RUHR
ECONOMIC PAPERS

Alexander Haering
Timo Heinrich
Thomas Mayrhofer

Exploring the Consistency of Higher-Order Risk Preferences

#688

Imprint

Ruhr Economic Papers

Published by

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

RWI Leibniz-Institut für Wirtschaftsforschung
Hohenzollernstr. 1-3, 45128 Essen, Germany

Editors

Editorial Office

Ruhr Economic Papers #688

Responsible Editor: Jochen Kluve

ISSN 1864-4872 (online) – ISBN 978-3-86788-798-4

Ruhr Economic Papers #688

Alexander Haering, Timo Heinrich, and Thomas Mayrhofer

Exploring the Consistency of Higher-Order Risk Preferences

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über: *http://dnb.d-nb.de* abrufbar.

Das RWI wird vom Bund und vom Land Nordrhein-Westfalen gefördert.

http://dx.doi.org/10.4419/86788798
ISSN 1864-4872 (online)
ISBN 978-3-86788-798-4

Alexander Haering, Timo Heinrich, and Thomas Mayrhofer[1]

Exploring the Consistency of Higher-Order Risk Preferences

Abstract

In this study we measure 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. A large majority of choices is consistent with a preference for either (i) combining "good" outcomes with "bad" ones or (ii) combining "good" outcomes with "good" ones. 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. We explore potential explanations for this framing effect.

JEL Classification: C91, D81

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

April 2017

1 Alexander Haering, RWI and IN-EAST School of Advanced Studies (UDE); Timo Heinrich, IN-EAST School of Advanced Studies (UDE); Thomas Mayrhofer, Harvard University and Stralsund University. *- We thank Jeannette Brosig-Koch, Louis Eeckhoudt, Glenn Harrison, Heike Hennig-Schmidt, Zwetelina Iliewa, Johanna Kokot, Gijs van de Kuilen, Isabel Marcin, Peter Martinsson, Sonja Opper, Hendrik Schmitz, Jan Siebert, Stefan Traub, Stefan Trautmann, and Toshio Yamagishi for very helpful comments. 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 and the 6th ACCER International Workshop 2017 in Duisburg. Furthermore, we thank Christine E. Looser, Haiying Qin, Weijing Le, Chaoliang Yang and Guanzhong Yang for their support in organizing the experiments. 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.* *- All correspondence to: Timo Heinrich, IN-EAST School of Advanced Studies, University of Duisburg-Essen, Geibelstr. 41, 47057 Duisburg, Germany, e-mail: timo.heinrich@ibes.uni-due.de*

I Introduction

The first researchers who studied people's attitudes towards risk quickly observed that their choices depend on the moments of the probability density distributions over payoffs, as suggested by Tintner (1942). Most people prefer higher expected payoffs, but people vary in their attitudes towards variance, skewness, and kurtosis (see, e.g., Coombs and Pruitt 1960, Alderfer and Bierman 1970, Cooley 1977). An aversion to variance is closely related to risk aversion (but is not identical to it, as pointed out by Rothschild and Stiglitz (1970)). Within the expected utility framework, risk aversion is captured by a negative second-order derivative of the utility function ($U^{\text{II}} < 0$), prudence by a positive third-order derivative ($U^{\text{III}} > 0$), and temperance by a negative fourth-order derivative ($U^{\text{IV}} < 0$). While prudence is closely related to skewness-seeking behavior, temperance is related to kurtosis aversion (see, e.g., Chiu 2005, 2010, Eeckhoudt and Schlesinger 2013, Mayrhofer 2017). In many economic models, preferences over skewness and kurtosis are usually assumed only *implicitly*, even though they drive many important economic decisions. In lifecycle savings models, for example, prudence and temperance determine how current savings are influenced by the riskiness of future income (Kimball 1990, 1992).[1]

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 Brocket and Golden 1987, Caballé and Pomansky 1996). Therefore, these utility functions assume risk aversion, prudence, temperance, and so on. Based on the binary lotteries of Eeckhoudt and Schlesinger (2006), Eeckhoudt, Schlesinger, and Tsetlin (2009) define mixed risk aversion outside the expected utility framework. In their definition, mixed risk aversion is a preference for combining "good" outcomes with "bad" ones. In 2013, Crainich, Eeckhoudt, and Trannoy (2013)

---

1 Based on these theoretical findings, several studies have aimed to identify precautionary savings behavior in the field by analyzing the relationship between the variance of labor income and savings behavior. Carroll and Kimball (2008) review this literature and point out that exogenous variation in risks is difficult to identify empirically. They note that estimates of the precautionary component of wealth lie in an implausible range from a few percent to as much as 50%, and conclude that the estimates are sensitive to elicitation procedures and subject to problems from unobserved heterogeneity (see also Geyer 2011, and Noussair, Trautmann, and van de Kuilen 2014). Other areas in which higher-order risk preferences have been theoretically shown to influence behavior are, for example, auctions (Esö and White 2004), bargaining games (White 2008), prevention (Eeckhoudt and Gollier 2005, Courbage and Rey 2006, 2016), medical decision making (Eeckhoudt 2002, Felder and Mayrhofer 2014, 2017) and research and development expenditures (Nocetti 2015).

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.[2] While mixed risk averters are risk-averse, prudent, and temperate, mixed risk-loving individuals are risk-loving, prudent, and intemperate, or, to put it more generally, (second-order) risk averters coincide in their choices with (second-order) risk lovers in the odd orders but differ in the even orders.

Recently, Deck and Schlesinger (2014) empirically considered individual types *across* multiple orders of risk, and made two major observations. First, in their data a non-negligible minority of individuals make consistently risk-loving choices. Second, in line with the theoretical prediction, they observed a consistent pattern of mixed risk-averse and mixed risk-loving behavior: risk averters tend to be temperate (risk apportionment of order 4) and to show risk apportionment of order 6, while risk lovers tend to choose in the opposite way. Furthermore, both types exhibit prudent (risk apportionment of order 3) and edgy (risk apportionment of order 5) behavior.[3]

In our paper we ask whether this dichotomy can be regarded as a widespread 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 the higher-order risk preferences (up to order 6) of people from three distinct subject pools around the globe. Furthermore, we study the effect of a tenfold increase in the stakes, and we also explore whether the dichotomous pattern of behavior is robust to a straightforward change in the framing of the lotteries.

---

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

[3] Deck and Schlesinger (2014) were not only the first to study mixed risk-averse and mixed risk-loving behavior empirically, 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 (called "bentness" by Ebert, Nocetti, and Schlesinger 2016) has rarely been studied, so far. 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 (2016) 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.

Eeckhoudt and Schlesinger (2006) provided the first definition of higher-order risk preferences outside the expected utility framework. Their definition is based on binary lotteries and lends itself to economic experiments that measure higher-order risk preferences directly. The 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 study by Deck and Schlesinger (2010) – in line with temperance (see Appendix A1 for a more detailed comparison).[4] 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 higher-order risk preferences will help to build more realistic economic models that can incorporate individual heterogeneity in a parsimonious way. To this end, however, it needs to be established whether the previous findings are robust and sufficiently general. With regard to 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). Additionally, it has been observed that displaying compound lotteries in a reduced form may impact the degree of risk aversion (see, e.g., Halevy 2007, and Harrison, Martínez-Correa, and Swarthout 2015). Even though these factors have been studied with respect to (second-order) risk aversion, there has been little work on higher-order risk preferences. Since risk aversion and higher-order risk preferences are related theoretically as well as empirically, we expect that the factors may also influence higher-order risk preferences.

---

[4] Originally, prudence and temperance were defined by Kimball (1990, 1992) within an expected utility framework. The term risk apportionment was introduced by Eeckhoudt and Schlesinger (2006), who defined higher-order risk preferences using preferences over lotteries. Their definition is equivalent to the original definition by Kimball (1990, 1992) but is not confined to expected utility. 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 in medical decision making (Krieger and Mayrhofer 2012, 2017). This has also been studied experimentally in social settings (Heinrich and Mayrhofer 2014) and across multiple domains (Ebert and van de Kuilen 2015, Deck and Schlesinger 2016).

The results we report in this paper suggest that a considerable proportion of people are risk-loving, and that choices can be explained rather well by a dichotomy of preference types. Using the methodology introduced by Deck and Schlesinger (2014), we identify similar degrees of mixed risk-averse and mixed risk-loving behavior in three distinct subject pools in China, in the USA and in Germany as well as before and after a tenfold increase in the stakes. In total, 62% of the participants can be classified as mixed risk-averse and 11% as mixed risk-loving. These results support the view that risk-loving preferences characterize a non-negligible proportion of people. We also discover that the dichotomy of choice patterns can be strengthened through the framing of the lottery.[5] When we display the lotteries in compound rather than reduced form, we observe significantly more prudent and temperate behavior *within* the same subjects.

The paper is structured as follows. In the next section we recapitulate the theoretical background and describe our hypotheses. In section three we present the experimental design. We summarize the findings of our experiments in section four and conclude in section five.

II Theoretical background and hypotheses

A Theoretical background

In this section we present the theoretical background to our experiment, which follows Deck and Schlesinger (2014). Deck and Schlesinger use different sets of lotteries in their experiment, which are based not on the assumptions of expected utility theory (EUT) but on the theoretical work by Eeckhoudt and Schlesinger (2006), Eeckhoudt, Schlesinger, and Tsetlin (2009) and Crainich, Eeckhoudt, and Trannoy (2013). The lottery sets consist of binary lotteries 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.

[5] A very recent study by Deck and Schlesinger (2016) that was conducted in parallel with ours makes a similar observation with respect to the framing of lotteries. These authors compare behavior in reduced and compound lotteries (see Sections II.B and V). Furthermore, they replicate the observations made in Deck and Schlesinger (2014) and also consider choices when the risks are non-monetary.

Figure 1 shows risk apportionment 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 of lottery pairs in Figure 1 illustrates a risk aversion (risk apportionment of order 2) task in which risk aversion is a preference for disaggregating harms, i.e. the sure losses k_1 and k_2. Disaggregating the harms reduces the spread between the two possible outcomes. This corresponds to a lower variance and thus a lower second-order risk. Lottery *A2* has a greater spread (and thus variance) than lottery *B2*. Therefore, 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*.

The second row of lottery pairs in Figure 1 shows a prudence task (risk apportionment of order 3). In this case the sure loss k_1 is replaced by a zero-mean background risk ε. Eeckhoudt and Schlesinger (2006) define prudence as a preference for disaggregating a sure loss and an additional zero-mean background risk. Therefore, a prudent individual would prefer lottery *B3* over lottery *A3*, while an imprudent individual would prefer the opposite.

The third row of lottery pairs in Figure 1 exemplifies a temperance task (risk apportionment of order 4). 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 temperance as a preference for disaggregating two independent zero-mean background risks. Thus, a temperate individual would prefer lottery *B4* over lottery *A4*, while an intemperate individual would prefer lottery *A4* over lottery *B4*.

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 degree risk than X_1. According to Ekern (1980), Y_1 has more n-th degree 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 degree 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 degree 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 degree risk is "risk apportionment of order $m + n$". An individual who is risk apportionment of order $m + n$ would choose lottery B over lottery A in Figure 2. This approach is more general, and can be used for all orders. For instance, if m and n are both 2 and an individual prefers less 4th degree risk, then the individual is risk apportionment of order 4 (temperate).

Moreover, 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. Deck and Schlesinger (2014) 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". Note that in the EUT framework, mixed risk aversion (defined by an increasing utility function with an alternating pattern of signs of successive derivatives) and mixed risk-loving behavior (defined by a utility function that has all derivatives strictly positive) would yield the same behavior.

Figure 1: Risk apportionment up to order 4 as lottery preferences

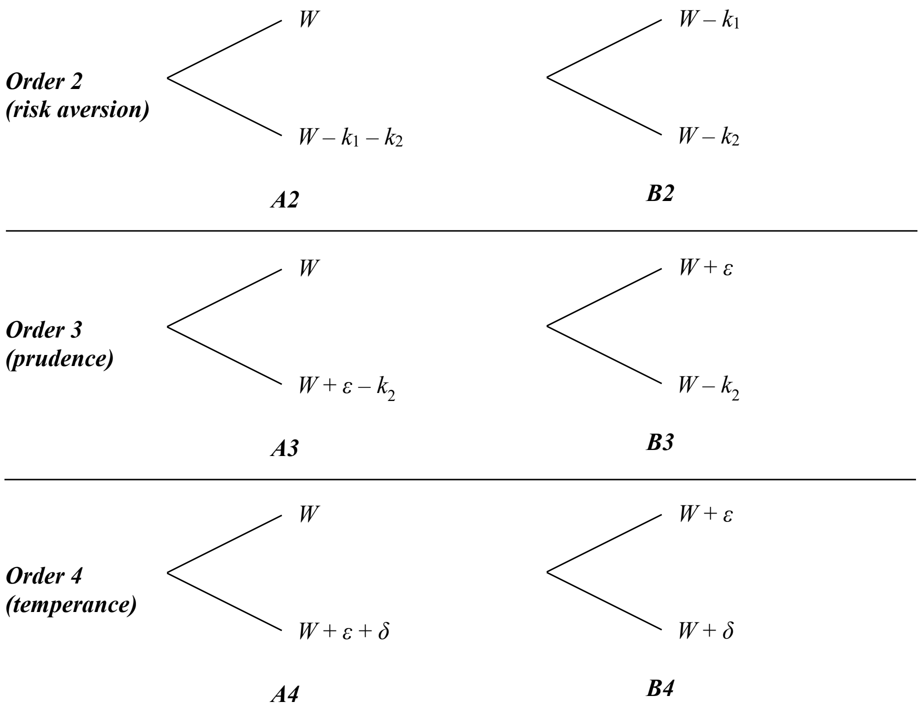
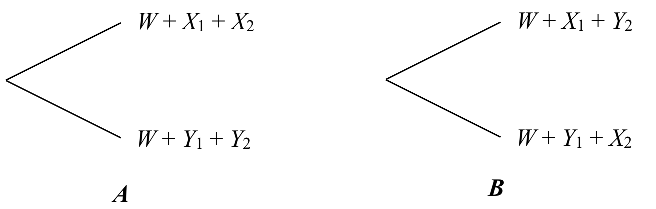


Figure 2: Risk apportionment in a general framework



B Hypotheses

Cross-country differences.—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, Henrich et al. 2001, Oosterbeek, Sloof, and van de Kuilen 2004, Herrmann, Thöni, and Gächter 2008). Measuring risk preferences in very different 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 large stakes experiments for (relatively) low cost.

The first reason for selecting China, the USA, and Germany for the current study is the economic importance of these countries for the world economy. They are the countries with the highest population in their respective continents, and with the largest economies in terms of total GDP. Furthermore, they are the three largest exporters in the world and are among the top ten countries in their respective regions with regard to wealth per adult. The second reason for selecting these countries is that they differ culturally. Taras, Rowney, and Steel (2009) survey 121 measures for examining culture. As they point out, the most prominent method for quantifying cultural traits is the one introduced by Hofstede (1980). He originally identified four dimensions that characterize a culture: power distance, individualism, masculinity and uncertainty avoidance. Of these four dimensions, individualism and uncertainty avoidance have been found to be associated with risk preferences.

In one of the first studies analyzing 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 finding by referring to the lower individualism in China: according to the cushion hypothesis, people from China are 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 US, and

China differ widely with respect to individualism: the USA ranks 1st, Germany 17th and China 52nd out of these 69 countries. In a comprehensive survey covering 53 countries, Rieger, Wang, and Hens (2015) observe that more uncertainty avoidance is associated with less risk taking. The data on uncertainty avoidance suggests smaller disparities: Germany ranks 40th, the USA 57th and China 63rd among the 69 countries. Thus, on the basis of both measures we would expect Chinese people to be the least risk-averse.[6]

While there are many international comparison studies on second-order risk aversion, nothing is known about differences in higher-order risk preferences.[7] 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:

[6] 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 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, 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 risk aversion in the gain domain. A similar observation is made in the large experimental study by Vieider et al. (2015b). These authors elicit the risk preferences of students in 30 countries and observe a positive correlation between (log) GDP per capita and risk aversion. Based on these correlations, we would expect Chinese people to be the least risk-averse. Further experimental comparisons between China and western countries have been conducted by Kachelmeier and Shehata (1992a) and Ehmke, Lusk, and Tyner (2010).

[7] Noussair, Trautmann, and van de Kuilen (2014) observe that those who are prudent are more likely to have a savings account, and those who are temperate have less risky investments. The financial decision making of households in the three countries we consider differs widely because of the differences in the political and institutional environments. Not surprisingly, the countries also differ starkly in related indicators. Net household savings as a share of a households disposable income in the year 2014 (OECD 2017) is lowest in the USA and highest in China. With respect to risky investments, Europeans tend to invest less in stocks than Americans (see, e.g., Christelis, Georgarakos, and Haliassos 2013). However, obtaining comparable data for household investments in China is difficult. With respect to the value of stocks traded as a percentage of GDP in 2014 (World Bank 2017), the USA has the largest and Germany the smallest stock market.

Hypothesis 1: *Chinese people make fewer mixed risk-averse and more mixed risk-loving choices than Americans and Germans.*[8]

Differences in stake sizes.—The use of monetary incentives has been an imperative in experimental research in economics, at least since the study by Smith (1976). However, stake size has frequently been subject to criticism – most prominently that formulated by Harrison (1989). Therefore, increasing the stake size provides a robustness check on previous results on higher-order risk preferences. It also moves the measurement of preferences to a payoff domain, which is relevant for important decisions in human life such as health or educational choices.

The theory of higher-order risk preferences outlined in the previous section does not depend on the size of the payoffs at stake or people's wealth – people adhere to either the mixed risk-averse or the mixed risk-loving pattern of decisions. Within EUT, however, 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 individuals are risk-loving when the stakes are small and 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 prudence and by Eeckhoudt, Gollier, and Schlesinger (1996) and Gollier and Pratt (1996) regarding temperance.[9]

Empirically testing these theoretical assumptions that determine the utility function of an individual has been a challenge, since it requires a considerable variation of the stake size. One approach to this challenge is to conduct experiments in developing countries, where large monetary incentives can be provided at lower cost than in developed countries (e.g. Carlsson et al. 2013). These studies typically observe more risk-averse choices with higher stakes when

[8] Please note that, thus, our null hypothesis is that there is no difference in mixed risk-averse choices and no difference in mixed risk-loving choices between Chinese people and Americans or between Chinese people and Germans.

[9] In what follows, when we speak about increasing/decreasing risk aversion, prudence and temperance, we mean increasing/decreasing *relative* risk aversion, prudence, and temperance.

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 risk aversion increases with real stakes but not with hypothetical stakes.[10]

There are only two experimental papers that consider the relationship between stake size and higher-order risk preferences.[11] Deck and Schlesinger (2010) confront subjects with ten choices, each between two lotteries, to study prudence and temperance. These lotteries have an overall expected payoff of $25.80. The comparison of two of the ten choice situations allows them to study the influence on prudence of a fivefold increase in the stake; two more comparisons of choice situations allow them to study the influence on temperance. Deck and Schlesinger (2010) find weak support for the hypothesis that prudence 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 prudent choices). Although they find mostly intemperate behavior in their subject population, intemperate behavior is less common when the stakes are higher.

Noussair, Trautmann, and van de Kuilen (2014) study the prevalence of prudence and temperance 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 (second-order) risk aversion and temperance increase when the hypothetical stakes are increased (from €70 to €10,500). However, they find no significant difference between the real monetary payoff treatments and a treatment with

---

[10] Harrison et al. (2005) use the same method as Holt and Laury (2002) but control for order effects. They reaffirm the general results found by Holt and Laury, but observe a significantly smaller effect. Another approach to studying risk preferences for large payoffs is to use game shows. One of the most prominent of the game shows used for estimating risk aversion over large stakes is “Deal or no Deal”. Post et al. (2008) study the choices of participants in the Netherlands, Germany, and the USA. They find behavior that could be explained by increasing risk aversion with increasing stake sizes.

[11] Results from survey data based on self-reported savings (Guiso, Jappelli, and Terlizzese 1992, 1996) and hypothetical investment questions (Eisenhauer and Ventura 2003) suggest increasing risk aversion and increasing prudence with increasing wealth.

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 prudence.

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 risk aversion increases with higher stakes. In addition, there is limited evidence indicating an increase in temperance.

Hypothesis 2: *The number of mixed risk-averse choices will increase and the number of mixed risk-loving choices will decrease when the stake size increases.*[12]

Differences through displaying reduced rather than compound lotteries.—It is well-known that the framing of lotteries can strongly influence decision making under risk, as captured in prospect theory (Kahneman and Tversky 1979) in which losses and gains are valued differently, for example. 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), while reduced and compound lotteries have not been compared directly. According to most theories of decision making, displaying actuarially equivalent lotteries as compound or reduced does not influence choices. However, Deck and Schlesinger (2014) note 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).

With respect to second-order risk aversion, it has been known for a while that reduced lotteries are often valued differently from 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).

Recent economic research on the reduction of compound lotteries has its roots in two classic behavioral patterns. First, Grether and Plott (1979) find that many people exhibit preference reversals, replicating an observation first made by Lichtenstein and Slovic (1971) and Lindman

[12] Therefore, our null hypothesis is that there is no difference in mixed risk-averse choices and no difference in mixed risk-loving choices when the stake size increases.

(1971).[13] As Holt (1986) points out, preference reversal might be caused by a violation of the independence axiom of expected utility theory. Karni and Safra (1987) make a similar argument. They show that preference reversals can be consistent with Quiggin's (1982) rank-dependent utility model, which does not assume independence. Segal (1988) shows that a violation of the reduction of compound lotteries axiom (ROCL) can generate preference reversals even if the independence axiom holds.[14]

Second, the classic examples by Ellsberg (1961) illustrate that many people have a non-neutral attitude to ambiguity. They are, instead, typically ambiguity-averse with respect to gains – that is, they prefer a gamble with known probabilities over a gamble with unknown probabilities (see e.g. Camerer and Weber 1992, and Trautmann and van de Kuilen 2015, for surveys). It has been suggested by Becker and Brownson (1964) and Smith (1969) that gambles with ambiguous probabilities should be viewed as two-stage lotteries: the probabilities of the outcomes are drawn in the first stage, before the outcomes are determined in the second stage. Several theoretical explanations have been proposed along these lines. Some assume that the ROCL holds, while others relax this assumption. Segal (1987, 1990) shows that ambiguity aversion can be explained by relaxing the ROCL and applying the rank-dependent utility model. Another stream of theoretical models is based on expected utility theory but assumes different preferences for first-

[13] When offered the choice between two lotteries, many people prefer a lottery with a high probability of a modest payoff over a lottery with a low probability of a large payoff. However, when asked to state selling prices for lottery tickets, they demand higher prices for the ticket of the latter lottery (see, e.g. Tversky and Thaler 1990, Hsee et al. 1999, and Seidl 2002, for surveys). In these experiments the mechanism proposed by Becker, DeGroot, and Marschak (1964) is used to elicit prices, which is theoretically incentive-compatible. In this mechanism, sellers of a lottery ticket state their willingness-to-accept. A price is randomly chosen and sellers receive this price in exchange for the lottery ticket if the price exceeds their willingness-to-accept. Otherwise they keep the lottery ticket. However, each price elicitation using this mechanism can be regarded as a two-stage lottery.

[14] 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 in some prominent papers 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.

and second-stage lotteries (see, e.g., Klibanoff, Marinacci, and Muerkji 2005, Nau 2006, Seo 2009). Recent empirical studies have largely corroborated the argument that ambiguity and preferences over compound risks are closely linked (Halevy 2007, Chew, Miao, and Zhong 2015, and Abdellaoui, Klibanoff, and Placido 2015).

Higher-order risk preferences are typically elicited using compound lotteries. To our knowledge, only Maier and Rüger (2012) and Deck and Schlesinger (2016) have used reduced lotteries to elicit higher-order risk preferences. Maier and Rüger (2012) observe, in the gain domain, 55% of choices to be risk-averse, 60% to be prudent, and 58% to be temperate. These percentages are at the lower end of the range of observed frequencies in other studies (see also Table A1 in the Appendix). In a study conducted in parallel to ours, Deck and Schlesinger (2016) find a significant framing effect between the compound and the reduced presentation of the lotteries. They observe less temperance and edginess in reduced lotteries but no difference in the frequency of prudent choices when lotteries are displayed in a reduced rather than a compound form.

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. There exists prior empirical evidence that this is often the case with respect to second-order risk preferences.

Hypothesis 3: *We will find fewer* n*-th order risk-averse choices when lotteries are displayed in a reduced rather than a compound framing.*[15]

III Experimental design

A Elicitation method

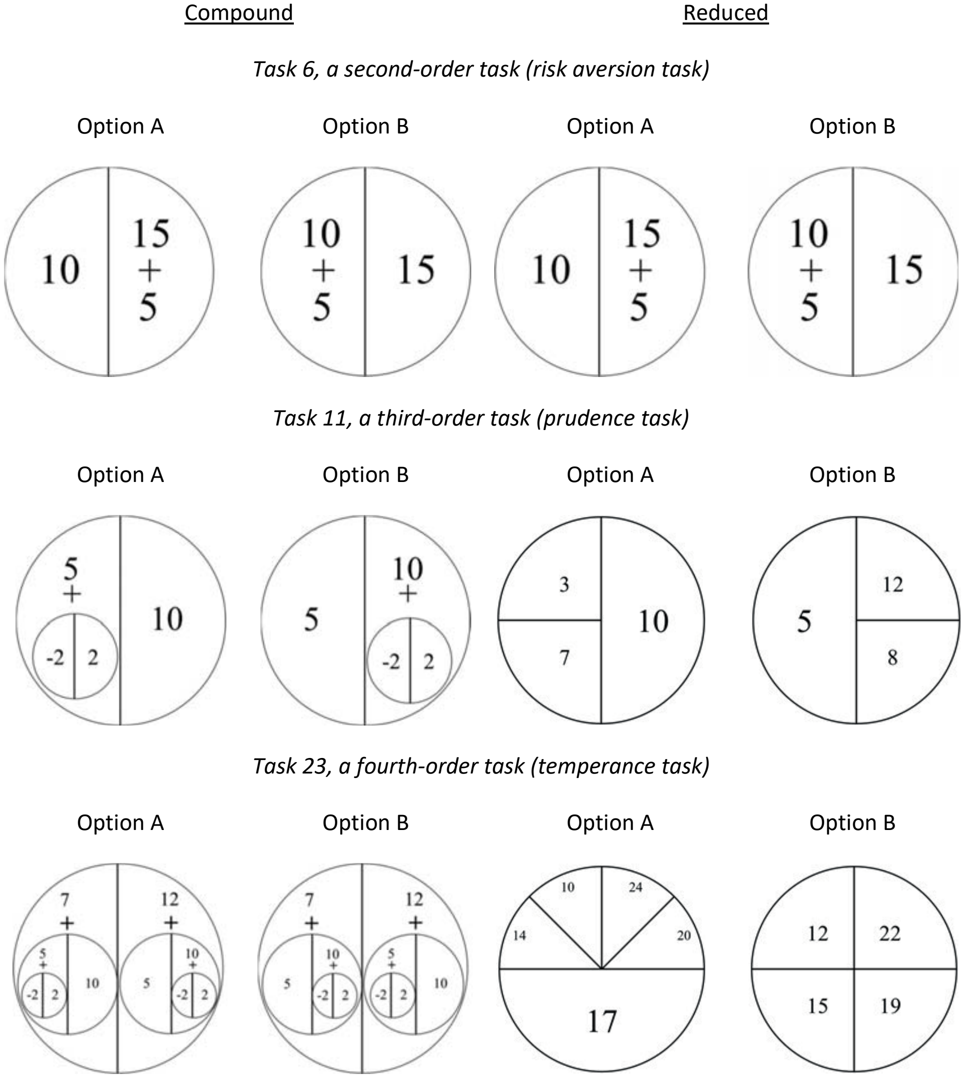
Our elicitation method and general experimental design follows Deck and Schlesinger (2014). In short, the main experiment contains 38 tasks (see Deck and Schlesinger 2014, 1922ff and Online Appendix O1), and each of these tasks involves choosing between Option A and Option B. Examples of the risk aversion, prudence, and temperance lotteries (i.e. lotteries of orders 2, 3 and 4) as presented to the participants are shown in Figure 3. Each option involves different

[15] Therefore, our null hypothesis is that the number of mixed risk-averse and mixed risk-loving choices does not change when lotteries are displayed in a reduced rather than a compound form.

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 risk aversion 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 II.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 risk-averse individual should choose the certain option over the risky one when the expected values are the same.

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.

In the prudence 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 prudence as a preference for disaggregating a sure loss and an additional zero-mean background risk. Therefore, a prudent individual should prefer Option B over Option A.

In the temperance task lottery (order 4, task 23), the outcomes in a 50-50 lottery may contain not just one but two other lotteries. For example, Option A involves two additional 50-50 lotteries. Moreover, the example shown is a composition of (1+3)th-order risk, since Option A can be written as [7 + 11A, 12 + 11B] where 11A and 11B refer to Options A and B of task 11 (shown in Figure 3 as the prudence task). Analogously, Option B of the temperance example can be written as [7 + 11B, 12 + 11A]. An individual who prefers lotteries with lower ($m + n$)-th degree risk is risk apportionment of order $m + n$. In this example, an individual who is risk apportionment of order $m + n$, which here means 1 + 3 = 4, and therefore an individual who is temperate, would choose Option B over Option A.

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.

B Experimental treatments

We conducted an economic laboratory experiment with sessions in China, the USA, and Germany. Subjects faced the 38 tasks in randomized order, and one of the tasks was randomly selected for payment.[16] In each task, the subjects had to choose between two lotteries, which revealed their risk preference up to order 6 (the instructions are shown in Appendix A2).

The treatments, the orders of the lotteries and the number of subjects are shown in Table 1. The 38 lottery pairs created by Deck and Schlesinger (2014) consist of compound lotteries that make the different combinations of "good" with "bad" or "good" with "good" outcomes salient. Lottery pairs that were displayed in the original compound form 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, as well as average payoffs in the local currency. These payments include a show-up fee of $8.50 which was adjusted using the same exchange rates.

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 reduced and compound lotteries, we ran additional sessions in Germany. 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 form). Each subject faced lotteries of two additional orders in the original compound form as well as in the reduced form. 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.

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

---

[16] 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 (2016, 1) point out, the random payment technique "is essentially the only incentive compatible mechanism".

compound lottery.[17] 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).

Furthermore, we were concerned about differences in the general level of cognitive ability and in numerical skills between the subject pools. Therefore, 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 risk-averse in the gain domain. It has been reported previously that higher cognitive ability is associated with lower 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 are significantly more prudent. 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 it 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 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

[17] The test of understanding consisted of two pages on the computer screen with two control questions each, as in Deck and Schlesinger (2014) and as shown in Online Appendix O2. All subjects were able to answer these four questions correctly. The computer screen, however, was locked after a subject had entered a wrong answer more than twice on one screen. The subjects were asked to raise their hand if this occurred, and were approached by an experimenter to explain any potential misunderstanding. This was the case for 9 subjects in CHN, 10 subjects in USA, 19 subjects in GER, 2 subjects in CHN 10x and 11 subjects in the Compound & Reduced treatment. We provide robustness checks in Online Appendices O4, O6 and O8 controlling for this.

experiment was computerized and programmed using zTree (Fischbacher 2007). Screenshots are provided in Online Appendix O2.

Table 1: Treatments and design

Treatment	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

Notes: [1] This equals \$0.47 (CHN), \$0.68 (GER) and \$4.67 (CHN 10x) at the time of the experiment. [2] This equals \$7.59 (CHN), \$20.18 (GER), \$92.24 (CHN 10x) and \$19.49 (Compound & Reduced) at the time of the experiment.

C Experimental conditions across countries

Ensuring similar experimental conditions is the main challenge in conducting cross-country experiments, as has been pointed out by Roth et al. (1991). In order to create similar conditions in the CHN, USA, and GER treatments, we surveyed the experimental literature (as outlined in Haering and Heinrich 2016) and followed best practice as described below.

To minimize currency effects we used experimental-currency-units (ECU). This means that 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). This way we avoided showing the subjects different nominal payoffs, which might have confounded our observations. In the economics literature there appears to be no consensus on how to select exchange rates to keep incentives constant across locations. Calculating exchange rates for experiments is complicated by the fact that the cost of living also varies widely within countries, while the relevant data on purchasing power is usually only available for major cities or at country level. Furthermore, local guidelines for payments to subjects have to be taken into consideration.

The experiments in our study took place in major metropolitan areas, but there was no reliable data on purchasing power available for all three cities (Tianjin, Boston, and Essen). However, the UBS Prices & Earnings survey (UBS 2014) provides data for Beijing, New York City, and Berlin. To determine the payoffs in our study we proceeded as follows. We used the relative cost of living from the UBS Prices & Earnings survey as a first proxy (this measure is also used by Özer, Zheng, and Ren 2014). The second proxy we used was 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). Including this 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 from the mean of the two proxy exchange rates. 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. The average payoffs in local currency are shown in Table 1.[18]

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

[18] Note that Vieider (2012) finds no influence of small variations in payoffs (+/- 20%) on second-order risk aversion.

Gächter 2008 for a similar approach). To ensure that the instructions were similar, we only used written instructions. These – as well as 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, Özer, Zheng, and Ren 2014).

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. In the USA and Germany we relied on existing databases. In both these countries this procedure was handled via ORSEE (Greiner 2015). In China, however, we had to build a database from scratch.[19] 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 the 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.[20]

---

[19] 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.

[20] There were two adjustments we had to make because of American regulations and which we control for in the 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 did not receive prior information in the recruitment e-mail or 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.

IV Results

A Summary statistics

Table 2 summarizes the characteristics of all the participants in all the treatments. In the sessions for country comparison, slightly more women than men participated in all three countries (CHN, USA and GER). We collected data first in China, then in the USA and lastly in Germany. Because we were not able to recruit subjects based on their age 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 students differs across countries. The age distribution of the German participants differs significantly from that of the Chinese and Americans subjects ($p \leq 0.041$, two-sided Mann-Whitney U tests). The proportion of female subjects does not differ significantly across the three subject pools ($p \geq 0.545$, Fisher's exact tests).

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 ($p \leq 0.007$).[21]

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

---

[21] Frederick (2005) observed that students in Princeton 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.

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: 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)

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.

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-th order) risk-loving choices as a measure of (n-th order) risk aversion – that is, the more (n-th order) risk-loving 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: we observe that 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).[22] 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-th order risk-loving choices in each country in all orders. In all countries we observe a general tendency of subjects to avoid the more risky lotteries. Comparing the distribution of choice frequencies to the distribution that would result from random choices yields significant differences in each order for each country ($p < 0.001$, Chi-squared tests). Furthermore, the number of n-th order risk-loving 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 ($p = 0.242$).[23]

Comparing the frequencies of n-th order risk-loving choices between countries using Mann-Whitney U tests suggests some differences: Chinese participants tend to be weakly[24] less risk-averse than participants from the USA ($p = 0.060$, two-sided) and from Germany ($p = 0.068$). They are weakly more edgy than Germans ($p = 0.092$) and significantly more edgy than Americans ($p = 0.031$). Additionally, Americans tend to be weakly less risk apportionment of order 6 than Chinese people ($p = 0.065$) and Germans ($p = 0.082$). None of the remaining differences are significant ($p \geq 0.180$).

[22] In Online Appendix O4 we provide robustness checks controlling for this choice pattern.

[23] 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 (for all subjects) and in Online Appendix O9.2 (separately for risk-loving and risk-averse subjects).

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

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
<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>USA</u>						
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
<u>GER</u>						
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.

To control for subject pool differences, we also run an ordinary least squares (OLS) regression for each order with the number of risk-loving 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).[25] For order 2

[25] 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 20). The latter two are influential. First, we observe a significant influence of our Chinese experimenter. In the GER session conducted by him, participants behaved in a more risk-averse way. Second, we find that in the GER sessions in which the subjects were provided with IRB information, the subjects behaved in a less risk-averse way. Because of a computer error we could not collect the CRT and BNT scores for eight subjects in CHN. Therefore, we also report additional regressions without controlling for CRT and BNT. In

the regression suggests that the Chinese subjects make more risk-loving choices than the German ones (while we observe no significant differences for higher orders). The Chinese country dummy indicates that the Chinese subjects make one more risk-loving choice than Americans, but the difference is not significant (β = 0.981, robust standard error = 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 standard error = 0.590, p = 0.910). However, the country dummies of China and Germany differ significantly (p = 0.009, two-sided Wald test).

Consistency of risk preferences.—Based on a preference for combining "good" with "bad" as described by Deck and Schlesinger (2014), individuals who are risk-averse should also be temperate (order 4) and should pick the risk apportionment option in order 6 (that is, they should exhibit mixed risk-averse behavior). In contrast, individuals who have a preference for combining "good" with "good" and are risk-loving should also be intemperate and choose the anti-risk apportionment option in order 6 (that is, they should exhibit mixed risk-loving behavior). Both mixed risk averters and mixed risk lovers should behave in a prudent (order 3) and edgy (order 5) manner (see Section III.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 go further and classify the subjects based on all orders and giving equal weight to all decisions that we collect from a subject.

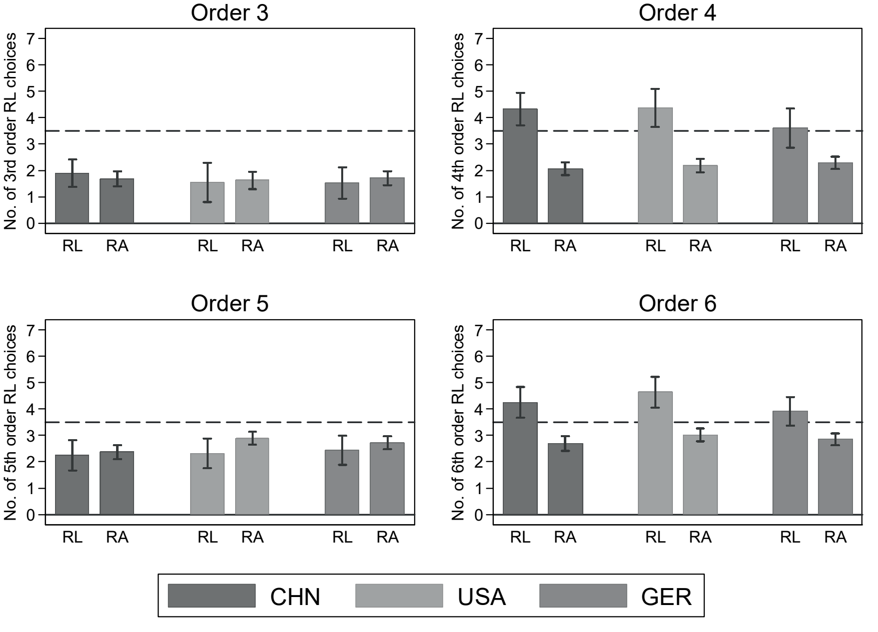
For the first step we classify subjects as risk-averse or risk-loving according to whether they make choices in line with this preference in the majority of their decisions (i.e. at least four out of seven times). With this classification scheme, 80% of the Chinese participants, 83% of the Americans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

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 relevant country or who did not answer the relevant question. This was the case for 26 subjects in USA and 7 subjects in GER. However, none of these further robustness checks suggests a different interpretation of our data.

Figure 4 displays the average number of *n*-th order risk-loving choices for risk averters ("RA") and risk lovers ("RL") across the three countries. In addition, it includes a dashed line at 3.5 indicating the number of risk-loving choices expected under random behavior, as well as 90% confidence intervals. 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 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 made by mixed risk averters and mixed risk lovers.

This impression is confirmed by non-parametric tests: risk averters and risk lovers in all countries are prudent and edgy ($p \leq 0.005$, two-sided one-sample Wilcoxon signed-rank tests), but only risk averters are also temperate and risk apportionment of order 6 ($p \leq 0.005$). All risk lovers are intemperate and anti-risk apportionment of order 6 in China and in the USA ($p \leq 0.072$) but not in Germany ($p \geq 0.259$). Comparing choices within the three countries between risk averters and risk lovers reveals no significant difference with respect to the number of prudent (order 3) or edgy choices (order 5) ($p \geq 0.135$, two-sided Mann-Whitney U tests). However, all comparisons with respect to the number of temperate (order 4) or risk apportionment of order 6 choices yield significant differences in all subject pools ($p \leq 0.007$).

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



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 choices can be classified as being consistent 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 38 choices adhere to the mixed risk-averse type, for example. If we can reject this hypothesis and most choices adhere to the pattern, we classify him as mixed risk-averse. The same procedure is applied for mixed risk-loving behavior.

Table 4 summarizes the proportion of subjects that can be classified into the two types. It lists the distributions for the usual three significance thresholds that can be applied to the results of binomial tests. The 1% threshold corresponds to 27 of 38 choices that have to be consistent with one type or the other, the 5% threshold corresponds to 26 choices, and the 10% threshold to 25 choices.[26] Under the strictest criterion, between 42% and 45% percent are mixed risk-averse across the countries. Between 6% and 9% are mixed risk-loving. If the criterion is relaxed, these percentages go up to between 60% and 64%, or 11% and 15%, respectively. The proportion of subjects consistent with one type or the other ranges from 51% to 76%.

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.

We can also use the individual behavioral patterns to compare the consistency across countries. For this, we count, 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 suggested by Figure 4, there appears to be no difference in the behavioral patterns across countries. Comparing the number of choices consistent with mixed risk-averse and mixed risk-loving preferences yields a similar impression: neither the differences in the

[26] 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 CHN and one subject in the USA. When applying the 10% threshold this is the case for four subjects in CHN, one subject in GER and five subjects in USA.

number of mixed risk-averse nor the differences in the number of mixed-loving choices are significant ($p \geq 0.396$, two-sided Mann-Whitney U tests). Taking the maximum of the two variables allows the overall consistency to be compared with the theory. This comparison also yields no significant differences ($p = 0.353$).

Running separate OLS regressions (see Appendix A4) with these three dependent variables provides additional evidence. 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 = 2.041, p = 0.560, two-sided) and Germany are insignificant (β = 0.689, robust standard error = 1.716, p = 0.689), but they differ weakly, suggesting that the Chinese subjects make fewer mixed risk-averse choices than the Germans ($p = 0.099$, two-sided Wald test). Second, in the regression of the number of mixed risk-loving choices on the country dummies, the dummies for China (β = 1.858, robust standard error = 1.630, $p = 0.255$) and Germany (β = 0.388, robust standard error = 1.473, $p = 0.792$) are insignificant. Again, the two dummies differ weakly ($p = 0.094$), meaning that the Chinese subjects make rather 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-th order) risk-loving 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 91.7 and therefore not significantly larger ($p = 0.494$, Fisher's exact test).

Table 5 shows the number of n-th order risk-loving choices under both incentive structures in all orders. The observed distribution is significantly different from the distribution of random choices ($p < 0.001$, Chi-squared tests). The number of risk-loving 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).

Table 5: n-th order risk-loving choices across stakes

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
CHN 10x						
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.

The previous evidence suggests that more risk-averse choices are made when the stakes increase. The data presented in Table 5 suggest a similar effect: with low incentives, participants make, on average, 1.971 decisions in a risky way, but with high incentives the average is 1.417 decisions. This difference is significant ($p = 0.045$, two-sided Mann-Whitney U test). Table 5 also indicates that there may be differences in the higher orders. In particular, in the remaining even orders (4 and 6) there appear to be fewer risky decisions when the stakes are high. However, for orders 3 to 6 none of the treatment differences are significant ($p \geq 0.109$). To control for subject pool differences within China, we run OLS regressions separately for each order, with the number of risk-loving 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). These regressions suggest a similar pattern: for order 2 there is a weakly negative effect of increased stakes on the number of risky choices (β = -0.616, robust standard error = 0.319, p = 0.055, two-sided), and there are no significant differences for the other orders ($p \geq 0.309$).

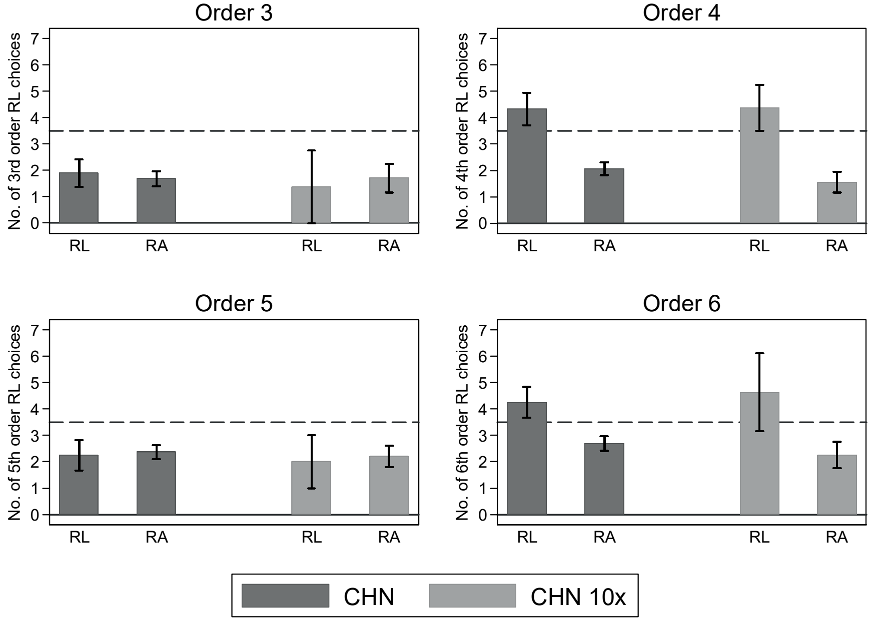
Consistency of risk preferences.—Above we describe the patterns of mixed risk-averse and mixed risk-loving behavior in three different subject pools around the world. This evidence was gathered using standard incentives. In this section we analyze whether these patterns prevail under large stakes. Under the high stakes in CHN 10x, we also expect risk averters to coincide in their choices with 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 anchor the analysis on the choices observed for order 2. Based on second-order risk aversion, we classify the subjects into risk-averse or risk-loving and compare their behavior for the higher orders. Classifying everyone with a majority of risk-averse choices as risk-averse yields a figure of 83% for 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-th order risk-loving choices for risk averters ("RA") and 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 risk-loving choices that would be expected under random behavior: for the odd orders, risk averters and risk lovers significantly favor the more (third- or fifth-order) risk-averse options ($p \leq 0.048$, two-sided one-sample Wilcoxon signed-rank tests), while for the even orders, only risk averters do so ($p \leq 0.001$). Risk lovers for order 4 weakly prefer the more risky options (p = 0.084) while there is no significant tendency for order 6 (p = 0.222). In addition, we compare the decisions of the two types. The number of prudent and edgy choices does not differ significantly between risk averters and risk lovers ($p \geq 0.622$, two-sided Mann-Whitney U tests). In addition, risk averters are significantly more temperate and risk apportionment of order 6 than risk lovers ($p \leq 0.008$).

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



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 suggested by the theory. 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.[27] Table 6 also lists the shares from the CHN treatment for comparison, in which between 51% and 79% could be classified in this way.

We also use this data to compare consistency between the two treatments. The number of choices consistent with mixed risk-averse behavior is weakly larger with high stakes (p = 0.065,

[27] 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.

two-sided Mann-Whitney *U* test). The number of choices consistent with mixed risk-loving behavior is weakly smaller (p = 0.099). Accordingly, the number of choices consistent with one type or the other does not differ significantly between treatments (p = 0.127).

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%	43%
$p < 0.05$	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 addition, we 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). These regressions control for differences in the subjects' characteristics in both treatments (the experimental procedures were exactly the same). However, 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 standard error = 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 standard error = 0.843, p = 0.276). Also, there appears to be no significant influence on the maximum of either dependent variable (β = 0.592, robust standard error = 0.909, p = 0.516).

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 find no 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.—In this section the first dependent variable we consider is the number of (n-th order) risk-loving choices. 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 experiment this share is also 97% (p = 1.000, Fisher's exact test).

Table 7 presents the number of n-th order risk-loving choices in this experiment. 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 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 risk-loving alternative for orders 3 to 5. The frequency distribution of choices differs significantly from the distribution expected under random behavior in the compound and in the reduced lotteries (p < 0.001, Chi-squared tests), except in order 6, where the difference is significant neither in the compound (p = 0.431) nor in the reduced lotteries (p = 0.532). Comparing choice frequencies to the 3.5 risk-loving choices that would be expected under random behavior yields a similar pattern but only for the compound lotteries. 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 ($p \geq 0.113$).

Our results suggest that using a compound lottery influences choices for orders 3 and 4: on average, the risk-loving option is chosen 0.845 times less often in the compound than in the reduced prudence lotteries, and it is chosen 1.014 times less often in the compound temperance lotteries. The difference is significant in both cases ($p \leq 0.001$, two-sided Wilcoxon tests). The differences with respect to edginess and risk apportionment of order 6 are insignificant ($p \geq$ 0.785). In addition, we run linear panel regressions with individual random effects separately for each order, with the number of risk-loving choices as the dependent variable. The regressions include a dummy indicating 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 also indicate fewer prudent choices in reduced lotteries (β = 0.845, robust standard error = 0.223, p < 0.001, two-sided) and fewer temperate choices (β = 1.014 robust standard error = 0.279, p < 0.001).

Table 7: n-th order risk-loving 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.

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 theoretical 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 but differ in the even orders. Yet the compound format might facilitate viewing a lottery as a combination of "good" and "bad" outcomes.

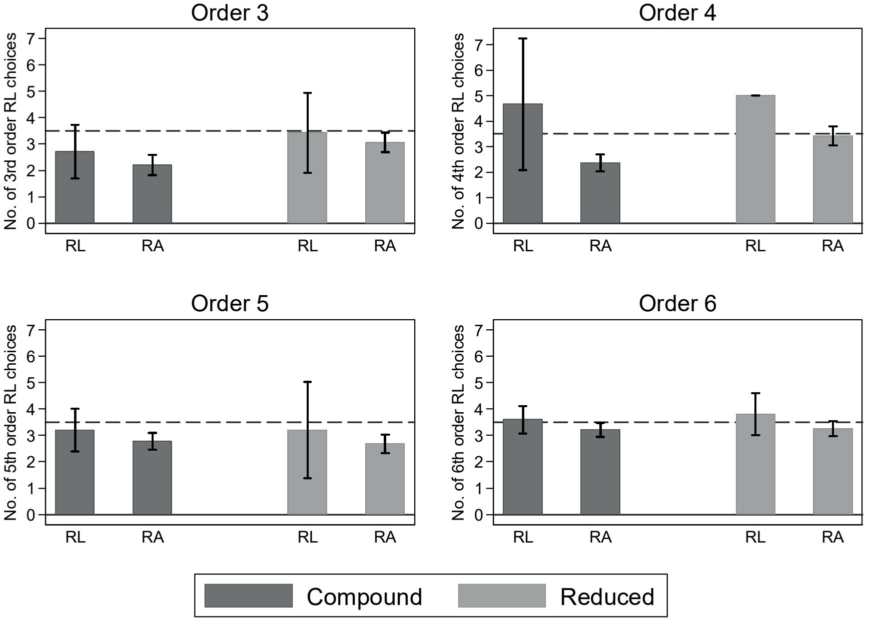
Even though the experiment presented in this section possesses a different data structure from the experiments above, we proceed with the same two steps when analyzing consistency. In the first step we compare the number of third-, fourth-, fifth- and sixth-order risk-loving choices made by those who predominately chose risk-averse options in order 2 with the choices of those who mainly chose risk-loving options in order 2. Classifying everyone in this way yields 93% risk-averse and 7% risk-loving subjects for the Compound & Reduced treatment.

Figure 6[28] displays the average frequency of risk-loving 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 risk averters, we replicate the previous findings using compound lotteries: for orders 3 to 5 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 see this preference ($p = 0.136$). When using reduced lotteries we still find at least a weak tendency of 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 benchmark ($p = 0.738$). Independent of the 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 ($p \geq 0.480$).[29]

[28] The extreme shape of the confidence intervals for order 4 is driven by the fact that only three subjects are classified as risk-loving, so that there is either a pronounced or a non-existent dispersion, depending on the subjects' choices.

[29] However, there are relatively few risk lovers (only between three and seven risk lovers for each order) compared to the cross country and the stake analyses. This might be driven by differences in the subject pool composition (cf. Table 2).

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



In addition, we compare the average number of choices of risk averters and risk lovers between types. In line with the theory, for the odd orders there are no significant differences when using compound ($p \geq 0.376$, two-sided Mann-Whitney *U* tests) or reduced lotteries ($p \geq 0.566$). For the even orders the difference between the two types is significant for the compound lotteries of order 4 ($p = 0.037$) but not for the compound lotteries of order 6 ($p = 0.495$). The differences for the reduced lotteries with even orders are both insignificant ($p \geq 0.149$).

Comparing the compound and reduced lotteries for risk-averse subjects, we find that the subjects choose the less risky lottery significantly less often in the reduced prudence and temperance lotteries ($p \leq 0.002$, two-sided Wilcoxon signed-rank tests) but not in the reduced edginess or risk apportionment of order 6 lotteries ($p \geq 0.785$). There are no significant differences with regard to risk-loving individuals ($p \geq 0.231$). This pattern suggests that some individuals exhibit preference reversals. On the individual level, 27% of the subjects make more risk-loving choices in the reduced than in the compound lotteries of order 3, while 10% make fewer risk-loving choices. These percentages are 32% and 15% for order 4, 20% and 19% for order 5 and 18% and 17% for order 6.[30]

In the second step we place the same weight on all the decisions gathered from the subjects in a treatment. In this experiment this means that we separate the decisions for orders 3 to 6 on the basis of the lottery format. While in the previous classification we could use all 38 decisions at once, we now have to use 10 choices for orders 1 and 2, and 14 choices from two of the higher orders. These 14 choices are from either the compound or the reduced lotteries – depending on the treatment. Note that because of the within-subjects design one subject can fall into different categories in either treatment.

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.[31] In addition, we compare the number of choices that are consistent with the two types between the two treatments. There are significantly more choices consistent with mixed risk aversion in the compound than in the reduced lotteries ($p < 0.001$, two-sided Wilcoxon test). With respect to mixed risk-loving behavior there is no significant difference ($p = 0.518$). The number of choices consistent with either pattern is significantly larger with compound lotteries ($p = 0.001$).

[30] 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 risk-loving choices differed between the two treatments (see the Online Appendix O7 for details on our estimation strategy and Appendix A6 for the regression results). 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.

[31] 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.

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.

We also 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 standard error = 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 treatment (β = 0.161, robust standard error = 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 standard error = 0.227, $p < 0.001$).

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 II.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 prudent and more temperate behavior in compound lotteries than we do in reduced lotteries. To gather some exploratory evidence, we also conducted a non-incentivized pen-and-paper survey following another experiment (unrelated to this study and not dealing with risk preferences).

In this survey we confronted one half of the participants with a compound and the other half with a reduced lottery pair. We selected the prudence lottery with three outcomes in which we observed the largest share of preference reversals within-subjects in the experiments described above. In task 11 (see Figure 3 in Section III.A) a zero mean risk (paying either 2 or -2) is added to either outcome (5 or 10) of a binary lottery, so that both lotteries have the same expected value and variance. Participants had to make a hypothetical choice between the two lotteries and were asked to provide a written answer to "*Why do you prefer the lottery you have selected? Please explain briefly how you have made your choice.*"

First, with respect to choices we find an even more pronounced framing effect in the survey than in the experiment. In the compound framing of the survey 80.6% of the 36 participants chose the more prudent lottery. Only 38.9% of the 36 participants did so in the reduced framing (p = 0.001, Fisher's exact test). In the experiment 70.4% of subjects chose the more prudent lottery in the compound framing, while only 49.3% did so in the reduced framing (p = 0.006, McNemar's test).

Second, based on prior considerations and a subsample of ten randomly selected answers we developed a classification scheme to analyze the content of the explanations provided. For three arguments the frequency differs significantly between framings ($p \leq 0.029$, Fisher's exact tests). 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 are also the most frequently used arguments of our classification scheme (see Online Appendix O10 for a complete list of categories).

The first two arguments suggest that one should choose the more prudent lottery (and all participants using these arguments did so). Each of these arguments was used by 53% of the participants in the compound framing and only by 17% and 25% in the reduced framing. This difference is significant in both cases ($p \leq 0.029$). The third argument suggests that one should choose the imprudent lottery (and all participants using this argument did so). It was used by 50%

of the participants in the reduced framing and by only 14% in the compound framing. Again the difference is significant ($p = 0.002$).

Overall it appears that the reduced display of lotteries leads the subject to focus on the payoff of the most likely outcome.

V Conclusion

In this study we analyze the consistency of higher-order risk preferences in a series of laboratory experiments. Based on the elicitation method introduced by Deck and Schlesinger (2014), we explore the role of country differences between China, the USA, and Germany, differences in stake size, and differences through displaying reduced rather than compound lotteries.

Overall, our findings confirm the previous findings that prudence (risk aversion of order 3) is more widespread than risk aversion (order 2) and temperance (order 4). Furthermore, the patterns of mixed risk-averse and mixed risk-loving behavior explain the choices in more complex lotteries (of orders 5 and 6) surprisingly well: those who tend to make risk-averse choices are also more likely to make temperate choices and more choices showing risk apportionment of order 6, while those who tend to make risk-loving choices exhibit the opposite choice pattern. However, both groups predominantly make prudent and edgy choices in orders 3 and 5. In the USA we replicate the findings of Deck and Schlesinger (2014). In addition, we identify a similar pattern in Germany and in China. A large majority of participants in all three countries can be classified as mixed risk averters, and a non-negligible share as mixed risk lovers. The former type seems to be more prevalent in Germany than in China, while the latter type is less common in Germany than in China.

Moreover, we provide the first analysis of the consistency of higher-order risk preferences with large monetary payoffs. In line with prior empirical evidence, we observe an increase in second-order risk aversion when the stakes are increased tenfold. We also observe the pattern of mixed risk-averse and mixed risk-loving behavior with high stakes. However, after controlling for differences in the subjects' characteristics, we find no significant difference in the number of mixed risk-averse and mixed risk-loving choices when the stakes increase.

Finally, we 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 significantly when reduced lotteries are used. Interestingly, a rather low prevalence of prudence and temperance has also been reported by Maier and Rüger (2012) when using reduced lotteries (cf. Appendix A1). However, they did not compare compound and reduced lotteries directly. In a study conducted in parallel with ours, Deck and Schlesinger (2016) compare behavior between the two formats directly. In a similar way to us, they conduct a within-subject analysis and, like our study, their data reveals a significant framing effect. Contrary to our results, they find fewer edgy choices and no change in the frequency of prudent choices when lotteries are displayed in a reduced rather than a compound form, but with respect to temperance they make a similar observation, observing less temperance in reduced lotteries.

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. Exploring potential explanations for this framing effect we conduct a non-incentivized follow-up study asking subjects to describe the determinants of their choices in a prudence task. Our results suggest that, in the reduced format, subjects focus mainly on the payoff of the lottery's most likely outcome.

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.

Alderfer, C. P., and H. Bierman. 1970. "Choices with Risk: Beyond the Mean and Variance." *Journal of Business* 43(3): 341-53.

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

Azrieli, Y, C. P. Chambers, and P. J. Healy. 2016. "Incentives in Experiments: A Theoretical Analysis." Working Paper 16-03, Ohio State University, Department of Economics.

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.

Becker, S. W., and F. O. Brownson. 1964. "What Price Ambiguity? Or the Role of Ambiguity in Decision-making." *Journal of Political Economy* 72(1): 62-73.

Becker, G. M., M. H. DeGroot, and J. Marschak. 1964. "Measuring Utility by a Single-response Sequential Method." *Behavioral Science* 9: 226-32.

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.

Brocket, 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.

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.

Camerer, C., and M. Weber 1992. "Recent Developments in Modeling Preferences: Uncertainty and Ambiguity." *Journal of Risk and Uncertainty* 5(4): 325-70.

Carlsson, F., Martinsson, P., Qin, P. and M. Sutter 2013. "The influence of spouses on household decision making under risk: an experiment in rural China." *Experimental Economics* 16(3): 383-401.

Carroll, C. D., and M. S. Kimball 2008. "Precautionary Saving and Precautionary Wealth." In *The New Palgrave Dictionary of Economics*, edited by N. S. Durlauf, and L. E. Blume, London: Macmillan: 579-84.

Chew, S. H., B. Miao, and S. Zhong. 2015. "Partial Ambiguity." Working Paper. National University of Singapore and Shanghai University of Finance and Economics.

Chiu, W. H. 2005. "Skewness Preference, Risk Aversion, and the Precedence Relations on Stochastic Changes." *Management Science* 51(12): 1816-28.

Chiu, W. H. 2010. "Skewness Preference, Risk Taking and Expected Utility Maximisation." *Geneva Risk and Insurance Review* 35(2): 108-29.

Christelis, D., D. Georgarakos, and M. Haliassos. 2013. "Differences in Portfolios across Countries: Economic Environment versus Household Characteristics." *Review of Economics and Statistics* 95(1): 220-36.

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.

Cooley, P. L. 1977. "A Multidimensional Analysis of Institutional Investor Perception of Risk." *Journal of Finance* 32(1): 67-78.

Coombs, C. H., and D. G. Pruitt 1960. "Components of Risk in Decision Making: Probability and Variance Preferences." *Journal of Experimental Psychology* 60(5): 265-77.

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. 2016. *On the Robustness of Higher Order Risk Preferences*. Working Paper, Chapman University.

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. 2016. *Greater Mutual Aggravation*. Netspar Discussion Paper No. 01/2016-002.

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. and H. Schlesinger, 2013. "Higher-order Risk Attitudes." In *Handbook of Insurance*, edited by G. Dionne, 41-57. New York: Springer.

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.

Eisenhauer, J. G., and L. Ventura. 2003. "Survey Measures of Risk Aversion and Prudence." *Applied Economics* 35(13): 1477-84.

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.

Geyer, J. 2011. "The Effect of Health and Employment Risks on Precautionary Savings." DIW Berlin Discussion Paper No. 1167.

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.

Guiso, L., T. Jappelli, and D. Terlizzese. 1992. "Earnings Uncertainty and Precautionary Saving." *Journal of Monetary Economics* 30: 307-37.

Guiso, L., T. Jappelli, and D. Terlizzese. 1996. "Income Risk, Borrowing Constraints, and Portfolio Choice." *American Economic Review* 86: 158-72.

Haering, A., and T. Heinrich. 2016. *Risk Preferences in China – Results from Experimental Economics*. Unpublished mimeo.

Halevy, Y. (2007). "Ellsberg Revisited: An Experimental Study." *Econometrica* 75(2): 503-36.

Harrison, G. W. 1989. "Theory and Misbehavior of First-price Auctions." *American Economic Review* 79(4): 749-62.

Harrison. G. W., E. Johnson, M. M. McInnes, and E. E. Rutström. 2005. "Risk Aversion and Incentive Effects: Comment." *American Economic Review* 95(3): 897-901.

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. 2014. "Higher-order Risk Preferences in Social Settings." Ruhr Economic Papers No. 508.

Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, and R. McElreath. 2001. "In Search of *homo economicus*: Behavioral Experiments in 15 Small-scale Societies." *American Economic Review* 91(2): 73-78.

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. 1986. "Preference Reversals and the Independence Axiom." *American Economic Review* 76(3): 508-15.

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. K., G. F. Loewenstein, S. Blount, and M. H. Bazerman. 1999. "Preference Reversals between Joint and Separate Evaluations of Options: A Review and Theoretical analysis." *Psychological Bulletin* 125(5): 576-90.

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.

Kahneman, D., and A. Tversky. 1979. “Prospect Theory: An Analysis of Decision under Risk.” *Econometrica* 47: 263-91.

Karni, E., and Z. Safra. 1987. “‘Preference Reversal’ and the Observability of Preferences by Experimental Methods.” *Econometrica* 55(3): 675-85.

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.

Klibanoff, P., M. Marinacci, and S. Muerkji. 2005. “A Smooth Model of Decision Making Under Ambiguity.” *Econometrica* 73(6): 1849-92.

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.

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. 2017. “Skewed Background Risks and Higher-order Risk Preferences: Prudent versus Temperate Behavior.” *Applied Economics Letters* 24(5): 338-41.

Nau, R. F. 2006. “Uncertainty Aversion with Second-order Utilities and Probabilities.”

Management Science 52(1): 136-45.

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).

OECD. 2017. "Household savings (indicator)", doi: 10.1787/cfc6f499-en (Accessed on 03 February 2017).

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.

Post, T., M. van den Assem, G. Baltussen, and R. Thaler. 2008. "Deal or no Deal? Decision Making under Risk in a Large-payoff Game Show." *American Economic Review* 98(1): 38-71.

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.

Rothschild, M., and J. E. Stiglitz. 1970. "Increasing Risk: A Definition." *Journal of Economic Theory* 2(3): 225-43.

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.

Seidl, C. 2002. "Preference Reversal." *Journal of Economic Surveys* 16(5): 621-55.

Seo, K. 2009. "Ambiguity and Second-Order Belief." *Econometrica* 77(5): 1575-1605.

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

Smith, V. L. 1969. "Measuring Nonmonetary Utilities in Uncertain Choices: The Ellsberg Urn." *Quarterly Journal of Economics* 83(2): 324-29.

Smith, V. L. 1976. "Experimental Economics: Induced Value Theory." *American Economic Review* 66(2): 274-79.

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.

Taras, V., J. Rowney, and P. Steel. 2009. "Half a Century of Measuring Culture: Review of Approaches, Challenges, and Limitations Based on the Analysis of 121 Instruments for Quantifying Culture." *Journal of International Management* 15(4): 357-73.

Tintner, G. (1942). "A Contribution to the Non-Statistic Theory of Choice." *Quarterly Journal of Economics* 56(2): 274-306.

Trautmann, S. T., and G. van de Kuilen. 2015. "Ambiguity Attitudes." In *The Wiley Blackwell Handbook of Judgment and Decision Making*, edited by G. Keren, and G. Wu, 89-116, West Sussex: John Wiley and Sons.

Tversky, A., and R. H. Thaler. 1990. "Anomalies: Preference Reversals." *Journal of Economic Perspectives* 4(2): 201-11.

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.

World Bank. 2017. "Stocks traded, total value (% of GDP)", doi: CM.MKT.TRAD.GD.ZS (Accessed on 17 February 2017).

Appendix A1 Comparison of related papers

Table A1: Comparison of related papers

Study	Location(s)	Average payoff	Payment	Framing	Lottery type	Share of risk-averse/prudent/temperate choices
Deck and Schlesinger (2010)	USA	\$25.56	1 of 10 choices	Direct choice Endowment & loss	Compound	– / 61% / 38%
Ebert and Wiesen (2011)	Germany	€18.50	1 of 34 choices	Direct choice Endowment & loss	Compound	– / 65% / –[1]
Maier and Rüger (2012)	Germany	–	1 of 84 choices	Direct choice Endowment & loss	Reduced	56% / 56% / 56%[2]
Deck and Schlesinger (2014)	USA	\$20.92	1 of 38 choices	Direct choice	Compound	74% / 77% / 58%
Ebert and Wiesen (2014)	Germany	€17.50[3]	1 of 120 choices	Choices in list Endowment & loss	Compound	66% / 88% / 75%[1]
Heinrich and Mayrhofer (2014)	Germany	€18.09	1 of 240 choices	Choices in list Endowment & loss	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	Direct choice	Compound	72% / 89% / 62%
Deck and Schlesinger (2016)	USA	\$21.66	1 of 52 choices	Direct choice	Compound Reduced	– / 73% / 64% – / 77% / 47%
This study (2017)	China Germany USA	¥47.06 / ¥571.92[4] €18.10 / €17.40[4] \$28.14	1 of 38 choices	Direct choice	Compound Reduced	75% / 76% / 64%[5] – / 56% / 50%

Notes: Table following Noussair et al. (2014). The dash (–) indicates that the values are not reported. [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%.

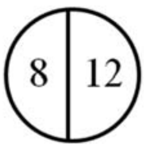
Appendix A2 Instructions (English version)

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.

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).

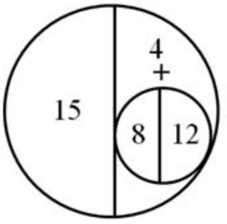
For ECU 1 you will receive $ 0.93.

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,

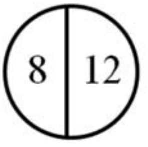


is a 50-50 lottery in which you would receive either ECU 8 or ECU 12, each with an equal chance. To determine the outcome of any 50-50 lottery, we will use a computerized random-number generator.

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

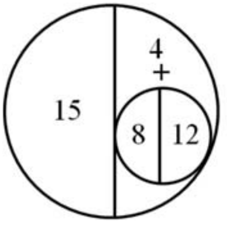


is a 50-50 lottery where you receive either ECU 15 or you receive ECU 4 plus the 50-50 lottery



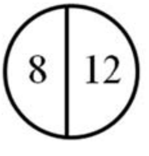
.

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 +



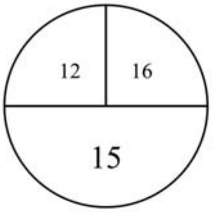
in the first 50-50 lottery.

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%.

Reduced:

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

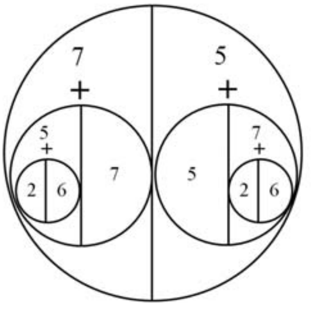
Like in the sample above



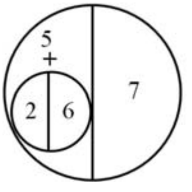
is a lottery in which you can either receive 12 ECU, 15 ECU or 16 ECU.

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.]

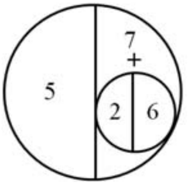
Let's look at a more complicated example.



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



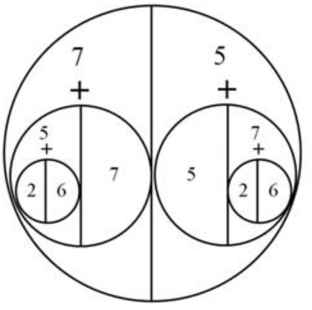
or you receive ECU 5 plus the 50-50 lottery



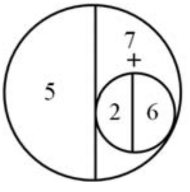
,

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,

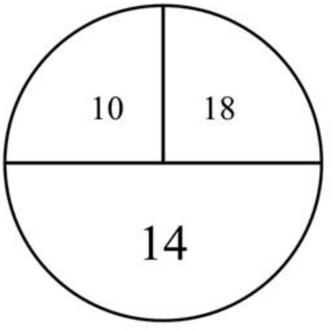
or 2) earning ECU 7 (in the first lottery) + ECU 7 (in the second lottery) which happens with a 0.5 x 0.5 = 0.25 = 25% chance,

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.

[**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%.]

Appendix A3 Summary of variables

Table A3: Summary of variables

Variable	Description
y_i / y_{it}:	
Order n	Subject's number of *n*-th order risk-loving 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)

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

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

Note: 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

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

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

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

Note: 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

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

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

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

Note: 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

Note: 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

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

Online Appendix

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

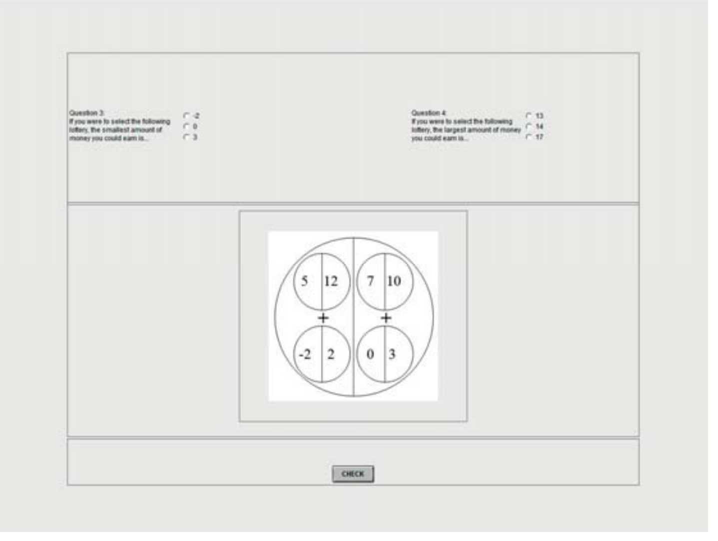
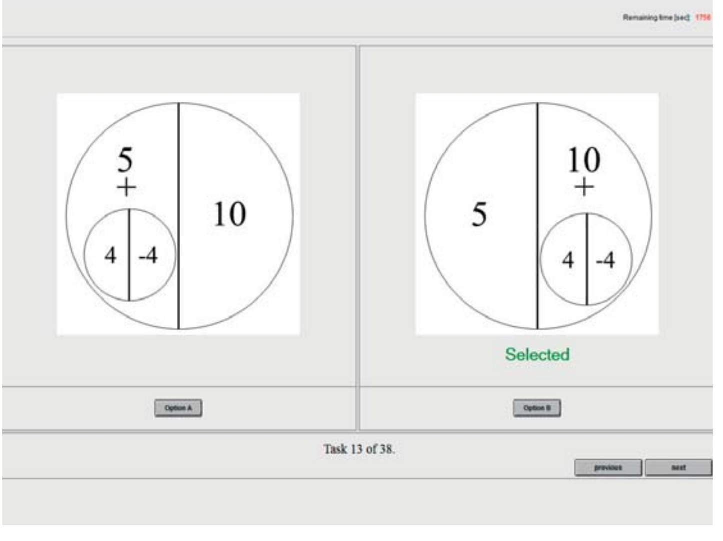


Figure O2.2: Lottery choice



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:

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

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,* China only) and individuals who made more than two mistakes on one of the pages of the test of understanding (*QuizWrong*). 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*-th order risk-loving 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

Note: 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

Note: 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:

$$y_i = \beta_0 + \beta_1 CHN10x_i + \boldsymbol{X}'_i\gamma + \boldsymbol{C}'_i\zeta + \varepsilon_i \tag{5.1}$$

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. $CHN10x_i$ is a dummy variable indicating subjects that received a tenfold increased payoff. The vector $\boldsymbol{X}'_i$ contains additional explanatory variables to investigate potential effects of an individual's demographics (*Female, Age 18-20, Age > 23*) or of cognitive and statistical (*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 pages of the test of understanding (*QuizWrong*). All variables are described in Table O3.1.

We estimate equation 5.1 without (see Appendix A5) and with $\boldsymbol{C}'_i$ (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

Note: 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

Note: 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:

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

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. $Reduced_{it}$ is a dummy variable indicating whether lotteries where displayed in reduced format in the respective treatment. The vector $\boldsymbol{X}'_{it}$ contains additional explanatory variables to investigate potential effects of an individual's demographics (*Female, Age 18-20, Age > 23*) or of cognitive and statistical skills (*CRT, BNT*). The vector $\boldsymbol{C}'_{it}$ 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 who made more than two mistakes on one of the test of understanding pages (*QuizWrong*) and interaction dummies with the $Reduced_{it}$ dummy. All variables are described in Table O3.1.

We estimate equations 7.1 and 7.2 without (see Appendix A6) and with $\boldsymbol{C}'_{it}$ (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:

$$P(y_i = 1 \mid \boldsymbol{X}'_i\, \boldsymbol{C}'_i) = \Lambda(\beta_0 + \boldsymbol{X}'_i\gamma + \boldsymbol{C}'_i\zeta) \tag{7.2}$$

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 $\boldsymbol{X}'_i$ and $\boldsymbol{C}'_i$ 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

Note: 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

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

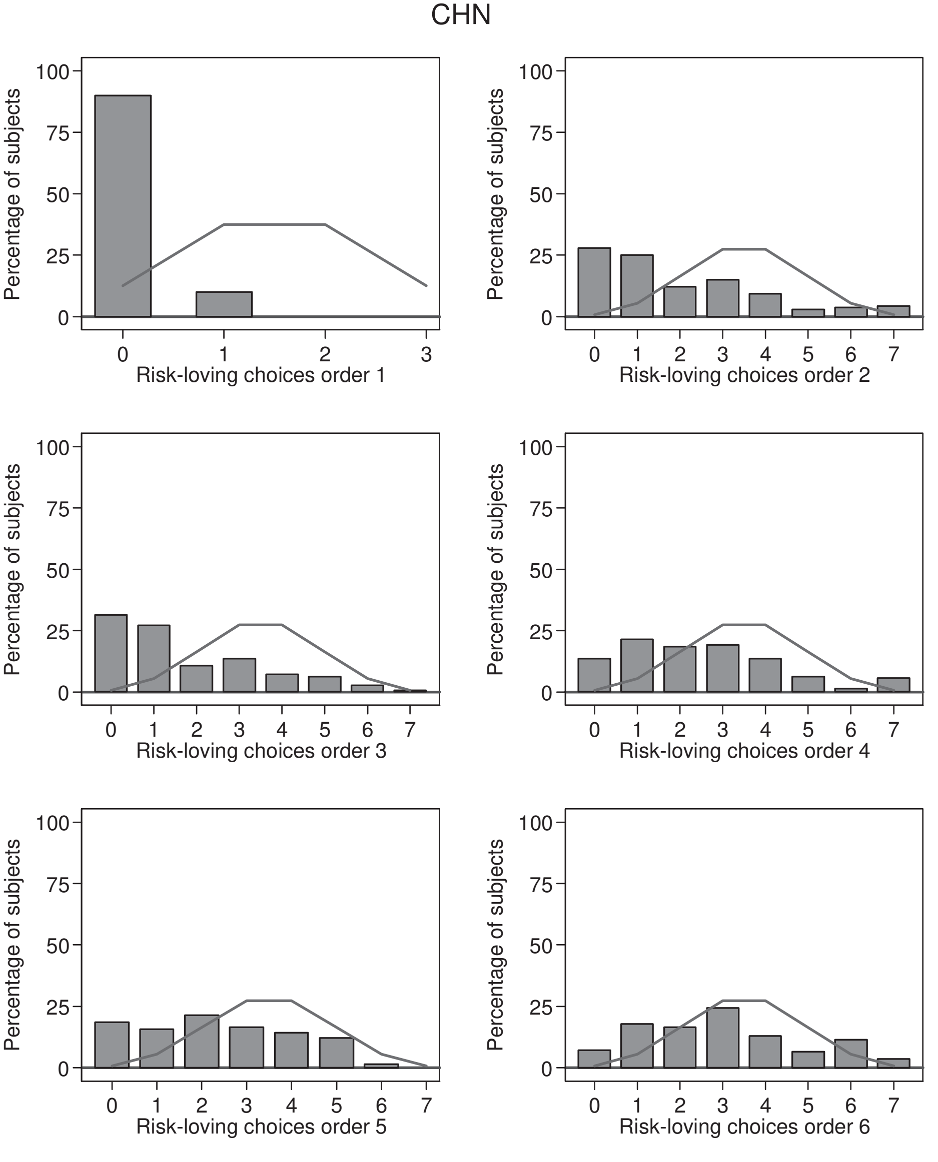
Table O8.3: Logit regression

	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

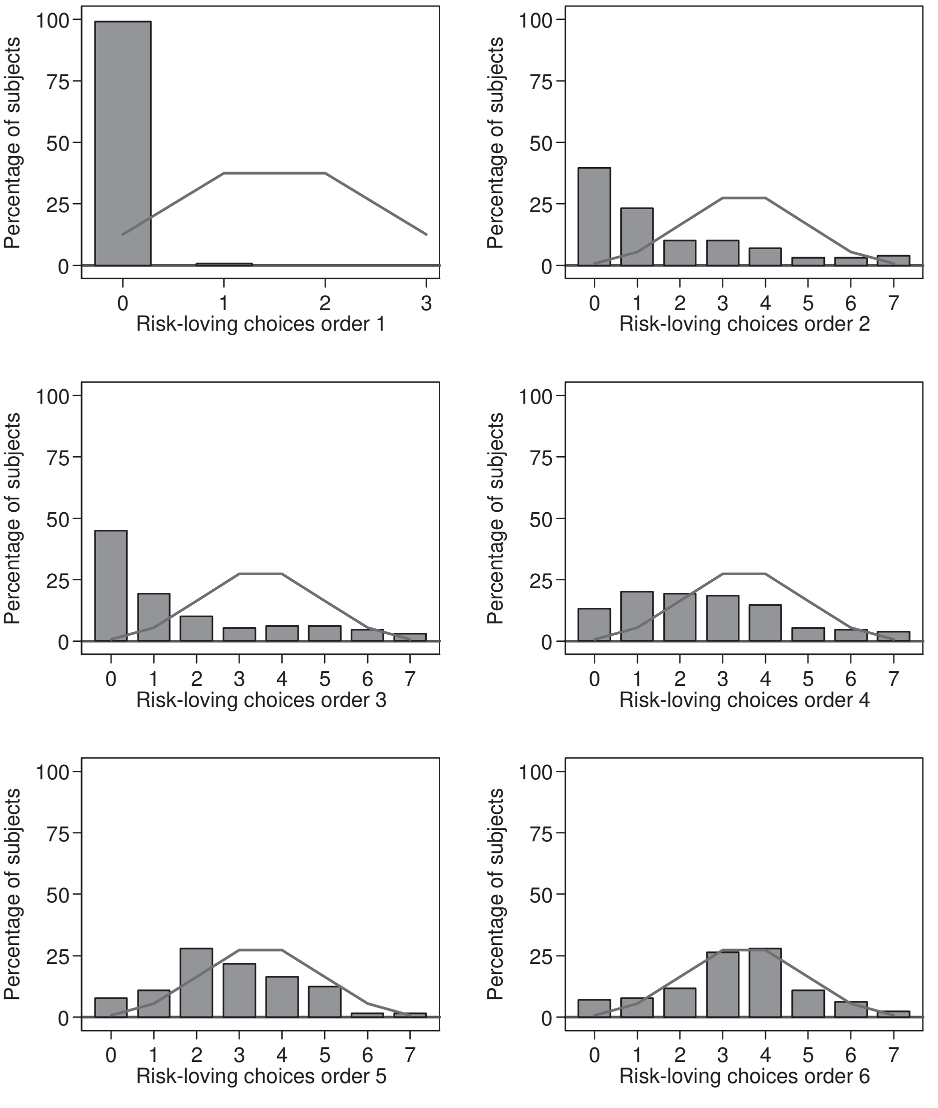
Note: Calculation of marginal effects: Delta-method, constant not reported, robust standard errors in parentheses, asterisks indicate the significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

O9 Distribution of choice frequencies within each order

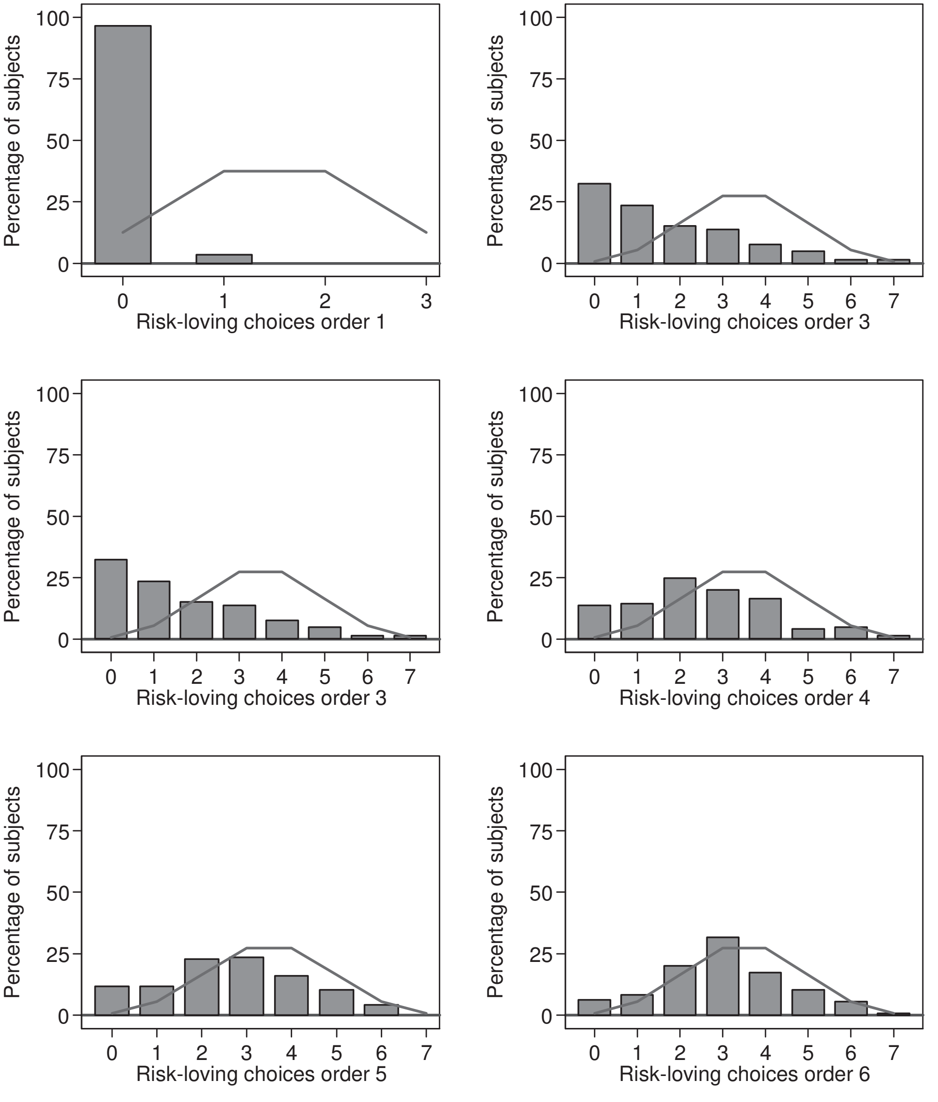
O9.1 All subjects



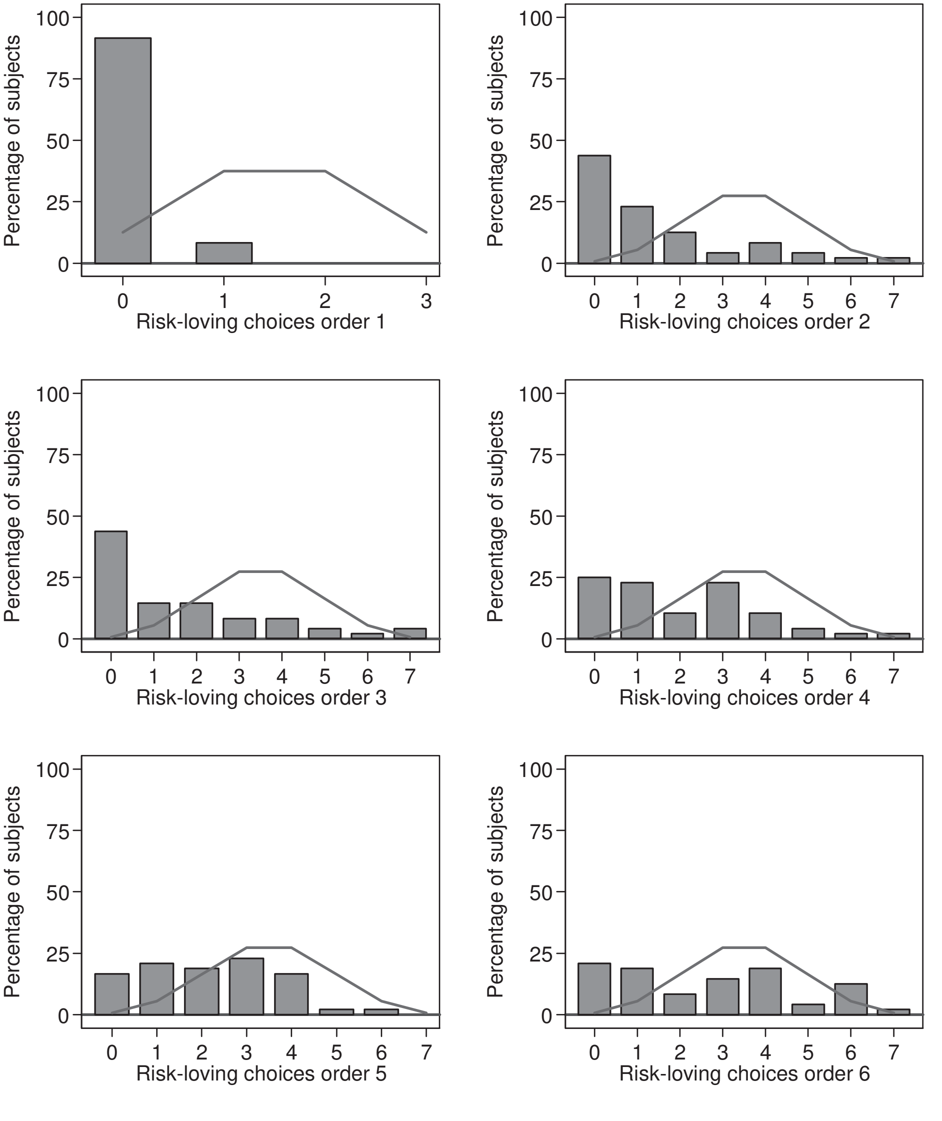
USA


Percentage of subjects
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

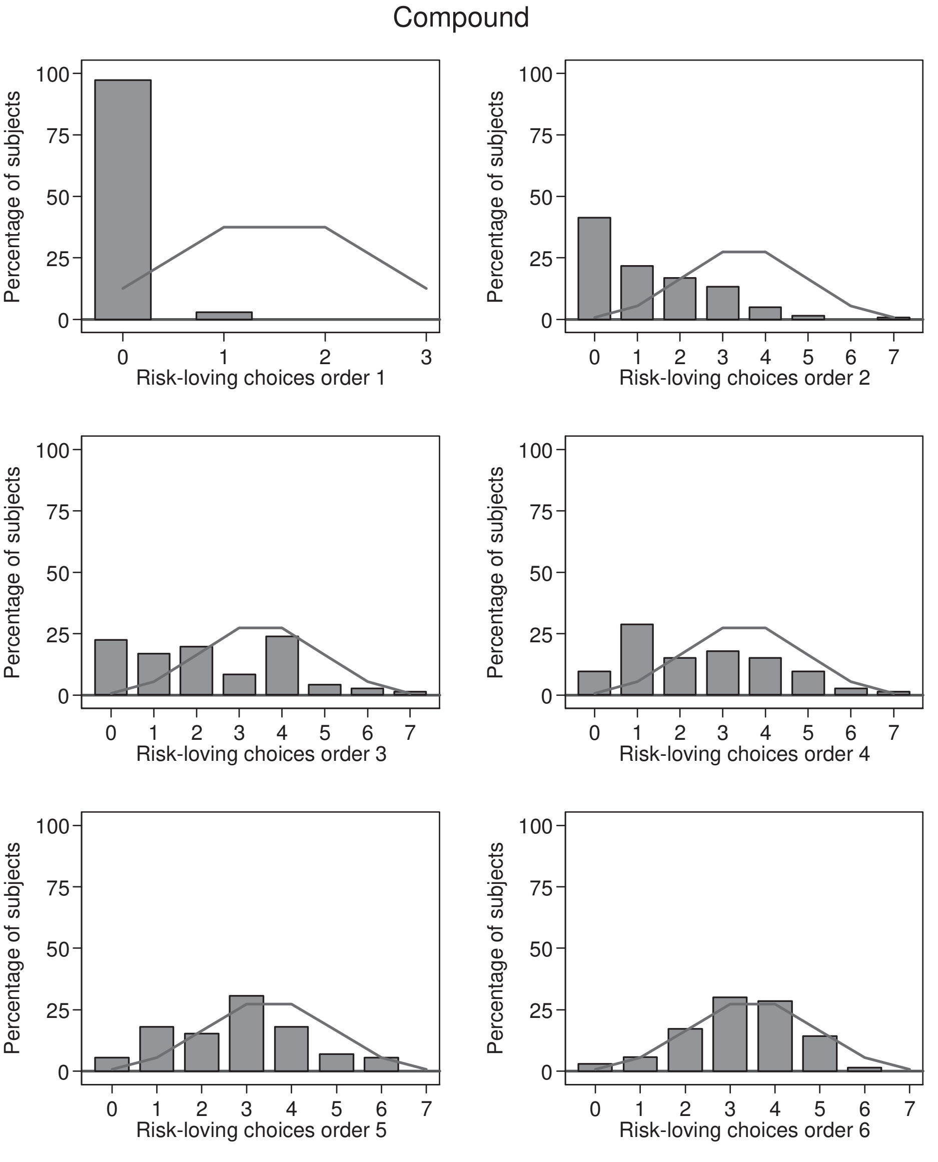
GER


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

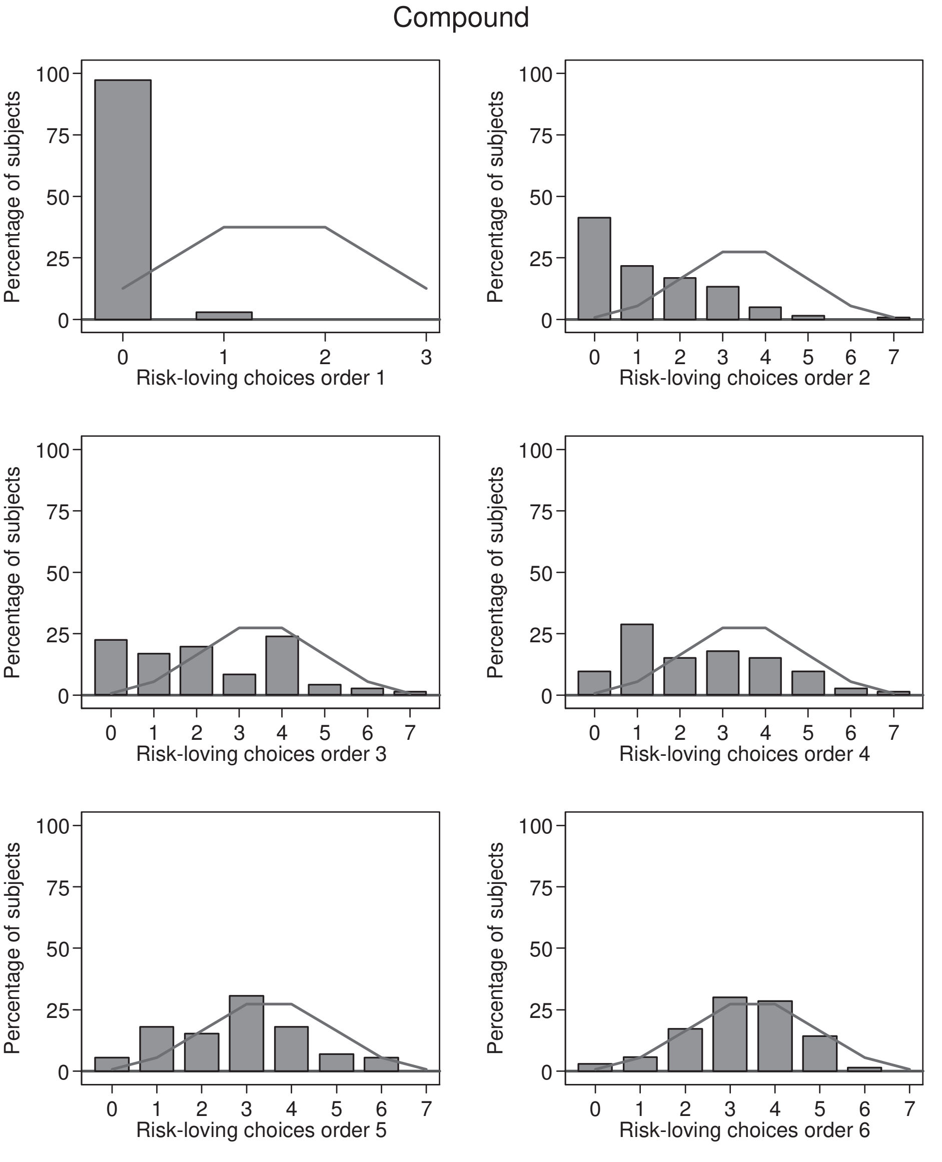
CHN 10x



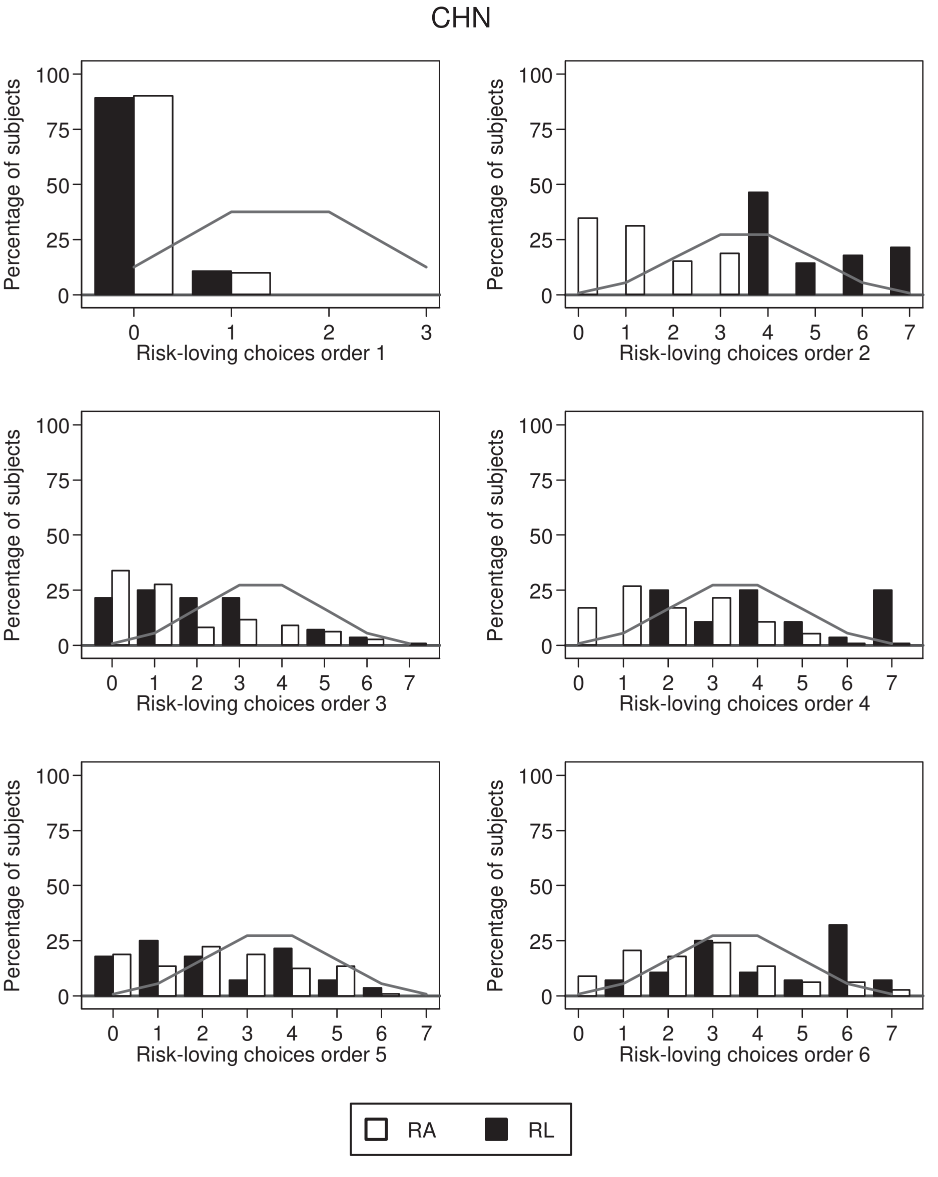
Percentage of subjects
100
75
50
25
0
0 1 2 3
Risk-loving choices order 1
0 1 2 3 4 5 6 7
Risk-loving choices order 2
Risk-loving choices order 3
Risk-loving choices order 4
Risk-loving choices order 5
Risk-loving choices order 6

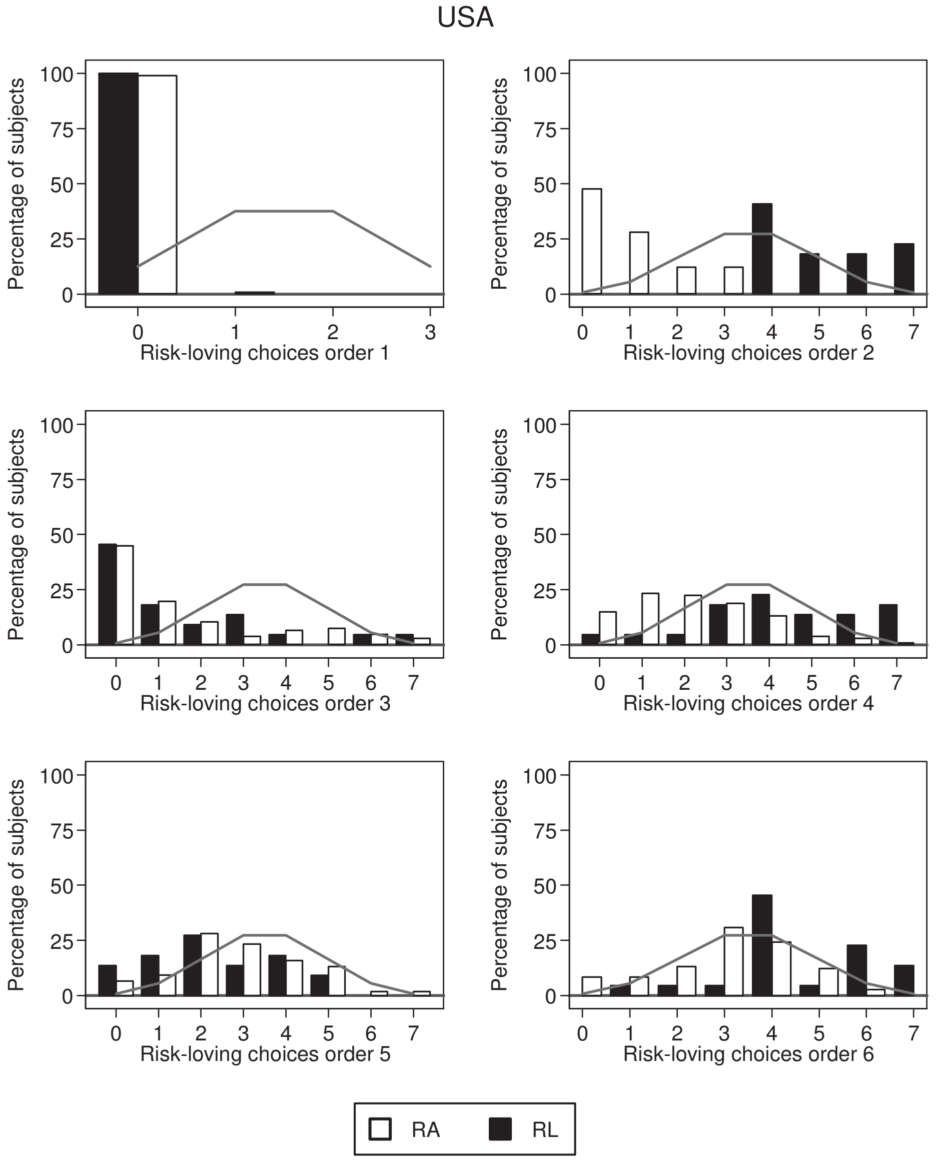


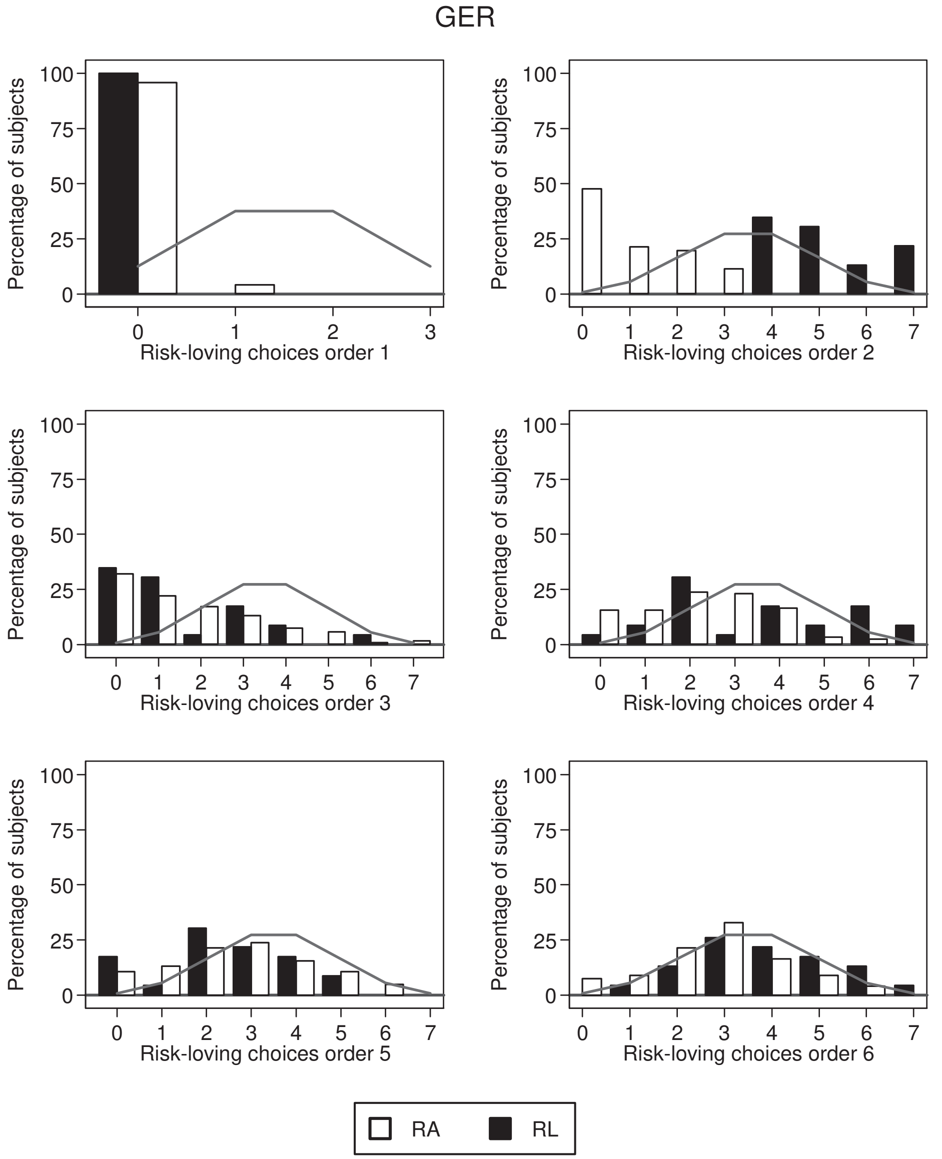
Compound
Percentage of subjects
100
75
50
25
0
Risk-loving choices order 1
Risk-loving choices order 2
Risk-loving choices order 3
Risk-loving choices order 4
Risk-loving choices order 5
Risk-loving choices order 6
0 1 2 3 4 5 6 7

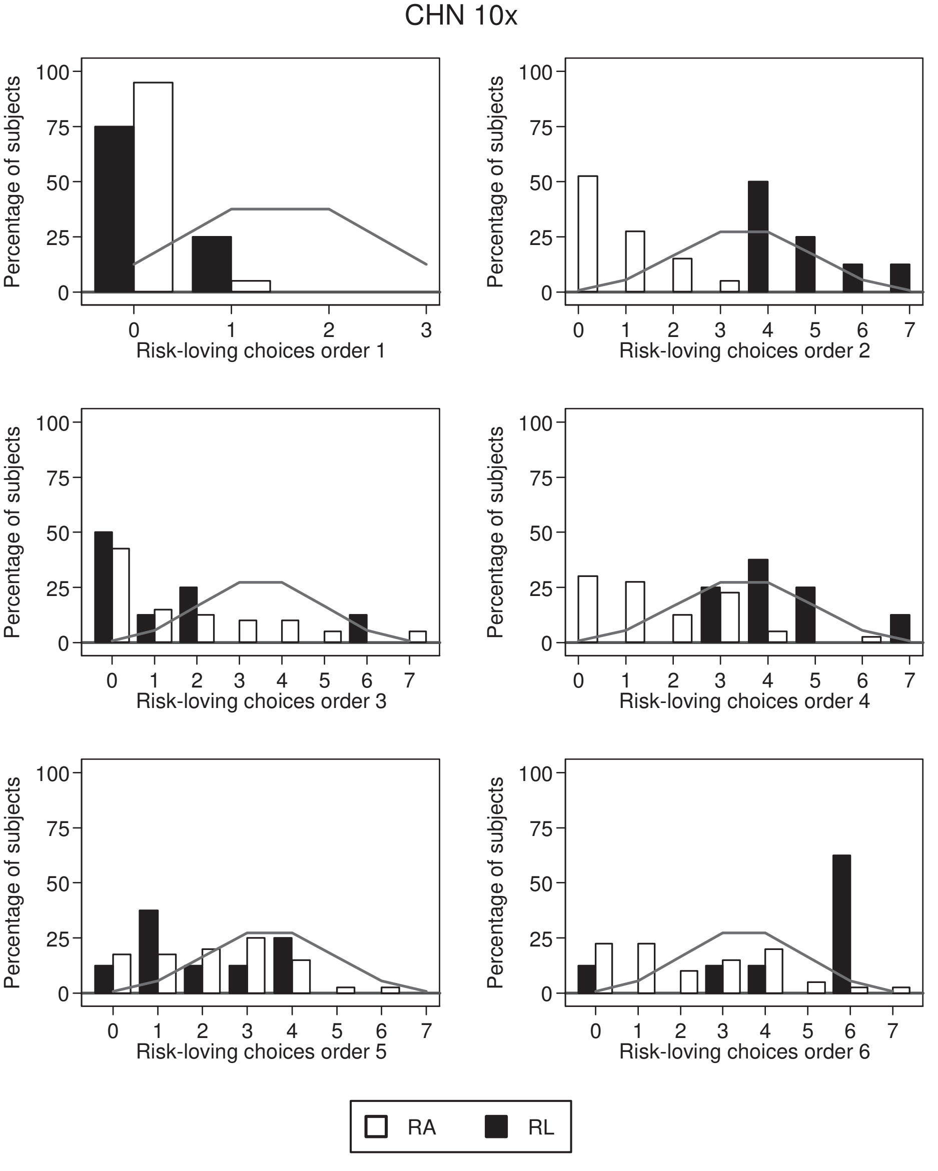

Compound
Percentage of subjects
100
75
50
25
0
0
1
2
3
Risk-loving choices order 1
0
1
2
3
4
5
6
7
Risk-loving choices order 2
Risk-loving choices order 3
Risk-loving choices order 4
Risk-loving choices order 5
Risk-loving choices order 6

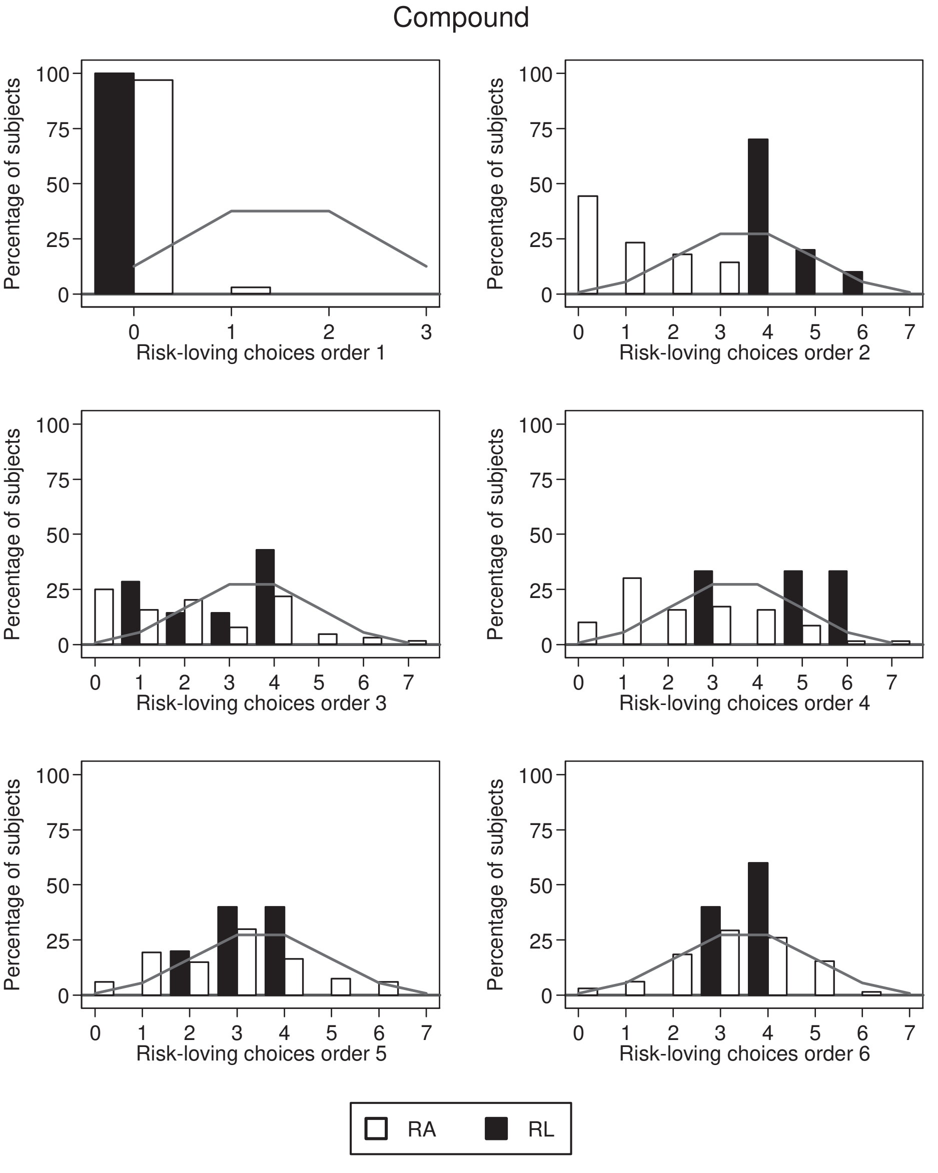
O9.2 Separated by risk-averse (RA) and risk-loving (RL) subjects

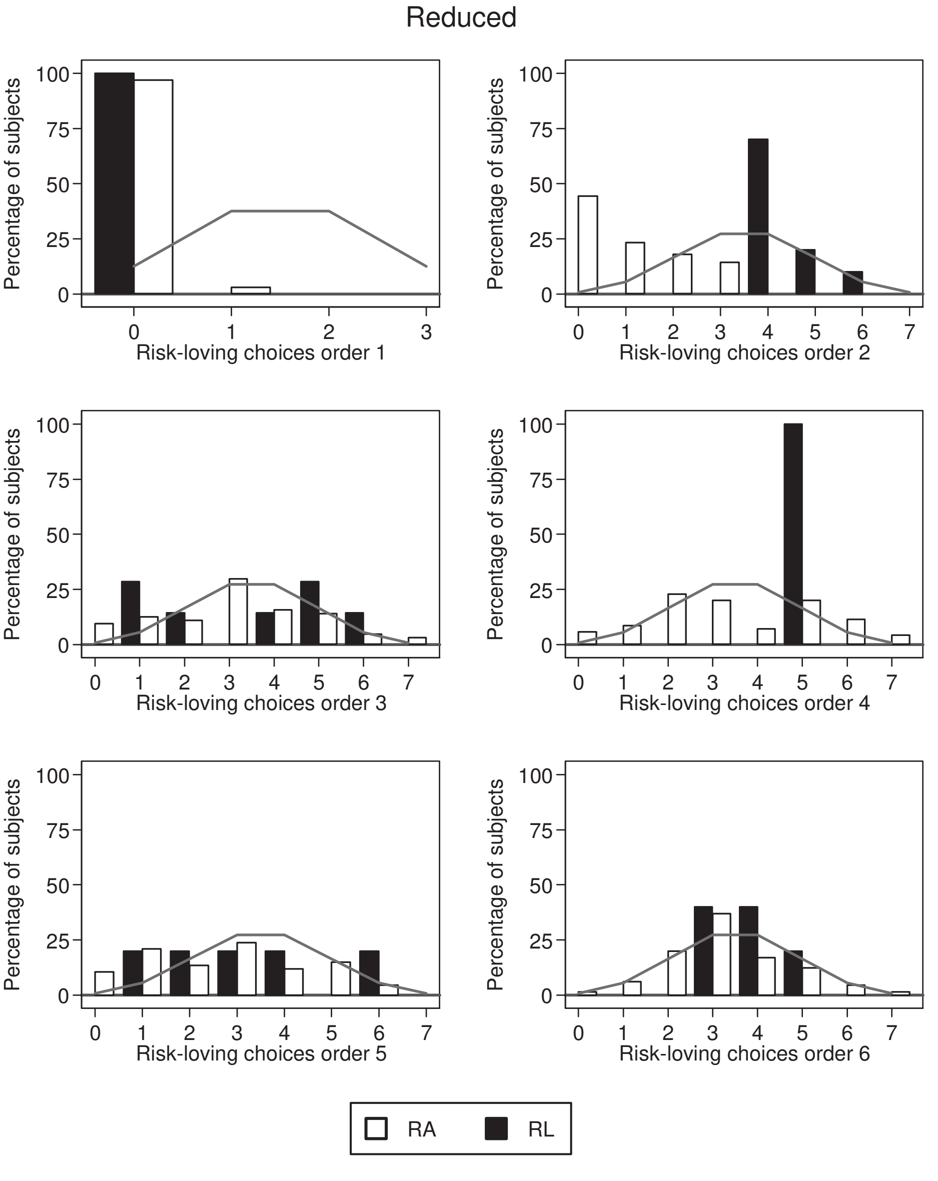



USA
Percentage of subjects
100
75
50
25
0
Risk-loving choices order 1
Risk-loving choices order 2
Risk-loving choices order 3
Risk-loving choices order 4
Risk-loving choices order 5
Risk-loving choices order 6
0
1
2
3
4
5
6
7
RA
RL


GER
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
RA
RL


CHN 10x
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


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
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

O10 Classification questionnaire compound (C) versus reduced (R)

Table O10: Questionnaire classification

				Compound			Reduced		
	Mean			Mean			Mean		
Variable	C	R	*p*-value	A	B	*p*-value	A	B	*p*-value
Decision: 1: imprudent choice; 0: prudent choice	0.194	0.611	**0.001**	–	–	–	–	–	–
Maximization of the largest potential payoff	0.528	0.167	**0.003**	0.000	0.655	**0.002**	0.000	0.429	**0.002**
Maximization of the smallest potential payoff	0.528	0.250	**0.029**	0.000	0.655	**0.002**	0.000	0.643	**0.000**
Maximization of the probability of the largest potential payoff	0.056	0.028	1.000	0.286	0.000	**0.033**	0.045	0.000	1.000
Minimization of the probability of the smallest potential payoff	0.000	0.028	1.000	0.000	0.000	–	0.045	0.000	1.000
Maximization of the payoff in the most likely outcome	0.139	0.500	**0.002**	0.714	0.000	**0.000**	0.818	0.000	**0.000**
Minimization of the payoff in the less likely outcome	0.000	0.028	1.000	0.000	0.000	–	0.045	0.000	1.000
Reference value	0.167	0.222	0.767	0.571	0.069	**0.008**	0.227	0.214	1.000
Maximization of the number of lager payoff	0.111	0.083	1.000	0.143	0.103	1.000	0.000	0.214	**0.051**
Minimization of the number of smaller payoff	0.000	0.056	0.493	0.000	0.000	–	0.045	0.071	1.000
Maximization of the sum of all payoffs	0.028	0.028	1.000	0.000	0.034	1.000	0.000	0.071	0.389
N	36	36		7	29		22	14	

Note: *p*-values of two sided Fisher's exact test with bold values indicating significant differences at least at the 0.10 level, C: compound R: reduced, A: imprudent choice B: prudent choice, all variables (except *Decision*) are equal to 1 if the explanation contains the respected argument and 0 otherwise.