Higher order risk preferences and economic decisions

Yilong Xu[1], Maarten Boksem[2], Charles N. Noussair[3], Stefan T. Trautmann[4], Gijs van de Kuilen[5], and Alan Sanfey[6]

[1]Utrecht School of Economics, Utrecht University, The Netherlands; [2]Erasmus University, The Netherlands; [3]University of Arizona, USA; [4]University of Heidelberg, Germany ; [5]Tilburg University, The Netherlands; [6]Radboud University, The Netherlands

October 30, 2025

Abstract

In theory, individuals' higher order risk attitudes of prudence and temperance influence saving and investment decisions. Prudent individuals save more when their future income becomes more uncertain, and temperate individuals prefer less risky investments in the presence of greater background risks. In a controlled experiment, we measure individuals' higher order risk attitudes directly, using two different elicitation methods. Participants then make saving and investment decisions under varying levels of background risk. We find strong effects of background risk on saving and investment. Moreover, individual prudence measures correlate with the strength of precautionary saving, while individual temperance measures do not do so with investment. The risk attitudes acquired with the two elicitation methods are strongly correlated with each other. The representative individual is risk averse and prudent, and neutral towards temperance.

KEYWORDS: high-order risks, precautionary saving, portfolio choice, risky decision-making

JEL CLASSIFICATIONS: C91; D15; D81; E21; E22; G51

Emails: y.xu3@uu.nl. mboksem@rsm.nl; cnoussair@arizona.edu; trautmann@uni-hd.de; g.v.d.kuilen@uvt.nl; a.sanfey@bsi.ru.nl. We thank Stephen Cheung, Brice Corgnet, David Crainich, Thomas Geelen, Ben Greiner, Pablo Guillen, Daniel Kárpáti, Fabrice Lelec, Alberto Prati, Frank Schilbach, Sebastian Schneider, Coen Teulings, Agnieszka Tymula, Robert Veszteg, seminar participants at Hanken School of Economics, IESEG iRisk, University of Queensland, University of Würzburg, Bilkent University, Bogazici University, University of Exeter, Radboud University, University of East Anglia, University of Sydney, Waseda University, Vienna University of Economics and Business, University of Goettingen, University of Southern Denmark, University of Montpellier, Maastricht University, Shandong University and conference participants at BFGA 2022, 2023 ESA Asia-Pacific meeting in Seoul, 2023 ESA World Meeting in Lyon, and BMC Frankfurt 2024, for helpful comments and suggestions.

1. Introduction

Second-order risk aversion in terms of concavity of the utility function is the most widely studied risk attitude in the literature, but it is far from a complete description of an individual's posture with regard to risk. For instance, properties of third- and fourth-order derivatives underpin the precautionary motives for savings and investment, characterizing how decision makers respond to background variables such as income risk or the riskiness of the overall macro-economic environment. Under the Von-Neumann-Morgenstern utility framework, the higher order attitude of *prudence* is defined as a positive third derivative ($u''' > 0$) of the utility function (Kimball, 1990). A key implication of prudence is the response of saving to future income risk (Leland, 1968), with an increase in background risk triggering greater (lower) savings on the part of prudent (imprudent) agents. Analogously, *temperance* is defined as $u'''' < 0$, a negative fourth derivative of the utility function (Kimball, 1992; Gollier and Pratt, 1996; Courbage and Rey, 2020). Temperance is a necessary condition for risk vulnerability and governs the response of investment portfolio choice to a change in background risk. Under greater background risk, an individual who engages in less risky (tempered) investment must be temperate.[1]

Micro-economic studies measuring higher order risk attitudes have shown that people are predominantly prudent, while the evidence for temperance is mixed (see the review by Trautmann and van de Kuilen, 2018). On the other hand, the macro-economic literature supports the empirical relevance of precautionary saving and tempering effects in investment decisions using observational data (e.g., Guiso et al., 1992, 1996; Guiso and Paiella, 2008; Fuchs-Schündeln and Schündeln, 2015; Bacon et al., 2020; see Lugilde et al. 2019 for a review), and, more recently, direct measures of subjective second-moment beliefs regarding macroeconomic uncertainty (Fermand et al., 2024) and administrative evidence on insurance under background-risk (Gropper and Kuhnen, 2025). However, the evidence based on observational data is innately indirect because of the lack of direct observation of higher order risk preferences. Thus, one cannot unambiguously attribute the observed decisions to prudent

[1] Under certain conditions on the utility function, temperance is also sufficient for tempered investment, e.g. under the widely-used CRRA utility functions.

and temperate preferences of the decision maker, as suggested by theory, rather than other latent factors. Moreover, most studies also use very indirect measures of background uncertainty, making it hard to interpret the observed effects (Lugilde et al. 2019). In contrast, studies that measure higher order risk preferences directly in population samples typically rely on self-reported or hypothetical measures of saving and background uncertainty (e.g., Noussair et al., 2014; Schneider and Sutter, 2020; Schneider et al., 2022). Given this lack of direct evidence for a linkage between preferences and precautionary behavior, the current paper aims to complement the evidence from field data, which has high external validity but suffers from measurement problems, with laboratory evidence from artificial tasks, but providing full control over the relevant background uncertainties and observation of the relevant preferences. This allows us to directly study the suggested mechanism underlying the empirically observed precautionary effects via higher order risk attitudes.

To this end, we devise two laboratory economic decision tasks to study savings and investment behavior. The experimental environment allows us to unambiguously alter the degree of income uncertainty and background risk and observe their influence on savings and investment decisions. We can thus offer *indirect* evidence about the incidence of prudent and temperate behavior, as in field data, but free from potential selection effects and measurement error of background risks. Moreover, we also elicit individuals' prudence and temperance preferences as measured with abstract lottery choice tasks. This allows us to also *directly* test whether more prudent and temperate individuals react more strongly to clearly defined background risks, which is typically not possible using field data. If the evidence supports the direct link between higher order risk attitudes and precautionary behavior, survey measures of these attitudes can be useful in predicting the population's reaction to increases in unavoidable background risk, e.g., in times of economic crisis.

The economic decision making tasks we administer mimic savings and investment decisions with various levels of income uncertainty. In the savings task, which is similar to that employed in Bostian and Heinzel (2011), we present individuals with a two-period decision in which, in the first period, they have an opportunity to save for, or to borrow from, the second period, to smooth their income. The realization of the second-period income is risky, with the variance of income differing among trials. A prudent individual saves more first-period income for the second period, the greater the riskiness of the second-period income. We thus interpret such a pattern of precautionary saving as indirect evidence for prudence. The opposite pattern reveals imprudence.

For the investment task, we use a design similar to Beaud and Willinger (2015). Individuals must allocate an endowment between a safe and a risky asset. However, they also receive a random shock to their income (background risk). The variance of the income shock, but not the realization of the shock, is known at the time of the investment decision. In this situation, a temperate individual holds a less risky portfolio as her background risk increases.

The two tasks measuring higher order risk preferences involve risk apportionment and premium measurement. Both tasks are based on the work of Eeckhoudt and Schlesinger (2006) who propose a behavioral definition distinguishing prudent from imprudent individuals. Suppose an individual is forced to accept a zero-mean risk. If she is more willing to accept the risk when her wealth is relatively high (low), she is said to be prudent (imprudent). Under the assumption of expected utility, the above behavior distinguishes individuals on the basis of the sign of $u'''(x)$.[2] Analogously, a behavioral definition of temperance can be formulated. An individual is temperate if, when facing two unavoidable independent risks that each arise only in one state, she prefers to have them occur in different states. An intemperate individual prefers to have them concentrated in the same state. Eeckhoudt and Schlesinger (2006) show that, if expected utility is assumed, this behavioral definition is equivalent to signing $u''''(x)$. Implementations of these behavioral definitions have been employed to classify individuals as (im)prudent and (in)temperate by, among others, Deck and Schlesinger (2010, 2014), Ebert and Wiesen (2011, 2014) and Noussair et al. (2014). In essence, prudence is identified by asking

[2] The behavioral definition can be applied to classify individuals as prudent or imprudent in an intuitive manner, even if the expected utility hypothesis is not assumed. There is a natural analogue for risk aversion. Risk aversion is defined as concavity of the utility function, or $u'' < 0$, or alternatively as a preference to decline an optional zero-mean risk. The second definition is behavioral because it is based on a decision rather than an unobservable utility function. Under the assumption of expected utility, the two definitions are equivalent because one would decline any zero-mean risk if and only if one had a concave utility function (leaving aside the indifference of risk neutral agents). However, it is clear that an agent who declines all zero-mean risks, even if not an expected utility maximizer, is in a very intuitive sense risk averse.

individuals to apportion an unavoidable risk to either a low or high wealth state. Temperance is identified in a task asking them to apportion an unavoidable risk between a state that already carries risk and a state that has no other risks.

Our experiment makes use of a within-person design, measuring preferences and economic decisions in much detail for each participant. In the first risk preference elicitation task of our protocol, subjects make a total of 33 pairwise choices between lotteries. These include 5 pairs for measuring risk aversion, 14 pairs for measuring prudence and another 14 pairs for measuring temperance. Risk aversion is measured with choices between a safe and a risky lottery. The prudence measurement choices involve assigning an unavoidable risk to either a high or a low wealth state. The temperance measurement decisions require assignment of an unavoidable risk to either a risk-free or already risky state. We count the number of instances in which an individual chooses the risky, prudent, or temperate options and use this count as a measure of strength of preference for higher order risk attitudes, following the extant literature (e.g. Deck and Schlesinger 2010, 2014, Noussair et al. 2014).

The second risk preference elicitation task is a premium-measurement protocol that builds on the pairwise lottery decisions just described. Using sequences of titrated pairwise decisions following the methodology of Abdellaoui et al. (2011), we iteratively change the expected payoff of the risky, imprudent, or intemperate lottery in a pair (keeping the higher order moments constant), until we approximately find a participant's point of indifference between the two lotteries. This method is known as the titration method or the bisection method (see also Schneider et al., 2022, who use the same method to find certainty equivalents). The resulting difference in expected payoff allows the measurement of a risk, prudence, and temperance premium for each individual.[3] We randomize the order of choices to disguise the sequential nature and dependence of later titration steps on earlier choices (see Section 2 for details). Importantly, the premium-based method allows us to directly make claims regarding the strength of risk preferences.

[3] Ebert and Wiesen (2014) also elicit risk, prudence, and temperance premia based on Eeckhoudt and Schlesinger (2006)'s risk apportionment task, but using a choice-list approach.

In our analysis, we first study the correlation between the two elicitation methods for (higher order) risk preferences to gauge whether the commonly used count measure can indeed be interpreted as a strength of preference. We then use the measured risk preferences to predict economic decisions. We hypothesize that the more prudent an individual's behavior in the lottery choice task, the more prudently she behaves in the economic decision tasks. That is, those who prefer to incur unavoidable risks in relatively high-income states and tend to have higher prudence premia are more likely to save more, the greater the future income risk that they face. Similarly, we hypothesize that those who invest in a more risk averse manner as the background risk in their environment increases, behaving in a more temperate manner, also prefer to disaggregate unavoidable risks across states and tend to have relatively high temperance premia.

Our results make several novel contributions of interest to the literature on financial decision making. First, we find strong evidence for precautionary saving and tempered investment in our framed economic saving and investment tasks. That is, we replicate the indirect evidence that the majority of individuals are prudent and temperate, found in the empirical literature in macroeconomics. However, the strong evidence of temperate investment is somewhat at odds with the relatively low prevalence of temperance typically observed in experimental studies in which temperance is directly measured.

Second, we show that for prudence, choice-based preference measures predict individuals' saving behavior in a manner consistent with the predictions of expected utility theory. This is direct evidence in support of the mechanism based on the effect of uncertainty on the marginal utility of wealth as first formulated in Leland (1968).[4] In contrast, tempered investment is not associated with measures of temperance. Given the clear evidence for tempered investment and the mixed evidence on temperance in studies using Eeckhoudt and Schlesinger (2006) risk apportionment tasks, our result suggests that these tasks may not successfully tap into the economic attitudes they are supposed to capture. There may be room for improvement in the design of risk apportionment protocols to measure temperance.

[4] The evidence is thus consistent with the predictions of the expected utility framework. We discuss the implications for non-expected utility analyses of saving under background risk in Section 6.

Third, the two risk-attitude elicitation methods are strongly correlated with each other at the level of the individual decision maker. These findings are especially noteworthy given the poor correlations across elicitation methods for second-order risk preferences ("risk aversion"), and the often-poor predictive power of these tools (Friedman et al., 2014, 2022). Our results support the strength-of-preference interpretation of the binary choice count measure widely used in the experimental literature on higher order risk preferences.

The paper proceeds as follows. In Section 2, we describe how we measure higher order risk attitudes. In Section 3 we describe the economic tasks and state our hypotheses. We provide details on the implementation of the experiment in Section 4. The results are reported in Section 5, and in Section 6 we offer a concluding discussion.

2. Measuring higher order risk attitudes

2.1 Theoretical Foundations

The seminal work by Rothschild and Stiglitz (1970) lays down a behavioral definition of risk aversion in terms of dislike of mean-preserving spreads. In other words, a person prefers a sure payment over non-degenerate lotteries that offer the same expected value. In a similar fashion, Eeckhoudt and Schlesinger (2006) offer a behavioral definition of the higher order risk attitudes of prudence and temperance, in terms of preferences for risk apportionment. In particular, prudent individuals prefer to confront an unavoidable zero-mean risk when their wealth level is high, rather than when it is low. Similarly, temperate individuals prefer to disaggregate two independent zero-mean risks across different states, rather than dealing with both risks in a single state.

Under the expected utility framework, the behavioral definition of prudence of the last paragraph is equivalent to a positive third derivative of the utility function, $u'''(x) > 0$, and thus a convex marginal utility (Kimball, 1990). The behavioral definition of temperance is equivalent to a negative fourth derivative of the utility function, $u''''(x) < 0$, and thus a concave second derivative of the utility function (Kimball, 1992).

The preference conditions developed by Eeckhoudt and Schlesinger (2006) are shown in Figures 1 and 2. In Figure 1, x and y are two monetary payoffs with $x > y$. The payoffs

$+z_1$ and $-z_2$, where $z_1 > 0$, and $z_2 > 0$, are the two outcomes of a zero-mean lottery that yields $+z_1$ with probability p, and outcome $-z_2$ with probability $1-p$. Let us denote this by $(p: z_1; (1-p): -z_2)$. In our experiment, we set $p = \frac{1}{2}$ and $z_1 = z_2$. A decision-maker who prefers more money to less and is risk-averse will dislike the zero-mean lottery $(p: z_1; (1-p): -z_2)$. The decision-maker is asked to choose between the two lotteries, L and R, that have the same expected value and the same variance. In both lotteries, an initial equal-chance risk determines whether the high amount x or the low amount y is received. Depending on which lottery is chosen, the decision-maker may additionally receive the outcome of the zero-mean lottery $(p: z_1; (1-p): -z_2)$ either in the state where the payoff is high (x), or in the state where the payoff is low (y).[5] That is, the tradeoff between lotteries L and R in Figure 1 is whether to bear the unavoidable zero-mean risk when income is high or when it is low. An alternative interpretation is that receiving the outcome y implies a relative loss of $k = x - y$. Choosing L (R) means that the risk averse decision-maker prefers to disaggregate (aggregate) the unpleasant loss of k and the unpleasant zero-mean lottery. In contrast, a risk-seeking decision maker may prefer to concentrate two events he views as favorable, the lottery and the higher payoff x. This would lead the risk-seeking individual to also choose Lottery L. It is shown in Eeckhoudt and Schlesinger (2006) that a preference for Lottery L in Figure 1 for every x, y and z implies that the decision-maker is *prudent* (has u''' > 0) under expected utility.

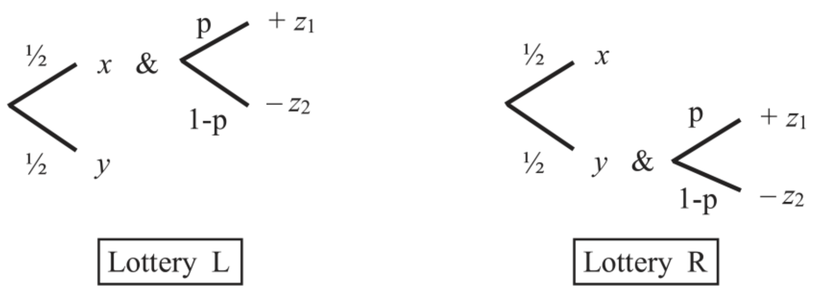

Figure 1. Risk apportionment task identifying prudence

[5] These outcomes x and y do not have to be in the gain domain. See Attema et al. (2019) using lotteries with hypothetical outcomes in wealth and health, or Brunette and Jacob (2019), who study the lotteries of Noussair et al. (2014) in the loss domain.

Figure 2 shows the structure of a lottery pair that is used to identify temperance. The decision-maker is asked to choose between the two alternatives, L or R, that have the same expected value, variance and skewness. In both lotteries, the initial wealth level is the certain amount x. The decision amounts to aggregating or disaggregating two independent zero-mean risks, $(p: y_1; (1-p): -y_2)$ and $(q: z_1; (1-q): -z_2)$. For a risk-averse decision-maker, being temperate implies that she prefers to disaggregate the two unavoidable zero-mean risks (lottery L), instead of aggregating them (lottery R). In the experiment, we choose $p = q = \frac{1}{2}$, $y_1 = y_2$, and $z_1 = z_2$. Eeckhoudt and Schlesinger (2006) show that, under expected utility, a preference for Lottery L for every x, y and z is equivalent to *temperance* (u'''' < 0).

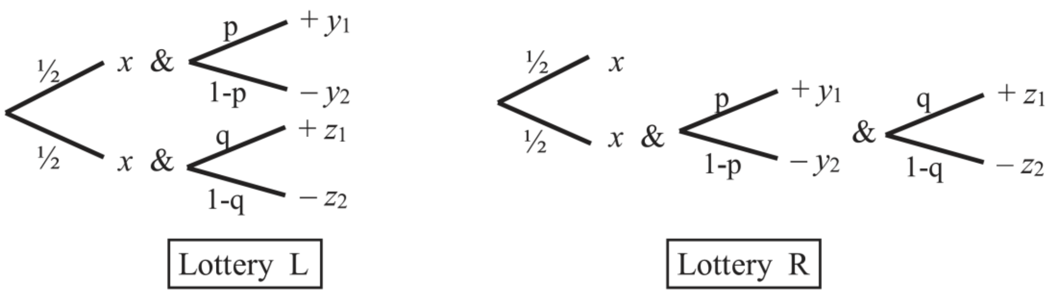

Figure 2. Risk apportionment task identifying temperance

2.2. Eliciting risk, prudence, and temperance preferences.

We employ two approaches to assess individuals' risk attitudes based on the theoretical framework discussed above. The first, the Binary Choice Count, measures preferences by counting the number of instances in which individuals choose risk-free/prudent/temperate lotteries in a set of independent choices (e.g., Deck and Schlesinger, 2010, 2014, Noussair et al., 2014). The second is a titration method that identifies risk premia through a sequence of pairwise lottery choices. The titration method has the advantage of directly measuring individuals' strength of risk, prudence, and temperance preferences, while the count method serves this purpose only under additional assumptions. However, the count method has been used widely in the literature: the comparison between the two methods allows us to assess the validity of the interpretation of the count as a measure of preference strength in influential previous work.[6] No study has yet attempted to provide such a test.

Binary Choice Count

In the Binary Choice Count method, participants made choices between the 33 pairs of lotteries shown in Table A1 in the appendix. There were 5 choices measuring risk attitude, which were naturally ordered in terms of strength of risk aversion (more risk averse individuals choose lower-payoff safe lotteries over a given risky lottery). The count of the number of risk averse choices provides a direct measure of the strength of risk aversion. There were 14 choices each measuring prudence and temperance, and a prudent individual should choose the prudent alternative in each of these choices (same for temperance). Following previous literature, we interpret the count of the number of prudent and temperate choices, respectively, as a measure of strength of preference. This interpretation is based on idea that a, say, prudent individual should optimally choose the prudent option in *each* choice problem. However, if her utility assessment of the two alternatives is prone to some noise, she may choose the less preferred option occasionally. This is more likely the closer the prudent and imprudent lotteries in a choice problem are in terms of utility. Thus, a stronger preference for prudence will then appear as a more consistent choice of the prudent option over the 14 choice problems. The number of prudent (temperate) choices can then be used as a measure of the strength of these higher order risk attitudes.

Titration method

The titration method directly measures the cash premia that individuals require to be indifferent between a risky and a safe, a prudent and imprudent, and a temperate and an intemperate lottery in a choice pair. To measure the second-order risk premium, we use choice RA4 in Table A1 that asks participants to choose between a sure amount of €35 and a risky lottery paying €5 or €65 with equal probability. For the prudence and temperance premia, we use all 14 pairs of lotteries (Prud 1-14 and Temp 1-14, in Table A1). To calculate a participant's overall prudence and temperance premia, we average over the 14 elicited prudence or temperance premia,

[6] The two tasks were administered on different days, see Section 4 for details on the implementation.

respectively.

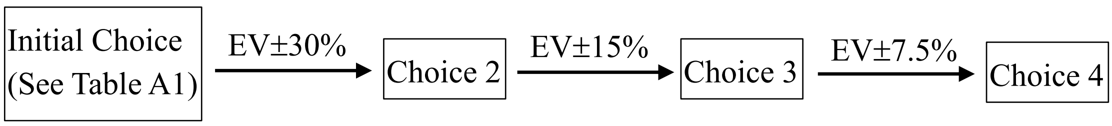

Figure 3. Titration steps. EV stands for Expected Value of Lottery R

The titration, illustrated in Figure 3, works as follows. Depending on the decision-maker's choice of Lottery L or Lottery R in a pair, we iteratively adjust the expected payoff of Lottery R (the risky, imprudent, intemperate lottery) to approximate the point where the decision maker is indifferent between the two lotteries. The titration process consists of three steps after the initial choice in a pair, as can be seen in Figure 3. In steps one to three of the titration, both outcomes x and y (in risk and prudence choices, see Figure 1) or outcome x (in temperance choices, see Figure 2) of Lottery R are increased or decreased by 30%, then a further 15% and then another 7.5%, with respect to the current expected value in each step. The direction of the adjustments in each step depends on the choice made by the decision maker before each step. If Lottery L (R) is chosen, the expected payoff of lottery R will subsequently be increased (decreased) to make it more (less) attractive. The zero-mean risks are not adjusted in the process. Thus, for each decision problem, the measurement of indifference is based on 4 choices.

Indifference values for x and y in lottery R are calculated as follows. If the decision maker chooses different lotteries in Choice 3 and Choice 4, the midpoint between the Choice-3 and Choice-4 values is taken as the indifference value for x and y. If the decision maker chooses the same lottery in Choice 3 and Choice 4, we take the final values in Choice 4 as indifference values. That is, we do not project values outside the range of values covered. Note that a slight misspecification of indifference values is possible using this method, and that additional titration steps would have allowed us to zoom into preferences more precisely. This would have come at a cost in terms of effort and time though, and pretests suggested that the 4-choice procedure provides robust indifference points.[7]

Once we find the indifference lotteries, we can calculate the premia, which we define as the difference between the Expected Value (EV) of the indifference lottery R and the EV of the original Lottery R in each decision problem, divided by the EV of the original Lottery R: $\frac{EV_{indiff}-EV_{original}}{EV_{original}}$. The normalization is warranted because the pairs of lotteries vary substantially in terms of their initial expected values (ranging from €24 to €72). We define a person's prudence and temperance premium as the average value over the respective 14 premia for each preference. For risk, there is only one decision problem, RA4, which defines the risk premium. We provide a numerical example in Appendix A.2.

In total, our participants make 116 decisions (29 initial pairs * 4 titration steps) in this part of the experiment. These pairwise lotteries are randomized to disguise the sequential nature and dependence of later titration steps on earlier choices. More specifically, lottery choices are chosen one at a time from a pool that initially consists of the 29 original lottery pairs (lotteries RA4; Prud 1-14; Temp 1-14, see Table A1). Once a choice is made, the new (=titrated) lottery pair is added to the pool. Thus, the pool again has 29 lottery pairs. Therefore, the chance that the titrated lottery pair will directly follow the previous linked lottery pair is 1/29. This procedure is repeated until each lottery in the initial pool has been titrated 3 times and all choices have been made. Participants are only told to make 116 risky decisions without being informed about the titration method working in the background. Thus, it is unlikely that participants observed the connection between lotteries.

[7] In a pre-test of the method, we examined whether indifference points are well-identified. To this end, we either added or subtracted €2 or €4 euros to the indifference values for lottery R, creating 4 new lotteries, two more attractive, and two less attractive than the indifference lottery. We then had subjects choose between the original lottery L and these new lotteries R. If the indifference points are well defined, participants should choose L in the two choices with an unattractive lottery R, and R in the two choices with an attractive lottery R. Thus, we would like to observe 50% of choices of R and of L. This was indeed the case, with participants choosing the prudent (temperate) lottery in 55% (50%) of the trials, suggesting that the titration method worked well.

3. Economic decision tasks and theoretical predictions

3.1 Precautionary savings

In this part of the experiment, participants make 48 two-period saving or borrowing decisions for different combinations of current and future uncertain incomes. In period 1, participants receive a fixed income of €X, where X takes a value from a set of eight alternatives: {33, 36, 39, 42, 45, 48, 51, 54}. In period 2, income is risky. Participants either receive the high income €*a* or the low income €*b*, with equal probability. We denote the distribution of income in period 2 by *a_b*. Risky period-2 incomes are drawn from the set of six alternatives {70_20, 66_24, 62_28, 58_32, 54_36, 50_40}. All risky period-2 incomes have the same expected value of €45. Participants are confronted with a total of 48 saving decision problems, with each period-1 income paired with each period-2 income distribution once. They may save some period-1 income for the next period, or borrow (dissave) from period 2 to increase their income in the first period. The 48 problems are presented in a random order that differs for each participant.

The saving decisions are implemented as follows. Of the two periods in a saving problem, exactly one is paid out, with both periods having an equal chance to be selected for payment. Random selection of one period induces "time" separability, as in the underlying life cycle decision model. It follows that income smoothing maximizes the expected marginal utility of payments. The shape of the marginal utility affects decision makers' allocation of payments over periods in the case of uncertain income, inducing precautionary saving for prudent decision makers. If period 1 is selected for payment, the decision maker receives €X minus any amount saved (or plus any amount borrowed). If period 2 is selected, the risky payoff €*a* or €*b* is determined by a random draw, which is then increased by the amount saved (or reduced by the amount borrowed). The maximum amount that can be saved or borrowed in a decision problem is €20, and the decision maker can specify their decision as any multiple of €2. The interest rate equals zero for both borrowing and saving.

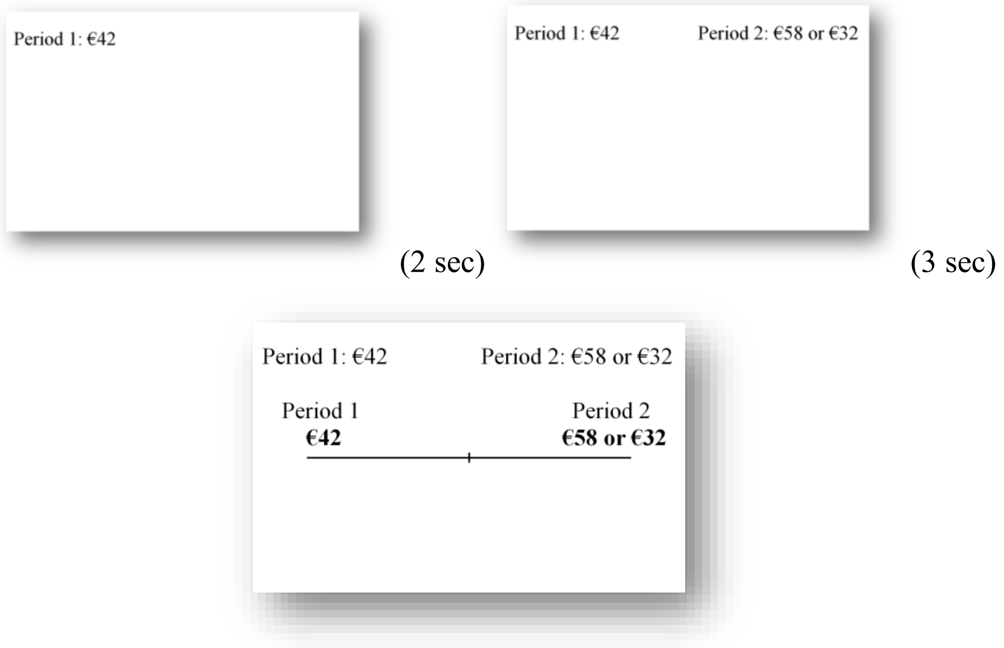

Figure 4. Presentation of saving/borrowing decisions

Figure 4 shows the interface for this task in the experiment. We first present period-1 income for two seconds, then we present the period-2 risky incomes for three seconds. A slider bar appears after three seconds, and participants can use a controller to move the slider left or right to transfer money between the two periods in steps of €2. The slider itself remains invisible (i.e., it is not preset) before participants start to make the decision using the controller to avoid influencing their decisions with the initial placement.

In the example shown in Figure 4, suppose that the maximum amount of €20 is borrowed, shifting income from period 2 to period 1 (done by moving the slider to the left end). Then period-1 income becomes €62 and period-2 income will be either €38 or €12. If, on the other hand, €10 is saved, shifting income from period 1 to period 2, then the period-1 income becomes €32, and the period-2 income will be either €68 or €42.

Under expected utility, if a decision-maker is risk-averse (has decreasing marginal utility, $u''(x)<0$), larger expected income differences between period 1 and period 2 predict more consumption smoothing to equalize the marginal utility of period-1 and period-2 consumption. Moreover, if a decision maker is prudent, which is equivalent to the third derivative of her utility function being positive ($u'''(x) > 0$) under expected utility (Kimball, 1990), then a higher level of period-2 income uncertainty predicts more savings. This is referred to as precautionary saving (Leland, 1968).

Our first hypothesis is an implication of risk aversion on the part of decision makers. If a majority of individuals are risk averse, we hypothesize that the representative individual will behave in a risk averse manner in the savings task and seek to smooth consumption.

***H1** (Consumption Smoothing)**:** There is a positive correlation between the income difference between the two periods (period-1 income minus period-2 expected income) and the amount saved.*

The second hypothesis is an implication of a tendency of individuals to be prudent. Regardless of whether they are risk averse or risk seeking[8] individuals engage in more precautionary saving, the greater is their future income risk.

***H2** (Precautionary Saving): There is a positive correlation between the riskiness of future income, measured by the spread of possible period-2 incomes, and the amount saved by decision-makers.*

Under expected utility, the curvature of an individual's marginal utility function should predict the strength of the effect of prudence on individuals' saving decisions. That is, the more convex an individual's marginal utility function is, the more her saving responds positively to increases in future income risk (Kimball, 1990). This is the basis of our third hypothesis.

***H3** (Heterogeneity in Precautionary Saving): Individuals' prudence preferences correlate with saving decisions. In particular, those who demand a higher prudence premium (or make a greater number of prudent choices in the binary choice count task) increase their savings more when period-2 income is riskier.*

While our hypotheses follow from established theoretical results in the literature, it may not be intuitively clear how these results map into our experimental choices. To aid intuition,

[8] Crainich et al. (2013) document that risk lovers are also prudent and accumulate precautionary savings, just like risk averters. Under expected utility, this is termed mixed risk aversion (Caballé and Pomansky, 1996). Evidence for mixed risk aversion has been found by Deck and Schlesinger (2014), Haering et al. (2020), and Mussio and de Oliveira (2022).

we derive a parametric example for the widely-used CRRA family in Appendix A.3, demonstrating all three hypothesized effects.

3.2 Investment decisions

In this part of the experiment, participants make 96 investment decisions. In each decision problem, before making their investment decision, they learn whether their baseline income is risk-free or risky, that is, whether background risk is present. Investments yield additional income on top of the baseline income. The investment task consists of a portfolio choice, to set the percentage of income invested into a safe and a risky asset.

Specifically, the baseline income has an expected value of either €15 or €20, and comes from a set of eight alternatives, either safe or risky: {15, 20_10, 25_5, 30_0; 20, 25_15, 30_10, 40_0}. For each expected income level, there are four levels of risk, where a larger spread implies a greater background risk for the investors. Investment decisions are operationalized as follows. The zero-risk safe asset in each trial is drawn from a set of two alternatives, {25, 30}, and the risky asset is drawn from a set of six alternatives, {50_20, 53_17, 56_14, 59_11, 65_5}. The combination of eight levels of background risk, with two levels of the safe investment's payoff and six different risky investments, implies a total of 96 different possible choice situations. Each participant experiences each situation exactly once. Situations are presented in a random order, which differs by participant.

Figure 5 shows the interface for this task. At first, the level of background risk is presented to the decision-makers for 2 seconds, and then the investment alternatives are shown for 3 seconds. Then they can make their portfolio allocation decision. Again, the slider remains invisible until they use the controller to select an allocation. In the example in Figure 5, suppose a decision-maker decides to invest 50% in the risky asset and 50% in the safe asset. Her income will be determined as follows: She will receive either €30 or €10 as baseline income. Additionally, she will receive €15 for sure from a 50%-share in the safe €30 asset, plus another €25 or €10 from her 50%-share in the risky asset.

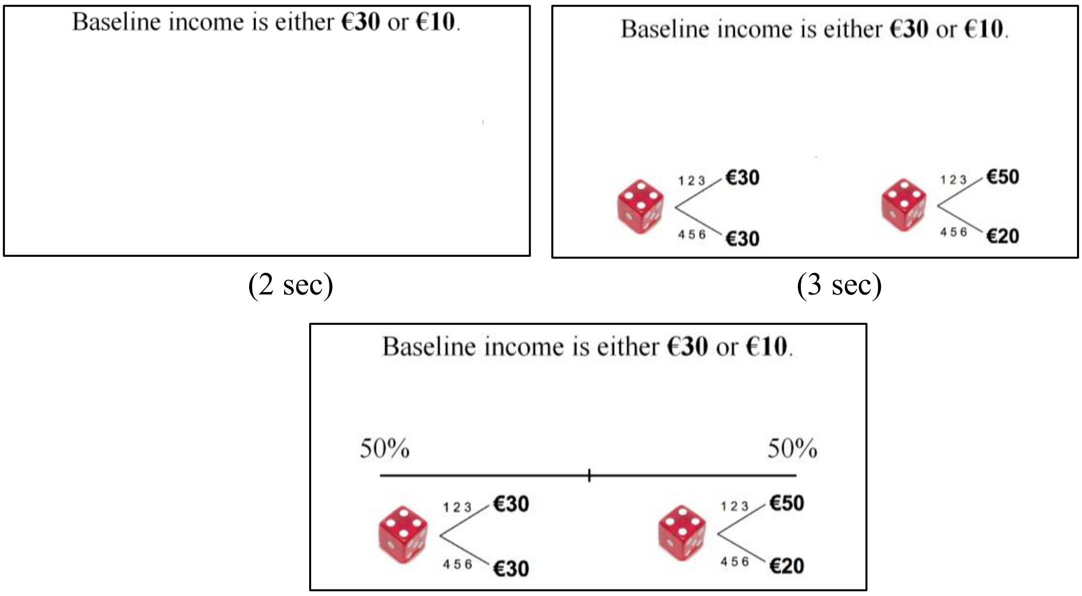

Figure 5. Presentation of investment decisions

Under expected utility, if the decision maker is temperate, the second derivative of her utility function is concave (Kimball, 1992). When facing a higher background risk (riskier baseline income), temperance is a necessary condition for the individual to reduce their exposure to the risky asset. For many widely used forms of the utility function, temperance is also sufficient for a tempering effect on risk taking (Gollier and Pratt, 1996). Assuming that the representative participant is temperate, we formulate our first hypothesis on the tempering effects of background risk.

***H4** (Tempering Effects): The level of background risk is positively correlated with the proportion of the investment allocated to the safe asset.*

Kimball (1992, 1993) and Gollier and Pratt (1996) derive how an individual's temperance preferences relate to their investment behavior. The exact relationship and the impact of the degree of temperance depend on several features of the utility function. Here we make a first attempt to test whether any tempering effects in investment (if observed in the investment decision task) can be explained by the strength of temperance as revealed in the preference measurement tasks. We formulate our second hypothesis on the tempering effects of background risk.

***H5** (Heterogeneity in Investment Decisions and Temperance): Individuals who have a stronger* *tendency to reduce investment in the risky asset as background risk increases also demand a higher temperance premium, or make a greater number of temperate choices in the count task.*

Appendix A.3 demonstrates Hypotheses H4 and H5 for the CRRA family with example calculations from our experimental investment choices.

4. Experimental Procedure and Lab Details

General Procedures. The experiment consists of three parts and is designed for within-subject analyses. Part 1 elicits individuals' risk attitudes using the titration method. Part 2 comprises the savings and investment decisions. Part 3 measures individuals' risk preferences using the binary choice count method. Part 1 took place on a different day than parts 2 and 3 (always at least one day apart). Because the study was part of a larger neuroeconomic project, participants made the Part-2 decision in an fMRI scanner. Prior to Part 2, there was another task in the scanner where subjects passively viewed several payoff-relevant lotteries but made no choices relevant to this paper. The current paper only discusses the behavioral data.

All choices were incentivized as follows. For the first session, each subject had a 10% chance of being selected for real payment, and if so, one of the 116 decisions in Part 1 was randomly chosen to be payoff relevant. Part 1 took approximately 1 hour and 15 minutes on average. For the second session, all subjects were paid and this session took 2 hours and 15 minutes on average. One task from Part 2, Part 3, or the passive lottery-viewing task was randomly selected for real payment. All payoffs depended on the participant's choices and the randomizations of lottery payoffs in the selected tasks. The average payment for our subjects was 47.29 euros for the whole experiment (including the 5 euro show-up fee).

The titration task, as well as the binary choice count task, was implemented with z-tree (Fischbacher, 2007). The saving and investment tasks inside the scanner were programmed in Presentation®.

Participants and Data Collection. A total of 51 participants participated in our study at the Donders Institute for Brain, Cognition and Behaviour at Radboud University, in Nijmegen, The Netherlands. Participants were students from various disciplines at Radboud University. They were recruited via the SONA system and immediately scheduled to participate in all three parts of the study when enrolling for the experiment.

The titration task was done with 3 to 4 individuals simultaneously in a computer room. Upon arrival, they were assigned a computer, waiting for the instructions. We made sure that the participants were aware that it was the first part of the study, and two other parts would take place on another day. We then handed a hardcopy of the instructions to the participants for them to read at their own pace. All questions were resolved privately before the experiment started. The titration task itself took about 45 minutes (excluding instructions and payments). After the participants made all their choices, the computer selected whether Part 1 of the study was payoff relevant (with a 10% chance). If this were the case, one of the lottery pairs out of the 116 decisions was randomly selected by the computer to be payoff relevant. The computer then rolled a virtual die to determine the payoffs of the relevant lottery, depending on the participant's choice.

Participants were then re-invited the next day or a few days later back to the lab on an individual basis, depending on the availability of the fMRI scanner, to participate in the remaining parts of the experiment. The participants received instructions for the economic decision tasks and read the instructions at their own pace. The experimenter resolved any remaining questions the participant had. We then demonstrated the experimental interface and subjects were asked to try out the controller for the sliders before entering the scanner. The decision tasks took approximately 50 minutes to finish inside the scanner.

After the scanning session, the participant went to a computer room to participate in the last part of the experiment, the binary choice count task. They had an opportunity to rest for 15 minutes before the start of Part 3 of the experiment. We handed out the instructions for the last part of the experiment and answered all remaining questions. This part of the experiment took approximately 15 minutes.

After participants had made their choices in Part 3, the payment process started. An online random number generator (www.random.org) determined which task was payoff relevant (out of all Part-2 and Part-3 decisions). We retrieved the participants' decision in the selected task and resolved the uncertainty in the involved lotteries using the same random number generator. The participants were then paid by bank transfer on the same day.

5. Results

5.1 Risk aversion, Prudence, and Temperance

In the risk-premium titration tasks, an individual is considered risk-averse, prudent, or temperate if their risk, prudence, or temperance premium, respectively, is positive. The elicited values of the premia indicate that the vast majority of participants are risk-averse (41 out of 51) and prudent (43 out of 51). About half of the participants are temperate (29 out of 51). On average, our participants demanded a 26% premium for the risky lotteries, an 11% premium for the imprudent lotteries, and a -0.10% premium for the intemperate lotteries. All, except for the temperance premium, are significantly different from zero at 1% level in the sample, according to a two-sided Wilcoxon Signed Rank test.[9] The results of the count method are consistent with these results. Out of 51 participants, 49 chose the risk-free option more than half of the time (i.e., choosing at least 3 risk-free options out of 5 pairs); 46 participants chose the prudent lottery more than half of the time (i.e., choosing at least 8 prudent lotteries out of 14 pairs); and 35 participants chose the temperate lottery more than half of the time (i.e., choosing at least 8 temperate lotteries out of 14 pairs).

The count numbers in the binary choice count method have typically been interpreted as indicating the strength of a person's risk aversion, prudence, and temperance (Trautmann and van de Kuilen, 2018). If this interpretation is warranted, the counts should be correlated with the premia in the titration task, which directly measure the strength of preference. Although Ebert and Wiesen (2014) and Schneider et al. (2022), among others, have expressed concerns over this interpretation of the count, we are not aware of any prior study explicitly testing the strength of preference interpretation of the count measure.

[9] Note that we observed only one participant choosing the same lottery four times, and only for one choice problem (a prudence question). This suggests that the four-step titration procedure is sufficient to effectively detect indifference values.

Table 1. Spearman's correlation coefficients between the risk attitudes measured in the two abstract tasks, the count and the premium measures

	# Safe Choices	# Prudent Choices	# Temperate Choices	Risk Premium	Prudence Premium
# Safe Choices	1.00				
# Prudent Choices	0.16	1.00			
# Temperate Choices	-0.04	0.10	1.00		
Risk Premium	0.44***	0.11	-0.13	1.00	
Prudence Premium	0.21	0.38***	0.17	0.34**	1.00
Temperance Premium	-0.04	-0.04	0.38***	0.20	0.37**

Note: *** p<0.01, ** p<0.05. The # of Safe, Prudent, and Temperate choices is the number of times those choices were made in Part 3 of the experiment. The premium measures are those obtained in the tasks in Part 1.

Table 1 shows the correlation between the count and premia measures in the highlighted cells. The first three rows and columns are correlations with the number of risk averse, prudent, and temperate choices in the count task. The last three rows show correlations with the premia measured in the binary choice titration protocol for each of the three risk attitudes. Figure 6 demonstrates the individual-level data graphically and provides Pearson's correlations. The figure also illustrates the distribution of the count and premia measures. It is clear that prudence and risk aversion are much more prevalent than temperance with most of the data indicating a premium greater than 0 and a count measure above 3 in the risk aversion task and 8 in the prudence task. There is substantial correlation between the two measures for each of the three risk attitudes, and most of the data lies in the Southwest and Northeast quadrants. The figure also shows that the temperance premia are, on average, approximately zero while most individuals make a majority of temperate choices under the count measure.

Note that the count and premia measures were elicited on different days. Despite the time lag and a different incentive structure, the elicited risk premia are strongly and significantly associated with the choice count measures for each of the three risk attitudes. Importantly, each premium is only associated with its corresponding risk attitude count and with neither of the other two. However, the prudence premia are associated with risk and temperance premia, while no such association is found for the count measure, suggesting that the two measures do differ from each other in some fashion. In the next subsection we test whether each of the two measurement approaches is successful in predicting participants' economic decisions. In our analysis, we at times refer to the two tasks as *abstract* tasks.

5.2. Explaining Precautionary Saving

We next study behavior in the saving task, using the regression specifications presented in Table 2. In our model, significant identification of effects stems from the within-person variation. Each of the 51 participants makes 48 saving decisions, resulting in a total of 2,448 observations within 51 clusters. The table contains two sections: the grey shaded area shows the effects of the background risk measures on saving (the smoothing and precautionary hypotheses), while the blue shaded area presents the (interaction) effects of the preference measures for risk aversion and prudence. That is, the grey area contains the indirect evidence for prudence through precautionary effects, and the blue area provides direct evidence through effects of directly measured prudence preferences controlling for the strength of risk aversion. The dependent variable in each specification is the amount a person saves in period 1. The unit of observation is the amount of saving of one individual in a single decision problem. Under expected utility, the difference between period 1 and period 2's expected income predicts saving for consumption smoothing purposes on the part of risk averse agents. The greater the income difference, the greater the saving. The income difference is indicated by the variable *Inc1 – Inc2*, and if the representative individual is risk averse, there will be a positive coefficient on this variable.

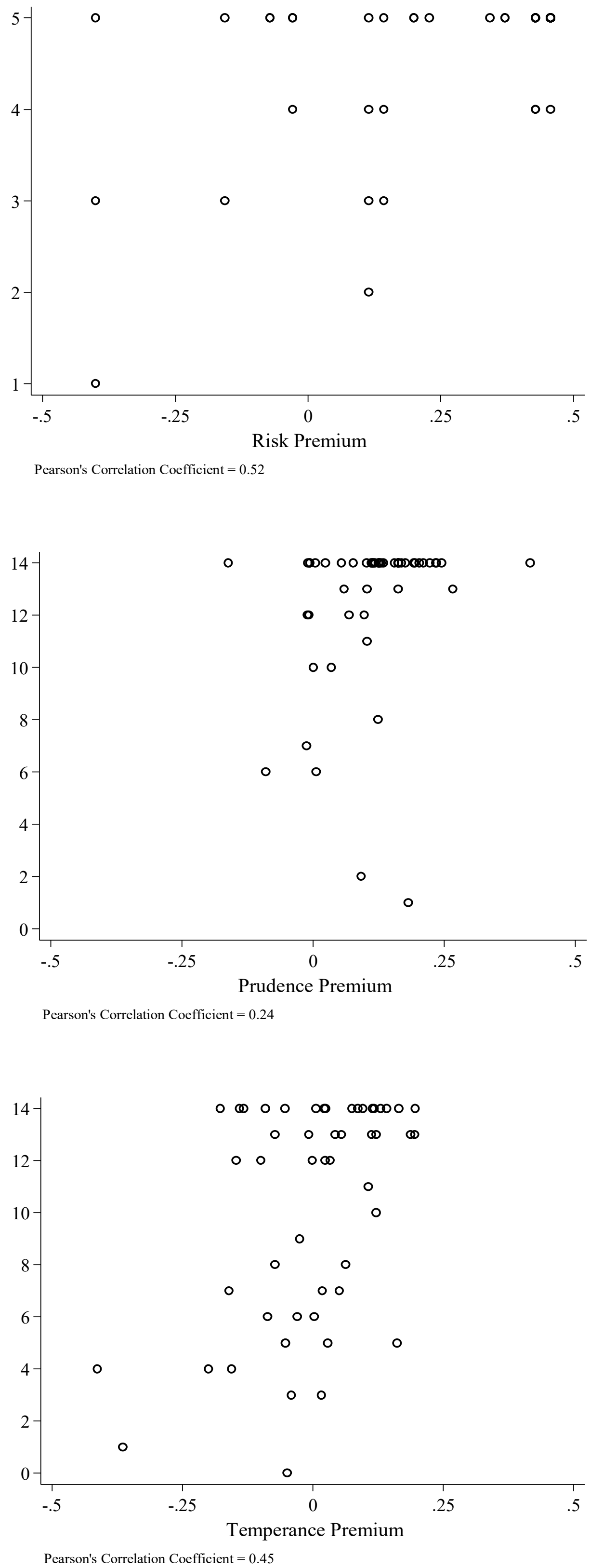

Figure 6. Distribution and Correlation for Count and Premia measures for Risk Aversion, Prudence and Temperance (each individual is the unit of observation).

Note: For risk aversion, there is substantial overlap of circles for participants making the same choices.

If people are prudent, the degree of riskiness of period-2 income will encourage additional precautionary saving. More prudent agents save more. Specifications (1) and (2) examine the predictive power of the prudence premium measured by the titration method, while specifications (3) and (4) test the predictive power of the prudent preference measured by binary choice count method. Specifications (1) and (3) use the period-2 risk level dummies *D.P2Risk2, …, D.P2Risk6* to capture the level of future income risk in increasing order (from 54_36 to 70_20, with *D.P2Risk1* for 50_40 being the excluded category). If individuals save in a prudent manner, the coefficients on the five variables should be positive and increasing (smallest for *D.P2Risk2* and largest for *D.P2Risk6*). Specifications (2) and (4) use the variable *Degree_P2risk*, which equals the difference between the possible higher and lower payoffs in period 2 (e.g., for the pair 70_20, *Degree_P2risk* = 50). If the representative individual behaves in a prudent manner, the coefficient on *Degree_P2risk* should be positive.

The coefficients in the second, blue-shaded part of the table reveal whether the Premium and Count measures predict precautionary saving in the manner that they would if they were measuring prudence attitudes accurately. The interaction terms *D.P2Risk[j]* x *PrudPrem*, with $j = \{1,\ldots,6\}$, are interaction terms between the prudence premium measured in the abstract titration task and a dummy variable for the period-2 risk level in a particular trial. If more prudent individuals, as measured in the premium task, save in a more prudent manner, the interaction terms should all be positive. The terms *D.P2Risk[j]* x *PrudCount* are interpreted analogously for the Count measure of prudence. We control for second order risk attitudes, age, and gender in all specifications.

The results provide strong evidence that the representative individual behaves in a prudent manner in the saving tasks and that behavior is related to decisions in the two prudence measurement tasks. First, we find clear evidence for consumption smoothing, with the coefficient on *Inc1 – Inc2* positive and significant in all specifications. Second, specifications (1) and (3) show that greater levels of the future income risk induce higher saving (see coefficients *D.P2risk2* to *D.P2risk6*). These categorical variables are jointly significantly different from zero, and pairwise comparisons show that the coefficients are significantly increasing with the level of period-2 income risk.[10] Specifications (2) and (4) show that the positive effect of risk on saving also significantly emerges when using the spread measure *Degree_P2risk.* These findings provide indirect evidence for prudence: behavior in the savings task is consistent with prudent preferences. The robustness of both the smoothing and precaution results also suggests that participants understood the structure of the game and made thoughtful decisions.

Third, the variable *RiskPremium* is not significant, indicating that savings are not predicted by the second order risk aversion measure. Fourth, the coefficients of the variables *D.P2Risk[j] x PrudPrem*, for $j = \{1,\ldots,6\}$ in equation (2) are all positive, and significantly positive for $j = 4 – 6$. This means that those who have a greater measured prudence premium exhibit a greater responsiveness of their saving to the level of background risk. The *D.P2Risk[j]* x *PrudCount* variables in specification (4) can be interpreted in a similar manner. They are all significantly positive, which means that those who are relatively prudent on the Count task save more on the savings task for each level of background risk. The effect of the spread-count interaction in specification (4) is positive but insignificant. However, overall, there is clear evidence that directly measured prudence is related to precautionary saving. We formulate our first three results.

Result 1 *(Consumption Smoothing)**:** We find strong evidence for consumption smoothing. The greater the difference between period 1 and period 2 income, the more people save for the future.*

Result 2 *(Precautionary Saving): We find strong indirect evidence for prudence as the riskiness of future income predicts precautionary saving. The riskier the future income is, the more people save.*

Result 3 *(Predicting Precautionary Saving): The effects of period-2 income uncertainty on savings are larger for more prudent decision-makers.*

---

[10] In specification (1), the coefficient of *D.P2risk[j]* is significantly larger than *D.P2risk[j-1]* at the 5% level for all *j*, except for the comparison between *D.P2risk4* and *D.P2risk5*. In specification (3), we find a similar increasing pattern in the estimated coefficients, though the increments are not significant at the 5% level, except for the comparison between *D.P2risk3* and *D.P2risk4*.

The results show both direct evidence (from the prudence preference tasks) and indirect evidence (future income uncertainty predicts precautionary saving) that most individuals behave prudently. We also observe that prudent behavior in direct measurement tasks is a predictor of precautionary saving.

Table 2. Predicting Precautionary Savings: Prudence Effects

Savings	(1) PrudPrem/ Risk Level	(2) PrudPrem/ Spread	(3) PrudCount/ Risk level	(4) PrudCount/ Spread
Inc1 – Inc2	0.36*** (0.03)	0.36*** (0.03)	0.36*** (0.03)	0.36*** (0.03)
D.P2risk2	1.09** (0.53)		1.60 (1.35)	
D.P2risk3	2.39*** (0.62)		2.53 (1.72)	
D.P2risk4	3.76*** (0.86)		5.06** (2.14)	
D.P2risk5	4.03*** (0.96)		4.64 (3.01)	
D.P2risk6	5.55*** (1.03)		7.88*** (1.81)	
[D.P2risk2=…=D.P2risk6=0]	p-val<0.0001		p-val<0.0001	
Degree_P2risk		0.13*** (0.02)		0.16***(0.05)
RiskPrem	-3.52 (2.18)	-3.18 (2.33)	-2.58 (2.15)	-2.15 (2.23)
Interaction Period 2 risk levels and Prud Prem:				
D.P2risk1 x PrudPrem	7.92 (6.37)			
D.P2risk2 x PrudPrem	7.47 (4.93)			
D.P2risk3 x PrudPrem	7.13 (5.08)			
D.P2risk4 x PrudPrem	9.30** (4.64)			
D.P2risk5 x PrudPrem	15.17*** (4.97)			
D.P2risk6 x PrudPrem	17.37*** (6.60)			
[D.P2risk1 x PrudPrem =… =D.P2risk2 x PrudPrem=0]	p-val=0.0012			
Interaction Period 2 risk levels and PrudCount:				
D.P2risk1 x PrudCount			0.44**(0.21)	
D.P2risk2 x PrudCount			0.40** (0.16)	
D.P2risk3 x PrudCount			0.42*** (0.13)	
D.P2risk4 x PrudCount			0.35*** (0.12)	
D.P2risk5 x PrudCount			0.46*** (0.17)	
D.P2risk6 x PrudCount			0.34** (0.15)	
[D.P2risk1 x PrudCount =… D.P2risk2 x PrudCount=0]			p-val=0.0466	
Degree P2risk x PrudPrem		0.27** (0.12)		
Degree_P2risk x PrudCount				0.00024 (0.0041)
Constant	-0.37 (2.17)	-1.11 (1.72)	-5.64 (3.47)	-0.53 (1.97)
# Observations	2,448	2,448	2,448	2,448
# Participants (clusters)	51	51	51	51

Notes: RiskPrem and *PrudPrem* in columns (1) and (2) are elicited risk preferences using the titration method. Risk Premium (RiskPrem) reflects the degree of risk aversion and Prudence Premium (PrudPrem) the degree of prudence, respectively. *Inc1 - Inc2* is the difference in expected value between period 1 and period 2 income. *D.P2risk[j]*are categorical variables measuring the level of Period 2 risk, *j={1,...,6}*. In column (3) and (4), we report the *PrudCount* variable, which is the number of prudent lotteries chosen when making 14 binary choices. It is a measure of the degree of prudence. *Degree_P2risk* is a monotonic measure of period 2 income uncertainty (spread of the period 2 income). We also control for age and gender but we do not find any effects. Entries are estimated coefficients from random effects panel regressions, with robust standard errors clustered at the individual level reported in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 3. Predicting Investment: Temperance Effects

Investment weight on Safe Asset	(1) TempPrem/ BGR Level	(2) TempPrem/ BGR Degree	(3) TempCount/ BGR Level	(4) TempCount/ BGR Degree
Risky Option Spread	1.02*** (0.10)	1.02*** (0.10)	1.02*** (0.10)	1.02*** (0.10)
Safe Option Value	1.76*** (0.20)	1.76*** (0.20)	1.76*** (0.20)	1.76*** (0.20)
D.BGR1	6.31*** (1.16)		5.98* (3.44)	
D.BGR2	11.50*** (1.66)		11.79** (4.60)	
D.BGR3	18.60*** (2.28)		22.67*** (6.01)	
D.BGR4	18.39*** (2.42)		24.87*** (5.81)	
[D.BGR1=...=D.BGR4=0]	p-val<0.0001		p-val<0.0001	
Degree_BGR		0.50*** (0.06)		0.65*** (0.15)
RiskPrem	23.17***(7.60)	23.22***(7.50)	22.55***(7.17)	22.95***(7.67)
Interaction BGR levels and Temp Prem:				
D.BGR0 x TempPrem	0.89 (19.82)			
D.BGR1 x TempPrem	-5.24 (16.81)			
D.BGR2 x TempPrem	-0.74 (15.49)			
D.BGR3 x TempPrem	-2.16 (15.15)			
D.BGR4 x TempPrem	-13.00 (13.37)			
[D.BGR0 x TempPrem =… =D.BGR4 x TempPrem=0]	p-val=0.40			
Interaction BGR levels and Temp Count:				
D.BGR0 x TempCount			0.72 (0.72)	
D.BGR1 x TempCount			0.75 (0.69)	
D.BGR2 x TempCount			0.69 (0.59)	
D.BGR3 x TempCount			0.32 (0.60)	
D.BGR4 x TempCount			0.09 (0.57)	
[D.BGR0 x TempCount =… =D.BGR4 x TempCount=0]			p-val=0.33	
Degree BGR x TempPrem		-0.22 (0.47)		
Degree BGR x TempCount				-0.01 (0.01)
Constant	-56.22*** (10.39)	-55.38*** (10.03)	-61.95*** (12.78)	-55.58*** (10.13)
# Observations	4,895	4,895	4,895	4,895
# Participants (clusters)	51	51	51	51

Notes: *RiskPrem* and *TempPrem* in columns (1) and (2) are elicited risk preferences using the titration method. They are short for Risk Premium, a measure of the degree of risk aversion, and Temperance Premium, a measure of the degree of temperance, respectively. *Risky Option Spread* is the difference between the possible higher and lower payoffs of the risky option in the investment task. *Safe Option Value* is the sure amount of money in the risk-free option. *D.BGR0,…,D.BGR4* are categorical variables measuring the level of background risk. *D.BGR0* indicates scenarios without background risk. In columns (3) and (4), we include the *TempCount* variable, which is the number of temperate lotteries chosen when making 14 binary choices measuring the degree of temperance. *Degree_BGR* is a monotonic measure of background risk, and is defined as the spread of the baseline income. We also control for age and gender. We find a strong gender effect that females invest more in the safe asset. Entries are estimated coefficients from random effects panel regressions, with robust standard errors clustered at the individual level reported in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

5.3. Investment Decisions and Background Risk

We next study the investment decisions under different levels of background risk. As with the savings task, we consider (i) whether behavior in the investment task implies temperance, and (ii) whether our abstract tasks explain temperate behavior in the sense that those who are relatively temperate in the abstract measurement tasks behave more temperately in the investment task. We again separately examine the predictive power of the temperance premium (titration method) and the number of temperate choices (binary choice count method). More precisely, we test if there are any interaction effects between the riskiness of the baseline income and the measured temperance preference on investment choices. We control for the value of the safe investment option relative to the risky option, the spread between high and low payoffs under the risky option, second order risk attitude, age, and gender in the regression models. Across all investment decisions, participants on average put approximately 60% weight on the safe asset, which is consistent with the representative participants being predominantly risk averse in the lottery task. Similar to the saving task, the within-person variation substantially contributes to the identification of effects in our regressions, with each of the 51 participants making 96 investment decisions, leading to 4896 decisions in 51 clusters.

The results are in Table 3. Specifications (1) and (2) examine the predictive power of the temperance premium, while specifications (3) and (4) test the predictive power of the temperance binary choice count. In specifications (1) and (3), the dummy variables *D.BGR[j]* for $j = \{0,\ldots,4\}$ are used to represent different levels of background risk. Specifically, the excluded category *D.BGR0* corresponds to scenarios with no background risk (i.e., a guaranteed payment), while *D.BGR1* through *D.BGR4* represent increasing levels of background risk, with risk spreads ranging from 10 to 40 euros. Specifications (2) and (4) use the continuous background risk spread measure *Degree_BGR*. The table again contains two sections: the grey shaded area shows the effects of the parameters of the investment options, i.e., the Safe Option Value, Risky

Option Spread, and the background risk measures on investment. The blue shaded area presents the interaction effects of the preference measures for risk aversion and temperance.

Table 3 reveals the following patterns. First, a relatively more valuable safe option as well as a higher-spread risky option lead to more weight on the safe option. This is consistent with monotonicity and risk aversion, that is, increasing and concave utility. It suggests that participants understood the structure of the game and made thoughtful decisions. Second, we find strong indirect evidence for temperance in the form of tempering effects both for the level of background risk in specifications (1) and (3) and the spread in specifications (2) and (4). In both specifications, the magnitudes of the coefficients increase as the background risk increases. In other words, higher background risk consistently and significantly leads to more investment in the safe asset.[11] Because of the necessity of temperance for risk vulnerability, the representative individual in our data clearly invests as if she were temperate.

Result 4 *(Tempering Effects): The level of background risk is positively associated with the proportion of the investment allocated to the safe asset.*

Looking at the effect of the preference measures, we find that the variable *RiskPrem* is positive and significant in all four regressions. This means that those who are more second-order risk averse in our abstract task also behave in a more risk averse manner in the investment decision. However, neither of the two abstract temperance measures significantly interacts with the background risk in influencing investment choices. For specifications (2) and (4), the coefficients on *Degree_BGR* x *TempPrem* and *Degree_BGR* x *TempCount* are not significantly different from zero, while for specifications (1) and (3), the coefficients for the risk level j are neither individually nor jointly significant. Temperance, as measured in our abstract tasks, does not predict

[11] For both specifications (1) and (3), pairwise comparisons between background risk dummies shows that D.BGR[j] is significantly greater than D.BGR[j-1] for all pairs, except for the comparison between D.BGR3 and D.BGR4.

temperate investment. The finding is robust to several robustness checks[12] and is summarized as Result 5.

Result 5 *(Predicting Investment Decisions with Temperance Measures): Individuals' degree of temperance does not aid in explaining investment decisions at the individual level.*

6. Discussion

Although risk aversion is the most widely studied risk preference in the literature, economic decisions also hinge in important ways on investors' higher order risk attitudes. An increase in the uncertainty of future income induces larger savings if an individual is prudent (Leland 1968; Sandmo, 1970; Kimball, 1990; Eeckhoudt and Schlesinger, 2008). Investors become more risk averse in the presence of background risk (Kimball, 1992, 1993; Eeckhoudt et al., 1996; Gollier and Pratt, 1996) if they are temperate.

Empirical contributions have put these theoretical relationships to the test (e.g., Guiso et al., 1992, 1996; Guiso and Paiella, 2008; Fuchs-Schündeln and Schündeln, 2015; Fermand et al., 2024). However, such empirical assessments are challenging due to self-selection into risk environments (e.g., occupations), the unobservability of the degree of background risk people are exposed to, and, last but not least, the unobservability of the relevant risk attitudes. We sidestep these difficulties by studying saving and investment decisions in a controlled laboratory setting that allows us both to exogenously and precisely vary background risk for saving and investment decisions, and to directly obtain independent measures of higher order risk attitudes. Risk attitudes

[12] The results do not change if we restrict the analyses to only strongly temperate (top 25%) and strongly intemperate individuals (bottom 25%). The null effect cannot be attributed to lack of variation for temperance either. Although people are on average neither temperate nor intemperate, the standard deviation of the elicited temperance premia (13%) across individuals is similar to the standard deviation of the prudence premia (10%). Additionally, given the results by Crainich et al. (2013) and Deck and Schlesinger (2014) that risk lovers can be intemperate, we also run the specifications in Table 3 on a subsample of risk averters only, still replicating the results shown in Table 3.

are measured in two different ways, which allows us to gauge the internal validity of the constructs (i.e., whether the two methods are correlated) as well as their external validity (whether they predict behavior in economic decisions).

Our analysis shows that the widely used count measures of higher order risk attitudes are significantly and substantially correlated with a titration-based risk-premium measure. This correlation is supportive of the strength of preference interpretation of the binary choice count measures of prudence and temperance employed in the literature, and suggests that the internal validity of these measures is high. As in previous studies, we find that people are predominantly risk averse and prudent. We also observe that the representative individual is neither temperate nor intemperate, a result consistent with previous studies (Deck and Schlesinger, 2010; Ebert and Wiesen, 2011; Noussair et al., 2014, Deck and Schlesinger, 2014; Baillon et al., 2018; Haering et al., 2017; Schneider and Sutter, 2020; and Schneider et al., 2022; see also the review by Trautmann and van de Kuilen, 2018). Our results suggest that both count and premium measures of preferences are suitable for empirical applications. The count measure, in particular, provides a simple measure that can readily be included in surveys of the general population (e.g., Noussair et al., 2014). However, the premium measure exhibits a clearer link with precautionary behavior as discussed next, and may also provide better precision.

The saving task provides strong indirect evidence of prudence. Controlled variation in background risk leads to an increase in precautionary saving under greater future income risk. The saving task also provides direct evidence that the channel for differences in precautionary saving is through prudence. The strength of individual-level prudence clearly moderates the effect of background risk on savings behavior in the predicted direction. The effect emerges for both abstract measures of prudence, but more strongly for the premium measure. In addition to the effect of prudence, we observe a strong tendency for consumption smoothing. Our first three hypotheses are supported. Our results corroborate Noussair et al. (2014), Schneider and Sutter (2020), and Schneider et al. (2022) who also find that more prudent people save more, either using a hypothetical savings question (Schneider and Sutter, 2020) or relating individual prudence levels to their real-life saving behavior based on survey answers (Noussair et al., 2014; Schneider et al., 2022).

In the investment task, participants behave in a clearly temperate manner, in that they take less risk when the background risk they face increases. This result is consistent with previous work by Beaud and Willinger (2020) and Strobl (2022), who also find evidence for risk vulnerability in experimental investment tasks. That is, there is robust evidence for a tempering effect of background risk on risky investment, implying that decision makers are temperate. In contrast to previous studies, we can directly observe whether the decision maker's preferences satisfy temperance. Contrary to the effect of prudence on precautionary saving, we fail to find any relationship between measured temperance on the lottery tasks and investment behavior. This result is robust to several robustness checks. Therefore, Hypothesis 4 is supported in the data while Hypothesis 5 is not.

Several explanations for this result are worth discussing. First, the measure of temperance might be too noisy, because the abstract risk preference measures become more complex as the order of the risk attitude increases (Deck and Schlesinger, 2014). However, this is unlikely to be the main driver of our result: the standard deviation of the elicited temperance premia across individuals is similar to that of the prudence premia (13% vs. 10%), and restricting our sample to those participants with very clear temperance preferences does not affect the result.

Second, on a theoretical level, the tempering effect of background risk requires a set of well-behaved properties of the utility function to induce people to be more risk averse, while the requirements are less strong for prudence. For instance, Gollier and Pratt (1996) show that individuals become more risk averse for every addition of a small and unfair (with $EV \leq 0$) background risk if and only if both absolute prudence and absolute temperance are larger than absolute risk aversion. These conditions require that both $u''' > 0$ and $u'''' < 0$, among other relationships. The sufficient conditions are even stronger for the addition of a greater independent background risk (section 4, Gollier and Pratt, 1996; for conditions of a more general effect of deteriorating background wealth on risk aversion, see Eeckhoudt et al., 1996). Thus, individuals may be temperate in the sense of Kimball (1992), but the set of conditions that induce more risk aversion may not be fully met. However, this explanation is at odds with the clear finding of tempering effects in the investment task.

Third, the behavioral measure of temperance that we employ may not identify the relevant attitudes sharply enough. That is, while for prudence there may be a close mapping from the risk allocation tasks to measure prudence to the underlying unobserved preferences, for the Eeckhoudt and Schlesinger (2006) measure of temperance the decisions in the risk allocation tasks may be distorted by other, behavioral aspects of the task. For instance, for some individuals, the two aggregated zero-mean risk lotteries may not be perceived as carrying high kurtosis, which temperate individuals dislike. Rather, it might be misperceived as a way of diversification, in the spirit of the famous Samuelson (1963) paradox. This interpretation would be consistent with the relatively mixed findings in the literature for temperance.

Despite the negative result for the behavioral temperance measure, there is overall strong support for the precautionary effects predicted within the expected utility context, and for the mechanism via prudence preference in the case of saving. How does this evidence square with the general view that expected utility is often violated by decision makers? Eeckhoudt et al. (2020) show that under Dual Theory (i.e., probability weighting), there exists no weighting function that implies consistent allocation preferences in our abstract tasks. Dhami et al. (2025) show that for Köszegi-Rabin loss averse preferences, there is neither a clear link between increases in background risk and precautionary saving, nor between the sign of the third derivative of the consumption utility and precautionary saving. That is, under non-expected utility, there are no clear predictions for our preference measurement and economic decision tasks, and our results neither support nor reject these models. Our results thus suggest that either violations of expected utility (for example, probability distortion) are not severe enough to distort the effects of the shape of the utility function. Alternatively, parametrizations of non-expected utility need to be specified that imply precautionary saving and tempered investment, and that can also rationalize the link to behavior in the prudence measurement task.

References

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

Attema, A. E., l'Haridon, O., and van de Kuilen, G. (2019). Measuring multivariate risk preferences in the health domain. *Journal of Health Economics*, 64, 15-24.

Bacon, P.M., Conte, A. and Moffatt, P.G. (2020). A test of risk vulnerability in the wider population. *Theory and Decision, 88,* 37–50.

Baillon, A., Schlesinger, H., and van de Kuilen, G. (2018). Measuring higher order ambiguity preferences. *Experimental Economics*, *21*(2), 233–256.

Beauchamp, J. P., Cesarini, D., and Johannesson, M. (2017). The psychometric and empirical properties of measures of risk preferences. *Journal of Risk and Uncertainty*, *54*(3), 203-237.

Beaud, M. and Willinger, M. (2020). Are people risk vulnerable? *Management Science 61,* 624–636.

Bostian, A. J. A., and Heinzel, C. (2012). Prudential saving: Evidence from a laboratory experiment. Working Paper.

Brunette, M., and Jacob, J. (2019). Risk aversion, prudence and temperance: An experiment in gain and loss. *Research in Economics*, *73*(2), 174-189.

Caballé, J., and Pomansky, A. (1996). Mixed risk aversion. *Journal of Economic Theory*, *71*(2), 485-513.

Castillo, M., Petrie, R., Cotla, C. R., and Torero, M. (2018). Risk preferences and decision quality of the poor. Working Paper.

Charness, G., Garcia, T., Offerman, T., and Villeval, M. C. (2020). Do measures of risk attitude in the laboratory predict behavior under risk in and outside of the laboratory? *Journal of Risk and Uncertainty*, *60*(2), 99-123.

Courbage, C. and Rey, B. (2020). On temperance and risk spreading. *Theory and Decision, 88,* 527–539.

Crainich, D., Eeckhoudt, L., and Trannoy, A. (2013). Even (mixed) risk lovers are prudent. *The American Economic Review*, *103*(4), 1529–1535.

Deck, C., and Schlesinger, H. (2010). Exploring higher order risk effects. *Review of Economic Studies*, *77*(4), 1403–1420.

Deck, C., and Schlesinger, H. (2014). Consistency of higher order risk preferences. *Econometrica*, *82*(5), 1913–1943.

Dhami, S., Hajimoladarvish, N. and Georgalos, K. (2025). Precautionary Savings, Loss Aversion, and Risk: Theory and Evidence. Working paper, Leicester.

Ding, X., Hartog, J., and Sun, Y. (2010). Can we measure individual risk attitudes in a survey? IZA Discussion Paper No. 4807.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3), 522-550.

Dohmen, T. J., Quercia, S., and Willrodt, J. (2019). Willingness to take risk: The role of risk conception and optimism. SOEP papers on Multidisciplinary Panel Data Research 1026.

Ebert, S., and Wiesen, D. (2011). Testing for prudence and skewness seeking. *Management Science*, 57(7), 1334–1349.

Ebert, S., and Wiesen, D. (2014). Joint measurement of risk aversion, prudence, and temperance. *Journal of Risk and Uncertainty*, 48(3), 231-252.

Eeckhoudt, L., Gollier, C., and Schlesinger, H. (1996). Changes in background risk and risk taking behavior. *Econometrica*, 64, 683-689.

Eeckhoudt, L., Laeven, R. and Schlesinger, H. (2020). Risk apportionment: The dual story. *Journal of Economic Theory*, 185, 104971.

Eeckhoudt, L., and Schlesinger, H. (2006). Putting risk in its proper place. *American Economic Review*, *96*, 280–289.

Eeckhoudt, L., and Schlesinger, H. (2008). Changes in risk and the demand for saving. *Journal of Monetary Economics*, *55*(7), 1329-1336.

Fermand, E., Kuhnen, C. M., Li, G., and Ben-David, I. (2024). Extrapolative uncertainty and household economic behavior. *Management Science*, 70, 5607-5625.

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

Friedman, D., Isaac, R. M., James, D., and Sunder, S. (2014). Risky curves: On the empirical failure of expected utility. Routledge.

Friedman, D., Habib, S., James, D., and Williams, B. (2022). Varieties of risk preference elicitation. *Games and Economic Behavior,* 133, 58-76.

Fuchs-Schündeln, N., and Schündeln, M. (2005). Precautionary savings and self-selection: evidence from the German reunification "experiment". *The Quarterly Journal of Economics*, *120*, 1085-1120.

Gollier, C., and Pratt, J. W. (1996). Risk vulnerability and the tempering effect of background risk. *Econometrica*, 1109-1123.

Gropper, M. J., & Kuhnen, C. M. (2025). Wealth and insurance choices: Evidence from US households. *The Journal of Finance*, *80*(2), 1127-1170.

Guiso, L., Jappelli, T., and Terlizzese, D. (1992). Earnings uncertainty and precautionary saving. *Journal of Monetary Economics*, *30*(2), 307-337.

Guiso, L., Jappelli, T., and Terlizzese, D. (1996). Income risk, borrowing constraints, and portfolio choice. *The American Economic Review*, 158-172.

Guiso, L., and Paiella, M. (2008). Risk aversion, wealth, and background risk. *Journal of the European Economic Association*, *6*(6), 1109-1150.

Haering, A., Heinrich, T., and Mayrhofer, T. (2020). Exploring the consistency of higher order risk preferences. *International Economic Review*, *61*(1), 283–320.

Hardeweg, B., Menkhoff, L., and Waibel, H. (2013). Experimentally validated survey evidence on individual risk attitudes in rural Thailand. *Economic Development and Cultural Change*, *61*(4), 859-888.

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 Newman, P., Milgate, M., and Falwell, J. (eds). *The New Palgrave Dictionary of Money and Finance* (London: Macmillan).

Kimball, M. S. (1993). Standard risk aversion. *Econometrica*, 61, 589-611.

Leland, H. (1968). Saving and uncertainty: The precautionary demand for saving. *The Quarterly Journal of Economics, 82*(3), 465-473. doi:10.2307/1879518

Lugilde, A., Bande, R., and Riveiro, D. (2019). Precautionary saving: a review of the empirical literature. *Journal of Economic Surveys*, *33*(2), 481-515.

Menkhoff, L., and Sakha, S. (2017). Estimating risky behavior with multiple-item risk measures. *Journal of Economic Psychology*, *59*, 59-86.

Mussio, I., and de Oliveira, A. C. (2022). The effect of additional background risk on mixed risk behavior. *Journal of Behavioral Economics for Policy*, *6*(S1), 85-92.

Noussair, C. N., Trautmann, S. T., and Van De Kuilen, G. (2014). Higher order risk attitudes, demographics, and financial decisions. *Review of Economic Studies*, *81*(1), 325–355.

Rothschild, M., and Stiglitz, J. E. (1970). Increasing risk: I. A definition. *Journal of Economic Theory*, *2*(3), 225–243.

Samuelson, P. A. (1963). Risk and Uncertainty: A Fallacy of Large Numbers. *Scientia* 98, 108-113

Sandmo, A. (1970). The effect of uncertainty on saving decisions. *The Review of Economic Studies*, *37*(3), 353-360.

Schneider, S. O., and Sutter, M. (2020). Higher order risk preferences: New experimental measures, determinants and field behavior. *MPI Collective Goods Discussion Paper*, (2020/22).

Schneider, S. O., Ibanez, M., and Riener, G. (2022). Measuring utility: An application to higher order risk preferences. Working Paper.

Strobl, R. (2022). Background risk, insurance, and investment behavior: Experimental evidence from Kenia. *Journal of Economic Behavior and Organization, 202*, 34-68.

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

Wakker, P. P. (2008). Explaining the characteristics of the power (CRRA) utility family. *Health Economics, 17*, 1329-1344.

Appendix

A.1. List of lotteries

Table A1. The list of lotteries

	Lottery L	Lottery R
RA 1	20	[65_5]
RA 2	25	[65_5]
RA 3	30	[65_5]
RA 4	35	[65_5]
RA 5	40	[65_5]
Prud 1	[(48+[10_-10])_32]	[48_(32+[10_-10])]
Prud 2	[(48+[5_-5])_32]	[48_(32+[5_-5])]
Prud 3	[(48+[15_-15])_32]	[48_(32+[15_-15])]
Prud 4	[(52+[10_-10)_28]	[52_(28+[10_-10])]
Prud 5	[(56+[10_-10])_24]	[56_(24+[10_-10])]
Prud 6	[(56+[6_-6])_24]	[56_(24+[6_-6])]
Prud 7	[(40+[10_-10])_24]	[40_(24+[10_-10])]
Prud 8	[(72+[10_-10])_48]	[72_(48+[10_-10])]
Prud 9	[(72+[15_-15])_48]	[72_(48+[15_-15])]
Prud 10	[(72+[20_-20])_48]	[72_(48+[20_-20])]
Prud 11	[(72+[25_-25])_48]	[72_(48+[25_-25])]
Prud 12	[(32+[5_-5])_16]	[32_(16+[5_-5])]
Prud 13	[(32+[10_-10])_21]	[32_(21+[10_-10])]
Prud 14	[(32+[5_-5])_21]	[32_(21+[5_-5])]
Temp 1	[(48+[10_-10])_(48+[10_-10])]	[48_(48+[10_-10]+[10_-10])]
Temp 2	[(48+[10_-10])_(48+[4_-4])]	[48_(48+[10_-10]+[4_-4])]
Temp 3	[(48+[10_-10])_(48+[17_-17])]	[48_(48+[10_-10]+[17_-17])]
Temp 4	[(48+[20_-20])_(48+[5_-5])]	[48_(48+[20_-20]+[5_-5])]
Temp 5	[(39+[10_-10])_(39+[10_-10])]	[39_(39+[10_-10]+[10_-10])]
Temp 6	[(39+[10_-10])_(39+[5_-5])]	[39_(39+[10_-10]+[5_-5])]
Temp 7	[(39+[9_-9])_(39+[12_-12])]	[39_(39+[9_-9]+[12_-12])]
Temp 8	[(39+[15_-15])_(39+[5_-5])]	[39_(39+[15_-15]+[5_-5])]
Temp 9	[(72+[15_-15])_(72+[15_-15])]	[72_(72+[15_-15]+[15_-15])]
Temp 10	[(72+[10_-10])_(72+[10_-10])]	[72_(72+[10_-10]+[10_-10])]
Temp 11	[(72+[10_-10])_(72+[25_-25])]	[72_(72+[10_-10]+[25_-25])]
Temp 12	[(72+[10_-10])_(72+[30_-30)]	[72_(72+[10_-10]+[30_-30])]
Temp 13	[(72+[20_-20])_(72+[20_-20])]	[72_(72+[20_-20]+[20_-20])]
Temp 14	[(48+[35_-35])_(48+[5_-5])]	[48_(48+[35_-35]+[5_-5])]

Notes: Each row represents a pairwise lottery choice task. [x_y] indicates an equiprobable lottery to receive either x or y; Choice of the Lottery L indicates risk aversion, prudence, or temperance.

A.2. Example of titration steps and premium calculation

Suppose that a decision-maker (DM) is confronted with the lottery pair Prud 1 in Table A1. We use Prud 1 as an example to demonstrate the titration process. Suppose that for the first two choices, the DM chooses Lottery L, the prudent option. For the third choice, she chooses Lottery R, the imprudent option. Since the last choice of the DM is used to determine the indifference lottery, we discuss the situations where the DM chooses Lottery L and Lottery R for the last choice separately. Her decisions are summarized in the last column in Table A2.

Table A2. Example of choices and titration steps

Titration Step	Lottery L	Lottery R	Decision
0 (original lottery)	[(48+[10_-10])_32]	[48_(32+[10_-10])]	L
1	[(48+[10_-10])_32]	[62_(42+[10_-10])]	L
2	[(48+[10_-10])_32]	[73_(48+[10_-10])]	R
3	[(48+[10_-10])_32]	[68_(44+[10_-10])]	If L: Final EV = 58
			If R: Final EV = 56

In this example, the DM first chooses Lottery L. To find the point of indifference, in the first step of the titration, the expected value (EV) of Lottery R is increased by 30% (see step 1 in Figure 3). This is done by adjusting the Lottery R's binary outcomes of either €48 or €32 to either €62 (=48*1.3, rounded) or €42 (=32*1.3, rounded). Thus, for the second choice, the decision maker is choosing between the original Lottery L and the adjusted Lottery R (see the third row of Table A2, Titration Step 1).

For the second choice, the DM chooses Lottery L again. Hence, the titration procedure further increases the binary outcomes of Lottery R by 15% (see step 2 in Figure 3). The new outcomes are either €73 (=62*1.15, rounded) or €48 (=42*1.15, rounded). The adjustment is made relative to the new baseline payoffs, not to the initial values. The decision maker now makes a choice between Lottery L and the twice-adjusted Lottery R (see the fourth row of Table A2).

Suppose that the DM chooses Lottery R for the third choice. Following the titration procedure, the EV of the chosen payoffs will be reduced by 7.5% to make it less attractive for the next choice, to find the point of indifference. Thus, the binary outcomes of Lottery R will be decreased by 7.5% to €68 (=73*0.925, rounded) and €44 (=48*0.925, rounded) respectively. This triple-adjusted Lottery R and the original Lottery L constitute the last choice faced by the DM.

After the fourth choice, we determine what the indifference lottery is, for which a DM is indifferent between the original Lottery L and the adjusted Lottery R. Suppose that the DM chooses Lottery R again in her fourth choice, the same as the previous choice. In this case, we will take Lottery R in the last row of Table A2 as the indifference lottery. That is, we do not extrapolate outside the range of values offered, as it is unclear how strongly we should do so. The EV of this triple-adjusted Lottery R equals €56 (= $\frac{1}{2}(€68 + €44)$). Comparison with the EV of the original Lottery R of €40 yields a prudence premium of 35% $= \frac{56-40}{40} * 100\%$ for Prud 1.

Suppose that the DM chooses Lottery L in her fourth choice instead. This means that the adjustment made for Lottery R in the previous step (a reduction of 7.5% in EV) was too strong, the true indifferent value lies between step-2 and step-3 lotteries. We choose the midpoint as the indifference value. In this example, the indifference lottery's high-state and low-state payoffs are the midpoint between €73 and €68 and the midpoint between €48 and €44. Thus, the EV of the indifference lottery is €58 (= $\frac{1}{2}[\frac{1}{2}(73 + 68) + \frac{1}{2}(48 + 44)]$) . Comparison with the original Lottery R yields a prudence premium of 45% $= \frac{58-40}{40} * 100\%$.

A.3. Example of calculations of precautionary saving and tempered investment

To illustrate the predictions following from the theoretical literature, we provide example calculations using our experimental parameters. For the utility function, we assume the widely used CRRA family as defined in Wakker (2008): $U(x) = x^r$ for $r > 0$ and $U(x) = -x^r$ for $r < 0$. We present calculations for the coefficients of risk aversion $r = 0.9$ (low risk aversion), $r = 0.1$ (medium risk aversion) and $r =$ –0.9 (high risk aversion). For this family, stronger risk aversion implies stronger prudence and temperance.

For the precautionary saving task, the agent chooses the savings s to maximize her intertemporal utility $U(s) = (y_1 - s)^r + 0.5\left(y_{2,high} + s\right)^r + 0.5\left(y_{2,low} + s\right)^r$, with y_1 being the period 1 income, and y_2 being the uncertain period 2 income.[13] Table A3 presents calculations of optimal savings for a low (33) vs. high (54) level of period 1 income, and a low (50_40) vs. high (70_20) level of period 2 background risk, for the three degrees of prudence. Note that the expected period 2 income is 45, implying a preference for borrowing in the low period 1 income and saving in the high period 1 income situation due to income smoothing.

Table A3. Precautionary Saving Example Calculations

Period-1 income	Period-2 background risk	r = 0.9	r = 0.1	r = –0.9
33 (low)	50_40 (low)	–5.82	–5.70	–5. 54
	70_20 (high)	–1.65	+0.59	+2.41
54 (high)	50_40 (low)	+4.64	+4.74	+4.86
	70_20 (high)	+7.94	+9.95	+11.78

Note: Entries show the amount of saving (+) or borrowing (–)

Table A3 illustrates the pattern we aim to test experimentally. First, we observe clear smoothing with consistent saving for the high period 1 income of 54, but both borrowing and saving for the low period 1 income of 33. Second, a precautionary effect is obtained, with more savings (or less borrowings) for higher prudence, and for larger background risk.

For the investment task, the agent chooses the portfolio share α of the safe asset to maximize her utility $U(\alpha) = 0.25(\beta_{low} + 30\alpha + \rho_{low}(1-\alpha))^r + 0.25\left(\beta_{low} + 30\alpha + \rho_{high}(1-\alpha)\right)^r + 0.25\left(\beta_{high} + 30\alpha + \rho_{low}(1-\alpha)\right)^r + 0.25\left(\beta_{high} + 30\alpha + \rho_{high}(1-\alpha)\right)^r$, with β_{low} and β_{high} the low and high outcomes of the background risk, and ρ_{low} and ρ_{high} the low and high outcomes of the risky investment.[14] Calculations are made for a safe asset with a payoff of 30.

[13] This holds for $r > 0$. For $r < 0$ intertemporal utility becomes $U(s) = -(y_1 - s)^r + (-0.5)\left(y_{2,high} + s\right)^r + (-0.5)\left(y_{2,low} + s\right)^r$.

[14] This holds for $r > 0$. For $r < 0$ intertemporal utility needs to be adjusted accordingly.

Table A4. Tempered Investment Example Calculations

Investment Risk	Background risk	r = 0.9	r = 0.1	r = –0.9
50_20 (low)	15_15 (low)	0	0	0.42
	40_0 (high)	0	0	0.57
65_5 (high)	15_15 (low)	0	0.71	0.87
	40_0 (high)	0	0.77	0.90

Note: Entries show the share on the safe asset (which has a payoff of 30)

Table A4 illustrates the pattern we aim to test experimentally. First, we find safer portfolios for stronger risk aversion (smaller r values), and safer portfolios if the risky asset is riskier. Moreover, for many parameter values, the agent would fully invest in the risky asset. Second, the tempering effect of background risk emerges if risks and risk aversion are sufficiently large to induce investment in the safe asset.