sided Mann-Whitney U tests, for age and test scores; $p = 0.401$, Fisher's exact test, for gender composition).

Moreover, Table 2 shows the summary statistics of the German participants, who were confronted with different lottery formats (Compound & Reduced). For this analysis, we did not stratify the selection of participants because we are interested in the within-subject comparison. In this experiment the proportion of women does not differ from that for the remaining German (GER) data ($p = 0.219$, Fisher's exact test). The subjects are older ($p = 0.048$, two-sided Mann-Whitney U tests), but neither the CRT nor the BNT scores differ between the two groups ($p \geq 0.627$).

<<Table 2 here>>

The sample recruited for the Follow-up Experiment does not differ significantly from the sample that participated in Compound & Reduced treatment ($p \geq 0.365$, two-sided Mann-Whitney U tests for age and test scores, $p = 0.909$ Fisher's exact test for proportion of female subjects).

B Higher-order risk preferences across countries

Aggregate risk preferences.—There were seven choices to be made for each order except the first. Following Deck and Schlesinger (2014), we use the number of n-RL choices as a measure of n-th order risk aversion – that is, the more n-RL choices, the lower the n-th order risk aversion. We assume that all participants prefer more money to less money. This assumption is supported by the data in treatments CHN, USA and GER: 90% of the subjects in China, 99% of the subjects in the USA and 97% of the subjects in Germany never choose a dominated payoff in order 1. The number is slightly smaller in China than in the USA and in Germany ($p = 0.001$, Fisher's exact test). This is similar to the figure of more than 92% observed by Deck and Schlesinger (2014). We expect participants to differ in their preferences in orders 2 to 6, but consider the aggregate data first before analyzing individual patterns.

Table 3 shows the number of *n*-RL choices in each country. In all countries, we observe a general tendency of subjects to avoid the more risky lotteries. The number of *n*-RL choices is always significantly lower than 3.5 (which is the expected average count with random behavior) ($p < 0.001$, two-sided one-sample Wilcoxon signed-rank tests). Only in the USA does the choice frequency for order 6 not differ significantly from 3.5 ($p = 0.242$).[14] The correlations of the individuals' share of *n*-RL choices between orders 2 to 6 are shown in Online Appendix O10 for all our treatments.

Comparing the frequencies of *n*-RL choices between countries only indicates a difference for order 2. To control for subject pool differences, we run an ordinary least squares (OLS) regression for each order with the number of *n*-RL choices as the dependent variable, dummies for China and Germany (so the USA acts as the baseline category) and various controls as independent variables (see Appendix A3 for an overview of the variables, Appendix A4 for the regression results and Online Appendix O3 for details on our estimation strategy).[15] For order 2, the regression suggests that the Chinese sub-

[14] The distributions of choice frequencies within each order across the three countries, across stakes and across lottery formats are shown in Online Appendix O9.1..

[15] The regression analyses include the demographic and test results listed in Table 2, as well as controls for the experimenter and the IRB (cf. footnote 13). The latter two are influential. First, we observe a significant influence of our Chinese experimenter. In the German session conducted by him, participants behaved in a more second-order risk-averse manner. Second, we find that in the German sessions in which the subjects were provided with IRB information, the subjects behaved in a less risk-averse manner. Because of a computer error, we could not collect the CRT and BNT scores for eight subjects in China. Therefore, we also report additional regressions without controlling for CRT and BNT. In Online Appendix O4, we present further robustness checks of this model. We asked subjects about their migration background. In an additional analysis, we use a control variable for those who were not born in the respective country or who did not answer the relevant question. This was the case for 26 subjects in the USA and 7 subjects in Germany. However, none of these further robustness checks suggests a different interpretation of our data. Furthermore, making dominated choices in order 1 is an

jects make more risk-loving choices than the German ones. The Chinese country dummy indicates no significant difference between Chinese and American subjects (β = 0.981, robust SE = 0.672, p = 0.145, two-sided). The same regression also yields no difference between the USA and Germany, as measured by the German country dummy (β = 0.066, robust SE = 0.590, p = 0.910). However, the country dummies of China and Germany differ significantly (p = 0.009, two-sided Wald test).

<<Table 3 here>>

Consistency of risk preferences.—Based on a preference for combining "good" with "bad" as described by Deck and Schlesinger (2014), individuals should be 2-, 4- and 6-RA: that is, they should exhibit mixed risk-averse behavior. In contrast, individuals who have a preference for combining "good" with "good" should be 2-, 4- and 6-RL: that is, they should exhibit mixed risk-loving behavior. Both mixed risk averters and mixed risk lovers should be 3-RA and 5-RA (see Section 3.A). In a first step, we follow Deck and Schlesinger (2014) and study consistency in the higher orders relative to order 2. In a second step, we classify the subjects based on all orders.

For the first step we classify subjects as second-order risk-averse or risk-loving according to whether they make choices in line with this preference in the majority of their seven decisions in order 2. With this classification scheme, 80% of the Chinese participants, 83% of the Americans and 84% of the Germans are classified as second-order risk-averse, and the remaining subjects as second-order risk-loving. In order to identify mixed risk-averse and mixed risk-loving behavior, we consider the behavior of the two groups for the higher orders.

Figure 4 displays the average number of n-RL choices for second-order risk averters ("RA") and second-order risk lovers ("RL") across the three countries. In addition, it includes a dashed line at 3.5

---

obvious error and can be viewed as a check for data quality. In the regressions presented in Online Appendices O4, O6, and O8, we provide robustness checks controlling for dominated choices.

indicating the number of n-RL choices expected under random behavior, as well as 90% confidence intervals.[16] Note that the confidence intervals are larger for the risk lovers because of the smaller number of observations. For the odd orders 3 and 5, the graph reveals a preference for the more risk-averse option for second-order risk averters and risk lovers. For the even orders 4 and 6 the two types differ, and only risk averters tend to prefer the less risky options in the higher orders. This is exactly the pattern that would be expected when decisions are mainly made by mixed risk averters and mixed risk lovers.

This impression is confirmed by non-parametric tests: second-order risk averters and risk lovers in all countries are 3-RA and 5-RA when comparing the number of n-RL choices to the benchmark of 3.5 ($p \leq 0.005$, two-sided one-sample Wilcoxon signed-rank tests), but only Second-order risk averters are also 4-RA and 6-RA ($p \leq 0.005$). Second-order risk lovers are instead 4-RL and 6-RL in China and in the USA ($p \leq 0.072$) but not in Germany ($p \geq 0.259$).

<<Figure 4 here>>

For the second step of the analysis, we consider choices of all orders at once, because – strictly speaking – the theory does not differentiate between any of the even orders or between any of the odd orders. All of a subject's individual choices can be classified as being consistent or inconsistent with mixed risk-averse behavior, for example. The classification yields a binary variable with 38 observations for each subject. Based on this, we can classify the subjects into types. Running a binomial test for each subject allows us to test the null hypothesis that half of his or her 38 choices adhere to the mixed risk-averse type, for example. If we can reject this hypothesis and most choices adhere to the

[16] The distributions of choice frequencies within each order across the three countries, across stakes and across lottery formats are shown Online Appendix O9.2 separately for risk-loving and risk-averse subjects.

pattern, we classify the subject as mixed risk-averse. The same procedure is applied for mixed risk-loving behavior.

<<Table 4 here>>

Table 4 summarizes the proportion of subjects that can be classified into the two types. It lists the distributions for three significance thresholds.[17] Under the strictest criterion, between 42% and 45% percent are mixed risk-averse and between 6% and 9% are mixed risk-loving across the countries. If the criterion is relaxed, these percentages go up to between 60% and 64%, or 11% and 15%, respectively. The proportion of all subjects consistent with one type or the other ranges from 51% to 76%.

Next, we use individual behavioral patterns to compare the consistency across countries, by counting for each subject, (*i*) the number of choices consistent with mixed risk-averse behavior, (*ii*) the number of choices consistent with mixed risk-loving behavior and (*iii*) the maximum of both. As already suggested by Figure 4, there appears to be no difference in the behavioral patterns across countries. Running separate OLS regressions (see Appendix A4) with these three dependent variables provides additional evidence with respect to Hypothesis 1. First, in the regression of the number of mixed risk-averse choices on country dummies, the dummies for China (β = -1.191, robust standard error (SE) = 2.041, p = 0.560, two-sided) and Germany are insignificant (β = 0.689, robust SE = 1.716, p = 0.689). But – reflecting the difference in 2-RA – the dummies differ weakly[18] from each other, suggesting that the Chinese subjects make somewhat fewer mixed risk-averse choices than the Germans

[17] Note that only the 1% threshold guarantees a mutually exclusive classification when the subjects make 38 decisions across orders 1 to 6. An individual may be classified as being consistent with respect to both types when applying the 5% or the 10% threshold. However, when applying the 5% threshold, this is only the case for one subject in China and one subject in the USA. When applying the 10% threshold, this is the case for four subjects in China, one subject in Germany and five subjects in USA.

[18] By weakly we mean significant at the 10% level.

(p = 0.099, two-sided Wald test). Second, in the regression of the number of mixed risk-loving choices on country dummies, the dummies for China (β = 1.858, robust SE = 1.630, p = 0.255) and Germany (β = 0.388, robust SE = 1.473, p = 0.792) are insignificant. Again, the two dummies differ weakly (p = 0.094), meaning that the Chinese subjects make somewhat more mixed risk-loving choices. Third, when using the maximum of the two variables for each subject as the dependent variable in an OLS regression, we find no significant country differences ($p \geq 0.436$).

Observation 1: *Between 51% and 76% of all subjects can be classified as adhering to either mixed risk-averse or mixed risk-loving behavior across countries. After controlling for procedural differences and the subjects' characteristics, the Chinese participants are found to make weakly fewer mixed risk-averse choices and weakly more mixed risk-loving choices than the Germans.*

C Higher-order risk preferences across stakes

Aggregate risk preference.—As in the previous analyses, we interpret the number of n-RL choices as a measure of n-th order risk aversion, and assume that all participants prefer more money to less. In CHN 90% never choose a dominated payoff in order 1. In CHN 10x this share is 92% and therefore not significantly larger (p = 0.494, Fisher's exact test).

Table 5 shows the number of n-RL choices under both incentive structures in all orders. The number of n-RL choices is significantly lower than would be expected under random behavior in all orders ($p \leq 0.010$, two-sided one-sample Wilcoxon signed-rank tests).

The previous evidence indicates that more second-order risk-averse choices are made when the stakes increase. The data presented in Table 5 suggest a similar effect: with regular stakes, participants make, on average, 1.971 decisions in a risky way, but under high stakes, the average is 1.417 decisions. To control for subject pool differences within China, we run OLS regressions separately for each order, with the number of n-RL choices as the dependent variable on a dummy for high stakes (and regular stakes as the baseline), as well as various controls (see Appendix A5 for the regression results and Online Appendix O5 for details on our estimation strategy). For order 2, there is a weakly negative effect of increased stakes on the number of risky choices (β = -0.616, robust SE = 0.319, p = 0.055, two-sided), and there are no significant differences for the other orders ($p \geq 0.309$).

<<Table 5 here>>

Consistency of risk preferences.—In this section we analyze whether the patterns of mixed risk-averse and mixed risk-loving behavior prevail under high stakes. Under the high stakes in CHN 10x, we also expect second-order risk averters to coincide in their choices with second-order risk lovers in the odd orders and to differ from them in the even orders.

Again, we analyze the consistency in two steps. In the first step we classify the subjects into second-order risk-averse or risk-loving and analyze their behavior in the higher orders. Classifying everyone with a majority of risk-averse choices as risk-averse yields 83% second-order risk-averse subjects in the CHN 10x, compared to 80% in CHN (the remaining subjects being classified as risk-loving).

Figure 5 displays the average number of n-RL choices for second-order risk averters ("RA") and second-order risk lovers ("RL") in the two treatments. For the odd orders 3 and 5 of CHN 10x, both types appear to favor the risk-averse option more frequently, while for orders 4 and 6 the two types appear to differ. This is the pattern suggested by the theory of mixed risk-averse and mixed risk-loving behavior.

Non-parametric tests also confirm this interpretation for CHN 10x, when comparing the number of choices to the 3.5 n-RL choices that would be expected under random behavior: second-order risk averters and risk lovers significantly favor the more 3-RA or 5-RA options ($p \leq 0.048$, two-sided one-sample Wilcoxon signed-rank tests), while only second-order risk averters favor the more 4-RA and 6-RA options ($p \leq 0.001$). Second-order risk lovers weakly prefer the more risky options (p = 0.084) for order 4, while there is no significant tendency for order 6 (p = 0.222).

<<Figure 5 here>>

In the second step, we consider the decisions for all orders jointly. Using binomial tests, we check whether each subject makes decisions that are in line with either of the patterns. Table 6 presents the resulting proportions of the subjects in the CHN 10x treatment. In this treatment, 58% of subjects can be classified as either mixed risk-averse or mixed risk-loving under the strictest threshold of 1% significance. This goes up to a total of 88% with the 10% significance threshold.[19] Table 6 also lists the shares from the CHN treatment for comparison, in which between 51% and 79% could be classified in this way.

<<Table 6 here>>

In addition, we again run separate OLS regressions with the number of choices consistent with (*i*) mixed risk-averse behavior, (*ii*) mixed risk-loving behavior and (*iii*) either type as dependent variables (see Appendix A5). When regressing the number of choices consistent with mixed risk-aversion on a treatment dummy and the control variables, we do not observe a significant effect of the treatment dummy of CHN 10x (β = 1.160, robust SE = 1.083, p = 0.286, two-sided). Also, when analyzing the number of choices consistent with mixed risk-loving behavior, we do not observe a significant influence of the treatment dummy (β = -0.921, robust SE = 0.843, p = 0.276). Also, there appears to be no significant influence on the maximum of either variable (β = 0.592, robust SE = 0.909, p = 0.516).

[19] Note that when applying the 10% threshold, one subject in CHN 10x is classified as being both mixed risk-averse and mixed risk-loving. In the other two cases, the resulting classifications are mutually exclusive.

Observation 2: *Between 58% and 88% of all subjects can be classified as adhering to either mixed risk-averse or mixed risk-loving behavior when the stakes are increased tenfold. After controlling for the subjects' characteristics, we do not find a significant difference in the number of mixed risk-averse or mixed risk-loving choices when the stakes increase.*

D Higher-order risk preferences across lottery formats

*Aggregate risk preferences.—*Again, we consider the number of *n*-RL choices first and, with respect to order 1, we assume that all participants prefer more money to less. In GER, 97% never choose a dominated payoff in order 1. In the Compound & Reduced treatment, this share is also 97% (p = 1.000, Fisher's exact test).[20]

Table 7 presents the number of *n*-RL choices in this treatment. Please note that only the lotteries of orders 3 to 6 were displayed in compound and reduced form. While the data for orders 1 and 2 is based on the choices of all participants, the data for the higher orders 3 to 6 is based on approximately half the sample (71 participants for order 3, 73 for order 4, 72 for order 5 and 70 for order 6). Each participant made choices for two of the higher orders and in both framings.

As before, there is a tendency of participants to prefer the less risky alternative for orders 3 to 5. Comparing choice frequencies to the 3.5 *n*-RL choices that would be expected under random behavior, the behavior for orders 3 to 5 differs significantly from the benchmark (p < 0.001, two-sided one-sample Wilcoxon tests), but the behavior for order 6 does not (p = 0.163). With respect to the reduced lotteries, however, the difference from the benchmark is only significant for order 5 (p < 0.001). It is weakly significant for order 3 (p = 0.075) and insignificant for orders 4 and 6 ($p \geq 0.113$).

[20] Comparing the number of *n*-RL choices between the compound lotteries of the Compound & Reduced treatment and the GER treatment reveals that participants in Compound & Reduced are significantly more 3-RA than participants in GER (p = 0.019, two-sided Mann-Whitney U test). In the remaining orders, we do not observe any significant differences ($p \geq 0.189$).

Our results suggest that the lottery format influences choices for orders 3 and 4. We run linear panel regressions with individual random effects separately for each order, with the number of *n*-RL choices as the dependent variable. The regressions include a dummy for choices in reduced lotteries, as well as various controls (see Appendix A6 for the regression results and Online Appendix O7 for details about our estimation strategy). The regressions indicate 0.845 less 3-RA choices in reduced lotteries (β = 0.845, robust SE = 0.223, $p < 0.001$, two-sided) and 1.014 less 4-RA choices (β = 1.014 robust SE = 0.279, $p < 0.001$).

<<Table 7 here>>

Consistency of risk preferences.—Above we reported a robust pattern of mixed risk-averse and mixed risk-loving behavior in three subject pools and under varying stakes based on the use of compound lotteries. However, the predictions we test are independent of the lottery format. They always suggest that second-order risk averters coincide in their choices with second-order risk lovers in the odd orders, while they differ in the even orders. Yet the compound format might facilitate viewing a lottery as a combination of "good" and "bad" outcomes.

Even though the treatment presented in this section possesses a slightly different data structure, we proceed in the same way as before. For the Compound & Reduced treatment, 93% of the subjects are classified as second-order risk-averse ("RA") and 7% as second-order risk-loving ("RL"). Figure 6 displays the average frequency of *n*-RL choices made by both types across orders 3 to 6. The pattern of choices is less clear cut than in the previous analyses. With respect to second-order risk averters, we replicate the previous findings using compound lotteries: for orders 3 to 5, second-order risk averters favor the less risky lotteries, if we compare their choices to the 3.5 benchmark ($p \leq 0.001$, two-sided one-sample Wilcoxon signed-rank tests). In the case of order 6, we do not observe this tendency ($p = 0.136$). When using reduced lotteries, we still find at least a weak tendency of second-order risk averters to favor less risky lotteries for orders 3, 5 and 6 ($p \leq 0.080$), but for order 4 their choices no longer differ significantly from the 3.5 benchmark ($p = 0.738$). Independent of the order, second-order risk lovers do not systematically favor one of the options ($p \geq 0.262$), in case of compound lotteries. In case of reduced lotteries, they do favor more risky lotteries for order 4 ($p = 0.083$), but not for any other order ($p \geq 0.480$). However, there are relatively few second-order risk lovers (only between three and seven for each order) compared to the treatments discussed above. This might be driven by differences in the subject pool composition (cf. Table 2).[21]

This pattern suggests that some individuals exhibit preference reversals. On the individual level, 27% of the subjects make more *n*-RL choices in the reduced than in the compound lotteries of order 3, while 10% make fewer *n*-RL choices. These percentages are 32% and 15% for order 4, 20% and 19% for order 5 and 18% and 17% for order 6. To shed some light on the drivers of this change in preferences, we also run a logit regression. The dependent variable in this regression is a dummy indicating whether a subject's number of *n*-RL choices differed between the two treatments (see Appendix A6 for the regression results and the Online Appendix O7 for details on our estimation strategy). As the explanatory variable, we use the characteristics of the subjects displayed in Table 2. This regression indicates that numeracy is somewhat associated with preference reversal: those with a higher score in the BNT are weakly less likely to switch (average marginal effect = 0.047, standard error = 0.027 p = 0.081). We do not observe any significant influences of the other control variables.

In a second step, we classify subjects as mixed risk-averse and mixed risk-loving. Participants made choices for two of the higher orders in both framings. While in the previous classification we could use all 38 decisions at once, we now rely on 24 choices to classify each participant: 10 choices from order 1 and 2 and 14 of 28 choices from two of the higher orders (either 14 from the compound or 14 from the reduced framing).

[21] Even though subjects in the compound lotteries of Compound & Reduced and GER do not differ with respect to the number of 2-RA choices, the share of subjects classified as second-order risk-averse ("RA") is significantly higher in the compound lotteries of Compound & Reduced than in GER ($p = 0.025$, Fisher's exact test).

<<Figure 6 here>>

Table 8 presents the percentage of subjects who are classified as either mixed risk-averse or mixed risk-loving based on the binomial tests. With the 1% significance threshold, 42% of the subjects can be classified as belonging to one of the two types, when using the compound lotteries. When using the reduced lotteries, only 28% of the subjects can be classified in this way. These shares go up to 59% and 55%, respectively, when applying the 10% significance threshold.[22]

<<Table 8 here>>

In addition, we compare the number of choices that are consistent with the two types between the two formats. We run separate linear random-effects panel regressions with the number of choices consistent with (*i*) mixed risk-averse behavior, (*ii*) mixed risk-loving behavior or (*iii*) either type as dependent variables. The regressions include the usual control variables (see Appendix A6). Considering the number of mixed risk-averse choices, we find a negative effect of the dummy indicating choices from reduced lotteries (β = -0.916, robust SE = 0.235, p < 0.001, two-sided). Conducting the same analysis for the number of mixed risk-loving choices, we do not find a significant influence of the format (β = 0.161, robust SE = 0.218, p = 0.460). On aggregate, the regression results also indicate that consistency with either type is smaller in the reduced lotteries (β = -0.811, robust SE = 0.227, p < 0.001).

[22] When the subjects make 24 decisions across orders 1 and 2, as well as two more orders $i, j \in \{3, 4, 5, 6\}$ with $i \neq j$, it depends on the combination of odd and even orders whether the classification of the two types is theoretically mutually exclusive. However, it is only when applying the 10% threshold that one subject in Compound and one in Reduced are classified as being mixed risk-averse *and* mixed risk-loving.

Observation 3: *Between 28% and 55% of all subjects can be classified as adhering to either mixed risk-averse or mixed risk-loving behavior when the lotteries are displayed in the reduced format. The number of mixed risk-averse choices increases significantly when the lotteries are displayed in the compound format, while the number of mixed risk-loving choices does not change.*

Explaining the framing effect.—In Section 2.B, we outlined previous results on the differences in choices observed between reduced and compound lotteries. To our knowledge, no previous study offers an explanation for why we observe more 3-RA and more 4-RA in compound than in reduced lotteries. To gather further evidence, we conducted a Follow-up Experiment varying the type of lottery choice (3-RA and 4-RA) as well as the framing (reduced and compound) between-subjects (cf. the experimental design described in Section 3.B).

First, with respect to choices in the 3-RA lottery, we find an even more pronounced framing effect in this Follow-up Experiment than in the Compound & Reduced treatment. In the compound framing of the Follow-up Experiment, 76% of the 58 participants choose the more 3-RA lottery. Only 45% of the 56 participants do so in the reduced framing ($p = 0.001$, Fisher's exact test). In the respective task of the Compound & Reduced treatment, 70% of subjects choose the more 3-RA lottery in the compound framing, while only 49% do so in the reduced framing ($p = 0.006$, McNemar's test). With respect to the 4-RA lottery, we do observe less 4-RA choices in the Follow-up Experiment, and we do not find a significant framing effect. In the Follow-up Experiment, 57% of the 54 participants choose the more 4-RA lottery and 43% of the 56 participants do so in the reduced framing ($p = 0.182$). In the Compound & Reduced treatment, 78% of subjects choose the more 4-RA lottery in the compound framing, while 51% do so in the reduced lottery ($p < 0.001$).

Second, to study the reasoning behind participants' choices, they were asked to send one written free-form message, together with their preferred choice, to the other participant in their group.

For three arguments from our classification scheme, the frequency differs significantly between framings in the 3-RA or the 4-RA lottery. These are:

(*i*) Maximization of the largest potential payoff.

(*ii*) Maximization of the smallest potential payoff.

(*iii*) Maximization of the payoff for the most likely outcome.

These arguments were reliably identified by two independent coders as indicated by values of Krippendorff's alpha above the commonly used threshold of 0.67 (see Krippendorff 2004 and Online Appendix O11 for the complete list of arguments).

With respect to the 3-RA lottery, the frequency of all of the three arguments differs between both framings ($p \leq 0.004$, Fisher's exact tests). The first two arguments suggest that one should choose the more 3-RA lottery (and all except one of the participants using these arguments do so). They are respectively used by 22% and 40% of the participants in the compound framing and only by 3% and 7% in the reduced framing. The third argument suggests that one should choose the 3-RL lottery (and all except one of the participants do so). It is used by 17% of the participants in the reduced framing and by only 3% in the compound framing. With respect to the 4-RA lottery, only the frequency of the second argument differs between both framings ($p = 0.023$). It suggests that one should choose the more 4-RA lottery (and all except one of the participants do so). It is used by 26% of the participants in the compound framing and only by 6% in the reduced framing.

Overall, it appears that the compound display of lotteries leads subjects to focus more on the smallest potential payoff in the 3-RA, as well as in the 4-RA lottery. This could drive the differences in choices we observe between compound and reduced lotteries for 3-RA and 4-RA.

Observation 4: *The most commonly used argument to justify 3-RA (prudent) and 4-RA (temperate) choices is the maximization of the smallest potential payoff. It is used significantly more often in compound than in reduced lotteries.*

5 Conclusion

In this study, we analyze the consistency of higher-order risk preferences. We contribute to this topic by exploring the role of country differences, the variation of stakes, and the framing of lotteries. In our American subject pool, we replicate the findings of Deck and Schlesinger (2014) and we identify a similar pattern in subject pools in Germany and in China. Across all three countries, a majority of participants can be classified as mixed risk averters or as mixed risk lovers (between 51% and 76% of all subjects depending on the significance level).

Existing evidence from non-incentivized and incentivized studies suggests that Chinese are more second-order risk-averse than Americans and Germans. We can only confirm this finding with respect to Chinese and Germans. We do not observe a significant difference in second-order risk aversion between Chinese and Americans. We have formulated our first hypothesis based on the assumption that differences in second-order risk aversion indicate differences in the underlying distribution of mixed risk averters and mixed risk lovers. In line with our first hypothesis, mixed risk averters are somewhat more common in Germany than in China, while mixed risk lovers are less common. Contrary to our first hypothesis, we do not observe differences in the prevalence of both types between our Chinese and American samples. However, it is important to note that these findings reflect the differences in second-order risk aversion in our sample. In fact, the evidence from incentivized studies on differences between Chinese and Americans is less clear cut than the evidence from non-incentivized studies (cf. Haering and Heinrich 2017). For example, in the first experimental comparison, Kachelmeier and Shehata (1992) also do not find significant differences between Chinese and Americans.

Moreover, we provide the first analysis of higher-order risk preferences with large monetary payoffs. We also observe a majority of choices to be in line with mixed risk-averse and mixed risk-loving behavior under high stakes. In line with prior evidence, we observe an increase in second-order risk aversion when the stakes are increased tenfold. However, contrary to our second hypothesis, we find no significant change in the number of mixed risk-averse and mixed risk-loving choices when the stakes increase.

A dichotomous population with respect to higher-order risk preferences may have important real-world implications. While mixed risk averters and mixed risk lovers coincide in their choices in the odd orders, they differ in the even orders. This means that a measurement of second-order risk aversion is not sufficient for estimating the prevalence of higher-order risk preferences per se. A measurement of second-order risk aversion will be indicative for temperance (4-RA) and 6-th-order risk aversion (6-RA) but not for prudence (3-RA) or 5-th-order risk aversion (5-RA). This is relevant for prevention decisions, for instance. From an economic perspective, prevention can be classified as self-protection or self-insurance. Self-protection lowers the *probability* of the occurrence of a loss, while the *size* of the loss is exogenous. In contrast, self-insurance aims at reducing the *size* of a loss while the *probability* of occurrence is exogenous (see Ehrlich and Becker 1972). In medicine, self-protection is known as primary prevention and self-insurance as secondary prevention.

While higher second-order risk aversion unambiguously leads individuals to choose higher levels of self-insurance, risk aversion is not sufficient to determine an individual’s level of self-protection (see Dionne and Eeckhoudt 1985 and Briys and Schlesinger 1990). Eeckhoudt and Gollier (2005) and Courbage and Rey (2006) show that more prudent individuals will expand less effort in self-protection.

An example for self-protection is the influenza vaccination. In general, the flu shot decreases the probability of getting the flu, but the harm of the flu itself is not affected. Prudent individuals try to avoid the worst outcome, which is facing the disutility that comes with the flu shot and still getting the flu. Therefore, more prudent individuals should be less likely to undergo an influenza vaccination. Indeed, Mayrhofer and Schmitz (2019) find that for high risk individuals, such as individuals over 60 years of age, prudence has a significant negative impact on the likelihood of undergoing influenza vaccination. Since both mixed risk averters and mixed risk lovers are prudent, they will expend the same effort for self-protection measures like flu shots. However, this is different with regard to self-insurance. For example, cancer screenings do not decrease the likelihood of getting cancer, but an early detection can lead to early treatment and thus less harm. In this case, mixed risk-averse individuals opt for screening more often or earlier than mixed risk lovers (see Felder and Mayrhofer 2014).

We also observe that subjects choose the prudent and temperate options less often, when the options are displayed in a reduced rather than a compound form. In the reduced lotteries, there is weak evidence that subjects generally behave prudently, and no evidence that they are generally temperate. In other words, the proportion of subjects who can be classified as mixed risk averters or mixed risk lovers decreases considerably, when reduced lotteries are used. This in line with our third hypothesis and the conjecture by Deck and Schlesinger (2014) who point out that compound lotteries may facilitate the interpretation of lotteries as combinations of “good” with “bad” or “good” outcomes.

To our knowledge, no previous study offers an empirical explanation for why we observe more prudent and more temperate behavior in compound lotteries than we do in reduced lotteries. To shed light on our findings, we conducted a follow-up experiment that was aimed at revealing the reasoning behind the subjects’ choices. Overall, it appears that the compound display of lotteries leads subjects to focus more on the smallest potential payoff in prudence as well as in temperance lotteries. This could drive the differences in choices we observe between compound and reduced lotteries for prudence and temperance.

As Abdellaoui, Klibanoff and Placido (2015) point out, different attitudes towards compound versus reduced risks might have big implications for marketing, policy and economics. For example, if people are less temperate with respect to reduced risks, they would invest more in risky assets if the associated risks are presented in reduced rather than compound form.

Appendix A1 Comparison of related papers

<<Table A1 here>>

Appendix A2 Instructions (English version)

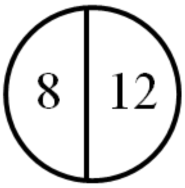
You are participating in a research study on decision making under uncertainty. At the end of the study you will be paid your earnings in cash and it is important that you understand how your decisions affect your payoff. If you have questions at any point during the study, please raise your hand and someone will assist you. Otherwise, please do not talk during this study and turn off your cell phone.

[**CHN, USA, GER, CHN 10x and Compound & Reduced:** In this study, there is a series of 38 tasks. Each of these tasks involves choosing between Option A and Option B. Once you have completed these tasks, one of the 38 will be randomly selected to determine your payoff. All values are given in experimental currency unit (ECU).**]**

[Follow-up Experiment: This there is one task. This task involves choosing between Option A and Option B. This decision can influence your payoff. All values are given in experimental currency unit (ECU).**]**

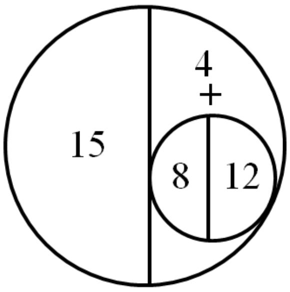
For ECU 1 you will receive $ 0.93.

Each option will involve amounts of money and possibly one or more 50-50 lotteries represented as a circle with a line through the middle. A 50-50 lottery means there is a 50% chance of receiving the item to left of the line and a 50% chance of receiving the item to the right of the line. For example,

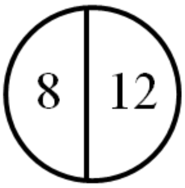


is a 50-50 lottery in which you would receive either ECU 8 or ECU 12, each with an equal chance. To determine the outcome of any 50-50 lottery, we will use a computerized random-number generator.

In some cases, one of the lottery outcomes in a 50-50 lottery may contain another lottery. For example,

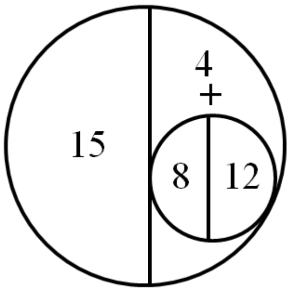


is a 50-50 lottery where you receive either ECU 15 or you receive ECU 4 plus the 50-50 lottery



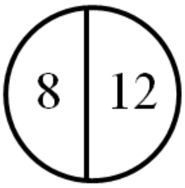
.

Continuing with the example,



,

there is a 50% chance that you would receive ECU 15 in the first 50-50 lottery and that would be it. There is also a 50% chance that you would receive ECU 4 +



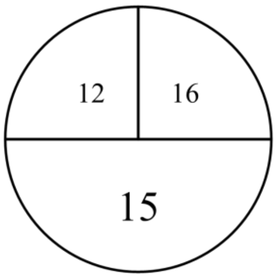
in the first 50-50 lottery.

Conditional on this outcome for the first 50-50 lottery, you would then have a 50% chance of receiving an extra ECU 8 and a 50% chance of receiving an extra ECU 12 in addition to the ECU 4. Therefore, the chance that you would end up with 4 + 8 = ECU 12 is 0.5 × 0.5 = 0.25 = 25%. The chance that you would end up with 4 + 12 = ECU 16 is 0.5 × 0.5 = 0.25 = 25%.

[Compound & Reduced and Follow-up Experiment:

The illustration of this option can also take place with the aid of a circle with different probabilities of the lottery results.

Like in the sample above



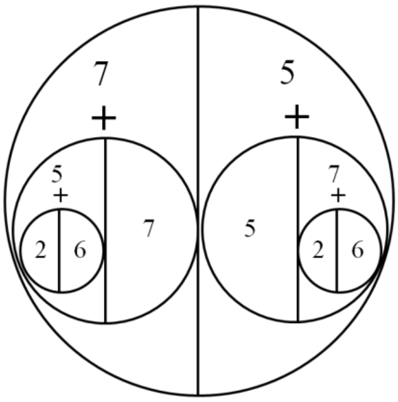
is a lottery in which you can either receive 12 ECU, 15 ECU or 16 ECU.

Again, there is a 50% probability that you will receive 15 ECU. In addition, the probability that you get 12 ECU or 16 ECU is 25% each.

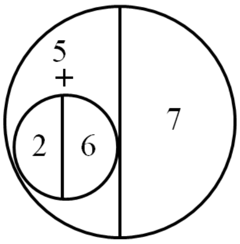
]

[CHN, USA, GER, CHN 10x and Compound & Reduced:

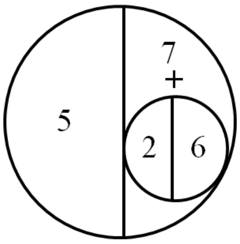
Let’s look at a more complicated example.



is a 50-50 lottery where you receive either ECU 7 plus the 50-50 lottery



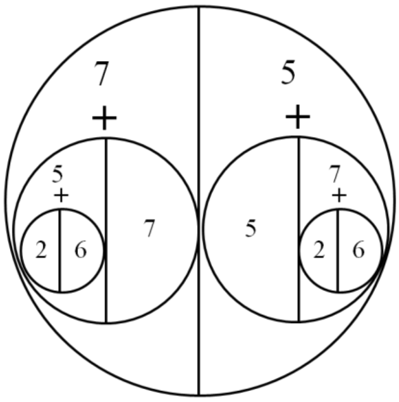
or you receive ECU 5 plus the 50-50 lottery



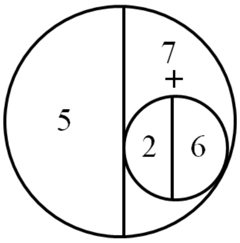
,

both of which include an additional 50-50 lottery.

In



you could earn ECU 10 if you get ECU 5 +



in the first lottery and then earn ECU 5 in the second lottery. This occurs with a 0.5 x 0.5 = 0.25 = 25% chance. Alternatively, you could earn ECU 14 with a 50% chance. Notice that you could earn ECU 14 in three ways:

by 1) earning ECU 7 (in the first lottery) + ECU 5 (in the second lottery) + ECU 2 (third lottery) which happens with a 0.5 x 0.5 x 0.5 = 0.125 = 12.5% chance,

or 2) earning ECU 7 (in the first lottery) + ECU 7 (in the second lottery) which happens with a 0.5 x 0.5 = 0.25 = 25% chance,

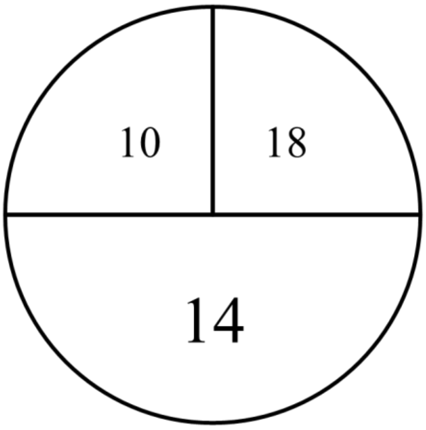
or 3) earning ECU 5 (in the first lottery) + ECU 7 (in the second lottery) + ECU 2 (third lottery) which happens with a 0.5 x 0.5 x 0.5 = 0.125 = 12.5% chance.

Finally there are two ways that you could earn ECU 18 which occurs with a 0.5 x 0.5 x 0.5 + 0.5 x 0.5 x 0.5 = 0.25 = 25% chance.

]

[**Compound & Reduced:**

This option can also be illustrated with the aid of a circle with different probabilities of the lottery results (see next page).



Just like in the example on the previous two pages, you can either receive 10 ECU, 14 ECU or 18 ECU. Again, there is a 50% chance that you will earn 14 ECU. In addition, the probability of receiving 10 ECU or 18 ECU is 25%.]

[Follow-up Experiment:

Choice of an option

You will be randomly assigned to another participant in the experiment as a partner with whom you will form a team. Your payoffs will be determined by the decisions of your team.

How does your team decision come about? Both team members will enter a final decision regarding the choice of Option A or Option B. However, only *one* of the decisions is chosen randomly and with equal probability as the team decision (with the help of a computerized random-number generator).

The chosen final decision counts for *both* team members. (Note that both team members receive the respective payoff of the task. If the pay-out of a lottery is 15 ECU, for example, both team members receive 15 ECU *each*).

Before you enter your final decision, you have the opportunity to influence the final choice of your partner: Before the decisions are entered, you will send a preferred choice together with a text message to your partner. Likewise, your partner sends a text message with his proposed option to you.

Only after both team members have received the other's text messages, they are allowed to enter their final decision. After both made their decision, one of the two decisions will be randomly chosen for the team. A questionnaire follows and finally your payoff is determined based on the chosen option.

Note: All participants of the experiment receive the same instructions.

Note for the text messages

The content of the text message is up to you. But be aware, that your text message is the only chance to persuade your partner of your proposed option. Thus, use the text message to explain your proposal.

It is forbidden, to provide any personal details such as e.g. name, age, address, field of study! If you violate the rules of communication, you can be excluded from the experiment without receiving your payoff. Every text message may include 420 characters at most (approximately 3 lines). Note: To send a typed text message you have to click 'send'.

]

Appendix A3 Summary of variables

<<Table A3 here>>

Appendix A4 Regression results on higher-order risk preferences across countries

<<Table A4.1 here>>

<<Table A4.2 here>>

Appendix A5 Regression results on higher-order risk preferences across stakes

<<Table A5.1 here>>

<<Table A5.2 here>>

Appendix A6 Regression results on higher-order risk preferences across lottery formats

<<Table A6.1 here>>

<<Table A6.2 here>>

<<Table A6.3 here>>

References

Abdellaoui, M., P. Klibanoff, and L. Placido. 2015. "Experiments on Compound Risk in Relation to Simple Risk and to Ambiguity." *Management Science* 61(6): 1306-22.

Arrow, K. 1965. *Aspects of the Theory of Risk-bearing*. Helsinki: Yrjö Jahnsson Foundation.

Azrieli, Y, C. P. Chambers, and P. J. Healy. 2018. "Incentives in Experiments: A Theoretical Analysis." *Journal of Political Economy* 126(4): 1472-1503.

Baillon, A. 2017. "Prudence with respect to ambiguity." *Economic Journal* 127(604): 1731-1755.

Baillon, A., H. Schlesinger, and G. van de Kuilen. 2017. "Measuring Higher Order Ambiguity Preferences." *Experimental Economics* 21: 233-256.

Baltussen, G., G. T. Post, M. J. van den Assem, and P. P. Wakker. 2012. "Random Incentive Systems in a Dynamic Choice Experiment." *Experimental Economics* 15(3): 418-43.

Bar-Hillel, M. 1973. "On the Subjective Probability of Compound Events." *Organizational Behavior and Human Performance* 9(3): 396-406.

Beattie, J., and G. Loomes. 1997. "The Impact of Incentives upon Risky Choice Experiments." *Journal of Risk and Uncertainty* 14(2): 155-68.

Binswanger, H. P. 1980. "Attitudes toward Risk: Experimental Measurement in Rural India." *American Journal of Agricultural Economics* 62(3): 395-407.

Binswanger, H. P. 1981. "Attitudes toward Risk: Theoretical Implications of an Experiment in Rural India." *Economic Journal* 91: 867-90.

Bohnet, I., F. Greig, B. Herrmann, and R. Zeckhauser. 2008. "Betrayal Aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States." *American Economic Review* 98(1): 294-310.

Bostian, A. J. and C. Heinzel 2012. "Prudential Saving: Evidence from a Laboratory Experiment." Working Paper, University of Virginia.

Brañas-Garza, P., P. Kujal, and B. Lenkei. 2015. "Cognitive Reflection Test: Whom, How, When." MPRA Working Paper No. 68049.

Brislin, R. W. 1970. “Back-translation for Cross-cultural Research.” *Journal of Cross-Cultural Psychology* 1(3): 185-216.

Briys, E., and H. Schlesinger. 1990. “Risk aversion and the propensities for self-insurance and self-protection.” *Southern Economic Journal* 57(2): 458–467.

Brockett, P. L., and L. L. Golden 1987. “A Class of Utility Functions containing all the Common Utility Functions.” *Management Science* 33(8): 955-64.

Buchan, N., and R. Croson. 2004. “The Boundaries of Trust: Own and Others’ Actions in the US and China.” *Journal of Economic Behavior & Organization* 55(4): 485-504.

Budescu, D. V., and I. Fischer. 2001. “The Same but Different: An Empirical Investigation of the Reducibility Principle.” *Journal of Behavioral Decision Making* 14(3): 187-206.

Burchardi, K. B., and S. P. Penczynski. 2014. "Out of your mind: Eliciting individual reasoning in one shot games." *Games and Economic Behavior* 84: 39-57.

Burks, S. V., J. P. Carpenter, L. Goette, and A. Rustichini. 2009. “Cognitive Skills Affect Economic Preferences, Strategic Behavior, and Job Attachment.” *Proceedings of the National Academy of Sciences* 106(19): 7745-50.

Caballé, J., and A. Pomansky 1996. “Mixed Risk Aversion.” *Journal of Economic Theory* 71: 485-513.

Cokely, E. T., M. Galesic, E. Schulz, S. Ghazal, and R. Garcia-Retamero. 2012. “Measuring Risk Literacy: The Berlin Numeracy Test.” *Judgement and Decision Making* 7(1): 25-47.

Courbage, C., and B. Rey 2006. “Prudence and Optimal Prevention for Health Risks.” *Health Economics* 15: 1323-27.

Courbage, C., and B. Rey 2016. “Decision Thresholds and Changes in Risk for Preventive Treatment.” *Health Economics* 25: 111-24.

Cox, J. C., V. Sadiraj, and U. Schmidt. 2015. Paradoxes and Mechanisms for Choice under Risk. *Experimental Economics* 18(2): 215-50.

Crainich, D., L. Eeckhoudt, and A. Trannoy. 2013. “Even (Mixed) Risk Seekers are Prudent.” *American Economic Review* 103(4): 1529-35.

Cubitt, R., R. Starmer, and R. Sugden. 1998. “On the Validity of the Random Lottery Incentive System.” *Experimental Economics* 1(2): 115-31.

Deck, C. and H. Schlesinger. 2010. “Exploring Higher Order Risk Effects.” *Review of Economic Studies* 77(4): 1403-20.

Deck, C., and H. Schlesinger. 2014. “Consistency of Higher Order Risk Preferences.” *Econometrica* 82(5): 1913-43.

Deck, C., and H. Schlesinger. 2017. “On the Robustness of Higher Order Risk Preferences.” *Journal of Risk and Insurance* 85(2): 313-333.

Dionne, G., and L. Eeckhoudt. 1985. “Self-insurance, self-protection and increased risk aversion.” *Economics Letters* 17: 39–42.

Dohmen, T., A. Falk, D. Huffman, and U. Sunde. 2010. “Are Risk Aversion and Impatience Related to Cognitive Ability?” *American Economic Review* 100(3): 1238-60.

Ebert, S. 2013. “Even (Mixed) Risk-Seekers are Prudent: Comment.” *American Economic Review* 103(4): 1536-37.

Ebert, S., D. Nocetti, and H. Schlesinger. 2017. “Greater Mutual Aggravation.” *Management Science* 64(6): 2809-2811.

Ebert, S., and G. van de Kuilen. 2015. “Measuring Multivariate Risk Preferences.” Working Paper, Available at SSRN 2637964.

Ebert, S., and D. Wiesen. 2011. “Testing for Prudence and Skewness Seeking.” *Management Science* 57(7): 1334-49.

Ebert, S., and D. Wiesen. 2014. “Joint Measurement of Risk Aversion, Prudence, and Temperance.” *Journal of Risk and Uncertainty* 48(3): 231-52.

Eeckhoudt, L. 2002. *Risk and Medical Decision Making.* Boston: Kluwer.

Eeckhoudt, L., and C. Gollier. 2005. “The Impact of Prudence on Optimal Prevention.” *Economic Theory* 26(4): 989-94.

Eeckhoudt, L., C. Gollier, and H. Schlesinger. 1996. “Changes in Background Risk and Risk Taking Behavior.” *Econometrica* 64(3): 683-89.

Eeckhoudt, L., and M. Kimball. 1992. “Background Risk, Prudence, and the Demand for Insurance.” In *Contributions to Insurance Economics*, edited by G. Dionne, 239-54, Netherlands: Springer.

Eeckhoudt, L. and H. Schlesinger, 2006. “Putting Risk in Its Proper Place.” *American Economic Review* 96 280-89.

Eeckhoudt, L. and H. Schlesinger, 2008. “Changes in Risk and the Demand for Saving.” *Journal of Monetary Economics* 55(7): 1329-36.

Eeckhoudt, L., H. Schlesinger, and I. Tsetlin. 2009. “Apportioning of Risks via Stochastic Dominance.” *Journal of Economic Theory* 144: 994-1003.

Ehmke, M., J. Lusk, and W. Tyner. 2010. “Multidimensional Tests for Economic Behavior Differences across Cultures.” *Journal of Socio-Economics* 39(1): 37-45.

Ehrlich, I. and G. Becker. 1972. „Market insurance, self-insurance, and self-protection.” *Journal of Political Economy* 80(4): 623–48.

Ekern, S. 1980. “Increasing Nth Degree Risk.” *Economics Letters* 6(4): 329-33.

Ellsberg, D. 1961. “Risk, Ambiguity, and the Savage Axioms.” *Quarterly Journal of Economics* 75(4): 643-69.

Esö, P., and L. White. 2004. “Precautionary Bidding in Auctions.” *Econometrica* 72(1): 77-92.

Falk, A., A. Becker, T. J. Dohmen, B. Enke, D. Huffman, and U. Sunde. 2015. *The Nature and Predictive Power of Preferences: Global Evidence*. IZA Discussion Paper (9504).

Fehr-Duda, H., A. Bruhin, T. Epper, and R. Schubert. 2010. “Rationality on the Rise: Why Relative Risk Aversion Increases with Stake Size.” *Journal of Risk and Uncertainty* 40: 147-80.

Felder, S., and T. Mayrhofer. 2014. “Risk Preferences: Consequences for Test and Treatment Thresholds and Optimal Cutoffs.” *Medical Decision Making* 34(1): 33-41.

Felder, S., and T. Mayrhofer. 2017. *Medical Decision Making: A Health Economic Primer*. 2nd ed. Berlin Heidelberg: Springer.

Fischbacher, U. 2007. “Z-tree: Zurich Toolbox for Readymade Economic Experiments.” *Experimental Economics* 10(2): 171-78.

Frederick, S. 2005. “Cognitive Reflection and Decision Making.” *Journal of Economic Perspectives* 19(4): 25-42.

Gollier, C. and J. W. Pratt. 1996. “Risk Vulnerability and the Tempering Effect of Background Risk.” *Econometrica* 64(5): 1109-23.

Greiner, B. 2015. “Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE.” *Journal of the Economic Science Association* 1(1): 114-25.

Grether, D. M., and C. R. Plott. 1979. “Economic Theory of Choice and the Preference Reversal Phenomenon.” *American Economic Review* 69(4): 623-38.

Grisley, W., and E. Kellog. 1987. “Risk-taking Preferences of Farmers in Northern Thailand: Measurements and Implications.” *Agricultural Economics* 1: 127-42.

Haering, A., T. Heinrich, and T. Mayrhofer. 2017. „*Exploring the consistency of higher-order risk preferences*”. Ruhr Economic Papers No. 688.

Harrison, G. W., J. Martínez-Correa, and J. T. Swarthout. 2015. “Reduction of Compound Lotteries with Objective Probabilities: Theory and Evidence.” *Journal of Economic Behavior and Organization* 119: 32-55.

Harrison, G. W., and J. T. Swarthout. 2014. “Experimental Payment Protocols and the Bipolar Behaviorist.” *Theory and Decision* 77(3): 423-38.

Heinrich, T., and T. Mayrhofer. 2018. “Higher-order Risk Preferences in Social Settings.” *Experimental Economics* 21(2): 434-56.

Heinrich, T., and J. Shachat. 2018. “The Development of Risk Aversion and Prudence in Chinese Children and Adolescents.” Mimeo.

Herrmann, B., C. Thöni, and S. Gächter. 2008. “Antisocial Punishment across Societies.” *Science* 319(5868): 1362-67.

Hofstede, G. 1980. *Culture’s Consequences: International Differences in Work-related Values*. Beverly Hills, CA: Sage Publications.

Hofstede, G., G. J. Hofstede, and M. Minkov. 2010. *Cultures and Organizations: Software of the Mind*. Volume 3, London: McGraw-Hill.

Holt, C. A., and S. K. Laury. 2002. "Risk Aversion and Incentive Effects." *American Economic Review* 92(5): 1644-55.

Holt, C. A., and S. K. Laury. 2005. "Risk Aversion and Incentive Effects: New Data without Order Effects." *American Economic Review* 95(3): 902-04.

Hsee, C., and E. U. Weber. 1999. "Cross-national Differences in Risk Preference and Lay Predictions." *Journal of Behavioral Decision Making* 12: 165-79.

Kachelmeier, S. J., and M. Shehata. 1992a. "Examining Risk Preferences under High Monetary Incentives: Experimental Evidence from the People's Republic of China." *American Economic Review* 82(5): 1120-41.

Kachelmeier, S. J., and M. Shehata. 1992b. "Culture and Competition: A Laboratory Market Comparison between China and the West." *Journal of Economic Behavior and Organization* 19(2): 145-68.

Kimball, M. S. 1990. "Precautionary Saving in the Small and in the Large." *Econometrica* 58(1): 53-73.

Kimball, M. S. 1992. "Precautionary Motives for Holding Assets." In *The New Palgrave Dictionary of Money and Finance*, edited by P. Newman, M. Milgate, and J. Falwell, 158-161, London: Macmillan.

Kocher, M. G., J. Pahlke, and S. T. Trautmann. 2015. "An Experimental Study of Precautionary Bidding." *European Economic Review* 78: 27-38.

Krieger, M., and T. Mayrhofer. 2012. "Patient Preferences and Treatment Thresholds under Diagnostic Risk: An Economic Laboratory Experiment." Ruhr Economic Papers No. 321.

Krieger, M., and T. Mayrhofer. 2017. "Prudence and Prevention: An Economic Laboratory Experiment." *Applied Economics Letters* 24(1): 19-24.

Krippendorff, K. 2004. *Content analysis: An introduction to its methodology*. Thousand Oaks, CA: Sage Publications.

Lajeri-Chaherli, F. (2004). "Proper Prudence, Standard Prudence and Precautionary Vulnerability." *Economics Letters* 82(1): 29-34.

Levitt, S. D., and J. A. List. 2007. “What do Laboratory Experiments Measuring Social Preferences Reveal about the Real World?” *Journal of Economic Perspectives* 21(2): 153-74.

Lichtenstein, S., and P. Slovic. 1971. “Reversals of Preference between Bids and Choices in Gambling Decisions.” *Journal of Experimental Psychology* 89(1): 46-55.

Lindman, H. R. 1971. “Inconsistent Preferences among Gambles.” *Journal of Experimental Psychology* 89(2): 390-97.

Maier, J., and M. Rüger. 2012. “Experimental Evidence on Higher-order Risk Preferences with Real Monetary Losses.” Working Paper, University of Munich.

Markowitz, H. 1952. “The Utility of Wealth.” *Journal of Political Economy* 60(2): 151-58.

Mayrhofer, T., and H. Schmitz. 2019. “*Prudence and Prevention: An Empirical Investigation*” mimeo.

Nocetti, D. C. 2015. “Robust Comparative Statics of Risk Changes.” *Management Science* 62(5): 1381-92.

Noussair, C. N., S. T. Trautmann, and G. van de Kuilen. 2014. “Higher Order Risk Attitudes, Demographics, and Financial Decisions.” *Review of Economic Studies* 81(1): 325-55.

OECD. 2015. “Purchasing Power Parities (PPP)”, doi: 10.1787/1290ee5a-en (Accessed on 02 March 2015).

Oosterbeek, H., R. Sloof, and G. van de Kuilen. 2004. “Cultural Differences in Ultimatum Game Experiments: Evidence from a Meta-analysis.” *Experimental Economics* 7(2): 171-88.

Özer, Ö., Y. Zheng, and Y. Ren. 2014. “Trust, Trustworthiness, and Information Sharing in Supply Chains Bridging China and the United States.” *Management Science* 60(10): 2435-60.

38-71.

Peter, R. 2017. “Optimal self-protection in two periods: On the role of endogenous saving.” *Journal of Economic Behavior & Organization* 137: 19-36.

Pratt, J. 1964. “Risk Aversion in the Small and in the Large.” *Econometrica* 32(1-2): 122-36.

Quiggin, J. 1982. “A Theory of Anticipated Utility.” *Journal of Economic Behavior & Organization* 3(4): 323-43.

Rieger, M. O., M. Wang, and T. Hens. 2015. “Risk Preferences around the World.” *Management*

Science 61(3): 637-48.

Roth, A. E., V. Prasnikar, M. Okuno-Fujiwara, and S. Zamir. 1991. “Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study.” *American Economic Review* 81(5): 1068-95.

Samuelson, P. A. 1952. “Probability, Utility, and the Independence Axiom.” *Econometrica* 20(4): 670-78.

Segal, U. 1987. “The Ellsberg Paradox and Risk Aversion: An Anticipated Utility Approach.” *International Economic Review* 28(1): 175-202.

Segal, U. 1988. “Does the Preference Reversal Phenomenon Necessarily Contradict the Independence Axiom?” *American Economic Review* 78(1): 233-36.

Segal, U. 1990. “Two-stage Lotteries without the Reduction Axiom.” *Econometrica* 58(2): 349-77.

Slovic, P. 1969. “Manipulating the Attractiveness of a Gamble Without Changing its Expected Value.” *Journal of Experimental Psychology* 79(1 Pt1): 139-45.

Starmer, C., and R. Sugden. 1991. “Does the Random-Lottery Incentive System Elicit True Preferences? An Experimental Investigation.” *American Economic Review* 81(4): 971-78.

Statman, M. 2008. “Countries and Culture in Behavioral Finance.” *CFA Institute Conference Proceedings Quarterly* 25(3): 38-44.

Sutter, M., M. G. Kocher, D. Glätzle-Rützler, and S. T. Trautmann. 2013. “Impatience and Uncertainty: Experimental Decisions Predict Adolescents’ Field Behavior.” *American Economic Review* 103(1): 510-31.

Trautmann, S. T. and G. van de Kuilen. 2018. “Higher order risk attitudes: A review of experimental evidence.” *European Economic Review* 103: 108-124.

UBS. 2014. “Prices and Earnings. Report,” UBS, Zurich. https://www.ubs.com/microsites/prices-earnings/open-data.html.

Vieider, F. M. 2012. “Moderate Stake Variations for Risk and Uncertainty, Gains and Losses: Methodological Implications for Comparative Studies.” *Economics Letters* 117(3): 718-21.

Vieider, F. M., T. Chmura, T. Fisher, T. Kusakawa, P. Martinsson, F. M. Thompson, and A. Sunday. 2015. “Within- versus Between-Country Differences in Risk Attitudes: Implications for Cultural Comparisons.” *Theory and Decision* 78(2): 209-18.

Vieider, F. M., M. Lefebvre, R. Bouchouicha, T. Chmura, R. Hakimov, M. Krawczyk, and P. Martinsson. 2015. “Common Components of Risk and Uncertainty Attitudes across Contexts and Domains: Evidence from 30 Countries.” *Journal of the European Economic Association* 13(3): 421-52.

Weber, E. U., and C. Hsee. 1998. “Cross-cultural Differences in Risk Perception but Cross-cultural Similarities in Attitudes towards Perceived Risk.” *Management Science* 44(9): 1205-17.

White, L. 2008. “Prudence in Bargaining: The Effect of Uncertainty on Bargaining Outcomes.” *Games and Economic Behavior* 62(1): 211-31.

Wik, M., T. Kebede, O. Bergland, and S. Holden. 2004. “On the Measurement of Risk Aversion from Experimental Data.” *Applied Economics* 36: 2443-51.

Figures

Figure 1: Lotteries for eliciting risk preferences up to order 4

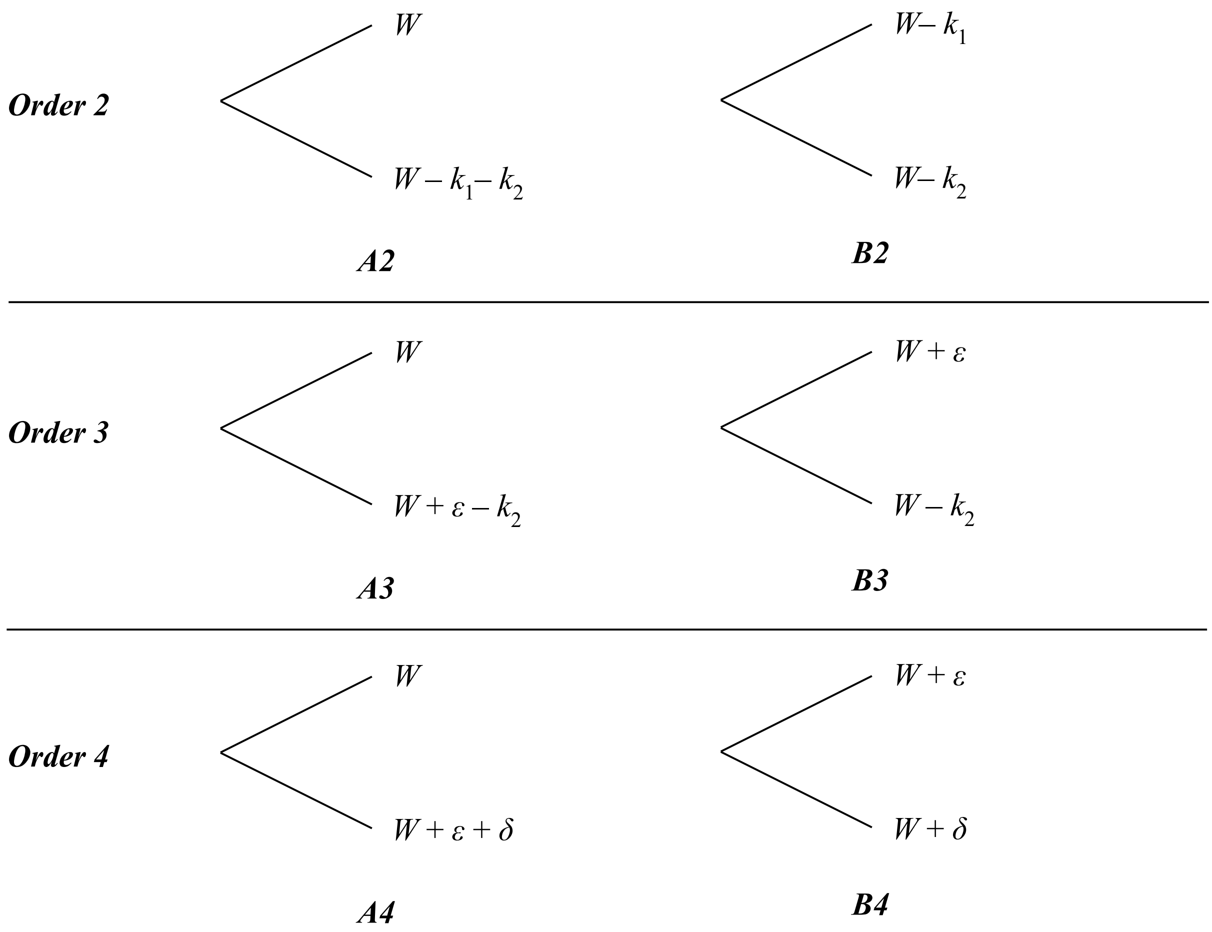
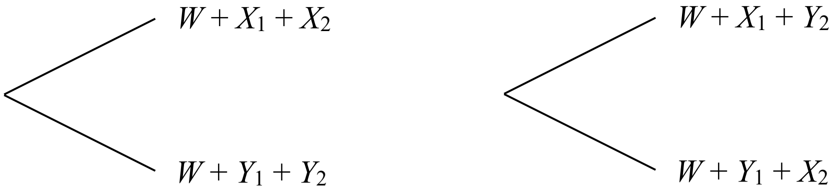
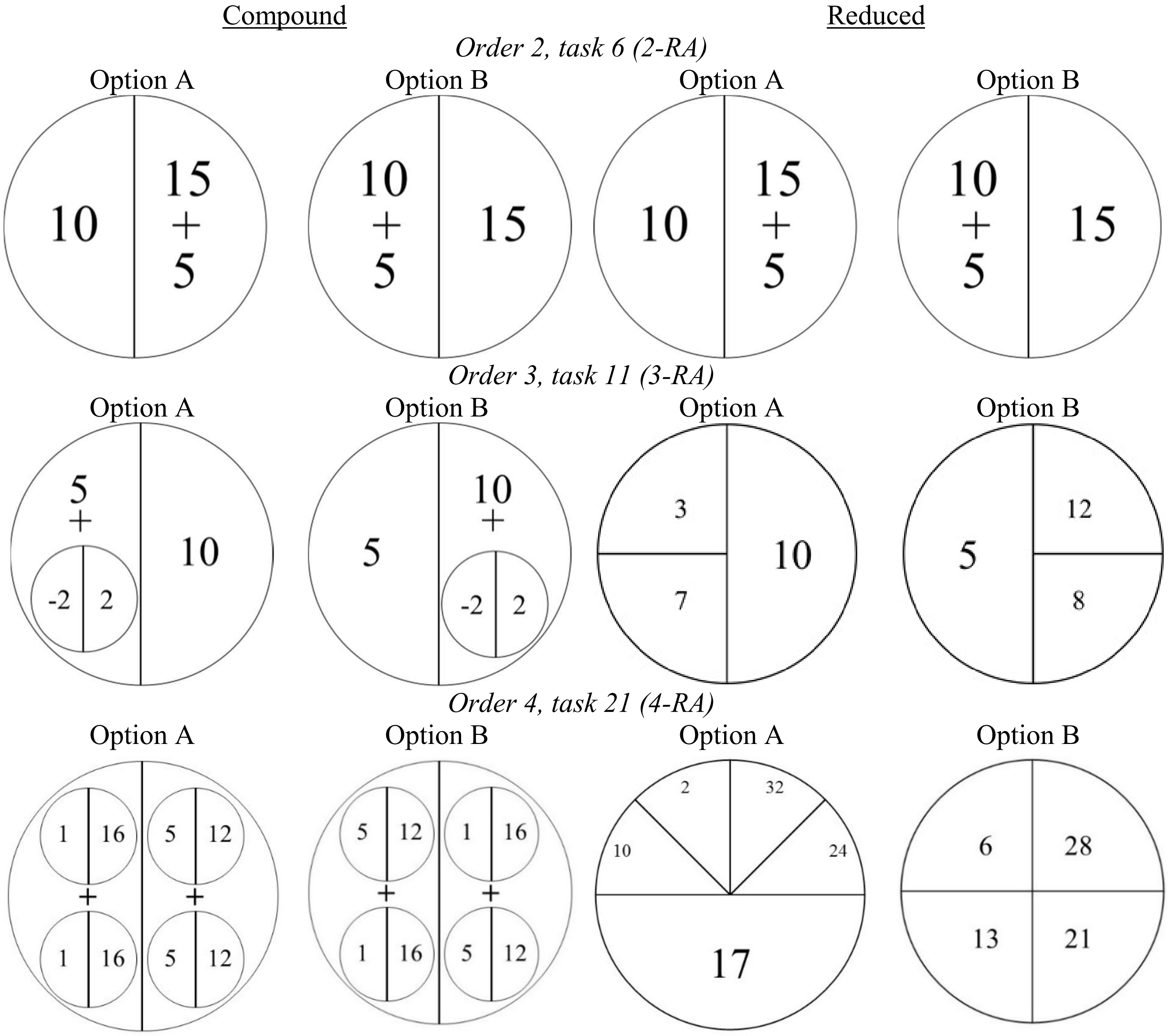


Figure 2: Lotteries for eliciting risk preferences in a general framework



A ***B***

Figure 3: Experimental examples of lotteries of orders 2, 3 and 4 as presented to participants



Note: First- and second-order tasks (here: task 6) do not differ in the compound and reduced presentation.

Figure 4: Average number of n-*th order risk-loving (*n-*RL) choices by risk-averse (RA) and risk-loving (RL) subjects*

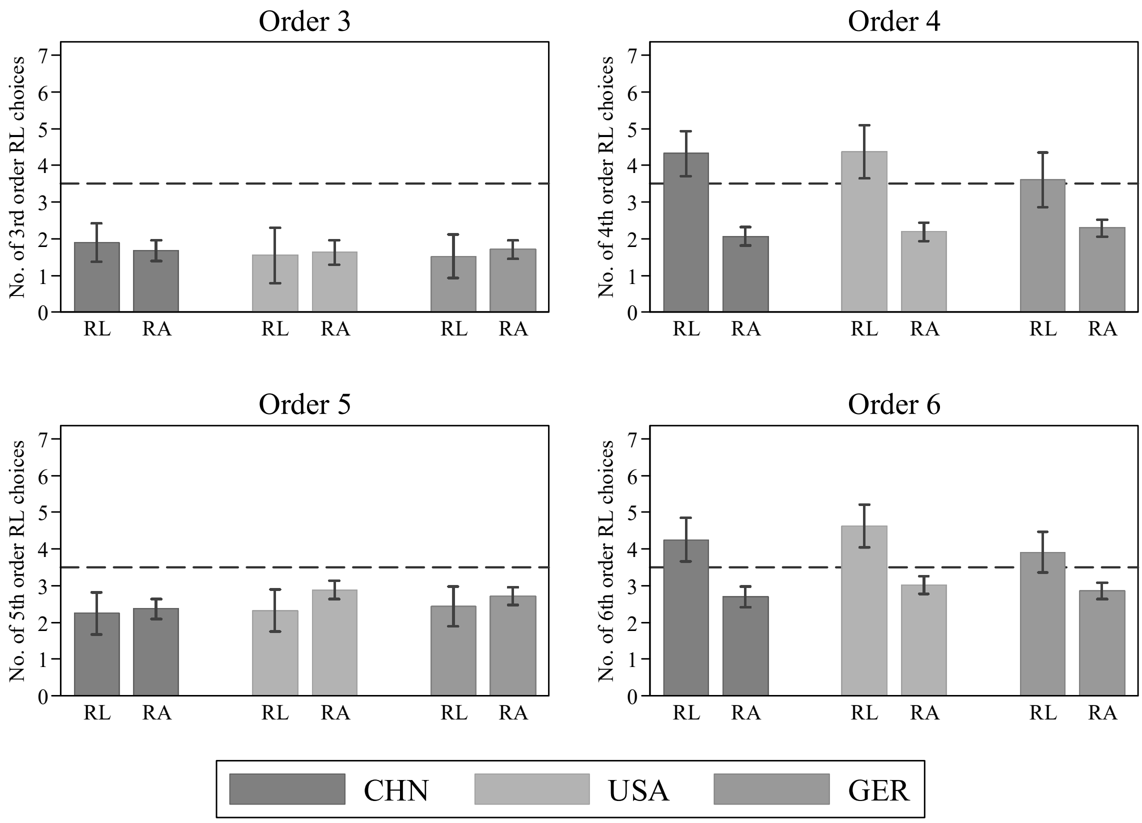


Figure 5: Average number of n*-th order risk-loving (*n*-RL) choices by risk-averse (RA) and risk-loving (RL) subjects*

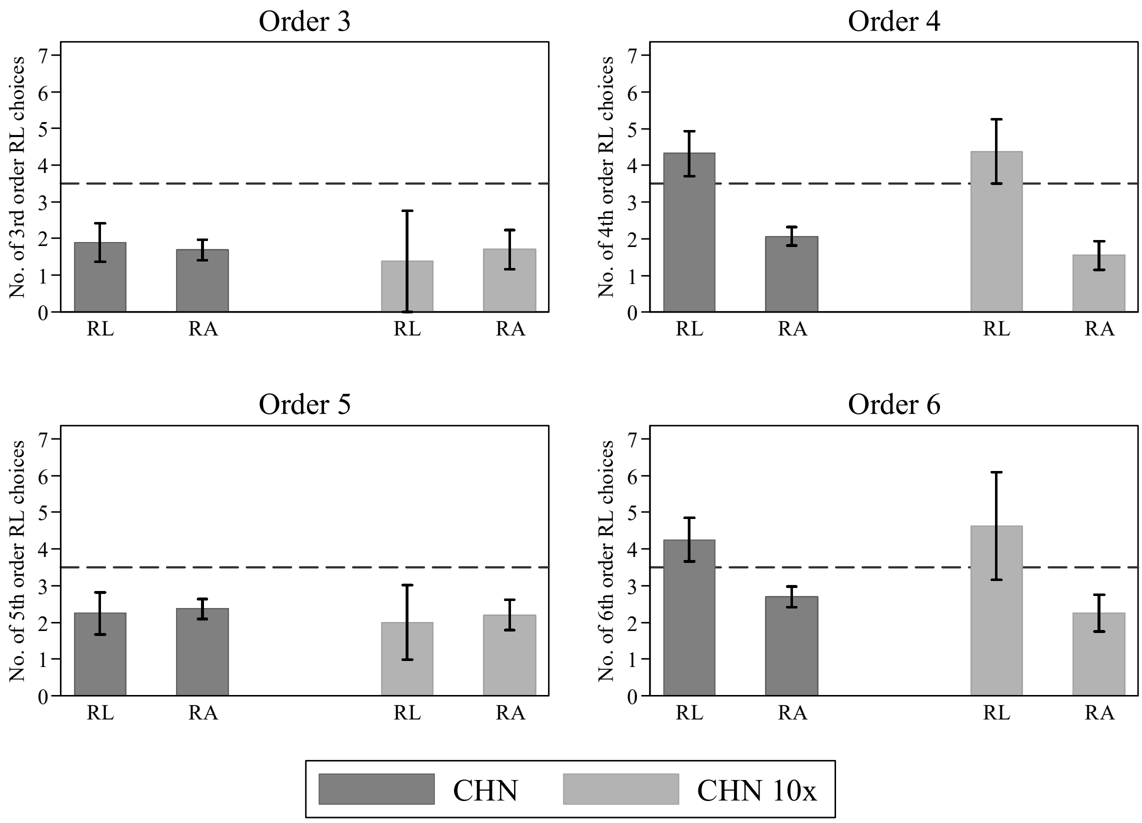
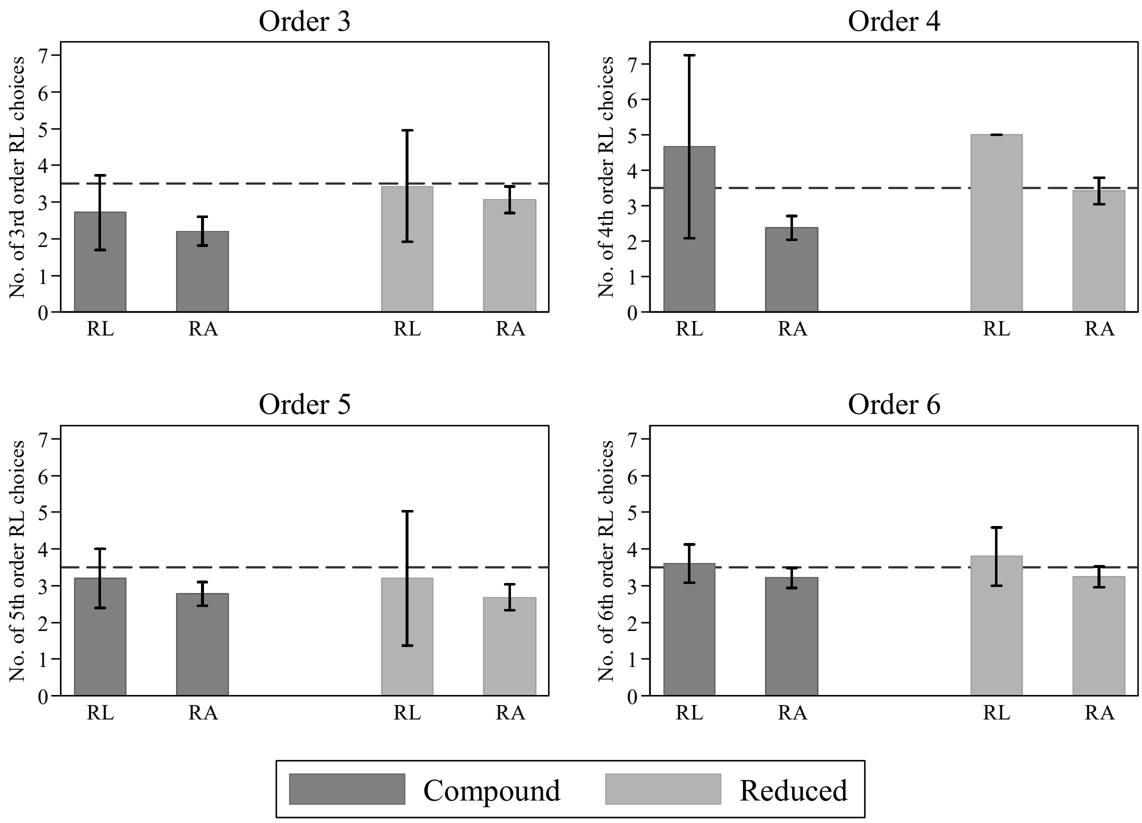


Figure 6: Average number of n*-th order risk-loving (*n*-RL) choices by risk-averse (RA) and risk-loving (RL) subjects*



Tables

Table 1: Treatments and design

	N	Lotteries by Deck and Schlesinger (2014) (order, C = compound, R = reduced)	ECU to local currency[1]	Average payoff local currency[2]
CHN	140	1; 2; 3C; 4C; 5C; 6C	2.90	47.06
USA	129	1; 2; 3C; 4C; 5C; 6C	0.93	28.14
GER	145	1; 2; 3C; 4C; 5C; 6C	0.61	18.10
CHN 10x	48	1; 2; 3C; 4C; 5C; 6C	29.00	571.92
Compound & Reduced	143	1; 2; *i*C; *i*R; *j*C; *j*R $i, j \in \{3, 4, 5, 6\}$, $i \neq j$	0.61	17.40
Follow-up Experiment	224	*i*C or *i*R $i \in \{3, 4\}$	0.61	12.05

[1] Equals \$0.47 (CHN), \$0.68 (GER) and \$4.67 (CHN 10x) at the time of the experiment. [2] Equals \$7.59 (CHN), \$20.18 (GER), \$92.24 (CHN 10x) and \$19.49 (Compound & Reduced) and \$13.43 (Follow-up) at the time of the experiment. All payments except for the Follow-up Experiment included a show-up fee of \$8.50, which was adjusted for China and Germany using the respective exchange rates.

Table 2: Summary statistics

	Demographics		Tests	
	Female	Age (SD)	CRT (SD)	BNT (SD)
CHN (*N* = 140)	57.9%	22.186 (2.337)	1.628 (0.840)	2.879 (1.254)
USA (*N* = 129)	62.0%	23.054 (5.039)	1.667 (1.106)	2.047 (1.262)
GER (*N* = 145)	61.4%	22.993 (2.835)	1.290 (1.154)	1.393 (1.144)
CHN 10x (*N* = 48)	50.0%	22.604 (2.574)	1.688 (0.879)	3.000 (1.187)
Compound & Reduced (*N* = 143)	68.5%	23.818 (3.320)	1.280 (0.982)	1.329 (1.099)
Follow-up Experiment (*N* = 224)	67.9%	24.470 (6.852)	1.201 (1.092)	1.277 (1.001)

N: number of participants, SD: standard deviation, CRT: Number of correct answers out of 3 in the Cognitive Reflection Test, BNT: Number of correct answers out of 4 in the Berlin Numeracy Test.

Table 3: n*-th order risk-loving choices across countries*

Order:	1	2	3	4	5	6
H_0:	1.5	3.5	3.5	3.5	3.5	3.5
CHN						
Mean	0.100***	1.971***	1.721***	2.514***	2.343***	3.007***
Std. Dev.	(0.301)	(1.952)	(1.763)	(1.868)	(1.691)	(1.879)
Median	0	1	1	2	2	3
USA						
Mean	0.008***	1.628***	1.612***	2.558***	2.791***	3.291
Std. Dev.	(0.088)	(1.957)	(2.063)	(1.849)	(1.560)	(1.622)
Median	0	1	1	2	3	3
GER						
Mean	0.034***	1.628***	1.676***	2.500***	2.676***	3.021***
Std. Dev.	(0.183)	(1.900)	(1.700)	(1.680)	(1.615)	(1.516)
Median	0	1	1	2	3	3

* p<0.10, ** p<0.05, *** p<0.01, two-sided, Wilcoxon signed-rank tests.

Table 4: Percentage of subjects who are classified as mixed risk-averse or mixed risk-loving

	Mixed risk-averse				Mixed risk-loving				Mixed risk-averse or -loving+			
Threshold	CHN	USA	GER	All	CHN	USA	GER	All	CHN	USA	GER	All
$p < 0.01$	42%	45%	42%	43%	9%	9%	6%	8%	51%	54%	48%	51%
$p < 0.05$	54%	53%	57%	55%	13%	12%	8%	11%	67%	65%	65%	66%
$p < 0.10$	64%	60%	63%	62%	15%	15%	11%	14%	79%	75%	74%	76%

Classification based on binomial tests with different significance thresholds: $p < 0.01$, $p < 0.05$ or $p < 0.10$, which represent 27, 26 or 25 consistent choices out of 38 possible choices. + In all countries the share of classified subjects is significantly different from the share that would be expected under random behavior ($p < 0.001$, Fisher's exact tests).

Table 5: n-*th order risk-loving (*n-*RL) choices across stakes*

Order:	1	2	3	4	5	6
H_0:	1.5	3.5	3.5	3.5	3.5	3.5
<u>CHN</u>						
Mean	0.100***	1.971***	1.721***	2.514***	2.343***	3.007***
Std. Dev.	(0.301)	(1.952)	(1.763)	(1.868)	(1.691)	(1.879)
Median	0	1	1	2	2	3
<u>CHN 10x</u>						
Mean	0.083***	1.417***	1.646***	2.021***	2.167***	2.646***
Std. Dev.	(0.279)	(1.820)	(2.005)	(1.780)	(1.521)	(1.521)
Median	0	1	1	2	2	3

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, two-sided Wilcoxon signed-rank tests.

Table 6: Share of subjects who are classified as mixed risk-averse or mixed risk-loving

	Mixed risk-averse			Mixed risk-loving			Mixed risk-averse or -loving[+]		
Threshold	CHN	CHN 10x	All	CHN	CHN 10x	All	CHN	CHN 10x	All
$p < 0.01$	42%	54%	45%	9%	4%	8%	51%	58%	53%
$p < 0.05$	54%	65%	57%	13%	13%	13%	67%	78%	70%
$p < 0.10$	64%	73%	66%	15%	15%	15%	79%	88%	81%

Classification based on binomial tests with different significance thresholds: $p < 0.01$, $p < 0.05$ or $p < 0.10$, which represent 27, 26 or 25 consistent choices out of 38 possible choices. [+] In both treatments, the share of classified subjects is significantly different from the share that would be expected under random behavior ($p < 0.001$, Fisher's exact tests).

Table 7: n-*th order risk-loving (*n-*RL) choices across lottery formats*

Order:	1	2	3	4	5	6
H_0:	1.5	3.5	3.5	3.5	3.5	3.5
Compound						
Mean	0.028***	1.266***	2.254***	2.466***	2.806***	3.243
Std. Dev.	(0.165)	(1.404)	(1.810)	(1.708)	(1.526)	(1.268)
Median	0	1	2	2	3	3
Reduced						
Mean			3.099*	3.479	2.722***	3.285
Std. Dev.			(1.790)	(1.872)	(1.730)	(1.342)
Median			3	3	3	3

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, two-sided Wilcoxon signed-rank tests.

Table 8: Percentage of subjects who are classified as mixed risk-averse or mixed risk-loving

	Mixed risk-averse			Mixed risk-loving			Mixed risk-averse or -loving+		
Threshold	Compound	Reduced	Mean	Compound	Reduced	Mean	Compound	Reduced	Mean
$p < 0.01$	41%	27%	34%	1%	1%	1%	42%	28%	35%
$p < 0.05$	50%	42%	46%	1%	1%	1%	51%	43%	47%
$p < 0.10$	57%	52%	56%	2%	3%	3%	59%	55%	59%

Classification based on binomial tests with different significance thresholds: $p < 0.01$, $p < 0.05$ or $p < 0.10$, which represent 19, 18 or 17 consistent choices out of 24 possible choices. + In both formats the share of classified subjects is significantly different from the share that would be expected under random behavior ($p < 0.001$, Fisher's exact tests).

Table A1: Comparison of related papers

Study	Location(s)	Average payoff	Payment	Elicitation method	Lottery type	Share of risk-averse/ prudent/temperate choices
Deck and Schlesinger (2010)	USA	\$25.56	1 of 10 choices	Binary choice	Compound	– / 61% / 38%
Ebert and Wiesen (2011)	Germany	€18.50	1 of 34 choices	Binary choice	Compound	– / 65% / –[1]
Maier and Rüger (2012)	Germany	–	1 of 84 choices	Binary choice	Reduced	56% / 56% / 56%[2]
Deck and Schlesinger (2014)	USA	\$20.92	1 of 38 choices	Binary choice	Compound	74% / 77% / 58%
Ebert and Wiesen (2014)	Germany	€17.50[3]	1 of 120 choices	Risk premia	Compound	66% / 88% / 75%[1]
Heinrich and Mayrho-fer (2014)	Germany	€18.09	1 of 240 choices	Risk premia	Compound	70% / 90% / 76%[1]
Noussair et al. (2014)	Netherlands	Real: 1/10 chance of €70.00[3] Hypothetical: €10,500.00[3]	1 of 17 choices	Binary choice	Compound	72% / 89% / 62%
Deck and Schlesinger (2017)	USA	\$16.66	1 of 52 choices	Binary choice	Compound Reduced	– / 73% / 64%[1] – / 77% / 47%[1]
Baillon et al. (2018)	Netherlands	€18.50[3]	1 of 30 choices	Binary choice	Reduced	84% / 71% / 43%
This study (2018)	China Germany USA	¥20.06 / ¥544.92[4] €12.50 / €11.80[4] \$19.64	1 of 38 choices	Binary choice	Compound Reduced	75% / 76% / 64%[5] – / 56% / 50%

Table following Noussair et al. (2014). The dash (–) indicates that the values are not reported. Average payoff does not include show-up fee. [1]: The amounts represent the share of subjects. [2]: The shares represent the choices across domains. In case of gains these values are 55% / 60% / 58% and in case of losses 57% / 55% / 54%. [3]: This value represents the expected payoff. [4]: The average payoffs in China represent the CHN / CHN 10x values and in case of Germany the GER / Compound & Reduced values. [5]: The shares represent the pooled choices for CHN, USA and GER treatments. In case of CHN 10x these values are 80% / 76% / 71% and in case of Compound 82% / 68% / 65%.

Table A3: Summary of variables

Variable	Description
y_i / y_{it}:	
Order n	Subject's number of *n*-RL choices in order *n*
No of MRA/MRL choices; MRA or MRL	Subject's number of mixed risk-averse/risk-loving choices in all orders; Subject's sum of mixed risk-loving and mixed risk-averse choices
Comp >/</= Redu	Dummy variable indicating that a subject's risk-loving choices in orders 3 to 6 are greater/smaller/equal in Compound compared to Reduced
$\boldsymbol{X}'_i$ / $\boldsymbol{X}'_{it}$:	
Exp.USA	Dummy variable indicating experimenter from USA
Exp.CHN	Dummy variable indicating experimenter from China
IRB	Dummy variable indicating the use of an IRB form
Female	Dummy variable indicating female subjects
Age 18-20	Dummy variable indicating subjects age 18 to 20
Age > 23	Dummy variable indicating subjects age 24 and above
CRT	Number of correct answers CRT (0 to 3)
BNT	Number of correct answers BNT (0 to 4)

Table A4.1: OLS regression

	Order 1		*Order 2*		*Order 3*		*Order 4*		*Order 5*		*Order 6*	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
CHN	0.073	0.078	0.981	1.100	0.136	0.244	0.615	0.719	-0.542	-0.430	-0.072	0.115
	(0.085)	(0.084)	(0.672)	(0.675)	(0.646)	(0.644)	(0.647)	(0.647)	(0.596)	(0.582)	(0.604)	(0.595)
GER	0.049	0.051	0.066	-0.016	-0.180	-0.188	0.102	0.098	-0.407	-0.425	-0.319	-0.344
	(0.058)	(0.059)	(0.590)	(0.595)	(0.584)	(0.573)	(0.572)	(0.563)	(0.491)	(0.489)	(0.515)	(0.506)
Exp.USA	0.029	0.028	-0.590	-0.561	-0.112	-0.107	-0.384	-0.385	-0.398	-0.393	-0.244	-0.235
	(0.048)	(0.048)	(0.460)	(0.463)	(0.481)	(0.484)	(0.427)	(0.424)	(0.339)	(0.339)	(0.420)	(0.409)
Exp.CHN	0.074	0.072	-0.892^{**}	-0.899^{**}	-0.458	-0.478	-0.662^{**}	-0.686^{**}	-0.420	-0.436	-0.568^{*}	-0.603^{*}
	(0.057)	(0.058)	(0.377)	(0.371)	(0.337)	(0.336)	(0.330)	(0.337)	(0.350)	(0.351)	(0.324)	(0.327)
IRB	0.029	0.028	0.448	0.433	-0.258	-0.267	0.480	0.468	-0.031	-0.043	0.064	0.045
	(0.036)	(0.036)	(0.405)	(0.416)	(0.329)	(0.327)	(0.385)	(0.383)	(0.381)	(0.380)	(0.323)	(0.324)
Female	-0.007	-0.007	-0.204	-0.304	0.593^{***}	0.622^{***}	0.086	0.090	0.318^{*}	0.297^{*}	-0.019	-0.001
	(0.021)	(0.021)	(0.208)	(0.203)	(0.185)	(0.177)	(0.188)	(0.184)	(0.173)	(0.167)	(0.176)	(0.172)
Age 18-20	0.008	0.002	-0.283	-0.250	-0.163	-0.122	-0.219	-0.256	0.086	0.063	-0.416^{*}	-0.401^{*}
	(0.022)	(0.022)	(0.240)	(0.233)	(0.228)	(0.224)	(0.231)	(0.223)	(0.215)	(0.211)	(0.219)	(0.216)
Age > 23	0.069^{**}	0.069^{**}	0.082	0.056	0.213	0.240	-0.215	-0.212	0.215	0.196	-0.201	-0.179
	(0.029)	(0.030)	(0.229)	(0.229)	(0.220)	(0.219)	(0.207)	(0.206)	(0.182)	(0.181)	(0.194)	(0.193)
CRT	-0.014^{*}		0.086		-0.106		-0.124		-0.056		-0.164^{*}	
	(0.008)		(0.098)		(0.097)		(0.091)		(0.085)		(0.088)	
BNT	0.002		0.104		0.073		0.064		0.055		0.128^{*}	
	(0.007)		(0.083)		(0.083)		(0.079)		(0.076)		(0.074)	
p-value *CHN=GER*	0.684	0.656	0.009	0.000	0.328	0.151	0.126	0.054	0.692	0.987	0.445	0.142
N	406	414	406	414	406	414	406	414	406	414	406	414
AIC	-114.688	-106.587	1696.322	1724.476	1652.363	1680.044	1640.922	1666.755	1552.701	1585.141	1580.230	1612.268
BIC	-70.619	-70.355	1740.391	1760.709	1696.433	1716.276	1684.992	1702.988	1596.770	1621.373	1624.300	1648.501

Constant not reported, robust standard errors in parentheses, asterisks indicate the significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4.2: OLS regression all orders

	No of MRA choices		*No of MRL choices*		*MRA or MRL*	
	(1)	(2)	(1)	(2)	(1)	(2)
CHN	-1.191	-1.826	1.858	2.041	0.492	0.053
	(2.041)	(2.028)	(1.630)	(1.647)	(1.802)	(1.785)
GER	0.689	0.823	0.388	0.300	1.180	1.107
	(1.716)	(1.712)	(1.473)	(1.440)	(1.513)	(1.482)
Exp.USA	1.697	1.653	-0.738	-0.709	1.882	1.904*
	(1.366)	(1.369)	(1.042)	(1.035)	(1.169)	(1.150)
Exp.CHN	2.927**	3.031***	-1.318	-1.345	2.017*	2.106*
	(1.156)	(1.165)	(0.900)	(0.891)	(1.049)	(1.087)
IRB	-0.731	-0.663	1.252	1.228	0.014	0.044
	(1.099)	(1.099)	(1.056)	(1.058)	(1.013)	(1.017)
Female	-0.768	-0.698	-1.042*	-1.127**	-1.485***	-1.632***
	(0.566)	(0.565)	(0.550)	(0.517)	(0.496)	(0.487)
Age 18-20	0.986	0.965	-0.849	-0.850	0.706	0.628
	(0.704)	(0.694)	(0.655)	(0.627)	(0.624)	(0.620)
Age > 23	-0.162	-0.170	-0.831	-0.841	-0.564	-0.648
	(0.637)	(0.638)	(0.559)	(0.555)	(0.554)	(0.550)
CRT	0.378		-0.026		0.585**	
	(0.286)		(0.241)		(0.257)	
BNT	-0.426*		0.166		-0.180	
	(0.244)		(0.215)		(0.215)	
p-value *CHN=GER*	0.099	0.015	0.094	0.033	0.500	0.290
N	406	414	406	414	406	414
AIC	2545.650	2595.508	2462.897	2501.033	2428.318	2477.862
BIC	2589.720	2631.741	2506.967	2537.266	2472.388	2514.095

Constant not reported, robust standard errors in parentheses,
asterisks indicate the significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A5.1: OLS regression

	Order 1		*Order 2*		*Order 3*		*Order 4*		*Order 5*		*Order 6*	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
CHN10x	-0.024	-0.035	-0.616*	-0.604*	0.011	-0.073	-0.314	-0.380	-0.106	-0.137	-0.111	-0.197
	(0.054)	(0.051)	(0.319)	(0.313)	(0.352)	(0.328)	(0.307)	(0.304)	(0.272)	(0.260)	(0.351)	(0.345)
Female	-0.055	-0.051	-0.236	-0.347	0.292	0.355	0.261	0.213	0.508*	0.450*	0.475	0.412
	(0.047)	(0.044)	(0.300)	(0.284)	(0.289)	(0.280)	(0.279)	(0.272)	(0.266)	(0.256)	(0.293)	(0.284)
Age 18-20	-0.026	-0.045	-0.370	-0.240	-0.075	-0.081	-0.256	-0.366	-0.051	-0.068	-0.604*	-0.583*
	(0.041)	(0.037)	(0.330)	(0.315)	(0.368)	(0.338)	(0.363)	(0.339)	(0.330)	(0.315)	(0.337)	(0.330)
Age > 23	0.129**	0.130*	0.120	0.242	0.188	0.222	-0.774**	-0.673**	0.086	-0.010	-1.023***	-0.887**
	(0.064)	(0.066)	(0.380)	(0.368)	(0.336)	(0.330)	(0.307)	(0.303)	(0.282)	(0.273)	(0.360)	(0.355)
CRT	-0.025		-0.018		-0.315*		-0.114		0.006		-0.248	
	(0.022)		(0.159)		(0.188)		(0.155)		(0.152)		(0.161)	
BNT	-0.010		0.055		0.131		-0.029		-0.007		0.160	
	(0.017)		(0.114)		(0.108)		(0.101)		(0.114)		(0.111)	
N	177	188	177	188	177	188	177	188	177	188	177	188
AIC	70.421	71.174	740.298	783.117	724.952	765.952	726.697	767.163	685.291	726.357	732.864	782.145
BIC	92.654	87.356	762.531	799.299	747.185	782.134	748.930	783.345	707.524	742.540	755.097	798.327

Constant not reported, robust standard errors in parentheses, asterisks indicate the significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A5.2: OLS regression all orders

	No of MRA choices		*No of MRL choices*		*MRA or MRL*	
	(1)	(2)	(1)	(2)	(1)	(2)
CHN10x	1.160	1.427	-0.921	-0.936	0.592	1.114
	(1.083)	(1.045)	(0.843)	(0.860)	(0.909)	(0.869)
Female	-1.245	-1.032	-0.245	-0.477	-1.205	-1.278^{*}
	(0.894)	(0.895)	(0.783)	(0.759)	(0.768)	(0.765)
Age 18-20	1.382	1.382	-1.078	-0.995	0.455	0.509
	(1.100)	(1.070)	(0.956)	(0.899)	(0.951)	(0.923)
Age > 23	1.274	0.975	-2.080^{**}	-1.661^{*}	-0.253	-0.316
	(1.074)	(1.052)	(0.891)	(0.877)	(0.916)	(0.885)
CRT	0.715		-0.046		0.606	
	(0.528)		(0.416)		(0.481)	
BNT	-0.301		0.072		-0.013	
	(0.369)		(0.266)		(0.307)	
N	177	188	177	188	177	188
AIC	1134.686	1204.294	1088.558	1155.179	1076.959	1140.411
BIC	1156.919	1220.476	1110.791	1171.362	1099.192	1156.593

Constant not reported, robust standard errors in parentheses,
asterisks indicate the significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A6.1: Random-effects GLS

	Order 3	*Order 4*	*Order 5*	*Order 6*
Reduced	0.845***	1.014***	-0.083	0.043
	(0.223)	(0.279)	(0.231)	(0.204)
Female	0.111	0.166	-0.442	-0.245
	(0.451)	(0.333)	(0.353)	(0.275)
Age 18-20	-0.682	-0.575	-1.108***	-0.450
	(0.563)	(0.397)	(0.376)	(0.377)
Age > 23	-0.560	-0.590*	-0.544	-0.146
	(0.406)	(0.329)	(0.362)	(0.248)
CRT	-0.160	-0.439**	-0.111	-0.067
	(0.197)	(0.199)	(0.130)	(0.127)
BNT	-0.027	-0.157	-0.103	-0.044
	(0.183)	(0.155)	(0.175)	(0.129)
N	142	146	144	140
N in group	71	73	72	70
χ^2	18.890	34.462	13.285	5.520

Constant not reported, robust standard errors in parentheses,
asterisks indicate the significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A6.2: Random-effects GLS all orders

	No of MRA choices	*No of MRL choices*	*MRA or MRL*
Reduced	-0.916***	0.161	-0.811***
	(0.235)	(0.218)	(0.227)
Female	0.645	-0.260	0.337
	(0.525)	(0.416)	(0.444)
Age 18-20	1.622**	0.504	1.384**
	(0.712)	(0.564)	(0.602)
Age > 23	0.915*	0.233	0.894*
	(0.546)	(0.432)	(0.461)
CRT	0.247	-0.087	0.291
	(0.270)	(0.214)	(0.228)
BNT	0.017	0.216	0.062
	(0.241)	(0.191)	(0.204)
N	286	286	286
N in group	143	143	143
χ^2	22.806	3.198	21.639

Constant not reported, robust standard errors in parentheses,
asterisks indicate the significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A6.3: Logit regression

	Comp > Redu	*Comp < Redu*	*Comp = Redu*
Female	-0.051	0.067	-0.013
	(0.081)	(0.089)	(0.062)
Age 18-20	-0.095	0.007	0.089
	(0.112)	(0.122)	(0.086)
Age > 23	-0.066	0.001	0.072
	(0.083)	(0.093)	(0.072)
CRT	-0.035	-0.004	0.038
	(0.042)	(0.046)	(0.032)
BNT	-0.015	-0.040	0.047*
	(0.038)	(0.041)	(0.027)
N	143	143	143
AIC	180.918	206.374	125.872
BIC	198.695	224.151	143.649

Calculation of marginal effects: Delta-method, constant not reported, robust standard errors in parentheses, asterisks indicate the significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

For Online Publication: Exploring the consistency of higher-order risk preferences – Online Appendix

O1 Choice tasks

Task	Order	Construction	Option A	Option B
1	1	–	20	20 + 10
2	1	–	2	2 + 5
3	1	–	[2 + [10, 20], 20]	[25, 27 + [–1, 1]]
4	2	–	[5, 10 + 5]	[5 + 5, 10]
5	2	–	[2, 4 + 8]	[2 + 8, 4]
6	2	–	[10, 15 + 5]	[10 + 5, 15]
7	2	–	[2, 4 + 3]	[2 + 3, 4]
8	2	–	[20, 40 + 30]	[20 + 30, 40]
9	2	–	[4, 10]	7
10	2	–	[1, 19]	10
11	3	–	[5 + [–2, 2], 10]	[5, 10 + [–2, 2]]
12	3	–	[10 + [–4, 4], 20]	[10, 20 + [–4, 4]]
13	3	–	[5 + [–4, 4], 10]	[5, 10 + [–4, 4]]
14	3	–	[2 + [1, –1], 4]	[2, 4 + [1, –1]]
15	3	–	[20 + [10, –10], 40]	[20, 40 + [10, –10]]
16	3	–	[8 + [2, –2], 10]	[8, 10 + [2, –2]]
17	3	–	[12 + [1, –1], 14]	[12, 14 + [1, –1]]
18	4	2 + 2	[[14, 20] + [14, 20], [10, 24] + [10, 24]]	[[10, 24] + [14, 20], [14, 20] + [10, 24]]
19	4	2 + 2	[[7, 10] + [7, 10], [5, 12] + [5, 12]]	[[5, 12] + [7, 10], [7, 10] + [5, 12]]
20	4	2 + 2	[8B + 7B, 8A + 7A]	[8A + 7B, 8B + 7A]
21	4	2 + 2	[[1, 16] + [1, 16], [5, 12] + [5, 12]]	[[5, 12] + [1, 16], [1, 16] + [5, 12]]
22	4	1 + 3	[14 + 12A, 24 + 12B]	[14 + 12B, 24 + 12A]
23	4	1 + 3	[7 + 11A, 12 + 11B]	[7 + 11B, 12 + 11A]
24	4	1 + 3	[1 + 11A, 18 + 11B]	[1 + 11B, 18 + 11A]
25	5	2 + 3	[[7, 10] + 11B, [5, 12] + 11A]	[[7, 10] + 11A, [5, 12] + 11B]
26	5	2 + 3	[5B + 12B, 5A + 12A]	[5B + 12A, 5A + 12B]
27	5	2 + 3	[8B + 11B, 8A + 11A]	[8B + 11A, 8A + 11B]
28	5	2 + 3	[[5, 12] + 11B, [1, 16] + 11A]	[[5, 12] + 11A, [1, 16] + 11B]
29	5	1 + 4	[5 + 19A, 7 + 19B]	[5 + 19B, 7 + 19A]
30	5	1 + 4	[1 + [5B + [7, 10], 5A + [5, 12]], 4 + [5A + [7, 10], 5B + [5, 12]]]	[1 + [5A + [7, 10], 5B + [5, 12]], 4 + [5B + [7, 10], 5A + [5, 12]]]
31	5	1 + 4	[1 + 20A, 20 + 20B]	[1 + 20B, 20 + 20A]
32	6	3 + 3	[11A + 11A, 11B + 11B]	[11A + 11B, 11B + 11A]
33	6	3 + 3	[11A + 12A, 11B + 12B]	[11B + 12A, 11A + 12B]
34	6	3 + 3	[12A + 14A, 12B + 14B]	[12A + 14B, 12B + 14A]
35	6	3 + 3	[16A + 16A, 16B + 16B]	[16A + 16A, 16B + 16B]
36	6	2 + 4	[[8, 12] + 19B, [5, 15] + 19A]	[[5, 15] + 19B, [8, 12] + 19A]
37	6	2 + 4	[[8, 12] + [5A + [7, 10], 5B + [5, 12]], [5, 15] + [5B + [7, 10], 5A + [5, 12]]]	[[5, 15] + [5A + [7, 10], 5B + [5, 12]], [8, 12] + [5B + [7, 10], 5A + [5, 12]]]
38	6	2 + 4	[[2, 4] + 20B, [5, 1] + 20A]	[[5, 1] + 20B, [2, 4] + 20A]

In this table [*X*, *Y*] denotes a lottery where there is a 50-50 chance of receiving *X* and a 50-50 chance of receiving *Y*. "Task" is the internal task reference number, and table entries of the form #A and #B denote the content of Option A and Option B, respectively, for Task #. "Order" refers to the risk-order being tested. "Construction" refers to the *m* and *n* chosen for decomposing ($m + n$)-th-order risk.

O2 Screenshots (English version)

Figure O2.1: Test of understanding

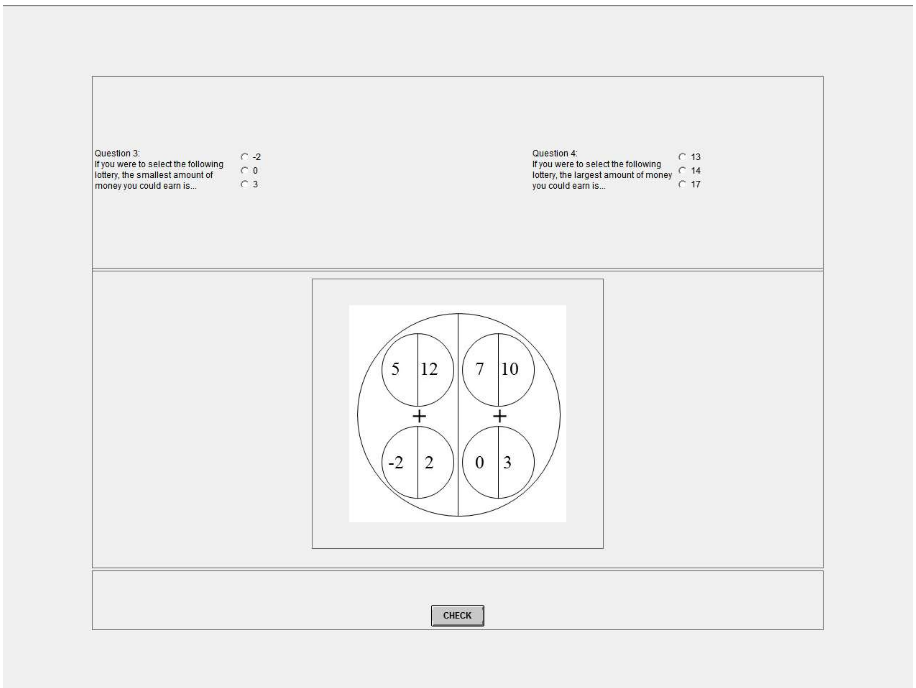
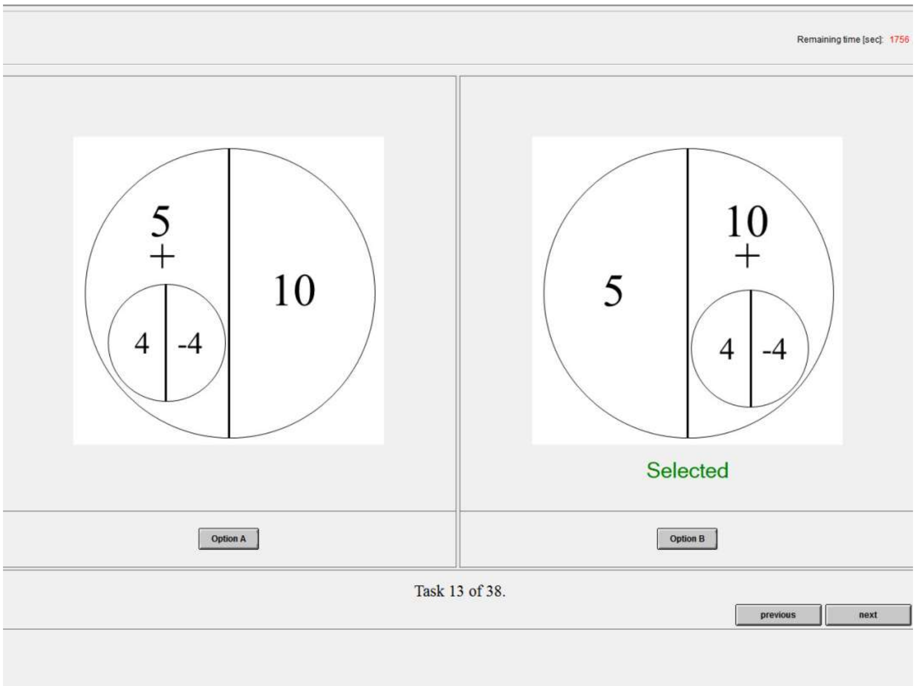


Figure O2.2: Lottery choice



O3 Details: regression results on higher-order risk preferences across countries

We estimate an OLS regression to investigate differences in higher-order risk preferences across China, the USA and Germany using the following equation:

$$y_i = \beta_0 + \beta_1 CHN_i + \beta_2 GER_i + \gamma \boldsymbol{X}'_i + \gamma \boldsymbol{C}'_i + \varepsilon_i \tag{3.1}$$

In equation 3.1 y_i represents a person's number risk-loving choices within one order n or the number of mixed risk-loving (MRA) or mixed risk-averse (MRL) choices across all orders. CHN_i and GER_i are both dummy variables, indicating subjects from China or from Germany. The vector $\boldsymbol{X}'_i$ contains additional explanatory variables to investigate potential effects of experimental procedures (*Exp.USA, Exp.CHN, IRB*), of individual's demographics (*Female, Age 18-20, Age > 23*) or of cognitive and statistical skills (*CRT, BNT*). The vector $\boldsymbol{C}'_i$ contains additional control variables to investigate the robustness of our results. It contains dummy variables for individuals who choose a dominated option in order 1 (*DominatedChoice*), individuals who were not born in the respective country or whose parents were also not born in the respective country (*Immigrant*), individuals with no CRT or BNT test results (*MissingTest*) and individuals who made more than two mistakes on one of the two pages of the test of understanding (*QuizWrong*). In order to ensure that every subject understands the procedure of the experiment, the screen was locked if a subject made more than two mistakes on one of the pages. The subjects were asked to raise their hand if this occurred and were approached by a local experimenter to explain any potential misunderstanding. This was the case for 9 subjects in China, 10 subjects in USA, 19 subjects in Germany, 2 subjects in CHN 10x and 11 subjects in the Compound & Reduced treatment. All variables are described in Table O3.1.

We estimate equation 3.1 without (see Appendix A4) and with $\boldsymbol{C}'_i$ (chapter O4). To avoid problems due to a correlation between the error terms ε_i between subjects in a specific country or from a particular session (heteroscedasticity), we use robust standard errors.

Table O3.1: Summary of variables

Variable	Description
y_i / y_{it}:	
Order n	Subject's number of *n*-RL choices in order *n*
No of MRA/MRL choices; MRA or MRL	Subject's number of mixed risk-averse/risk-loving choices in all orders; Subject's sum of mixed risk-loving and mixed risk-averse choices
Comp >/</= Redu	Dummy variable indicating whether a subject's number of risk-loving choices in orders 3 to 6 in Compound are greater/smaller/equal than in Reduced
$\boldsymbol{X}'_i$ / $\boldsymbol{X}'_{it}$:	
Exp.USA	Dummy variable indicating experimenter from USA
Exp.CHN	Dummy variable indicating experimenter from China
IRB	Dummy variable indicating the use of an IRB form
Female	Dummy variable indicating female subjects
Age 18-20	Dummy variable indicating subjects age 18 to 20
Age > 23	Dummy variable indicating subjects age 24 and above
CRT	Number of correct answers CRT (0 to 3)
BNT	Number of correct answers BNT (0 to 4)
$\boldsymbol{C}'_i$ / $\boldsymbol{C}'_{it}$:	
DominatedChoice	Dummy variable indicating subjects that choose a dominated option in order 1
Immigrant	Dummy variable indicating subjects (or subjects whose parents) were not born in the respective country or who did not answer the question
MissingTest	Dummy variable indicating subjects with no CRT or BNT test results (China only)
QuizWrong	Dummy variable indicating more than 2 mistakes on one of the test of understanding pages
"*A*" x "*B*"	Dummy variable indicating interaction between variables "*A*" and "*B*"

O4 Robustness: regression results on higher-order risk preferences across countries

Table O4.1: OLS regression

	Order 1		*Order 2*		*Order 3*		*Order 4*		*Order 5*		*Order 6*	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
CHN	0.059	0.061	0.671	0.804	0.303	0.388	0.490	0.559	-0.517	-0.459	-0.076	0.066
	(0.081)	(0.080)	(0.663)	(0.675)	(0.633)	(0.637)	(0.648)	(0.653)	(0.602)	(0.590)	(0.608)	(0.601)
GER	0.031	0.030	-0.222	-0.290	-0.070	-0.078	-0.030	-0.057	-0.391	-0.412	-0.342	-0.383
	(0.051)	(0.051)	(0.569)	(0.582)	(0.572)	(0.567)	(0.577)	(0.571)	(0.496)	(0.496)	(0.517)	(0.508)
Exp.USA	0.019	0.018	-0.693	-0.638	-0.161	-0.154	-0.440	-0.441	-0.403	-0.400	-0.255	-0.251
	(0.044)	(0.044)	(0.424)	(0.434)	(0.479)	(0.484)	(0.422)	(0.421)	(0.344)	(0.343)	(0.419)	(0.407)
Exp.CHN	0.068	0.067	-0.917^{**}	-0.907^{**}	-0.643^{**}	-0.647^{**}	-0.675^{**}	-0.697^{**}	-0.441	-0.452	-0.544	-0.581^{*}
	(0.060)	(0.060)	(0.391)	(0.381)	(0.326)	(0.326)	(0.332)	(0.334)	(0.353)	(0.354)	(0.331)	(0.334)
IRB	0.017	0.016	0.326	0.330	-0.297	-0.303	0.413	0.398	-0.036	-0.046	0.048	0.025
	(0.034)	(0.034)	(0.410)	(0.426)	(0.307)	(0.307)	(0.392)	(0.391)	(0.381)	(0.381)	(0.324)	(0.325)
Female	-0.010	-0.014	-0.233	-0.359^{*}	0.599^{***}	0.623^{***}	0.067	0.047	0.318^{*}	0.293^{*}	-0.030	-0.028
	(0.021)	(0.021)	(0.204)	(0.200)	(0.180)	(0.174)	(0.188)	(0.184)	(0.174)	(0.168)	(0.176)	(0.173)
Age 18-20	0.006	-0.001	-0.300	-0.273	-0.181	-0.132	-0.228	-0.278	0.084	0.043	-0.416^{*}	-0.421^{*}
	(0.020)	(0.021)	(0.242)	(0.236)	(0.224)	(0.220)	(0.232)	(0.225)	(0.215)	(0.212)	(0.220)	(0.217)
Age > 23	0.061^{**}	0.060^{*}	0.019	0.000	0.060	0.100	-0.247	-0.251	0.198	0.196	-0.186	-0.163
	(0.030)	(0.031)	(0.229)	(0.230)	(0.212)	(0.211)	(0.211)	(0.210)	(0.187)	(0.186)	(0.195)	(0.194)
CRT	-0.007		0.158		-0.074		-0.083		-0.052		-0.152^{*}	
	(0.008)		(0.097)		(0.096)		(0.093)		(0.089)		(0.092)	
BNT	0.002		0.114		0.063		0.069		0.054		0.130^{*}	
	(0.007)		(0.082)		(0.083)		(0.079)		(0.076)		(0.074)	
DominatedChoice			-0.048	-0.102	1.892^{***}	1.711^{***}	-0.145	-0.134	0.194	0.130	-0.503	-0.508
			(0.364)	(0.354)	(0.426)	(0.437)	(0.440)	(0.420)	(0.360)	(0.350)	(0.329)	(0.319)
Immigrant	0.006	0.006	-0.592^{**}	-0.572^{**}	1.164^{***}	1.170^{***}	-0.164	-0.158	0.154	0.160	0.000	0.012
	(0.031)	(0.031)	(0.288)	(0.285)	(0.383)	(0.379)	(0.299)	(0.303)	(0.305)	(0.304)	(0.289)	(0.288)
MissingTest		0.050		0.452		0.280		0.576		0.983		0.763
		(0.121)		(0.463)		(0.683)		(0.456)		(0.687)		(0.596)
QuizWrong	0.113^{*}	0.119^{*}	1.146^{***}	0.959^{***}	-0.019	0.027	0.670^{*}	0.711^{**}	0.007	0.037	0.290	0.355
	(0.064)	(0.062)	(0.362)	(0.360)	(0.291)	(0.289)	(0.343)	(0.337)	(0.259)	(0.249)	(0.263)	(0.253)
p-value *CHN=GER*	0.656	0.630	0.016	0.001	0.236	0.116	0.125	0.059	0.715	0.885	0.419	0.164
N	406	414	406	414	406	414	406	414	406	414	406	414
AIC	-120.328	-111.905	1688.990	1721.875	1626.611	1658.334	1642.476	1668.785	1558.179	1589.865	1584.062	1616.151
BIC	-68.246	-63.595	1745.079	1774.211	1682.700	1710.671	1698.565	1721.121	1614.267	1642.201	1640.151	1668.488

Constant not reported, robust standard errors in parentheses, asterisks indicate the significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table O4.2: OLS regression all orders

	No of MRA choices		*No of MRL choices*		*MRA or MRL*	
	(1)	(2)	(1)	(2)	(1)	(2)
CHN	-0.871	-1.359	1.299	1.500	0.580	0.381
	(2.013)	(2.019)	(1.578)	(1.620)	(1.786)	(1.771)
GER	1.055	1.221	-0.133	-0.240	1.288	1.298
	(1.693)	(1.707)	(1.426)	(1.404)	(1.513)	(1.488)
EXP.USA	1.953	1.883	-0.824	-0.776	2.030*	2.070*
	(1.330)	(1.348)	(0.978)	(0.975)	(1.171)	(1.146)
EXP.CHN	3.220***	3.283***	-1.053	-1.086	2.318**	2.394**
	(1.155)	(1.151)	(0.890)	(0.875)	(1.041)	(1.066)
IRB	-0.455	-0.404	1.120	1.101	0.162	0.212
	(1.087)	(1.092)	(1.044)	(1.053)	(0.986)	(0.991)
Female	-0.721	-0.575	-1.113**	-1.256**	-1.483***	-1.578***
	(0.564)	(0.563)	(0.540)	(0.515)	(0.496)	(0.488)
Age 18-20	1.041	1.061	-0.848	-0.883	0.744	0.715
	(0.702)	(0.691)	(0.657)	(0.632)	(0.618)	(0.609)
Age > 23	0.155	0.118	-0.671	-0.710	-0.281	-0.386
	(0.638)	(0.641)	(0.551)	(0.547)	(0.558)	(0.554)
CRT	0.203		0.050		0.492*	
	(0.294)		(0.244)		(0.265)	
BNT	-0.430*		0.197		-0.173	
	(0.244)		(0.212)		(0.217)	
DominatedChoice	-2.390**	-2.096**	-3.782***	-3.585***	-3.199***	-3.026***
	(1.054)	(1.061)	(1.205)	(1.162)	(0.954)	(0.935)
Immigrant	-0.563	-0.613	-2.074***	-2.048**	-0.906	-0.913
	(0.986)	(0.972)	(0.795)	(0.804)	(0.816)	(0.829)
MissingTest		-3.053		0.528		-3.321**
		(1.904)		(1.016)		(1.485)
QuizWrong	-2.094**	-2.090**	2.118**	1.960**	-0.649	-1.006
	(0.824)	(0.811)	(0.935)	(0.923)	(0.743)	(0.732)
p-value *CHN=GER*	0.094	0.019	0.103	0.037	0.485	0.353
N	406	414	406	414	406	414
AIC	2541.896	2591.780	2450.380	2491.365	2423.786	2470.473
BIC	2597.985	2644.116	2506.469	2543.702	2479.875	2522.809

Constant not reported, robust standard errors in parentheses,
asterisks indicate the significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

O5 Details: regression results on higher-order risk preferences across stakes

We estimate an OLS regression to investigate differences in higher-order risk preferences across stakes using the following equation:

$$y_{\mathrm{i}} = \beta_0 + \beta_1 CHN10x_{\mathrm{i}} + \gamma \boldsymbol{X}'_{\mathrm{i}} + \zeta \boldsymbol{C}'_{\mathrm{i}} + \varepsilon_{\mathrm{i}} \tag{5.1}$$

In equation 5.1 y_i represents a person's number of risk-loving choices within one order n or the number of mixed risk-loving (MRA) or mixed risk-averse (MRL) choices across all orders. $CHN10x_{\mathrm{i}}$ is a dummy variable indicating subjects that received a tenfold increased payoff. The vector $\boldsymbol{X}'_i$ contains additional explanatory variables and the vector $\boldsymbol{C}'_{\mathrm{i}}$ contains additional control variables (cf. section O3 and Table O3.1).

We estimate equation 5.1 without (see Appendix A5) and with $\boldsymbol{C}'_{\mathrm{i}}$ (chapter O6). To avoid problems due to a correlation between the error terms ε_{i} between subjects from a particular session (heteroscedasticity), we use robust standard errors.

O6 Robustness: regression results on higher-order risk preferences across stakes

Table O6.1: OLS regression

	Order 1		*Order 2*		*Order 3*		*Order 4*		*Order 5*		*Order 6*	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
CHN10x	-0.021	-0.031	-0.648**	-0.635**	0.064	-0.002	-0.354	-0.419	-0.077	-0.115	-0.131	-0.226
	(0.054)	(0.052)	(0.310)	(0.307)	(0.326)	(0.309)	(0.310)	(0.308)	(0.277)	(0.266)	(0.360)	(0.356)
Female	-0.059	-0.058	-0.243	-0.362	0.382	0.417	0.264	0.219	0.542**	0.490*	0.459	0.395
	(0.047)	(0.045)	(0.299)	(0.289)	(0.284)	(0.272)	(0.289)	(0.282)	(0.272)	(0.261)	(0.299)	(0.293)
Age 18-20	-0.029	-0.047	-0.421	-0.294	-0.041	-0.008	-0.277	-0.399	-0.014	-0.072	-0.628*	-0.621*
	(0.039)	(0.037)	(0.330)	(0.314)	(0.363)	(0.337)	(0.369)	(0.346)	(0.335)	(0.320)	(0.343)	(0.340)
Age > 23	0.118*	0.117*	-0.026	0.135	-0.092	-0.049	-0.791**	-0.676**	0.070	0.011	-1.031***	-0.878**
	(0.065)	(0.067)	(0.393)	(0.390)	(0.322)	(0.317)	(0.332)	(0.331)	(0.293)	(0.286)	(0.371)	(0.368)
CRT	-0.020		0.027		-0.250		-0.115		0.003		-0.243	
	(0.022)		(0.163)		(0.178)		(0.162)		(0.156)		(0.165)	
BNT	-0.007		0.067		0.159		-0.037		-0.005		0.160	
	(0.018)		(0.111)		(0.106)		(0.100)		(0.116)		(0.112)	
DominatedChoice			0.762	0.603	1.887***	1.714***	0.284	0.245	0.257	0.190	-0.019	-0.077
			(0.471)	(0.467)	(0.486)	(0.493)	(0.503)	(0.473)	(0.424)	(0.406)	(0.474)	(0.462)
Immigrant	0.024	0.029	3.145***	3.203***	0.222	0.303	1.877***	1.983***	-1.402***	-1.331***	1.124***	1.331***
	(0.045)	(0.048)	(0.359)	(0.335)	(0.385)	(0.367)	(0.397)	(0.369)	(0.346)	(0.314)	(0.402)	(0.385)
MissingTest		0.035		0.622		-0.029		0.236		0.750		0.276
		(0.089)		(0.638)		(0.552)		(0.568)		(0.547)		(0.721)
QuizWrong	0.142	0.159	0.713	0.652	0.481	0.513	-0.216	-0.143	-0.261	-0.226	0.167	0.155
	(0.143)	(0.137)	(0.719)	(0.696)	(0.665)	(0.726)	(0.683)	(0.666)	(0.401)	(0.397)	(0.576)	(0.575)
N	177	188	177	188	177	188	177	188	177	188	177	188
AIC	70.003	71.956	737.351	782.223	711.000	755.907	729.219	771.523	688.028	729.009	736.432	787.404
BIC	95.412	94.611	765.937	808.114	739.585	781.799	757.805	797.415	716.613	754.901	765.017	813.296

Constant not reported, robust standard errors in parentheses, asterisks indicate the significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table O6.2: OLS regression all orders

	No of MRA choices		*No of MRL choices*		*MRA or MRL*	
	(1)	(2)	(1)	(2)	(1)	(2)
CHN10x	1.145	1.397	-1.120	-1.163	0.565	1.049
	(1.040)	(1.025)	(0.838)	(0.851)	(0.872)	(0.848)
Female	-1.405	-1.158	-0.444	-0.655	-1.383*	-1.441*
	(0.895)	(0.896)	(0.813)	(0.791)	(0.776)	(0.772)
Age 18-20	1.380	1.393	-1.270	-1.233	0.412	0.445
	(1.096)	(1.068)	(0.974)	(0.923)	(0.952)	(0.921)
Age > 23	1.870*	1.457	-1.827**	-1.380	0.238	0.150
	(1.105)	(1.084)	(0.914)	(0.933)	(0.923)	(0.891)
CRT	0.578		-0.084		0.507	
	(0.526)		(0.437)		(0.484)	
BNT	-0.344		0.036		-0.042	
	(0.362)		(0.261)		(0.303)	
DominatedChoice	-4.171***	-3.675**	-2.117	-2.132*	-3.794***	-3.718***
	(1.554)	(1.552)	(1.307)	(1.257)	(1.217)	(1.217)
Immigrant	-4.967***	-5.489***	7.326***	7.545***	-3.035***	-3.270***
	(1.215)	(1.165)	(1.051)	(0.954)	(1.056)	(1.002)
MissingTest		-1.855		0.412		-1.049
		(2.008)		(1.878)		(1.491)
QuizWrong	-0.882	-0.950	0.444	0.377	-0.099	-0.326
	(1.908)	(1.932)	(1.972)	(1.950)	(1.735)	(1.795)
N	177	188	177	188	177	188
AIC	1129.652	1201.445	1088.011	1156.382	1071.640	1136.164
BIC	1158.238	1227.337	1116.596	1182.273	1100.225	1162.056

Constant not reported, robust standard errors in parentheses,
asterisks indicate the significance level: * $p< 0.10$, ** $p < 0.05$, *** $p < 0.01$.

O7 Details: regression results on higher-order risk preferences across lottery formats

We estimate a random-effects GLS regression to investigate differences in higher-order risk preferences across lottery formats by the following equation:

$$y_{\mathrm{it}} = \beta_0 + \beta_1 Reduced_{\mathrm{it}} + \gamma \boldsymbol{X}'_{\mathrm{it}} + \zeta \boldsymbol{C}'_{\mathrm{it}} + \upsilon_{\mathrm{it}} \tag{7.1}$$

In equation 7.1 y_{it} represents a person's number of risk-loving choices within one order n or the number of mixed risk-loving (MRA) or mixed risk-averse (MRL) choices across all orders in treatment t. Each participant made choices in two treatments. $Reduced_{\mathrm{it}}$ is a dummy variable indicating whether lotteries where displayed in reduced format in the respective treatment. The vector $\boldsymbol{X}'_{\mathrm{it}}$ contains additional explanatory variables and the vector $\boldsymbol{C}'_{\mathrm{it}}$ contains additional control variables (cf. section O3 and Table O3.1).

We estimate equations 7.1 and 7.2 without (see Appendix A6) and with $\boldsymbol{C}'_{\mathrm{it}}$ (chapter O8). Note that the error term in equation 7.1 υ_{it} has two components, the individual error term τ_{i} (the so called random individual effect) and the regression error term φ_{it}. To avoid problems due to a correlation between the error term υ_{it} (in case of equation 7.1) between subjects from a particular session (heteroscedasticity), we use robust standard errors.

To investigate potential factors that might explain different behavior in the two lottery formats, we also estimate the following logistic regression by maximum likelihood:

$$P(y_{\mathrm{i}} = 1 \mid \boldsymbol{X}'_{\mathrm{i}}\ \boldsymbol{C}'_{\mathrm{i}}) = \Lambda(\beta_0 + \gamma \boldsymbol{X}'_{\mathrm{i}} + \zeta \boldsymbol{C}'_{\mathrm{i}}) \tag{7.2}$$

In equation 7.2 $\Lambda(\bullet)$ is the logistic cumulative density function. Here, y_{i} is a dummy variable indicating that whether a subject's number of risk-loving choices in orders 3 to 6 is greater, smaller or equal in the Compound than in the Reduced treatment. The vectors $\boldsymbol{X}'_{\mathrm{i}}$ and $\boldsymbol{C}'_{\mathrm{i}}$ contain variables described above (and in Table O3.1).

O8 Robustness: regression results on higher-order risk preferences across lottery formats

Table O8.1: Random-effects GLS

	Order 3	*Order 4*	*Order 5*	*Order 6*
Reduced	0.836***	0.945***	-0.190	-0.137
	(0.237)	(0.348)	(0.271)	(0.219)
Female	0.358	0.129	-0.434	-0.179
	(0.439)	(0.334)	(0.324)	(0.308)
Age 18-20	-0.571	-0.563	-1.094***	-0.536
	(0.475)	(0.365)	(0.368)	(0.391)
Age > 23	-0.389	-0.556*	-0.471	-0.104
	(0.416)	(0.326)	(0.363)	(0.248)
CRT	-0.092	-0.374*	-0.093	-0.028
	(0.199)	(0.208)	(0.130)	(0.132)
BNT	0.062	-0.157	-0.027	-0.114
	(0.178)	(0.155)	(0.182)	(0.135)
DominatedChoice	1.378	-3.051***	-0.267	-1.525*
	(1.944)	(0.431)	(0.420)	(0.854)
Reduced x *DominatedChoice*	0.260	4.055***	0.190	1.669***
	(1.377)	(0.348)	(0.271)	(0.441)
Immigrant	-1.501***	-0.253	-0.280	-1.003***
	(0.529)	(0.549)	(0.626)	(0.331)
Reduced x *Immigrant*	0.664	-0.245	0.190	0.137
	(1.142)	(0.690)	(0.674)	(0.432)
QuizWrong	1.618	0.386	0.739*	-0.546
	(1.020)	(0.500)	(0.399)	(0.708)
Reduced x *QuizWrong*	-0.288	0.483	1.190*	1.403
	(1.182)	(0.583)	(0.718)	(0.936)
N	142	146	144	140
N in group	71	73	72	70
χ^2	296.522	–	–	824.255

Constant not reported, robust standard errors in parentheses,
asterisks indicate the significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table O8.2: Random-effects GLS

	No of MRA choices	*No of MRL choices*	*MRA or MRL*
Reduced	-0.731***	0.030	-0.648***
	(0.257)	(0.240)	(0.249)
Female	0.467	-0.317	0.153
	(0.515)	(0.420)	(0.431)
Age 18-20	1.641**	0.364	1.380**
	(0.696)	(0.568)	(0.583)
Age > 23	0.697	0.134	0.664
	(0.540)	(0.441)	(0.452)
CRT	0.096	-0.089	0.153
	(0.265)	(0.216)	(0.222)
BNT	-0.104	0.139	-0.051
	(0.239)	(0.195)	(0.200)
DominatedChoice	-0.199	-3.183**	-0.528
	(1.636)	(1.374)	(1.414)
Reduced x *DominatedChoice*	-2.416*	1.791	-2.060
	(1.422)	(1.329)	(1.381)
Immigrant	0.580	-0.467	0.765
	(0.960)	(0.806)	(0.830)
Reduced x *Immigrant*	-0.102	0.303	-0.185
	(0.844)	(0.789)	(0.820)
QuizWrong	-2.217**	-0.428	-2.003**
	(1.012)	(0.850)	(0.875)
Reduced x *QuizWrong*	-1.413	0.716	-1.164
	(0.882)	(0.825)	(0.857)
N	286	286	286
N in group	143	143	143
χ^2	42.233	9.929	42.353

Constant not reported, robust standard errors in parentheses,
asterisks indicate the significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

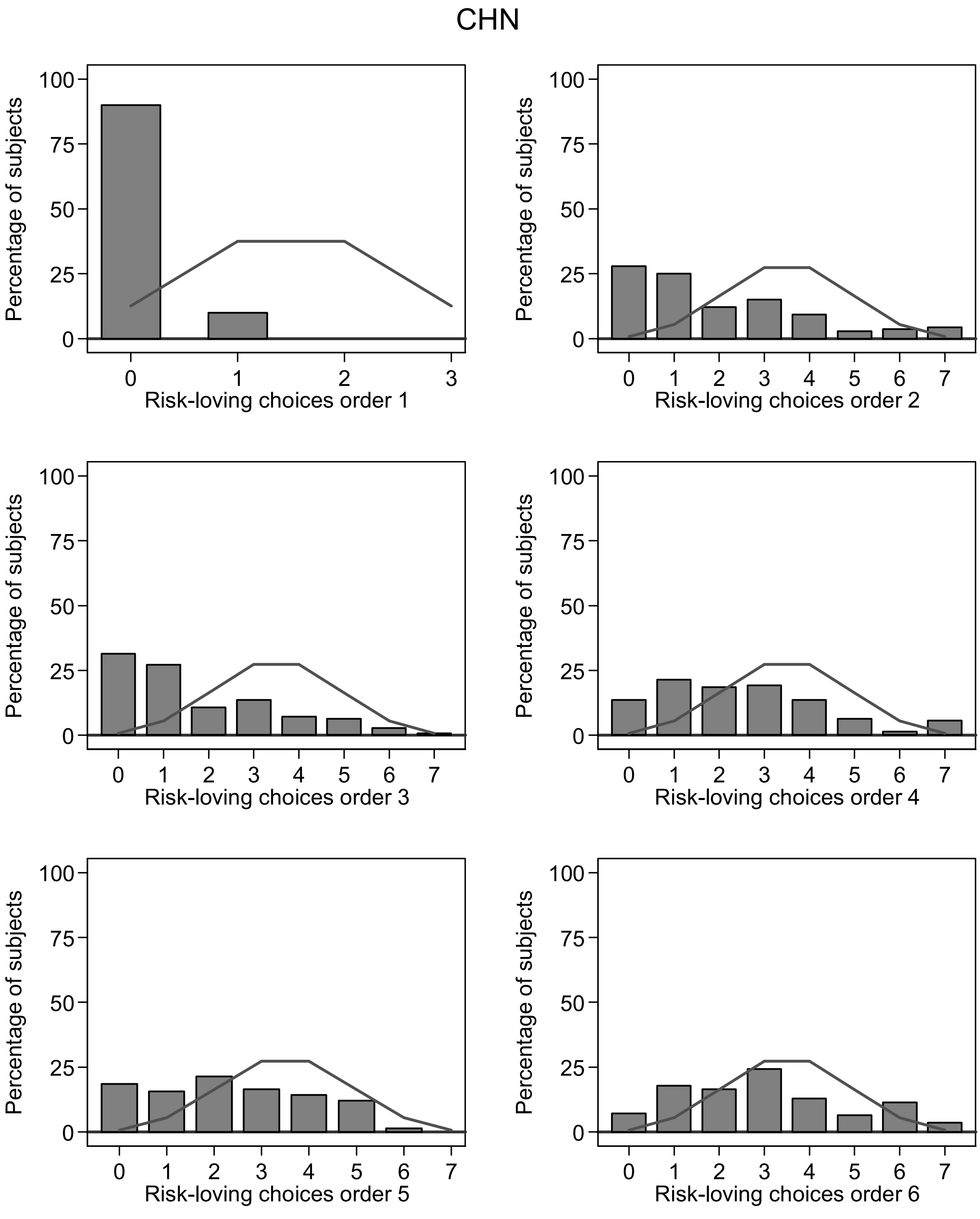
Table O8.3: Logit regression

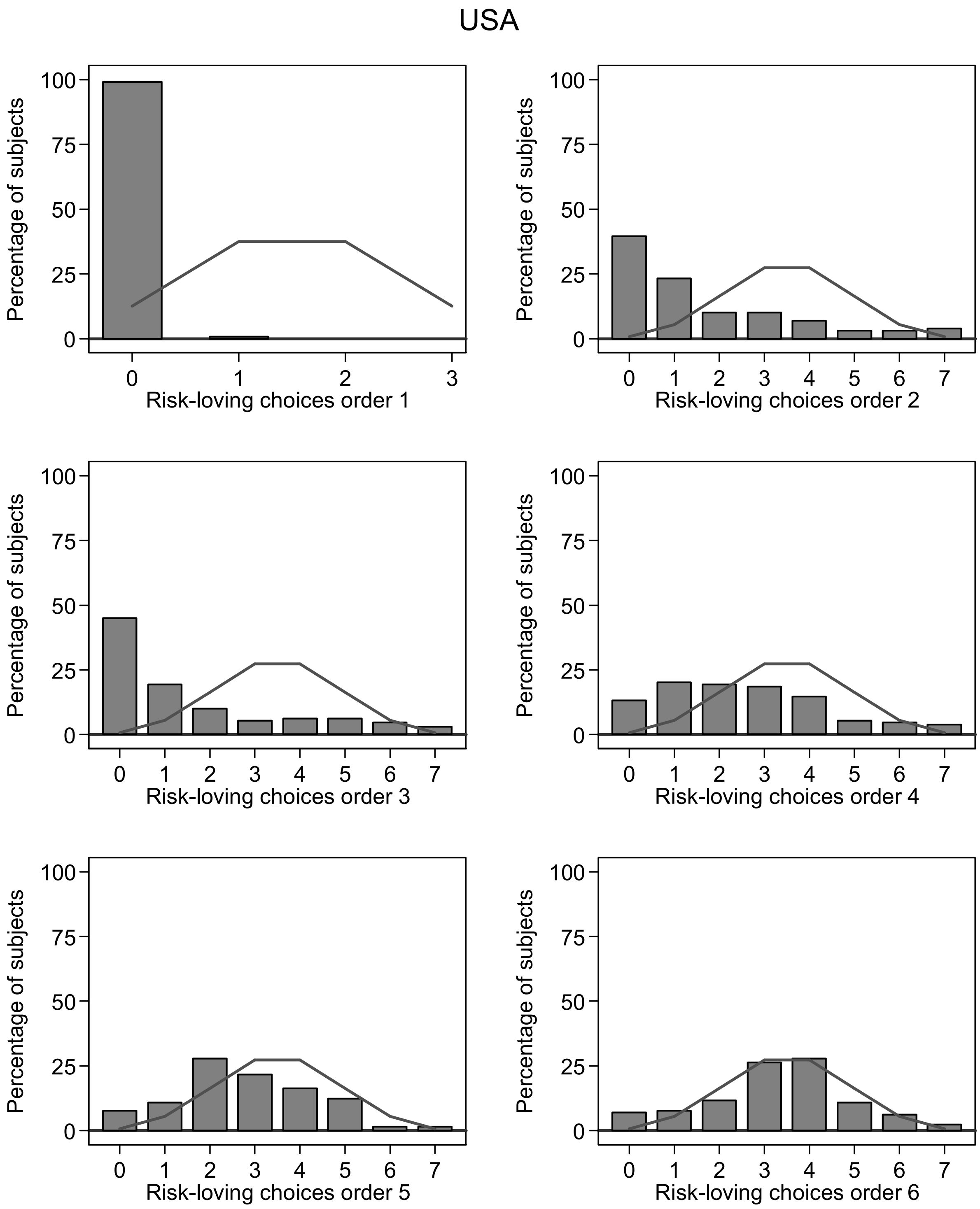
	Comp > Redu	*Comp < Redu*	*Comp = Redu*
Female	-0.074	0.091	-0.024
	(0.083)	(0.092)	(0.069)
Age 18 - 20	-0.122	0.032	0.105
	(0.114)	(0.124)	(0.094)
Age 24 -	-0.100	0.036	0.073
	(0.086)	(0.098)	(0.079)
CRT	-0.039	0.002	0.040
	(0.043)	(0.047)	(0.035)
BNT	-0.029	-0.021	0.045
	(0.040)	(0.042)	(0.030)
Immigrant	0.064	0.027	-0.114
	(0.135)	(0.153)	(0.145)
QuizWrong	-0.051	0.176	–
	(0.153)	(0.178)	
N	139	139	129
AIC	180.527	204.944	123.911
BIC	204.003	228.420	143.929

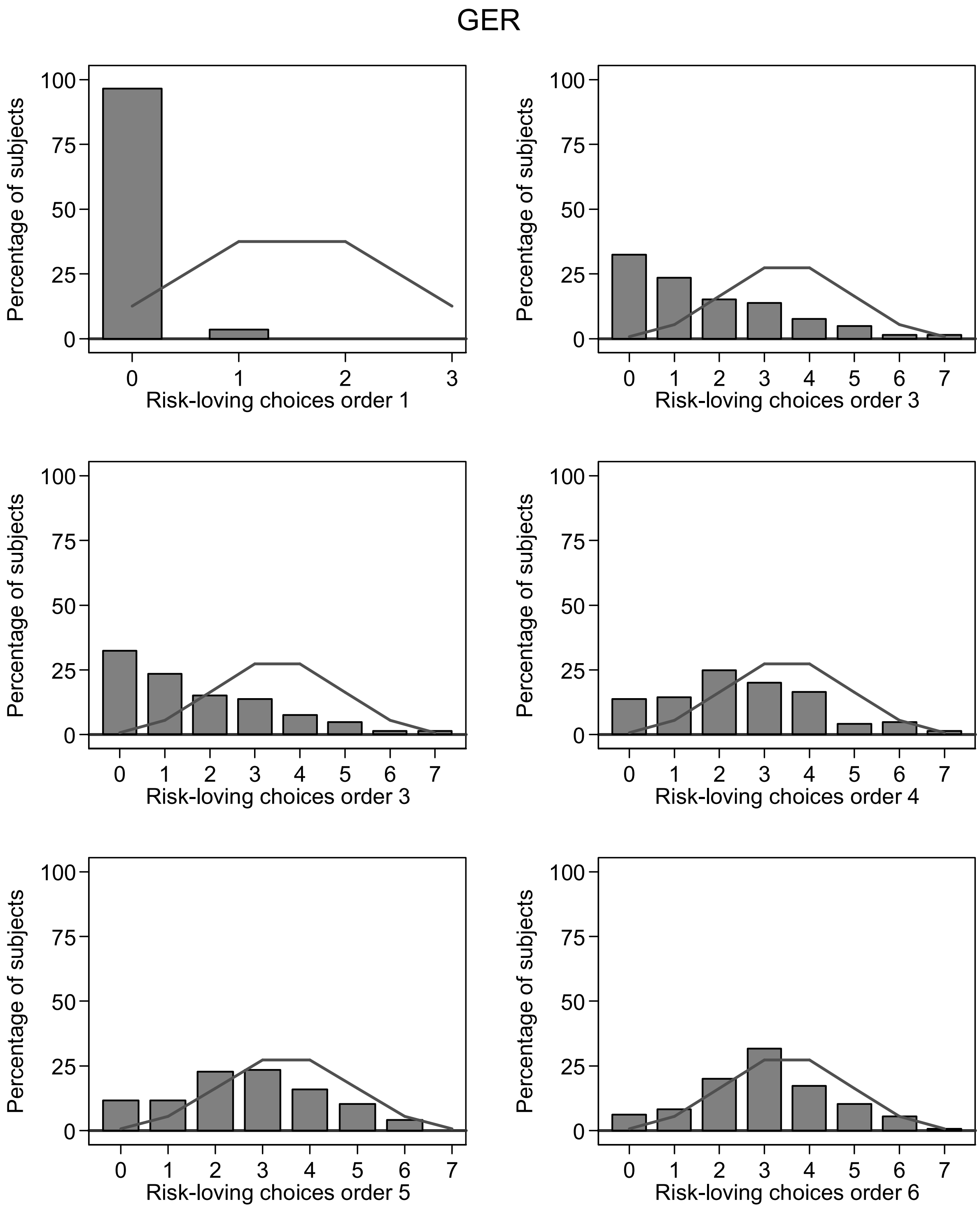
Calculation of marginal effects: Delta-method, constant not reported, robust standard errors in parentheses, asterisks indicate the significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

O9 Distribution of choice frequencies within each order

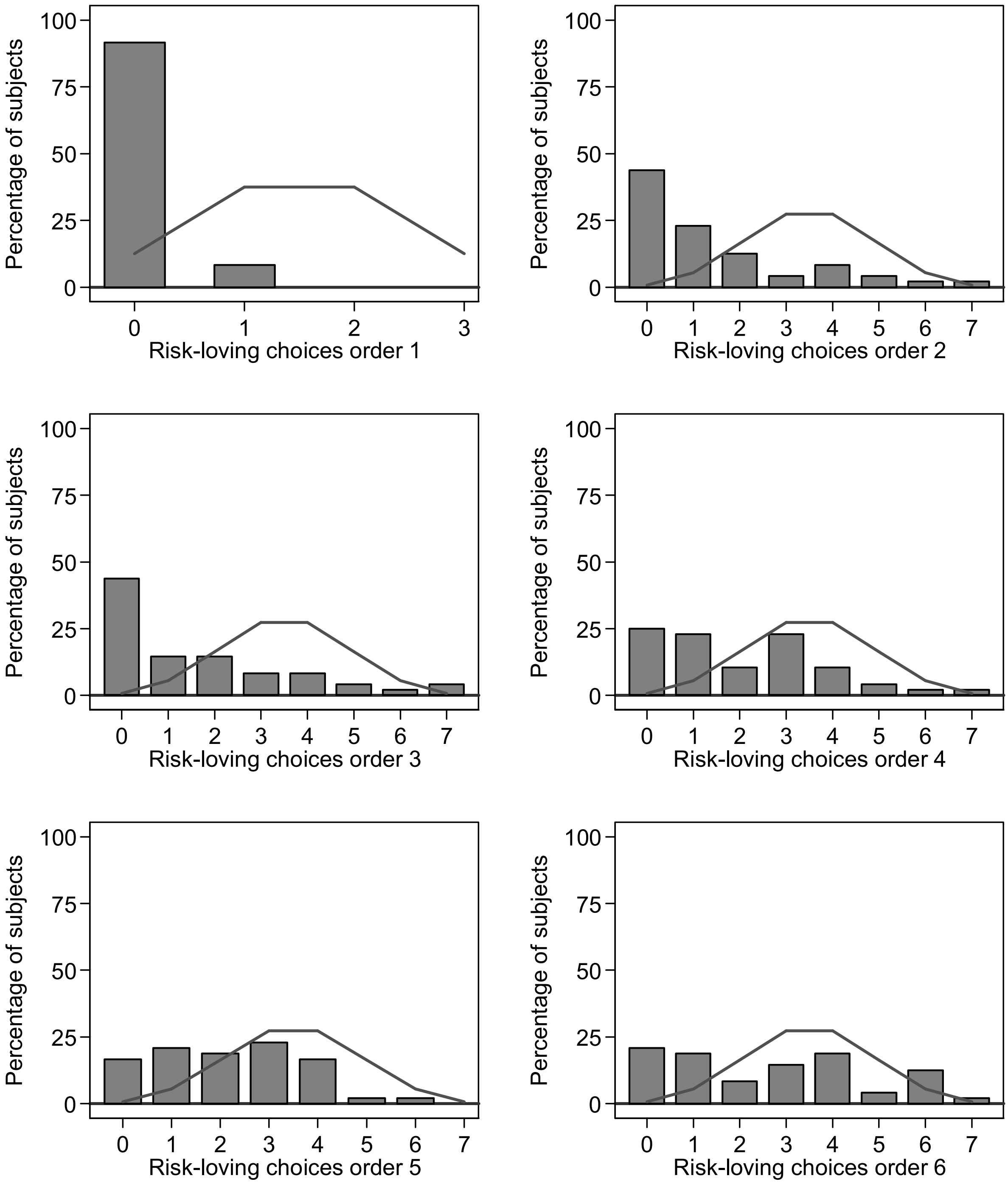
O9.1 All subjects




USA
Percentage of subjects
Risk-loving choices order 1
Risk-loving choices order 2
Risk-loving choices order 3
Risk-loving choices order 4
Risk-loving choices order 5
Risk-loving choices order 6

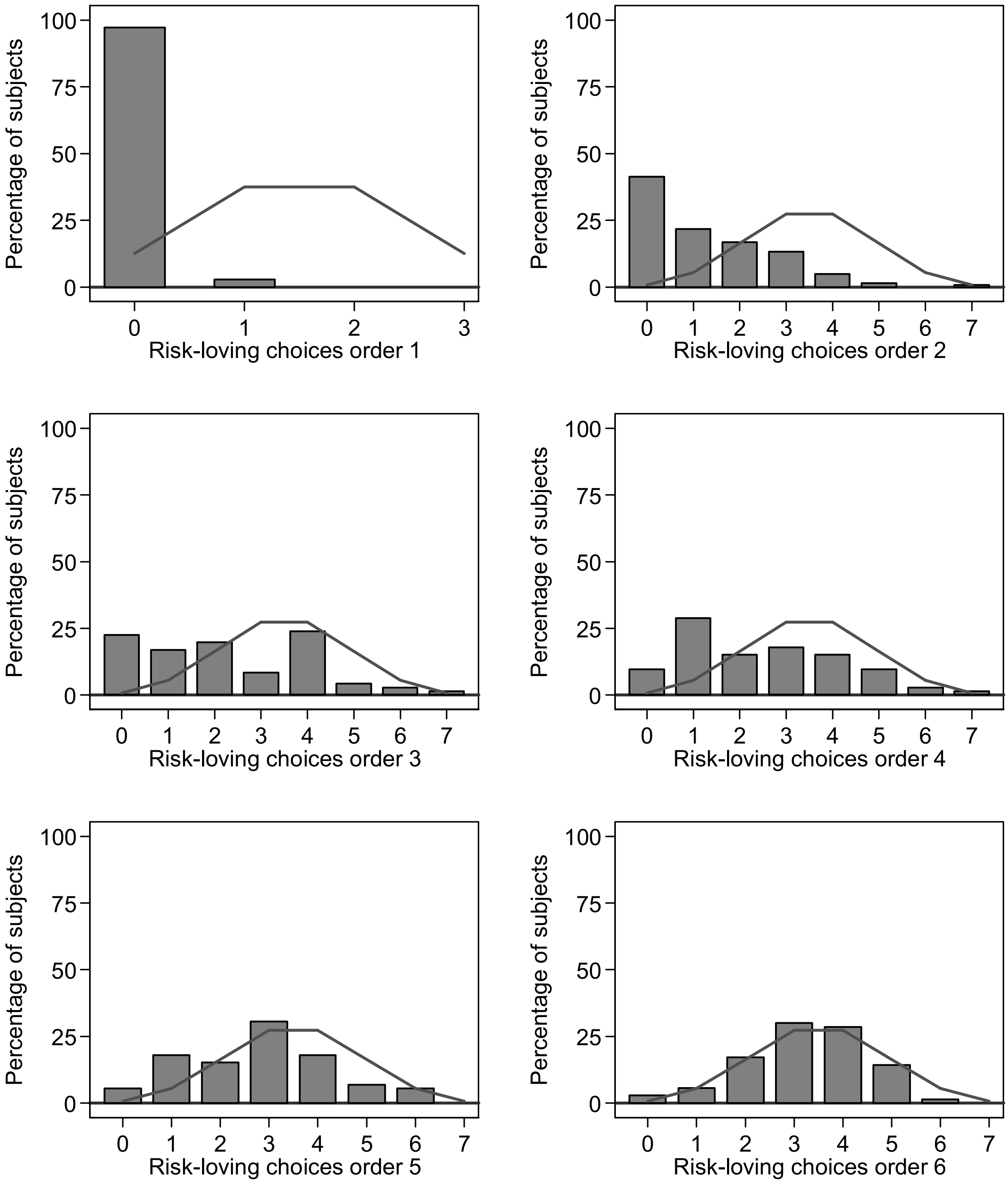

GER
Percentage of subjects
100
75
50
25
0
0
1
2
3
Risk-loving choices order 1
Percentage of subjects
0 1 2 3 4 5 6 7
Risk-loving choices order 3
Percentage of subjects
0 1 2 3 4 5 6 7
Risk-loving choices order 3
Percentage of subjects
0 1 2 3 4 5 6 7
Risk-loving choices order 4
Percentage of subjects
0 1 2 3 4 5 6 7
Risk-loving choices order 5
Percentage of subjects
0 1 2 3 4 5 6 7
Risk-loving choices order 6

CHN 10x

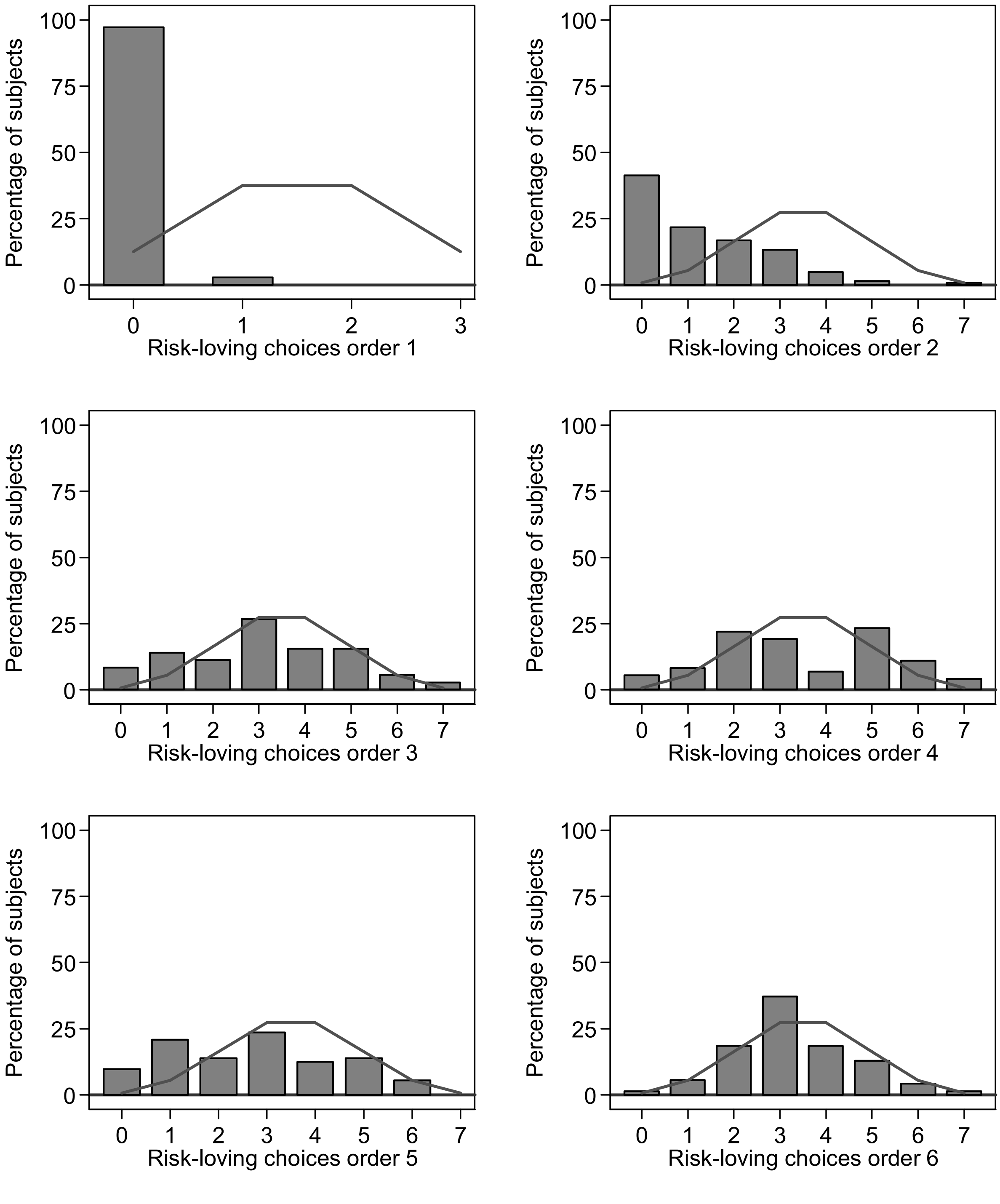


Percentage of subjects
100
75
50
25
0
Risk-loving choices order 1
Risk-loving choices order 2
Risk-loving choices order 3
Risk-loving choices order 4
Risk-loving choices order 5
Risk-loving choices order 6

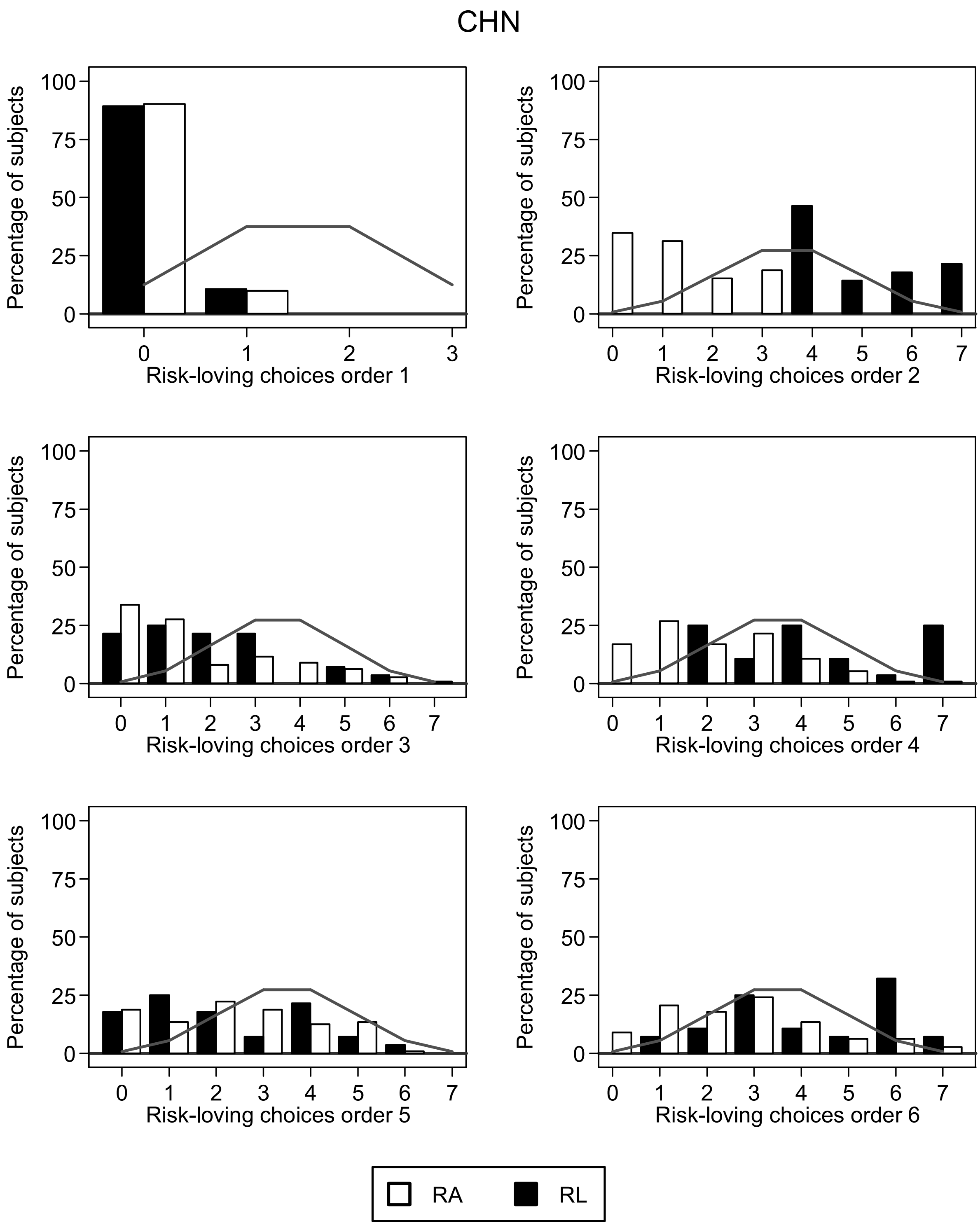
Compound


Percentage of subjects
100
75
50
25
0
Risk-loving choices order 1
Risk-loving choices order 2
Risk-loving choices order 3
Risk-loving choices order 4
Risk-loving choices order 5
Risk-loving choices order 6
0 1 2 3
0 1 2 3 4 5 6 7

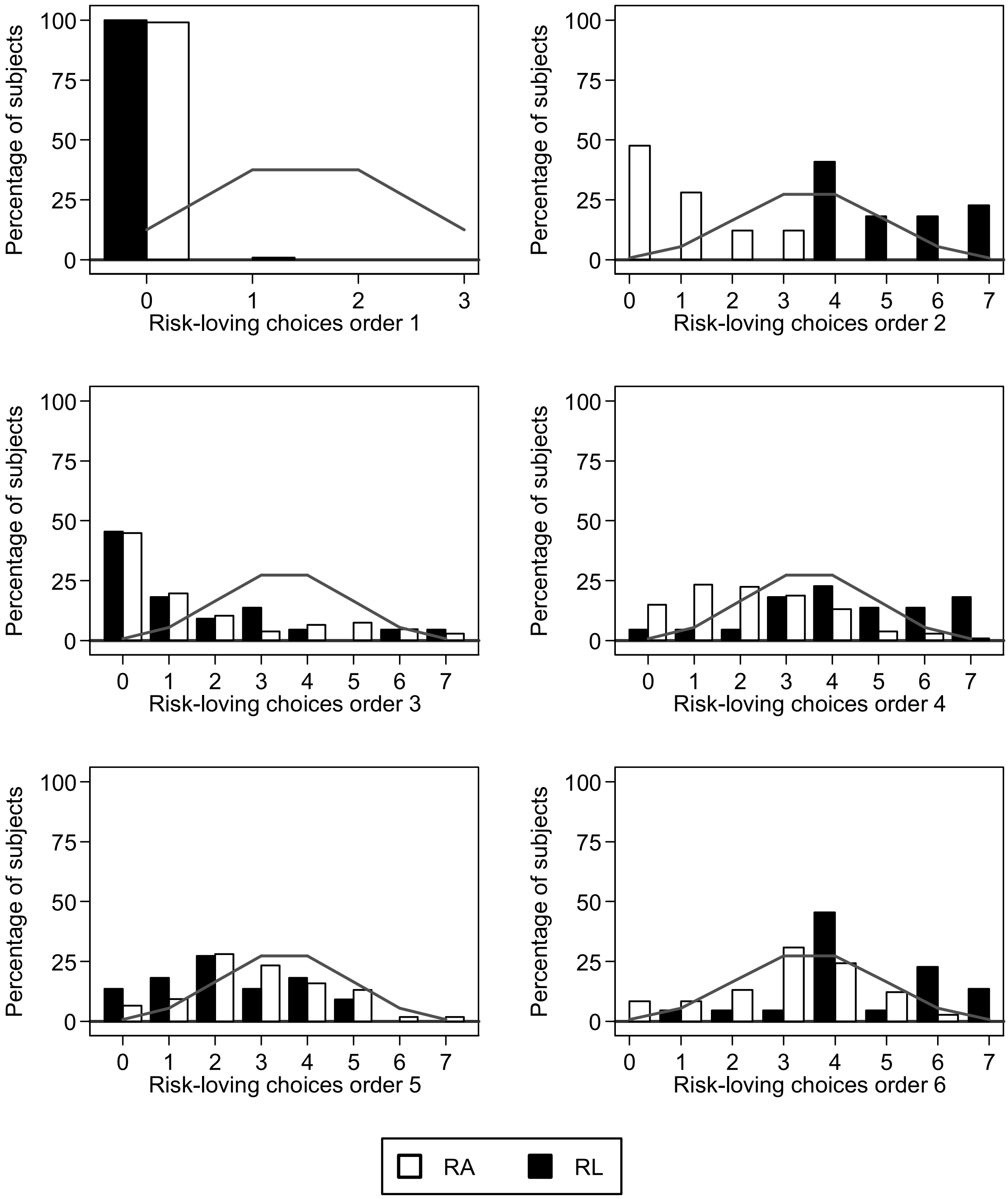
Reduced


Percentage of subjects
100
75
50
25
0
Risk-loving choices order 1
Risk-loving choices order 2
Risk-loving choices order 3
Risk-loving choices order 4
Risk-loving choices order 5
Risk-loving choices order 6

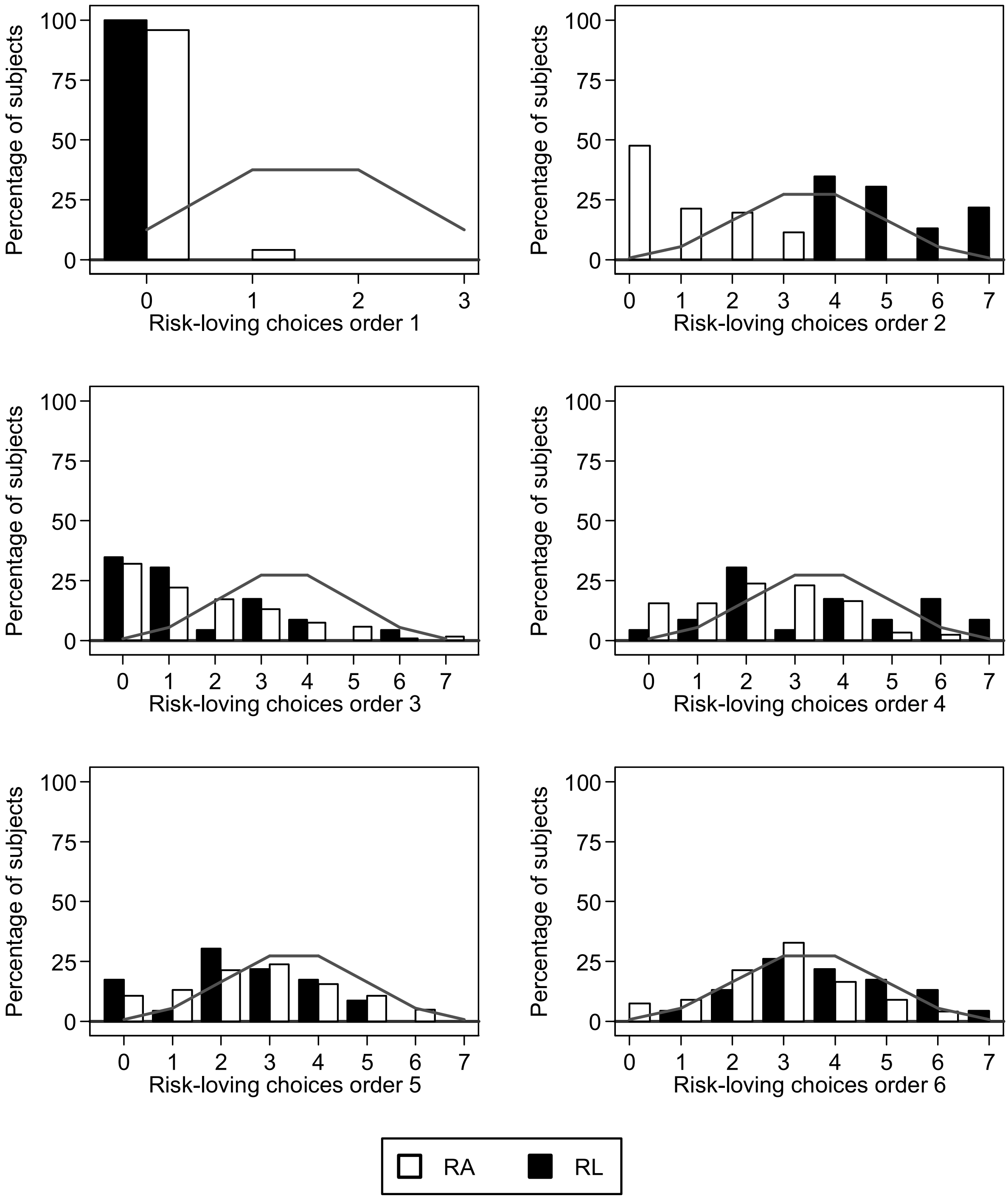
O9.2 Separated by risk-averse (RA) and risk-loving (RL) subjects



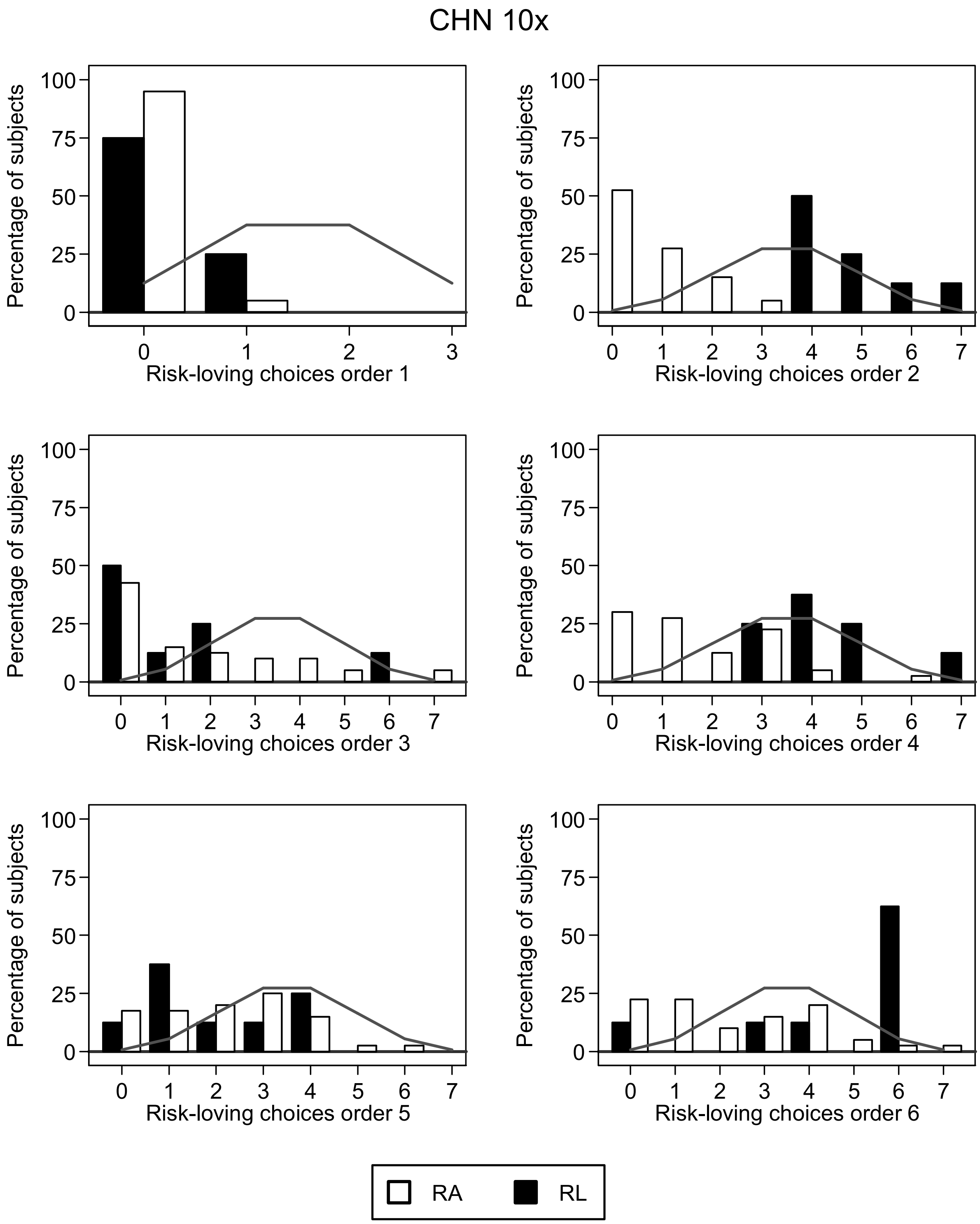
USA


Percentage of subjects
100
75
50
25
0
0 1 2 3
Risk-loving choices order 1
Risk-loving choices order 2
0 1 2 3 4 5 6 7
Risk-loving choices order 3
Risk-loving choices order 4
Risk-loving choices order 5
Risk-loving choices order 6
RA
RL

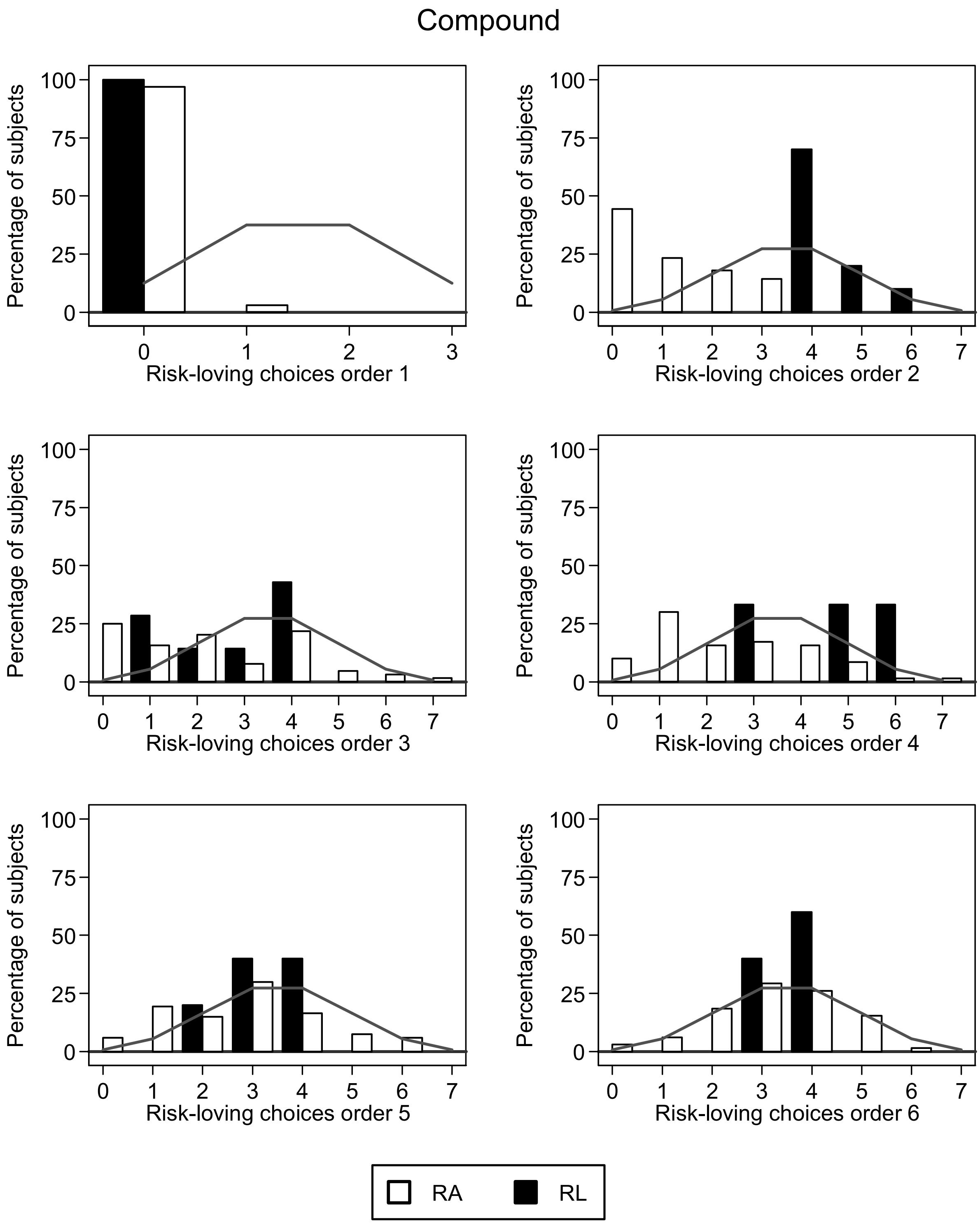
GER



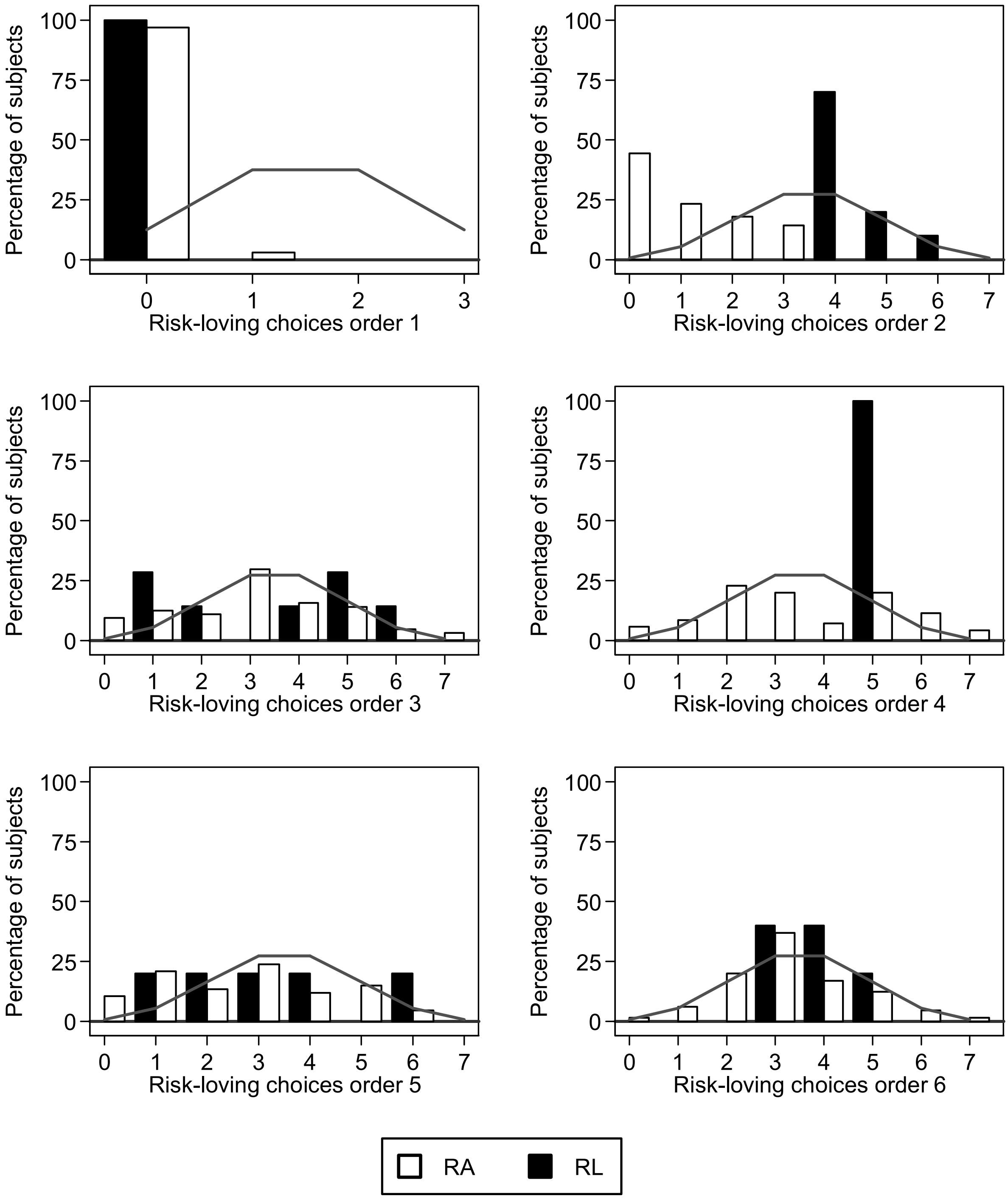
Percentage of subjects
100
75
50
25
0
0 1 2 3
Risk-loving choices order 1
0 1 2 3 4 5 6 7
Risk-loving choices order 2
Risk-loving choices order 3
Risk-loving choices order 4
Risk-loving choices order 5
Risk-loving choices order 6
RA
RL



CHN 10x
Percentage of subjects
0
25
50
75
100
Risk-loving choices order 1
Risk-loving choices order 2
Risk-loving choices order 3
Risk-loving choices order 4
Risk-loving choices order 5
Risk-loving choices order 6
RA
RL


Compound
Percentage of subjects
Risk-loving choices order 1
Risk-loving choices order 2
Risk-loving choices order 3
Risk-loving choices order 4
Risk-loving choices order 5
Risk-loving choices order 6
RA
RL

Reduced



Percentage of subjects
100
75
50
25
0
0
1
2
3
Risk-loving choices order 1
0
1
2
3
4
5
6
7
Risk-loving choices order 2
Risk-loving choices order 3
Risk-loving choices order 4
Risk-loving choices order 5
Risk-loving choices order 6
RA
RL

O10 Correlations of the individuals' share of *n*-th order risk-loving (*n*-RL) choices

The tables below report the correlations of the individuals' share of *n*-RL choices between orders 2 to 6 for all our treatments as in Deck and Schlesinger (2014). If we assume that decisions within a sample were made by equal proportions of mixed risk averters and mixed risk lovers, we would expect the share of *n*-RL choices to be only positively correlated between even orders (all subjects who are 2-RA are also 4-RA) and between odd orders (all subjects who are 3-RA are also 5-RA), but uncorrelated between even and odd orders (subjects who are 3-RA are not necessarily 2-RA). However, if there are only mixed risk averters in a sample, we would expect a positive correlation between *all* orders (all subjects who are 2-RA are also 3-RA *and* 4-RA). If there are only mixed risk lovers in a sample, we would expect a positive correlation between even orders and between odd orders (all subjects who are 2-RL are also 4-RL). Also, we would expect a negative correlation between even and odd orders (all subjects who are 2-RL are also 3-RA). In other words, the correlation of *n*-RL choices between even and odd orders is driven by the underlying sample composition and, thus, cannot be readily compared between samples.

Table O10.1: Correlation of individual n-th order risk-loving (n-RL) choices between orders 2 to 6 across countries

CHN	Order 2	Order 3	Order 4	Order 5
Order 3	0.123			
	(0.148)			
Order 4	0.608	-0.068		
	(0.000)	(0.428)		
Order 5	-0.008	0.245	0.108	
	(0.926)	**(0.004)**	(0.205)	
Order 6	0.439	0.142	0.511	0.067
	(0.000)	**(0.095)**	**(0.000)**	(0.431)
USA				
Order 3	-0.028			
	(0.751)			
Order 4	0.548	0.178		
	(0.000)	**(0.044)**		
Order 5	-0.126	0.302	0.016	
	(0.156)	**(0.001)**	(0.853)	
Order 6	0.488	0.273	0.526	0.099
	(0.000)	**(0.002)**	**(0.000)**	(0.266)
GER				
Order 3	-0.023			
	(0.787)			
Order 4	0.437	0.239		
	(0.000)	**(0.004)**		
Order 5	-0.015	0.437	0.131	
	(0.861)	**(0.000)**	(0.115)	
Order 6	0.388	0.186	0.495	0.281
	(0.000)	**(0.025)**	**(0.000)**	**(0.001)**

p-values reported in parenthesis and bold if $p < 0.100$.

Table O10.2: Correlation of individual n-th order risk-loving (n-RL) choices between orders 2 to 6 for CHN 10x

CHN 10x	Order 2	Order 3	Order 4	Order 5
Order 3	0.123			
	(0.405)			
Order 4	0.680	0.288		
	(0.000)	**(0.047)**		
Order 5	-0.033	0.327	0.180	
	(0.822)	**(0.023)**	(0.222)	
Order 6	0.492	0.250	0.639	0.316
	(0.000)	**(0.086)**	**(0.000)**	**(0.029)**

p-values reported in parenthesis and bold if $p < 0.100$.

Table O10.3: Correlation of individual n-th order risk-loving (n-RL) choices between orders 2 to 6 across lottery formats

Compound	Order 2	Order 3	Order 4	Order 5
Order 3	0.318			
	(0.007)			
Order 4	0.302	0.345		
	(0.009)	**(0.099)**		
Order 5	0.126	0.220	0.415	
	(0.291)	(0.302)	**(0.039)**	
Order 6	0.098	0.346	0.200	-0.172
	(0.420)	(0.106)	(0.350)	(0.432)
Reduced				
Order 3	0.199			
	(0.096)			
Order 4	0.041	0.426		
	(0.732)	**(0.038)**		
Order 5	0.017	-0.235	0.440	
	(0.886)	(0.270)	**(0.028)**	
Order 6	0.292	0.417	-0.016	0.329
	(0.014)	**(0.048)**	(0.942)	(0.125)

p-values reported in parenthesis and bold if $p < 0.100$.

O11 Classification messages Follow-up Experiment

Table O10: Message classification 3-RA

Variable	Alpha	Mean C	Mean R	*p*-value	Compound Mean A	Compound Mean B	Compound *p*-value	Reduced Mean A	Reduced Mean B	Reduced *p*-value
Maximization of the largest potential payoff	0.860	0.216	0.027	**0.004**	0.000	0.284	**0.025**	0.016	0.040	1.000
Maximization of the smallest potential payoff	0.887	0.397	0.071	**0.000**	0.000	0.523	**0.000**	0.000	0.160	**0.014**
Maximization of the probability of the largest potential payoff	0.616	0.026	0.089	0.267	0.071	0.011	0.428	0.161	0.000	**0.058**
Minimization of the probability of the smallest potential payoff	0.742	0.000	0.027	0.239	0.000	0.000	-	0.048	0.000	0.497
Maximization of the payoff in the most likely outcome	0.674	0.026	0.170	**0.004**	0.036	0.023	0.428	0.290	0.020	**0.007**
Minimization of the payoff in the less likely outcome	1.000	0.000	0.000	-	0.000	0.000	-	0.000	0.000	-
N	114	58	56		14	44		31	25	

Alpha gives Krippendorff's alpha for inter-coder reliability; C: compound R: reduced, A: imprudent choice B: prudent choice, all variables are equal to 1 if at least one of the two coders classified the message as containing the respective argument and 0 otherwise; *p*-values result from Fisher's exact tests and are reported in bold if $p < 0.100$.

Table O11: Message classification 4-RA

					Compound			Reduced		
		Mean			Mean			Mean		
Variable	Alpha	C	R	*p*-value	A	B	*p*-value	A	B	*p*-value
Maximization of the largest potential payoff	0.860	0.056	0.071	1.000	0.087	0.032	0.569	0.094	0.042	0.627
Maximization of the smallest potential payoff	0.887	0.259	0.063	**0.023**	0.043	0.419	**0.002**	0.016	0.125	0.153
Maximization of the probability of the largest potential payoff	0.616	0.037	0.045	0.679	0.000	0.065	0.502	0.031	0.063	1.000
Minimization of the probability of the smallest potential payoff	0.742	0.139	0.134	1.000	0.217	0.081	0.264	0.156	0.104	1.000
Maximization of the payoff in the most likely outcome	0.674	0.000	0.054	0.243	0.000	0.000	-	0.094	0.000	0.252
Minimization of the payoff in the less likely outcome	1.000	0.000	0.000	-	0.000	0.000	-	0.000	0.000	-
N	110	54	56		23	31		32	24	

Alpha gives Krippendorff's alpha for inter-coder reliability; C: compound R: reduced, A: intemperate choice B: temperate choice, all variables are equal to 1 if at least one of the two coders classified the message as containing the respective argument and 0 otherwise; *p*-values result from Fisher's exact tests and are reported in bold if $p < 0.100$.