Measuring Utility: An Application to Higher Order Risk Preferences

Sebastian O. Schneider
Max Planck Institute for Research on Collective Goods, Bonn,
sschneider@coll.mpg.de

Marcela Ibanez
University of Göttingen, mibanez@gwdg.de

Gerhard Riener
University of Düsseldorf, riener@dice.hhu.de

April 6, 2023

Abstract

We present the first experimental method to non-parametrically elicit utility functions and associated measures of (higher order) risk preferences. The method yields well-known theoretically-derived utility-based measures of intensities, such as the Arrow-Pratt measure for risk aversion, and analoguous measures for prudence and temperance. Unlike parametric alternatives, the method is free of assumptions about the shape of the utility function, and particularly of the commonly made but provenly inappropriate ones for the study of higher order risk preferences. The method's ability to account for decision errors is illustrated in a simulation exercise, where it performs comparable to parametric fitting techniques. In accompanying laboratory and online experiments, we validate our method and find significant relations to other methods. Finally, we apply our method in a sample of the poor population in Bogotá, Colombia to test the precautionary saving model by Leland (1968). We find strong support for the model by showing that income risk is associated with increases in savings for prudent individuals.

JEL Classifications:

Keywords: Non-parametric utility elicitation, method, higher order risk preferences, prudence, temperance, non-parametric intensity measures, validation, experiment, precautionary saving

1 Introduction

Characterizing the utility function is a cornerstone of microeconomics and has a long tradition in empirical economics (Deaton and Muellbauer, 1980). Structural parametric approaches are typically applied to estimate key utility parameters, such as the degree of risk aversion, or the degrees of the higher order risk preferences prudence and temperance. While this approach has proven suitable for modelling aggregate behavior, widely-used forms of parametric utility functions, such as the power utility family, are not flexible enough to model behavior at the individual level (Abdellaoui, Barrios, and Wakker, 2007; Holt and Laury, 2002). For the study of higher order risk attitudes, none of the commonly used parametric utility functions is flexible enough to model individual behavior: The relation between intensities of the different orders that are implied by the expo-power function and the power function does not match empirical evidence (Noussair, Trautmann, and Kuilen, 2014), and some combinations of (higher order) risk attitudes that are repeatedly observed empirically cannot be described by these functions (e.g., Deck and Schlesinger, 2010; Ebert and Wiesen, 2011, 2014; Maier and Rüger, 2011).[1]

To overcome these liminations, we develop a non-parametric (experimental) elicitation method for utility functions that we use to compute associated intensity measures of risk preferences as applied in theoretical work: the Arrow-Pratt coefficient of risk aversion, as well as the analoguous measures for prudence and temperance. We illustrate the characteristics of our method in comparison to well-known parametric utility functions using simulations, and validate the resulting measures using (variants of) the established risk apportionment method (Eeckhoudt and Schlesinger, 2006) in laboratory and online experiments with several hundred participants. Finally, we relate our measure of prudence to income risk and household saving of a sample of 650 urban poor individuals in Bogotá, Colombia, and provide support for theoretical models of precautionary saving going back to Leland (1968), for which we provide an extension.

While risk aversion can be described as the extent to which individuals dislike mean preserving increases in risks (Rothschild and Stiglitz, 1970), prudence is the propensity to prepare and forearm oneself in the face of

[1] In this respect, already Fuchs-Schündeln and Schündeln (2005) note that specifications generally allowing for risk aversion (risk seeking behavior) rule out combination with imprudence (prudence). We derive these limitations in Appendix C for the exponential (CARA), the power (CRRA) and the expo-power utility family.

risks, i.e., the degree to which an individual reacts to risk if it enters the decision situation (Kimball, 1990). Temperance in turn is the extent to which individuals dislike additional risk in the presence of an uninsurable background risk. Both in theory and in empirical work, the relation of prudence with precautionary saving has probably gained most attention (e.g., Dynan, 1993; Kimball, 1990; Leland, 1968; Noussair, Trautmann, and Kuilen, 2014). There have been attempts to infer prudence parameters from observational saving data (e.g., Dynan, 1993), or studies relating direct individual prudence classification measures to savings (e.g., Noussair, Trautmann, and Kuilen, 2014), but so far, no study has investigated the relation of income risk, direct individual intensity measures of prudence, and saving. Therefore, we use this application to illustrate our method—developed for, and first used in this and its companion paper (Schneider and Sutter, 2021), which focuses on the empirical relevance of higher order risk preferences compared to other measures of risk preferences.

The method can take any mapping of experimentally elicited utility levels as provided by, e.g., the certainty equivalent method, the trade-off method (Wakker and Deneffe, 1996) or the lottery equivalent method (McCord and Neufville, 1986), and provides an estimate of an unconstrained smooth utility function. It builds on penalized-spline (P-spline) regression (Eilers and Marx, 1996) to obtain a parameter-free estimate of a continuous and p-times differentiable utility function including its derivatives from the same model. To this end, we extend the toolbox of P-spline regression for the needs of utility functions and incorporate value constraints, develop a solution to jointly smooth different orders of derivatives, and establish a way to determine a data-driven minimum for the smoothness parameter, which in turn is the result of an optimization using machine learning methods.

The smooth utility functions and their derivatives allow constructing measures of the intensity of (higher order) risk preferences—risk aversion, prudence, temperance and higher orders—without imposing a relationship between for example risk aversion and prudence or risk aversion and temperance as commonly applied parametric forms do. Moreover, we can measure intensities of higher order risk preferences without relying on the strong assumption that the degree of a higher order risk preferences is reflected by the consistency of choices, as often assumed using the standard risk apportionment method (Deck and Schlesinger, 2010; Eeckhoudt and Schlesinger, 2006; Noussair, Trautmann, and Kuilen, 2014).

The method we propose is the first one to experimentally elicit continuous utility functions non-parametrically and to allow the non-parametric computation of utility-based intensity measures; in particular the Arrow-Pratt measure of risk aversion (Pratt, 1964), the intensity measures of prudence by Kimball (1990) and Crainich and Eeckhoudt (2008) and the generalization of the latter measure to any order of risk preference—including temperance—by Denuit and Eeckhoudt (2010).

We validate the measures resulting from our method in an online experiment and in the laboratory. We relate the measures with the elicitation methods by Noussair, Trautmann, and Kuilen (2014), Deck and Schlesinger (2010), and Ebert and Wiesen (2014). We observe significantly correlations with the compensation premia for prudence and temperance elicited with the method by Ebert and Wiesen (2014). Moreover, we find that the number of prudent and temperant choices in the risk apportionment methods as implemented by Deck and Schlesinger (2010) and Noussair, Trautmann, and Kuilen (2014) significantly predicts our intensity measures, and it explains up to about a third of its variation. This suggests that the count measure resulting from risk apportionment choice tasks approximates higher order risk intensities to a degree that might be deemed sufficient for certain applications.

Finally, since the relation between risk-seeking behavior, prudence and savings remains unexplored, we provide an application of our method to test the precautionary savings model due to Leland (1968). Using data from a financial survey with a sample of 693 individuals living in Bogota, Colombia, we use our method to estimate intensity of higher-order risk preferences. The empirical evidence provides unique support for the investigated model (Leland, 1968). There is a strong positive correlation between savings and prudence and this relation becomes even stronger when we include an exogenous measure of income risk—the ratio of shut-down to existing businesses in 2013 in Bogotá in the sector an individual was usually employed.

Our study adds to the extended research on household savings by theoretically and empirically establishing the link between risk aversion, savings and prudence. Previous studies used survey data to determine the share of savings that is due to income risk and in that way assess the importance of prudence (Dynan, 1993; Fagereng, Guiso, and Pistaferri, 2017a; Guiso, Jappelli, and Terlizzese, 1992).However, those studies face the problem that the data does not allow for separately identifying individual risk aversion and prudence as they rely on parametric utility forms, where usually the degree of risk aversion has an implication for the degree of prudence (see Appendix B for an illustration of this point and its consequences). Our measures are free from such restrictive assumptions, and are thus ideally suited for this application.

To deal with the limitations of inferring preferences—here in particular prudence—from observed behavior another branch of studies began only relatively recently using experimental measures of risk aversion, prudence and temperance (Deck and Schlesinger, 2010; Ebert and Wiesen, 2011, 2014; Maier and Rüger, 2011; Tarazona-Gomez, 2004). While those studies make a methodological contribution in the measurement of prudence and temperance, they focus on a student population and do not examine saving decisions. A notable exception is Noussair, Trautmann, and Kuilen (2014), who investigate financial-decision making in the general Dutch population. However, they—and most of the previously mentioned studies—face the problem that the risk-apportionment tasks that they apply cannot inform about theory-based intensity measures of prudence but simply classify individuals as prudent or imprudent and similarly for temperance. We provide measures of prudence and temperance from a non-student sample in a development context, thereby extending the work of Cardenas and Carpenter (2013) to higher order risk preferences. In addition, our study provides intensity measures of higher order risk preferences, which have never been (properly) measured with an adult population outside the laboratory.

An exception among the studies measuring higher order risk preferences is also Ebert and Wiesen (2014), introducing a method that builds on risk apportionment tasks, but asks for the compensation premium that makes individuals indifferent between the prudent (temperant) option and the imprudent (intemperant) one plus the premia. We add to this literature by presenting the first method to elicit theoretically-derived utility-based intensity coefficients of (higher order) risk preferences, such as the analoguous measures to the Arrow-Pratt measure for risk aversion. Our method is easy to understand for subjects and parsimonious with respect to the number of choice tasks needed. Moreover, it yields the intensity measures used in theoretical work, and thus opens up completely novel opportunities for theory testing.

The rest of the paper is organized as follows: Section 2 gives an overview over higher order risk preferences, and Section 3 explains the non-parametric method for the elicitation and estimation of the utility functions, including the derived intensity measures. Section 4 illustrates some key characteristics

(a) The Imprudent Option (b) The Prudent Option

Figure 1. An Abstract Lottery Pair for Elicitation of Prudence

of the method in simulations, while Section 5 presents the validation study. Section 6 finally presents the application study, and Section 7 concludes the paper.

2 Theoretical Background

With the behavioral definitions of prudence, temperance and even higher orders of risk preferences introduced by Eeckhoudt and Schlesinger (2006) an easy and concise way to measure those attitudes experimentally has been established. While we rely on the classical definitions of higher order risk preferences, the behavioral definitions have evoked a renaissance of at least prudence in mainstream empirical economics and the resulting elicitation method has become the canonical method for higher order risk preferences. We briefly summarize the behavioral as well as the classical definitions and the experimental toolbox related to higher order risk preferences.

Risk Aversion

Consider the mean-zero variable ε, independent of any other variable affecting an individual's initial wealth level x, and a deterministic reduction of wealth of size k, $k > 0$. Define the lotteries $B_2 = [0]$ as receiving an amount of 0 for sure, and $A_2 = [\varepsilon]$ as receiving ε with probability 1. With that notation, an individual is risk averse if and only if A_2 is preferred over B_2 for all initial wealth levels x and for all k and all ε.

Eeckhoudt and Schlesinger (2006) show that this definition coincides with the classical definition in an expected utility framework: A negative second derivative of the von-Neumann-Morgenstern utility function, i.e. $u'' < 0$, assuming u twice differentiable, is equivalent to risk aversion. Obviously, it is also equivalent to aversion of mean-preserving spreads (Rothschild and Stiglitz, 1970), or, in the more general terminology of stochastic dominance due to Ekern (1980), an aversion to second degree risk.

The classical measure for the degree of risk aversion is the Arrow-Pratt measure $r = -u''/u'$, and Pratt (1964) shows that, when comparing two individuals, the individual with the globally higher degree of risk aversion will always have the higher risk premium, where also the reverse is true. Locally, that is, for a fixed level of wealth, and for small risks, r is proportional to the risk premium. More precisely, r is the risk premium times twice the inverse of the variance for an infinitesimal, mean-zero risk.

Prudence

As for risk-aversion, we consider two lotteries: Define $B_3 = [-k, \varepsilon]$ and $A_3 = [\varepsilon - k, 0]$, where the two entries in the brackets represent the two equiprobable outcomes of the lotteries. If for all x and all k and all ε, B_3 is preferred over A_3, the individual is defined as prudent.

In the classical expected utility characterization, this is equivalent to $u''' < 0$, i.e., marginal utility being a convex function, assuming u is sufficiently differentiable (Eeckhoudt and Schlesinger, 2006).

Already the seminal work by Leland (1968) and Sandmo (1970) linked a positive third derivative of the von Neumann-Morgenstern utility function to a demand for precautionary savings. In further developments, Kimball (1990)—who also coined the term “prudence”—introduced $-u'''/u''$ as a measure for the intensity of prudence that indicates the strength of the precautionary saving motive in a two-period model of consumption and saving.

Via the characterization of the utility function, prudence is also equivalent to downside risk aversion as defined by Menezes, Geiss, and Tressler (1980). It is moreover equivalent to third-degree risk aversion in the more general framework by Ekern (1980). Thus, the measure u'''/u', originally a measure for the degree of downside risk aversion becomes interesting, especially as it is defined independently of the second derivative of the utility function. Crainich and Eeckhoudt (2008), building on earlier work by Modica and Scarsini (2017) and Keenan and Snow (2002), advocate for this measure and show that it is proportional to both, a premium m paid in the best states of the world to compensate for the misalignment of “harms” in A_3, i.e., $\tilde{A}_3 = [\varepsilon + k, m]$, and the utility premium, i.e., the difference of utility between A_3 and B_3, in monetary equivalents.

Moreover, u'''/u' is also a measure of skewness aversion: Modica and Scarsini (2017) show that the increase in premium that is due to an increase in skewness is proportional to this measure. This is particularly interesting, as Holzmeister et al. (2019) in a large-scale survey experiment with more than 2,000 finance professionals and more than 4,000 lay people in nine countries find that skewness drives investors' risk-perception, whereas variance does not influence their perception of risk. Thus, in light of these results, if a good measure for the willingness to accept financial risks is needed, u'''/u' is likely to be a very good candidate.

In Appendix A, we show that this measure is a rationale for the intensity of the precautionary demand for saving, building on the—compared to Kimball (1990)—more general framework of saving and consumption by Leland (1968). In particular, one prediction emerging from our model is that also risk lovers may save a non-trivial fraction of their income proportional to the intensity measure of prudence u'''/u', which extends the finding by Crainich and Eeckhoudt (2008) in the direction of allowing risk lovers to save a non-trivial fraction of their income—proportional to the measure of prudence u'''/u'.

Temperance

Instead of the deterministic reduction of wealth $-k$, and in addition to the first mean-zero variable ε_1, consider a second mean-zero variable ε_2, independent of ε_1. Define the lotteries $B_4 = [\varepsilon_1, \varepsilon_2]$ and $A_4 = [0, \varepsilon_1 + \varepsilon_2]$. An individual is temperant if lottery B_4 is preferred over A_4 for all x and all ε_i, $i = 1, 2$.

As for prudence, temperance can be characterized via a negative fourth derivative of the utility function in an expected utility setting: $u^{(iv)} < 0$ equals temperance, thus temperance equals an aversion to fourth-degree risk.

A measure for temperance or equivalently for a dislike of fourth-order risk as defined by Ekern (1980) is proposed by Denuit and Eeckhoudt (2010), who extend the work by Modica and Scarsini (2017) from third-order risk to arbitrary orders. The intuition of that measure, $(-1)^{n+1}u^{(n)}/u'$, is that the premium for a risk that has more nth degree risk than another risk, which is equal in all other aspects, should be the higher one, if an agent dislikes that increase. They show that their proposed measure is proportional to the increase in premium due to the increase in nth degree risk; in particular this is of course true for fourth-order risk and the measure $-u^{iv}/u'$ is a measure for temperance and dislike of kurtosis alike.

Higher-Order Preferences

Higher order preferences of any order can be defined similarly using preferences over pairs of lotteries with an iterative construction algorithm (Eeckhoudt and Schlesinger, 2006; Eeckhoudt, Schlesinger, and Tsetlin, 2009). Deck and Schlesinger (2014) have studied fifth-order and sixth-order attitudes experimentally and note for the latter that "behavior at this order is approaching random choice" and conclude that focusing on the first four orders seems reasonable.

Elicitation of Intensities of Higher Order Risk Preferences

Whereas the elicitation method building on preferences over pairs of binary lotteries allows classifying individuals according to their risk type for each order, theoretical measures of intensity of higher order risk preferences cannot be inferred from these choices without relying on parametric estimation. Following earlier work in the risk literature, Deck and Schlesinger (2010) and Noussair, Trautmann, and Kuilen (2014) argue that decision errors are less likely the stronger the attitude is pronounced and interpret the number of prudent (temperant) choices in above defined lotteries as measure of strength. Yet, little is known whether this is a reasonable approximation, in particular also because few real-world implications of higher order risk preferences have been studied empirically, with the notable exception of Noussair, Trautmann, and Kuilen (2014).

Mirroring the idea of the risk premium, which may be used as a straight forward measure of risk aversion, Ebert and Wiesen (2014) introduce a method based on risk apportioning to elicit intensities via premia—the premia that individuals ask for in order to accept the imprudent or the intemperant option. While this method—originally implemented with 5 choice lists of 20 comparisons each—allows the elicitation of prudence and temperance intensities, utility-based coefficients such as the Arrow-Pratt coefficient of risk aversion, the intensity measures of prudence by Kimball (1990) or the one advocated for by Crainich and Eeckhoudt (2008) or the coefficient of temperance by Denuit and Eeckhoudt (2010) cannot be derived, unless again relying on parametric estimation.

Mixed Risk Averters And Precautionary Saving

Building on work by Eeckhoudt, Schlesinger, and Tsetlin (2009) on "mixed risk averters" who like to combine good with bad outcomes, Crainich, Eeckhoudt, and Trannoy (2013) elegantly demonstrated the equivalence between a preference for combining good with good outcomes and a positive second, third and fourth derivative, corresponding to risk loving, prudent and intemperant behavior; this pattern they attribute to "mixed risk lovers" who agree with mixed risk averters on the sign of uneven derivatives, but disagree on even derivatives. Using a simple two period-model assuming time-separable utility, they show that prudent risk lovers devote all their income to saving in the presence of a future income risk. Empirical support for the existence of mixed risk averters and mixed risk lovers is given by Noussair, Trautmann, and Kuilen (2014), who study the prevalence of risk aversion, prudence and temperance in a student sample and the general population in the Netherlands. They find about 15 percent of their sample showing risk-loving behavior, and that "prudence is more prevalent than temperance".

What is more, they find that "[t]he degree of temperance seems to be more closely related to the degree of risk aversion than prudence is". In their student sample, the rank correlation between prudence and risk aversion is not significant (and negative), while the rank correlation between risk aversion and temperance is significant and positive.[2] Although they pool risk-averse and risk-loving subjects in their analysis, they find a strong connection between prudence and saving and wealth, which is robust to controlling for risk aversion levels and temperance, suggesting that this connection also holds for the 15 percent showing risk-loving behavior. However, results are not presented for the risk-loving and the risk-averse sub-samples separately.

When considering risk lovers while investigating the relationship between prudence and saving, the intensity measure of prudence put forward by Kimball (1990), $-u'''/u''$ is inapplicable; instead, u'''/u', the measure advocated for by Crainich and Eeckhoudt (2008) has to be applied.

3 Methodology

In this section, we introduce the procedure for the joint elicitation of (higher order) risk preferences. To this end, we combine well-established methods to elicit utility points with the statistical approach of P-spline regression. The experimental elicitation and estimation procedure consists of the following three steps:

[2] When pooling student sample and the general population, both rank correlations are positive and significant, however, the correlation between risk aversion and prudence is still lower than that between risk aversion and temperance. Moreover, temperance significantly increases with risk aversion, which is not the case for prudence.

1. Elicitation of utility points using any suitable method, such as the certainty equivalent method or the trade-off method (Wakker and Deneffe, 1996).
2. Estimation of differentiable, individual utility functions and their derivatives based on elicited utility points using penalized spline (P-spline) regression approach (Eilers and Marx, 1996) tailored to our needs.
3. Derivation of higher order risk preference (intensity) measures based on differentiable utility functions at the individual level.

For its non-parametric character, this procedure can be seen as the natural completion of non-parametric elicitation methods for utility *points* to a non-parametric elicitation method for utility *functions*.

3.1 Elicitation of Utility Points

In the expected utility (EU) framework, one established method to non-parametrically elicit utility points is the trade-off method (Wakker and Deneffe, 1996), which we will use to illustrate our procedure.[3] The method elicits payoffs x_i that imply indifference between two outcome gambles. Denote a binary lottery with $(x, p; y)$, where x is the upside, occurring with probability $p > 0$ and y is the downside with corresponding probability $(1 - p)$. The participant first states the value x_1 that makes her indifferent between $(x_1, p; r)$ and $(x_0, p; R)$ where $x_1 > x_0$ and $R > r$. Then the participant is asked for the value $x_2 > x_1$ that makes her indifferent between $(x_2, p; r)$ and $(x_1, p; R)$. Assuming EU, and denoting the utility[4] of a monetary outcome x with $U(x)$, these two indifferences imply that:

$$\begin{aligned} pU(x_1) + (1-p)U(r) &= pU(x_0) + (1-p)U(R) \quad \text{and} \\ pU(x_2) + (1-p)U(r) &= pU(x_1) + (1-p)U(R). \end{aligned}$$

From these equations, we derive

$$p(U(x_1) - U(x_0)) = (1-p)(U(R) - U(r)) = p(U(x_2) - U(x_1)).$$

[3] Note that for the elicitation of utility *functions* as introduced in this paper, it is also possible to use the less complex certainty equivalent method (see below) or the probability equivalent method.

[4] Note that for this method, no further specification of the utility function is needed.

Since $p > 0$, we can conclude that

$$U(x_1) - U(x_0) \;=\; U(x_2) - U(x_1).$$

This equality of utility differences yields utility points and by repeating this iterative process one can elicit the desired number of utility points.[5]

Note that in cases subjective probability distortions are of less concern or when the trade-off method might be too complex, certainty equivalents can be elicited by specifying p, usually setting $p = 0.5$ and by asking for the value $x_p \in (x_{\min}, x_{\max})$ that makes the decision maker indifferent between receiving $(x_{\min}, p; x_{\max})$ and x_p. If we normalize U so that $U(x_{\min}) = 0$ and $U(x_{\max}) = 1$, then the elicited indifference means $pU(x_{\max}) = p = U(x_p)$, and in case $p = .5$, $x_{.5}$ is the utility mid-point between $x_{\min}$ and $x_{\max}$.

3.2 P-Spline Interpolation and Error Correction for Utility Functions and Their Derivatives

How do penalized spline (P-spline) regressions connect these utility points? A first non-parametric approach is linear interpolation (see e.g., Abdellaoui, 2000; Abdellaoui, Bleichrodt, and Paraschiv, 2007; Etchart-Vincent, 2004; Fennema and Van Assen, 1998). This approach is suited if distances between subsequent utility points are small, decision or measurement errors are unlikely, precision of interpolation is of lower priority, and if enough points are elicited. It is relatively 'costly' in terms of the required points per derivative, e.g., for computing a fourth derivative, at least five utility points are required. Moreover, and important for applications, linear interpolation generally does neither establish a differentiable function nor account for defectively elicited points (due to measurement or decision errors).

Penalized spline regression establishes differentiable functions by smoothing the data in a 'global' way, thus incorporating all information available. As introduced here, this results in estimates for the utility function and its derivatives in one single fit. Therefore, there is no need to additionally smooth the derivatives or compute them numerically. Similar to the parametric approach, this method is parsimonious. One point between the fixed limit points is enough to determine the sign of the third derivative.

[5] This procedure can be extended to elicit utility on the loss domain, to elicit probability weighting (both in the gain and the loss domain) and to elicit loss aversion, as suggested by Abdellaoui (2000) and Abdellaoui, Bleichrodt, and Paraschiv (2007).

Spline Regression Spline regression generalizes the concept of a conventional linear regression. Instead of using only the x-values of elicited utility points in a regression, also higher powers of those values are added—up to degree p. Moreover, and contrarily to polynomial regression, the coefficients on those 'basis functions' (i.e., $1, x, x^2, x^3, \ldots, x^p$) are allowed to vary between pre-defined subintervals of the interval of interest.[6] This is achieved by exchanging (or extending) the global basis for a local basis, that is a set of basis functions that are only piecewise defined (i.e. each of them is different from zero only on a certain subinterval of the interval of interest).

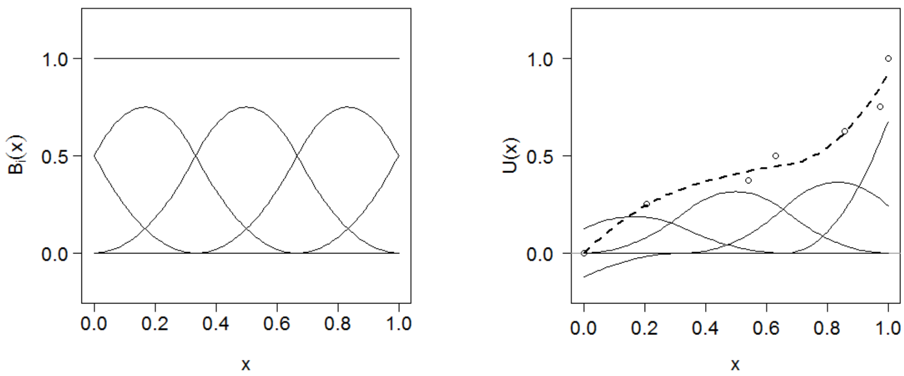
The result—i.e., the estimated function—is a smooth combination of piecewise polynomial functions of degree p, that is, with common implementation, $(p-1)$ times continuously differentiable. Figure 2 shows an exemplary local basis (called *B-spline* basis) with three subintervals and a regression on these basis functions, where for illustration the basis functions are scaled by the corresponding regression coefficients.[7]

Although spline regression possesses several advantages over other function fitting techniques, one challenge is that the fitted curve depends on the choice of the boundaries (called *knots*) of the subintervals on which the local basis functions are defined.[8] An additional challenge is the degree of the function to be fitted: To obtain a $p-1$ times continuously differentiable curve, one needs B-spline basis functions of degree p. Thus, in order to have the fourth derivative at least quadratic, B-spline basis functions of degree 6 are necessary. With barely more than 6 elicited utility points, a pure spline regression approach with local bases of degree 6 is impossible.

[6] Polynomial regression and interpolation, which leads to an estimated polynomial function of order p, roots in the strong theoretical foundation of the Stone-Weierstrass theorem. The Stone-Weierstrass theorem states that every real, continuous function defined on a closed interval can be uniformly approximated arbitrarily close by a polynomial function. In practice, however, the order is limited by available data points and the data will be underfit. Furthermore, even when the order of polynomial functions is high, interpolation quality can be very poor for some functions (e.g., *Runge's phenomenon*, see Runge, 1901) and at the boundaries of the interval under study, the interpolation function becomes unstable. In general, a global basis approach often lacks the flexibility to adequately adjust the degree of curvature to different, possibly asymptotically constant regions. The resulting function would thus in many cases either underfit or overfit the data.

[7] A short introduction to spline regression using a B-spline basis is given in Appendix D.1; for more details see De Boor (1987).

[8] More specifically, a higher amount of knots allows a higher flexibility, which, however, sometimes results in overfitting the data. Moreover, knot placement also influences the fitted curve considerably.



(a) B-spline Basis of Degree $p = 2$ (b) Regression on a B-spline Basis

Figure 2. Illustration of Simple (Non-Penalized) Spline Regression with B-splines of Second Degree Where the Interval $[0, 1]$ is Divided Into Three Intervals

P-spline Regression P-spline regression (Eilers and Marx, 1996) solves these challenges and in addition to smoothing the utility function itself (as spline regression also might do), it smoothens at least one derivative. P-spline regression combines the regression approach on a B-spline basis using an excessive number of (usually equally spaced) knots (i.e. B-spline basis functions) with penalties (usually on the curvature) to prevent the fitted curve from oscillating or fluctuating more than needed when allowing for a high flexibility by choosing a large amount of knots.[9] Technically, these penalties increase the number of conditions for the equation system to solve, and hence higher degrees can be combined with choosing a high amount of knots. Formally, the d-order differences of the B-spline coefficients are penalized by adding these differences to the objective function. We will refer to this summand as the dth-order penalty. When $\boldsymbol{D_d}$ denotes the matrix representation of the dth difference operator $\triangle^d$ defined as $\triangle^d a_j := \triangle(\triangle^{d-1} a_j)$ with $\triangle a_j := (a_j - a_{j-1})$, the objective function of the P-spline regression writes

$$Q_B(\boldsymbol{\alpha}) \;=\; \|\boldsymbol{y} - \boldsymbol{B\alpha}\|^2 + \omega\|\boldsymbol{D_d \alpha}\|^2, \tag{1}$$

[9] When introducing P-splines, Eilers and Marx state that generally the number of B-splines is moderate (10-20). In more recent work, they note that "the size of the basis can be anywhere from 10 to over 1000", depending on the application (Eilers and Marx, 2010). In our application, using 15 inner knots leads to under-fitting in a considerable share of cases, which is why we chose 20. This choice leads to the desired flexibility without overfitting the data.

where the goal is to minimize Q_B by choice of $\boldsymbol{\alpha}$. As with conventional spline regression using a B-spline basis, the design matrix $\boldsymbol{B}$ consists of the $k+p-1$ B-splines evaluated at the x-values of the N given data points from one individual (see Appendix D.1 for details), and $\boldsymbol{y}$ denotes the vector of length N containing the respective y-values, i.e. utility levels in our setting. The objective function (1) is minimized by $\hat{\boldsymbol{\alpha}} = (\boldsymbol{B}'\boldsymbol{B} + \omega \boldsymbol{D}_d' \boldsymbol{D}_{\boldsymbol{d}})^{-1}\boldsymbol{B}'\boldsymbol{y}$.

The non-negative tuning parameter ω allows controlling the smoothness of the predicted function. Choosing $\omega = 0$ results in the classical linear regression of $\boldsymbol{y}$ on $\boldsymbol{B}$ as formulated for example in (15) in Appendix D.1, with the well-known solution $\hat{\boldsymbol{a}} = (\boldsymbol{B}'\boldsymbol{B})^{-1}\boldsymbol{B}'\boldsymbol{y}$. Therefore, as laid out above, a low value of ω will result in a fitted function that overfits the data and possibly oscillates considerably. Depending on the order of the penalty, a linear ($d = 2$) or quadratic ($d = 3$) or more general, a polynomial function of order $d - 1$ will be the result in the limit of an increasing ω (Eilers and Marx, 2010).

Choice of the Penalty Term In this study—and more generally in the context of the elicitation of higher order risk preferences—interest lies in smoothing the utility function itself, but also in smoothing the third and possibly higher derivatives of the utility function. Moreover, we would like to have a continuous utility function with continuous derivatives with a suitable interpolation quality at least for the derivatives of interest.

Here, we use an approach to jointly smooth the third and fourth derivative, suited for the joint elicitation of prudence and temperance, which is defined by a negative fourth derivative. This requires the balanced use of penalties of orders $d = 3$ and $d = 4$, as laid out in Appendix D.3. Penalization of multiple orders has been applied before in the P-spline literature, remarkably in studies with a focus on the quality of interpolation (in chemistry for signal regression and to predict temperature based on ozone levels, see Aldrin, 2006; Marx and Eilers, 2002). However, both studies rely on visual inspection for determining the relation between the two penalties. Therefore, we develop a data-driven approach to jointly use multiple orders of penalties, which we present in Appendix D.3 together with the details on its implementation and on the choice of the penalty orders.

Shape Constraints Monotonicity, or more precisely, utility as a monotone increasing function of monetary units, is a common assumption for utility functions. We follow the approach introduced by Bollaerts, Eilers, and Mechelen (2006) to incorporate a monotonicity constraint in P-splines regression. In the spirit of P-splines, this constraint is approximately enforced using a discrete, asymmetric penalty. We elucidate this approach and its implementation in Appendix D.4.

Value Constraints Due to using the certainty equivalent or the trade-off method for elicitation of utility, the points $(0, 0)$ and $(1, 1)$ are fixed and non-defective. They should therefore be exactly predicted, thus the interpolating spline function has to meet the following conditions: $f(0) = 0$ and $f(1) = 1$. To our knowledge, to date there exists no solution for this in the P-Spline literature. In the spirit of the penalty idea behind P-splines and the implementation of the shape constraint above, we implement these constraints by iteratively increasing weights at the points $x = 0$ and $x = 1$, thus increasing the penalty on deviations from this points until the conditions are (approximately) met.[10]

3.2.1 Choosing the Degree of Smoothness.

With perfect fidelity to the data and only few data points, the fitted P-spline function generally might still resemble a linearly interpolated function. In some cases though, it will oscillate heavily and fail to describe the true overall shape of the function. These cases are examples of overfitting: Fidelity to the data is high, but the quality of predictions is most likely poor. On the other hand, when smoothness is overweighted, the resulting function might miss important changes in curvature and also fail to describe the overall shape of the function, i.e. the function is underfitting the data.

Choosing the Degree of Smoothness by Optimizing Predictive Quality When studying intensities, the precise shape of the utility curve and its derivatives are needed. Then, we wish to have a curve that perfectly balances smoothness and fidelity to the data. How should we choose a value for ω, the smoothing parameter or penalty weight? We apply a leave-k-out cross validation (CV)—an objective data-driven decision criteria that focuses on the quality of prediction and that is robust to overfitting the data in case of

[10] The same effect can probably be realized with a penalty term in an iterative approach such as the one presented by Bollaerts, Eilers, and Mechelen (2006) to incorporate shape constraints. However, the existence of a solution to the minimization problem for the modified objective function has to be investigated, see Bollaerts, Eilers, and Mechelen (2006).

correlated observations (Arlot and Celisse, 2010, Chapter 8.1).[11] Using cross validation, the model is fit with only a part of the data and the remainder, the k points left out, is used to compute prediction errors. According to CV, the smoothness parameter that minimizes the average prediction error is the preferred one.

Leave-k-out CV can be seen as a mean of error correction: The more points left out when fitting the model, the more important becomes the predictive quality and the higher the smoothness parameter will be, in case some points deviate from the common trend. Since reversal rates of one third are common in choice tasks as the ones applied in this study,[12] we perform 'leave-at-least-$1/3N$-out' cross validation, which results in leave-3-out CV, in case the maximum number of utility points was elicited for the individual under study.[13]

However, overfitting the data is still possible, if the distance between points is large, that is in case of sparse information per knot. The reason is that the penalized derivative of the function can change over wide intervals, thus the change needed from knot to knot to predict every point exactly may only be marginal, and is thus not sufficiently penalized. Therefore, we develop and apply a way to determine a data-driven minimum for the penalty parameter to rule out this reason of overfitting. We discuss our choice of the data-driven decision criteria and present the developed approach to rule out overfitting in case of sparse information per knot in Appendix D.5.

3.3 Intensity Measures of Risk Aversion and Prudence

Having established continuous utility functions from the elicited utility points, we can now apply well-known intensity measures of risk aversion and prudence.

Degree of Risk Aversion We measure the degree of risk aversion by the well-known and widely used Arrow-Pratt measure of (absolute) risk aversion, defined as $\rho(x) = -u''(x)/u'(x)$ (Pratt, 1964). For this measure, we compute the mean over the interval $[0, 1]$ based on 1000 points.

[11] Due to the chain structure of the experiment applied, the measurement error of single utility points might be correlated.

[12] See e.g., Abdellaoui (2000), Abdellaoui, Bleichrodt, and Paraschiv (2007), Etchart-Vincent (2004), and Fennema and Van Assen (1998).

[13] Note that we excluded the points $(0, 0)$ and $(1, 1)$ for computing the average prediction error.

Naturally, steep increases are associated with a higher intensity than a constant, slow increase. We therefore summarize the measure of risk aversion over the interval $[0, 1]$ by taking its mean to capture such steep parts of the second derivative, even if vast parts of this derivative are actually zero. The median in such a case would be zero, which certainly is not the right measure of risk aversion in comparison with individuals exhibiting a steady, but slow increase in the measure of risk aversion.

Degree of Prudence We compute the measure by Kimball (1990), commonly stated as $-u'''/u''$, for (strictly) risk-averse individuals and $\pi = u'''/u'$, the measure by Crainich and Eeckhoudt (2008), for all individuals. As for the degree of risk aversion, we aggregate this information by averaging these measures over the interval $[0, 1]$.

4 Measurement Error, Error Propagation and Error Correction

To inform about the characteristics of the proposed method with respect to measurement error, its sensitivity to error propagation and its ability to correct decision errors made by participants, we conduct a simulation exercise. We assume the true intensity measures to be known, as we assume that utility follows a well-known and well-studied parametric form, the so-called expo-power utility (Abdellaoui, Barrios, and Wakker, 2007; Holt and Laury, 2002). In this case, the intensity measures are directly linked to the parameters of the utility function.

We focus on, and limit our scope to risk aversion in this simulation for two reasons: First, commonly used parametric utility forms are not suited to study higher order risk preferences (Noussair, Trautmann, and Kuilen, 2014). Therefore, there is no simple or straight-forward way to simulate a data generating process or true shape of the utility curve when considering intensity measures of prudence or temperance. When investigating error propagation or error correction with our method, however, a "true" utility curve or at least utility points are required as a starting point. To investigate risk aversion, the power or the expo-power family are fairly standard and widely used (see, e.g., Abdellaoui, Barrios, and Wakker, 2007; Holt and Laury, 2002). This is not only helpful for simulating the "true" utility curves, but also to have an alternative benchmark model to pitch our method to. Second, the intensity measures of prudence and temperance resulting from our method are constructed analoguously to the Arrow-Pratt measure of risk aversion, and in particular, result from the same estimated utility curve. It is thus fair to assume that measurement error, error propagation or error correction affect intensity measures of higher order risk preferences analoguously. Therefore, focusing on risk aversion offers a clean approach of simulating the data generating process, and is in addition sufficent to illustrate these characteristics of the method.

4.1 Simulation Design

We generate data according to the expo-power utility function, $u(x) = (1 - \exp(-\alpha x^{1-r})/\alpha^{-1}$. The expo-power function matches empirical evidence better than the power function, as it allows the combination of increasing relative risk and decreasing absolute risk aversion (e.g., Holt and Laury, 2002). In addition, it is an attractive choice since it is more flexible than the one-parameter power function, $u(x) = x^{1-r}(1-r)^{-1}$, which implies constant relative risk aversion.

We draw parameters a and r of the expo-power function from normal distributions, with mean values and standard deviation informed by aggregate estimates of these parameters in the literature. We draw 500 values of r and a according to the estimates in Noussair, Trautmann, and Kuilen (2014), as they result from studying higher order risk preferences as well, and in addition from a general population. Additional 500 values of r and a are drawn according to the estimates reported in Holt and Laury (2002), for being, to our knowledge, the first study reporting estimates of the expo-power family, and in order to have lower values of risk aversion represented in the simulation as well.

We then determine the "true" certainty equivalents for all the lotteries used to elicit utility points with the certainty equivalent method, and chose, for every certainty equivalent, the one that comes closest to the "true" certainty equivalent from the list of the possible values of certainty equivalents after three "staircases" (iterations of the bisection approach).

Based on these utility points, we then fit the utility curve using our method, the power utility function, and the expo-power family. Finally we compute the Arrow-Pratt measure as described above using our method, or using the obtained parameters of the two parametric functions. For the power family, the Arrow-Pratt measure of absolute risk aversion is given by $AR_{\text{CRRA}} = rx^{-1}$, and for the expo-power family, it is $AR_{\text{Expo-Power}} =$

$rx^{-1} + ((1-r)\alpha)x^{-r}$. Just as for our method, we evaluate these measures for 100 evenly spaced monetary values in the interval from 0 to 140 to span the full range of incentives used, and then take the average.

4.1.1 Measurement Error. For this simulation exercise, we denote with measurement error the loss compared to perfect correlation in the correlation coefficient between (i) the intensity measure corresponding to the true shape of the utility curve, and (ii) the intensity measure obtained through elicitation of utility points using certainty equivalents in combination with one of the considered interpolation strategies to estimate utility functions. By assuming that the data generating process follows a parametric function – the expo-power function – with certain parameters, we are able to compute average Arrow-Pratt and higher order risk intensity measures without any measurement error, simply by plugging in the drawn parameters in the given formulas for AR, and averaging over the levels of wealth. In case this true measure would be obtained after utility point elicitation and subsequent estimation of the utility function with one of the considered methods, the correlation coefficient would be 1. Any deviation in the correlation coefficient from 1 can thus be attributed to measurement error.

4.1.2 Error Propagation. In general, iterative elicitation methods might suffer from error propagation, that is, a decision or measurement error at an early stage of the elicitation procedure carries over to later stages, and thus enters the final measurement several times. For example, in the certainty equivalents method, the certainty equivalent corresponding to a utlity of $u = 0.5$ will be used as high and low lottery outcome in subsequent choice tasks. If this value is measured with an error, it will have consequences for later stages as well, and thus affect the final intensity measure in various ways. Yet, iterative methods such as the bisection or staircase method have been praised for higher precision to begin with, as they lower the cognitive demand required by participants. In particular in development settings, with children, or with non-standard participant pools in general, this might be of non-negligible importance. Ex ante, it is thus unclear which effect will be dominating, as this will – to a large degree – depend on the sample. Nevertheless, it might be useful to derive a possible upper bound of measurement error: It is arguably not very likely that *every* participant makes a mistake at the *very first stage* of the procedure. Thus, most likely, measurement error imposed by error propagation will be smaller, and we therefore use this scenario as worst case scenario to create a possible upper bound of

Table 1. Simulation Results: Measurement Error, Error Propagation and Error Correction

	$AR_{\text{Expo-Power}}$	AR_{CRRA}	AR_{Linear}	$AR_{\text{Schneider et al.}}$
Correlation ρ with AR_{DGP}	0.95	0.97	0.37	0.96
Δ in ρ caused by error	0.04	0.00	0.38	0.00

Notes: This table shows in the first row the Pearson correlation coefficients between the "true" Arrow-Pratt measure of risk aversion according to the assumed data generating process, AR_{DGP}, and the Arrow-Pratt measures obtained via elicitation of certainty equivalents and subsequent estimation of a utility curve according to the respective methods. The entries in the second row denote the difference between the correlation coefficients in the first row and correlation coefficients obtained when (simulated) measurement error is introduced.

measurement error due to error propagation for each method to estimate a utility curve.

4.1.3 Error Correction. Least squares methods, linear and non-linear, minimize the total sum of squares. This means that their predictions result in the overall best fit given *all* the information available, without focusing too much on a single observation. For utility curves, this means that the overall shape of certainty equivalents is relevant for determining the predicted utility function, or equivalently when fitting a parametric curve, for estimating the parametric function's parameter. This feature thus allows error correction to a certain degree, as a single certainty equivalent that does not perfectly fit to the overall shape suggested by the other certainty equivalents will not change the fitted curve too much (as opposed to, e.g., linear interpolation). As the method we suggest is a least squares method as well, combined with a penalty for flexibility, it shares this feature with parametric utility functions that are fitted using a non-linear least squares approach. By estimating a utility curve with our method based on erroneous data, generated as described above, and fitting the two discussed parametric functions to the same data, we can compare the methods with respect to their ability to correct for minor to medium decision errors.

4.2 Simulation Results

We start by assessing measurement error due to elicitation of utility points and subsequent prediction of a utility curve based on these points using the different prediction methods. For each of the 1000 draws of a parameter pair for the expo-power utility curve, we compute the "true" Arrow-Pratt measure of risk aversion. Likewise, for each of these draws, we build Arrow-Pratt measures as described above. We then estimate the (Pearson) correlation with the "true" measures – the higher the correlation coefficients, the lower the measurement error. Results are reported in the first row of Table 1. Generally, with exception of linear interpolation, the approaches work well with correlation coefficients of up to 0.97 for the CRRA function, despite being more restrictive than the actual data generating process (for depending on one parameter only). Maybe a bit surprising, the expo-power function performs worse than the CRRA function. This suggests that the shape of the curve is less sensitive to substitution effects between the two parameters determining its form than the Arrow-Pratt measure is. Our method performs comparable to the two well-known methods, despite its greater flexibility and ability to incorporate prudent risk seekers and risk averters alike. That is, the flexibility in our method does not come at the cost of precision in the sense of lower measurement error.

How does error propagation affect the measurement error, and to which degree are the interpolation methods able to smoothen out these mistakes, thus correct these errors? We compute Arrow-Pratt measures for every method and every draw of parameter pairs, where now the elicited certainty equivalents are defective, as described above. We estimate (Pearson) correlation coefficients with the true Arrow-Pratt measures, and compare these coefficients with those obtaine before, i.e., without assuming decision errors in the first decision stage. This difference, which is the sum of error propagation and error correction, is reported in the second row of Table 1. The CRRA and our method perform best, without any difference in the first two digits of the correlation coefficients. The expo-power function still performs relatively well, but here again, we note that the Arrow-Pratt measure based on the parameters of the expo-power function seems to be more sensitive to changes in these parameters, relative to the shape of utility.

4.3 Summary

In total, we have seen that our method performs comparable to two well-known and widely used parametric utility functions with respect to measurement error when measuring Arrow-Pratt risk aversion coefficients for risk averse subjects, whose utility function can be described by an expo-power function.

Error propagation is, in general, of minor concern for mild to moderate degrees of decision errors made by *every single* participant as we have assumed here. The CRRA utility family, as well as our method, are even able to smoothen out these mistakes completely, in the sense that measures are correlated with the "true" measure to the *same degree* as without introducing decision errors. The measure obtained using linear interpolation turns negative once decision errors are simulated, which illustrates the limits of such an approach for computing intensity measures when decision errors cannot be ruled out.

5 Validation

We validate our method in laboratory and online experiments using the risk apportionment methods for elicitation of prudence and temperance as implemented in Deck and Schlesinger (2010) and Noussair, Trautmann, and Kuilen (2014), as well as the extension to obtain prudence and temperance premia proposed by Ebert and Wiesen (2014).

5.1 Laboratory Experiment

The laboratory session included the risky decision sets and a real-effort task in addition to a survey. By participation in the experiment and in the real-effort task, subjects earned money. In addition, subjects earned money from the risk tasks, resulting in an average earning of €27.16 (minimum was €21 and maximum was €34).

Sessions lasted around one hour, and were scheduled at the same time and weekday (a Wednesday), to avoid differential weekday effects. The dates were selected to avoid holidays or examination period in Bonn.

5.1.1 Risk tasks. We presented subjects with the risk tasks in blocks that were preceded by the corresponding instructions and comprehension questions. Following, e.g., Noussair, Trautmann, and Kuilen (2014), the order of blocks was chosen in ascending order of complexity to help subjects accustom themselves to the tasks. Subjects were always presented one decision at a time. All risk tasks consisted of a decision between two options without the possibility to express indifference. Whether the risk-averse, prudent or temperate options, respectively, was shown on the left- or on the right-hand side of the screen was randomized in each task. No lotteries were resolved before the end of the experiment and all randomization was performed by the computer. After participants had completed all decisions, one was selected randomly and paid out according to the decision taken.[14] All monetary amounts in the risk tasks were expressed in experimental dollars ($), that were converted to Euro at an exchange rate of €1 = $10.

Certainty equivalents The first block consisted of the choice tasks for elicitation of utility points using the certainty equivalence method as laid out in Section 3.1 with $p = 0.5$ and $(x_{\min}, x_{\max}) = (0, \$140)$. One of the two options in this block always offered a sure payment and the other option offered a lottery with two equally likely outcomes. To illustrate the equiprobable nature of the lottery, we applied an animated coin with a black and a white side that was rotating, see Figure 3.

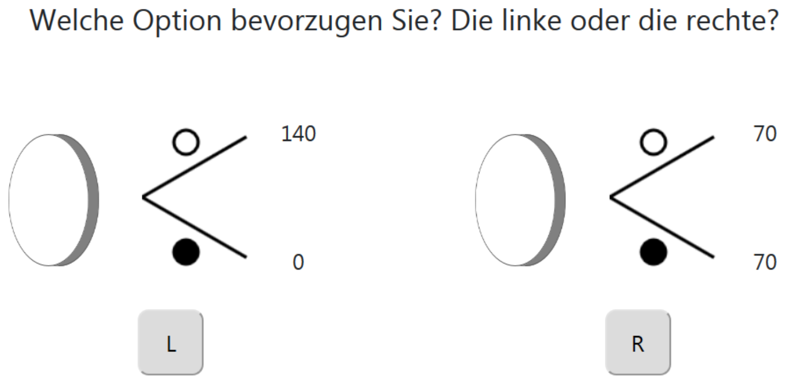


Figure 3. Decision Screen Example For The Elicitation of Certainty Equivalents Using the Rotating Coin Design (in German)

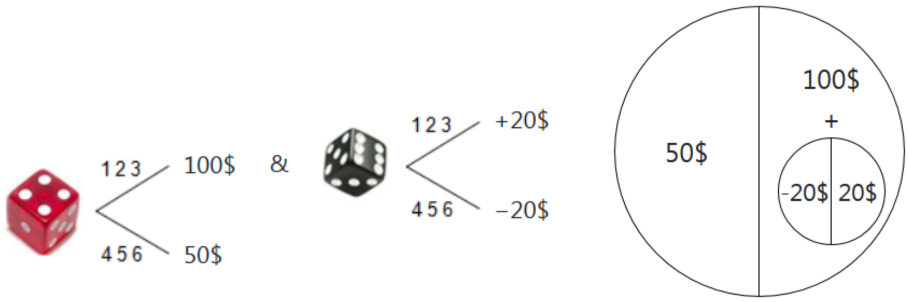
In case the subject preferred the safe option (the risky option), the amount of the safe option was increased (decreased), and the subject was again asked for a decision. This procedure is known as the bisection or staircase method.[15] We repeated it three times to approximate indifference for each of the 6 utility points we elicited.

Risk apportionment The second block contained five decision tasks to elicit prudence and the third one contained the five tasks for elicitation of temperance. All tasks in the second and the third block were using the risk

[14] The random problem selection mechanism does not alter behavior (Starmer and Sugden, 1991) and, moreover, it is the only incentive compatible mechanism when subject perform several tasks (Azrieli, Chambers, and Healy, 2018).
[15] See, e.g., Falk et al. (2018) for a recent example.

apportionment method described in Section 2 going back to Eeckhoudt and Schlesinger (2006) and thus contained equiprobable lotteries only.

Roughly half of the subjects was presented the decision situation in the "dice-design" applied by Noussair, Trautmann, and Kuilen (2014), see Figure 4(a), with the general population in the Netherlands, the other half was confronted with the "fortune-wheel-design" applied by Deck and Schlesinger (2014), see Figure 4(b), with a student sample in the laboratory. Assignment to the design was randomized at the individual level. Noussair, Trautmann,



(a) Dice-Design by Noussair, Trautmann, and Kuilen (2014)

(b) Fortune-Wheel-Design by Deck and Schlesinger (2014)

Figure 4. A Prudent Option in the Risk Apportionment Method (Eeckhoudt and Schlesinger, 2006)

and Kuilen (2014) use (up to) three different colored dice to highlight the independence of the different (equiprobable) risks represented by independent throws of the dies, which they deem crucial for the interpretation of the compound risks in terms of prudence and temperance. In a similar vein, Deck and Schlesinger (2014) use different sized circles divided in half to represent the equiprobable nature of the lotteries, thereby mimicking fortune-wheels, where the outcomes are written on the corresponding part of the fortune-wheel.

Figure 4 shows the prudent option of a lottery pair. Here, the unavoidable mean-zero risk represented by the black die or the small fortune wheel is added to the state of higher wealth (i.e., added to \$100), whereas in the imprudent option, it is added to the state of lower wealth (i.e., \$50). An individual choosing the shown option and expressing the will to throw the black die or spin the small fortune-wheel when the first lottery—throwing the red die or spinning the big fortune wheel—resulted in a payment of \$100 instead of when it resulted in a payment of only \$50 is classified as prudent. Similarly, an individual expressing the will to combine throwing two dies or spinning two fortune-wheels in the same state of the world would be classified as intemperate, whereas the one preferring to disaggregate them, and throwing one die or spinning one fortune-wheel in each state of the world, would be classified as temperate.

The decision situations varied in terms of the level of wealth that the "harms" were added to and the size of the "harms". A table containing all decision situations, exemplary decision situations as well as the (translated) instructions and control questions can be found in Appendix F.[16]

Prudence and Temperance Premium With a sub-sample consisting of about 50% of our participants, we additionally elicited prudence and temperance premia using the method introduced by Ebert and Wiesen (2014) building on the risk apportionment method (Eeckhoudt and Schlesinger, 2006) used in the second and third block. An exemplary decision situation to elicit a prudence premium is shown in Figure 5.[17]

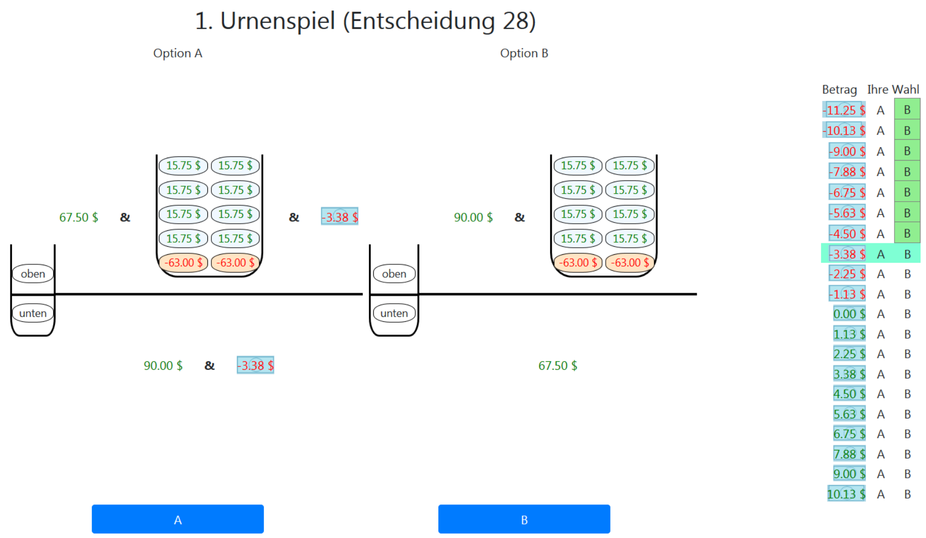


Figure 5. Elicitation of an "Imprudence-Premium" as introduced by Ebert and Wiesen (2014): In addition to the imprudent option, an increasing amount of money is offered for compensation of the "misalignment of the harms".

[16] Instructions and control questions were very similar to the original ones applied by the corresponding authors and adapted only where needed. Decision situations were selected and scaled in a way to make them have equal means.

[17] Note that the prudence and temperance premia as applied by Ebert and Wiesen (2014) differ from the one in Crainich and Eeckhoudt (2008) that is mentioned in Section 2: Here and in Ebert and Wiesen (2014), for the imprudent and intemperate option, the premium is paid in *both* states of the world, whereas in Crainich and Eeckhoudt (2008) it is paid just in the best state of the world, i.e. the lower state of the left option.

Subjects are presented with a lottery pair, where (compound) lotteries are represented by (combinations of) urns that contain balls. The lottery pair is accompanied by a list of increasing amounts on the right-hand side, where one amount at a time is added to the imprudent or intemperate option of the lottery pair ("Option A" in the example in Figure 5) to ask for preference in the specific situation. By highlighting the chosen option for each amount, consistency is enhanced. In total, subjects were asked to make 100 decisions; 20 for three decision situations each to elicit prudence and 20 for two decision situations each to elicit temperance. Compensation amounts ranged from $-11.25 to $10.13 in steps of $1.125 and were the same for temperance and prudence. A table containing all decision situations as well as the (translated) instructions and control questions can be found in Appendix F.

5.1.2 Real-effort tasks. We asked participants to complete two types of real-effort tasks, and rewarded them for completion. In the first real-effort task, subjects had to pick one out of six suggested solutions to solve a puzzle, and in the second, the so-called symbol-correspondence task, they were asked to select the symbol that corresponded to the digit that was randomly presented. The relevant correspondence table was shown throughout the task above the presented digit (see Appendix F for examples).

5.1.3 Experimental procedures. We recruited 58 participants via the recruitment platform ORSEE (Greiner, 2015) from the database of volunteers of the DecisionLab at the Max-Planck-Institute for Research on Collective Goods in Bonn, Germany.

Subjects conducted all decisions visually isolated from the other participants on a computer at which they were seated upon arrival. The experiment was computerized and administered using the experimental platform oTree (Chen, Schonger, and Wickens, 2016). Instructions with examples were given on screen and subjects had to answer comprehension questions before proceeding with their decisions; see Appendix F for details.

Payment All payments were made by electronic SEPA bank transfer which was initiated immediately after the experiment, thus it was guaranteed by law that they arrived at the same day.[18] Subjects were informed before registration that agreeing to receiving the payment via electronic bank transfer

[18] The allowed timing for clearance in Germany is governed by German civil law (BGB, §675s) to be one day for transfers within the European Economic Area (EEA)

was a prerequisite for participation. Thus, we ensured that the probability subjects attach to receiving the money is the same across sessions.

Subjects entered their International Bank Account Number (IBAN)—the standard for identifying bank accounts in the European Union and other countries—themselves, which is used for the bank transfer. Therefore, subjects could be confident that problems with identifying numbers in handwriting or the like were ruled out.

Business cards of the first author were distributed to participants, and they were instructed to contact him if there were difficulties with their payment. This measure was taken to signal that the experimenters themselves were in charge of the bank transfers.

5.2 Online Experiment

The online experiment follows closely the laboratory experiment. In particular, we used the same experimental tasks for elicitation of higher order risk preferences and the same implementation in oTree. We will therefore focus here only on the implementation differences. In total 527 subjects, who were recruited via mTurk, completed the online experiment. Payment was administered via mTurk. Subjects had a window of three days to complete the task that we first posted on May 17, 2019. On June 3, 2019, we finished data collection.

5.3 Validation Results

In this subsection, we pool results from the laboratory and the online experiment. We highlight notable differences between the laboratory and the online experiment in the text, if applicable, and present results from all analyzes conducted separately for laboratory and online experiments in Appendix E.1.

We first classify subjects into risk averse and risk loving, prudent and imprudent, and temperant and intemperant according to the measures as resulting from our method. Table 2 reports the distribution of our subjects. The largest part of the subjects are risk averse and prudent (63.11), followed by prudent risk seekers, amounting to 18.54 percent. Notably, the sample of students in the laboratory is considerably more risk averse than the online sample: Almost 88 percent behave risk averse in the laboratory, while only 75 do so online (see Table 10 in Appendix E.1).

Table 2. Distribution of Risk and Prudence Attitudes

		Risk averse	Risk loving	All
Non-prudent	%	12.75	5.61	18.35
Prudent	%	63.11	18.54	81.65
All	%	75.86	24.14	100.00

5.3.1 Validation: Relation with (Higher Order) Risk Premia. We validate the intensity measures derived from our method (see Section 3) by correlating them with the elicited compensation premia using the risk apportionment method, as introduced by Ebert and Wiesen (2014). For the risk premium, we use the indifference value from the first three certainty equivalence decisions, i.e., the certainty equivalent to the lottery $(140, 0.5; 0)$, divide it by the high outcome, and deduct this from 1.[19] As laid out in Section 2, the theoretical intensity measures of prudence and temperance are proportional to various premia, although, strictly speaking, not in a linear way. Moreover, the compensation premia that are elicited with the method by Ebert and Wiesen (2014) are defined slightly differently from those that the theoretical intensity measures are shown to approximate. Therefore, it is not necessarily to be expected to be able to explain a high share of variance.

Table 3 reports the results. All theoretical measures are significantly correlated with the corresponding premia ($p < 0.001$). In particular, this is true for prudence and temperance, which are elicited by completely different decision tasks and elicitation method—the tasks due to Ebert and Wiesen (2014), building on the risk apportionment method (Eeckhoudt and Schlesinger, 2006).

Correlations are much higher in our student sample, with correlation coefficients of up to .81 between the intensity measure of prudence elicited with our method and the prudence premium, and a still sizeable coefficient of .7 between the temperance measures, see Table 11 in Appendix E.1. This might raise the question whether results are driven by cognitive abilities required to solve the increasingly complex compound lotteries involved in the risk apportionment method. While cognitive abilities explain a relatively high share of variation in the prudence and temperance premia, correlation coefficients between the measures of higher order risk preferences remain sig-

[19] A value of 0.5 in this measure corresponds thus to risk neutrality, whereas values greater than 0.5 indicate risk aversion.

nificant, showing that the relation in the methods exists above and beyond the cognitive abilities required to grasp the tasks.

Table 3. Predicting (Higher Order) Risk Premia with Utility-Based Intensity Measures

	Risk Premium		Prudence Premium (EW)		Temperance Premium (EW)	
	(1)	(2)	(1)	(2)	(1)	(2)
AP Risk ($-u''/u'$)	0.661***	0.475***				
	(0.026)	(0.022)				
CE Prudence (u'''/u')			0.529***	0.334***		
			(0.055)	(0.041)		
DE Temperance ($-u^{(iv)}/u'$)					0.209***	0.113**
					(0.042)	(0.039)
Cognitive Abilities		0.517***		0.550***		0.600***
		(0.012)		(0.021)		(0.021)
Num.Obs.	1106	1106	368	368	368	368
R2	0.437	0.670	0.271	0.590	0.045	0.459
R2 Adj.	0.436	0.669	0.269	0.587	0.043	0.456

Notes: This table reports results from OLS regressions, where the dependent variables are the risk premium (column 1), the prudence premium (column 2) and the temperance premium (column 3). Explanatory variables are the utility-based intensity measures of higher order risk preferences, i.e., the Arrow-Pratt coefficient of risk aversion, the Crainich-Eeckhoudt coefficient of prudence and the Denuit-Eeckhoudt coefficient of temperance, as resulting from our method. Prudence and temperance premia are elicited using the method by Ebert and Wiesen (2014). Risk premium is the certainty equivalent of the lottery $(0, 0.5, 140)$ subtracted from 1. All measures are expressed in standard deviations. Choices pooled across the online experiment and the student (laboratory) sample. For separate results, see Table 11 in Appendix E.1.
+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

5.3.2 Predictive Quality of the Count Measures Based on Risk Apportionment.

Counting the number of prudent choices in risk apportionment task à la Eeckhoudt and Schlesinger (2006) is sometimes used to approximate utility-based intensity measure for prudence and temperance as used in theoretical work (e.g., Noussair, Trautmann, and Kuilen, 2014). The underlying—yet empirically unproven—assumption is that higher levels of prudence or temperance make decision errors less likely, and thus result in choosing the prudent or temperant option, respectively, more consistently, i.e., more often. We test whether this assumptions is justified, i.e., whether these 'count measures' predict utility-based intensity measures used in theoretical work. We elicit the intensity measures using the just validated method, and examine whether a high number of prudent and temperant choices is associated with higher levels of the Crainich-Eeckhoudt measure of prudence or the Denuit-Eeckhoudt measure of temperance, respectively.

Tables 4 reports the results from OLS regressions of utility-based intensity measures on the number of prudent (temperant) choices. When pooling the "dice" and the "fortune-wheel" design of the risk apportionment method in the spirit of Noussair, Trautmann, and Kuilen (2014) and Deck and Schlesinger (2010), one more prudent choice increases the Crainich and Eeckhoudt (2008) measure of prudence by .15 standard deviations, and the number of prudent choices explains 31 percent of the variation in the theoretical, utility-based prudence measure. In the laboratory experiment, the "dice-design" as applied by Noussair, Trautmann, and Kuilen (2014) even explains 39 percent of the variation, while in general, the share of explained variance is similar across implementation designs and sample, see Table 12 in Appendix E.1.

Turning to temperance, again pooling choices in the "dice design" and the "fortune-wheel design", we find that one more temperant choice in the risk apportionment tasks is associated with an increase of .07 standard deviations of the Denuit and Eeckhoudt (2010) measure of temperance. The variation of the theoretical intensity measure that is explained by these choices amounts to about 5 percent when pooling the online and the laboratory experiment. Notably, as for prudence, the "dice-design" applied in the laboratory experiment explains 14 percent of the variation, i.e., almost double the variation when pooling the implementation designs in the lab, or thrice the explained variation in the online experiment, Table 12 in Appendix E.1.

5.4 Summary

We have succesfully validated our elicitation and estimation method for the utility-based intensity measures of risk aversion, prudence and temperance using the prudence and temperance premia as elicited in the spirit of Ebert and Wiesen (2014) building on the well-known risk apportionment method (Eeckhoudt and Schlesinger, 2006), and a simple measure of risk premium. Although the premia are theoretically proportional to the utility-based intensity measures, this relationship is not linear, see Section 2. Thus, theory rules out that the variation of the premia that is explained by the intensity measures resulting from our method cannot be 1. Nevertheless, for prudence, this share is sizeable, and the variation explained in the laboratory amounts to .42, whereas it is .17 for temperance. Even when controlling for cognitive abilities, we have been able to show significant correlations.

Table 4. Predicting Utility-Based Intensity Measures With Counting Measures

	Prudence			Temperance		
	Pooled	N	DS	Pooled	N	DS
Coef.	0.153***	0.159***	0.147***	0.068***	0.059***	0.079***
	(0.007)	(0.011)	(0.010)	(0.011)	(0.014)	(0.016)
Num.Obs.	798	404	394	798	404	394
R2	0.306	0.292	0.323	0.049	0.039	0.062
R2 Adj.	0.305	0.291	0.321	0.048	0.037	0.060

Notes: This table reports results from OLS regressions, where the dependent variables are the utility-based intensity measures of higher order risk preferences, i.e., the Crainich-Eeckhoudt coefficient of prudence (columns 1-3) and the Denuit-Eeckhoudt coefficient of temperance (columns 4-6), as resulting from our method (see Section 3). Explanatory variables are the number of prudent and temperant choices, respectively, in the risk apportionment tasks as implemented by Noussair, Trautmann, and Kuilen (2014) with the "dice-design" (columns 2 and 4, labelled 'N') and by Deck and Schlesinger (2010) with the "fortune-wheel design" (columns 3 and 6, labelled 'DS'). Prudence and temperance intensity measures are expressed in standard deviations. Choices pooled across the online experiment and the student (laboratory) sample. For separate results, see Table 12 in Appendix E.1.
+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In addition, we have been able to provide empirical evidence for the yet unproven assumption that a higher number of prudent or temperant choices in the risk apportionment tasks implies a higher intensity of prudence or temperance, respectively. From our laboratory results, it seems that the "dice-design" due to (Noussair, Trautmann, and Kuilen, 2014) works better, whereas in the online experiment, the "fortune-wheel" design explains a higher share of variaton. As our laboratory sample was recruited in Germany, but the online sample was restricted to participants from the US, we cannot rule out cultural differences as the underlying explanation of these observed differences.

6 Application to Savings in Bogotá

We now turn to an application of the method using microdata. We measure preferences in a population of poor subjects in Bogotá, Colombia, and relate them to household saving.

6.1 Data

We collected our data as part of a larger study on savings for the old age. The survey gathered financial information for a sample of 1200 subjects who are beneficiaries of the social protection program SISBEN. The program targets the population in low social strata. We recruited the participants in a two-step procedure. First, we selected neighborhoods based on shares of the population belonging to the required strata. Neighborhoods that contained a large fraction of low income individuals and that were assessed as safe for the team to visit were included in the study. In the second step, enumerators visited randomly selected households and verified whether they belonged to the target group. If this requirement was not fulfilled, the enumerators visited the neighboring household. We oversampled older females to obtain a better picture of the potential vulnerabilities women are exposed to in the old age. Interviews took place in October and November 2013 and lasted on average 90 minutes.

The survey consisted of 16 sections on general demographics, wealth, general savings, pension savings, financial literacy, health behavior, expectations on future income and hypothetical questions on psychological traits. The experiment was conducted with a sub-sample of 693 participants. A team of enumerators conducted the experiment on tablet computers. To meet safety demands, the experiments were conducted in a public space (e.g., community houses) that was easily accessible for participants. The experiment lasted around 20 minutes.

6.1.1 Net Savings.

Savings In the survey we asked detailed questions on a comprehensive range of saving devices: housing, savings plans, savings and checking accounts, certificates of deposit, mutual funds, loans given out and savings in Colombian Pesos or other currencies. We use the sum of these assets to construct a savings measure denoted S_h. This variable intends to capture all liquid assets in the household.[20]

Debt Debt is defined as all financial liabilities a household has against other households, enterprises and financial institutions, including moneylenders. We denote total debt as D_h.

[20] As it is possible to withdraw money from all of these saving instruments in case of emergencies, we interpret all of these savings as liquid assets.

Net Savings Following Noussair, Trautmann, and Kuilen (2014) and Fuchs-Schündeln and Schündeln (2005) we use *net savings* as our main variable of interest, which we calculate as the difference between S_h and D_h.

6.1.2 Income Uncertainty Based on Economic Activity. In order to obtain measures for the financial riskiness of the sector subjects were working in we collected measures from the Commercial Register in Bogotá in 2013.[21] We calculate for 14 sectors the empirical probability of firm closure within each sector. We match this then with the sector subjects are usually working in, which has the same resolution as the official statistics. This measure of working sector dynamics proxies the degree of income uncertainty and hence measures the idiosyncratic risk which subjects might want to hedge using precautionary savings. As unemployment insurance is negligible in our sample, firm closure in the sector constitutes a major threat to personal income. This constitutes a substantial and unavoidable background risk.

In this sense our sample and the economic circumstances resemble the economy described in Aiyagari (1994). There are idiosyncratic earnings uncertainties, effectively no insurance markets and our subjects are borrowing constraint. Background risk, defined as uninsurable idiosyncratic shocks, is arguably substantial and not mitigated by wage insurance on the firm level as found by Fagereng, Guiso, and Pistaferri (2017b).

Our measure is closely related for example to the measure of income uncertainty applied by Fagereng, Guiso, and Pistaferri (2017a) when studying precautionary saving in Norway: They infer firm performance volatility from companies' balance sheets and use this to instrument their measure of income uncertainty. This limits their sample to employees of private firms with balance sheets available to the public and is arguably not suited in our context. We therefore use the next aggregation level, namely working sectors, since on this level, we are able to link individuals to secondary data on economic performance.

The arguments in favor of exogeneity of firm performance volatility given by Fagereng, Guiso, and Pistaferri (2017a) are valid also for our measure: First, they argue, firm shocks are hard to avoid for most workers, and second, firms pass over this variation onto their workers' wages. Clearly, in our case, negative working sector dynamics will be handed over to employees

[21] Data was processed and made available in an 'Overview of Indicators' by the Knowledge Management Board of the Chamber of Commerce (in Spanish: *'Tablero de Indicadores', Dirección de Gestión de Conocimiento*), and we use working sectors as provided in ISIC Rev. IIIa A.C. by DANE, based on the economic activity.

due to absence of formal contracts as well. Thus, payment can be adjusted easily, but also employment itself is more uncertain in working sectors with a higher share of closed businesses.

Time Preferences We followed the experimental design by Andersen et al. (2008) to elicit time preferences: Participants decided on receiving an amount x in 30 days or an amount $x(1 + r/12)$ with $r > 0$ in 60 days. Values of r were increased gradually and subjects usually switched from choosing x in 30 days to $x(1 + r/12)$ in 60 days for some r according to their time preferences. Using this switching point, we calculate a lower and an upper bound for the interest rate. This interest rate can also be interpreted as impatience, as people were deciding about the timing of receiving money.

We repeated the task with a higher delay of payment: Subjects now decided about receiving the lower amount in 180 days or receiving the higher amount in 210 days. Similarly to the case for the near future time frame, we deduce a lower and an upper bound of interest rate or impatience from the switching point.

The difference between both interest rates or impatience for the two time frames informs about consistency in interest rates. For individuals deciding consistently, the impatience to receive a monetary amount 30 days earlier should be unaffected by shifting the date of the earlier payment by 150 days. A lower impatience in the more distant future corresponds to an increasingly patient subject.

Other Data In the style of comparable, previous studies on precautionary saving (e.g., Fuchs-Schündeln and Schündeln, 2005; Noussair, Trautmann, and Kuilen, 2014), we control for other socioeconomic factors within our analysis such as age, gender, number of adult household members, number of children in a household, education and income (we use average per capita household income). To these, we add further characteristics that have been found to be important in explaining savings:[22]

We measured financial literacy within the survey (Lusardi and Mitchell, 2011) and include its result in the analysis.[23] Van Rooij, Lusardi, and Alessie

[22] Some studies focused on the likelihood of saving, others on the amount of saving. Since we use the same control variables for estimating the likelihood and the amount of saving, we include a variable (if possible) that has been found to either affect the likelihood of saving or the amount or both.

[23] In total, we were asking 18 questions on financial literacy concerning interest rate, asset classes, basic math and financial math. The variable included in the regression corresponds to the number of correctly answered questions.

(2012) find a positive correlation between financial literacy and accumulated savings. Devaney, Anong, and Whirl (2007) and Fisher and Montalto (2010, 2011) report short term planning and saving horizons (i.e. time preference for the present) having a negative effect on the likelihood of saving and net wealth. In our analysis, we use an experimental measure of impatience, time inconsistency with respect to impatience and the planning horizon with respect to financial decisions. Furthermore, we calculated the BMI from weight and height of subjects. The BMI serves as a proxy for temptation and self-control (Hofmann, Friese, and Roefs, 2009; Moffitt et al., 2011). It also serves as a proxy for health status that is positively associated with the likelihood of saving (Fisher and Montalto, 2010, 2011).

6.2 Results

We first give a characterization of our sample in terms of socioeconomic characteristics, risk and time measures. Then we use these measures in order to explain net savings, and finally we model individual precautionary savings motives, i.e., income uncertainty, and relate them to net savings.

6.2.1 Descriptive Statistics. We obtain the full set of household characteristics from 680 subjects of which 72% were female. The mean age was 49 years and spanned from 24 to 87 years. The median education level is primary school. In the financial literacy test, subjects answered on average 9.29 questions out of 18 correctly. The average BMI is slightly above 25, the threshold to mild overweight. The average income per household member is 319 thousand Colombian Pesos (COP) and the average debt is 1.64 million Pesos, which leaves an average of -1.36 million Pesos in net savings.

55 percent of our sample has neither savings nor any debt and average net-savings amount to $-1,361,000$ COP (710 USD). Around 85 percent of our population has no savings and average savings in the sample are 276,000 COP (140 USD), which is less than a month's average per capita household income. Around 27 percent of those reporting non-zero savings save exclusively in cash, another 20 percent save exclusively using other savings technologies. About 34 percent are saving exclusively for housing, of which roughly the half uses a special fund, whereas the other half uses any form of saving device. Average debt amounts to 1,637,000 COP (850 USD) and 38 percent of our sample hold positive debt. Table 5 provides an overview of the demographic characteristics.

Table 5. Summary Statistics

	Mean	Med.	s.d.	Min	Max	Obs.
Male	0.28	0.00	0.45	0.00	1.00	680
Age	48.87	49.00	13.43	24.00	87.00	680
Education	2.48	2.00	0.70	1.00	4.00	680
Financial literacy	9.29	10.00	3.39	0.00	16.00	680
BMI	25.75	25.51	4.31	12.89	42.97	680
Adult HH members	2.85	3.00	1.41	1.00	12.00	680
Children HH members	1.21	1.00	1.29	0.00	7.00	680
Income	3.19	2.77	2.25	0.01	18.00	680
Savings (100k)	2.76	0.00	15.18	0.00	200.00	680
Debt (100k)	16.37	0.00	61.85	0.00	588.04	680
Net savings (100k)	-13.61	0.00	63.94	-588.04	187.00	680
Zero net-savings	0.55	1.00	0.50	0.00	1.00	680

6.2.2 Risk Aversion, Prudence and Time Preferences. Table 6 shows the Arrow-Pratt coefficient of risk aversion, measures of prudence and time preferences. The mean annual interest rate r subjects asked to receive an amount $x(1 + r/12)$ in 60 days instead of an amount x in 30 days is 29.6 percent. This figure is in the range of estimates from experiments with the general population in Denmark (Harrison, Lau, and Williams, 2002). On average, this interest rate or mean impatience stays approximately constant when the timing is changed to receiving the monetary amounts in 180 or 210 days, respectively.

Table 6. Summary Statistics of Risk and Time Measures

	Mean	Med.	s.d.	Min	Max	Obs.
Risk Aversion (A&P)	0.03	-0.01	1.17	-2.43	2.86	588
Prudence (C&E)	7.24	7.61	6.28	-0.50	25.70	588
Impatience	29.60	22.00	15.41	16.00	52.00	693
Increase in patience	0.13	0.95	16.35	-38.90	36.95	693

We obtained a full set of risk preference measures for 588 subjects. Table 7 gives an overview over the classification of risk and prudence as measured based on optimally smoothed P-spline regression. We observe that all combinations of risk aversion and prudence attitudes are present, confirming previous findings by Noussair, Trautmann, and Kuilen (2014) that even risk lovers can be prudent.[24] 48 percent of the subjects exhibit risk aversion

[24] Crainich, Eeckhoudt, and Trannoy (2013) theoretically show that prudent risk lovers devote all their income to saving.

and roughly 60 percent are prudent. The most unlikely combination, with below 3 percent, is being risk loving and imprudent. So utility functions that require risk-loving subjects to be imprudent are not sufficiently flexible to describe our data.

Table 7. Classification of Risk Aversion and Prudence

	Risk averse pct	Risk loving pct	Mixed pct	Total pct
Imprudent	4.76	2.55	0.68	7.99
Prudent	29.25	26.87	1.87	57.99
Mixed	14.46	19.05	0.51	34.01
Total	48.47	48.47	3.06	100.00

Notes: This table reports the share of risk-averse and prudent individuals using the measures described in Section 3. The measures were computed using optimally smoothed spline functions, evaluated at and averaged over 1000 points in the support. Risk neutrality and prudence neutrality is a probability zero event, so none of our subjects was classified, so we omit this category.

6.2.3 Precautionary Savings and Prudence. We now turn to the relationship between our experimental measure of prudence and wealth. As laid out in the theoretical model in Appendix A, we expect to observe a positive relationship between our measure of the strength of prudence and people's accumulated wealth—given present or past income uncertainty. We run linear regressions on net savings on the Crainich and Eeckhoudt-measure of prudence attitudes.[25] We use several sets of control variables motivated by previous studies (e.g., Noussair, Trautmann, and Kuilen, 2014). The results can be found in Table 8.

In all specifications including both the risk averse and the risk loving we find a significant positive relationship between prudence and wealth. When analyzing both subgroups independently, we find a significant positive relationship only for the risk-averse sub-sample. When excluding those that neither save nor are indebted, however, the relationship is significantly positive also for the risk-loving (see Table 15 in Appendix E).

6.2.4 Income Risk. We now turn to the analysis of the impact of individual income risk, prudence and net savings. Several previous empirical studies have identified prudence parameters from consumption volatility

[25] Robustness checks applying the measure by Kimball (1990) can be found in Appendix E.

Table 8. Net Savings and Prudence (C&E)

	Full Sample			Risk Averse		Risk Loving	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Prudence (C&E)	1.140***	1.161***	1.114**	1.707***	1.468***	0.560	0.651
	(0.422)	(0.422)	(0.431)	(0.641)	(0.559)	(0.601)	(0.724)
Risk Aversion (A&P)		3.329*	3.291		0.637		6.038
		(2.002)	(2.134)		(3.120)		(5.325)
Controls	No	No	Yes	No	Yes	No	Yes
Observations	567	567	554	270	267	279	271

Notes: This table reports the results of ordinary least squares regressions on net savings. Prudence (C&E) is the Crainich-Eckhoud measure of prudence. Risk Aversion (A&P) is the Arrow-Pratt measure of risk aversion. The controls are time preferences, gender, age, financial literacy, body mass index (BMI), household members (adults and children), income as measured as the average income per household member, planning horizon and education. Coefficients of controls are reported in the full regression results in Appendix E. We account for potential heteroskedasticity by robust standard errors. Results of t-tests indicated at following significance levels * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

(Dynan, 1993; Fagereng, Guiso, and Pistaferri, 2017a; Guiso, Jappelli, and Terlizzese, 1992). Fagereng, Guiso, and Pistaferri (2017a) instrument consumption volatility with firm specific shocks that pass through to wages. We construct a similar measure of income risk by looking at firm closures by sector, see Section 6.1. Figure 6 shows the distribution of our measure of income risk aggregated at the *Localidad* level.

We regress net income on the prudence measure interacted with the probability of firm closure. The empirical model can be written as follows:

$$\begin{aligned} W \ &= \alpha + \beta_1 \text{ Prudence } + \beta_2 \text{ Shock } + \\ &\quad \beta_3 \text{ Prudence } \times \text{ Shock } + \boldsymbol{\beta_4 X} + \epsilon \end{aligned} . \tag{2}$$

This allows us to answer the question, whether prudent subjects who are confronted with a higher background risk accumulated higher levels of wealth as predicted by our theoretical framework. Results are presented in Table 9.

Column (1) in Table 9 shows the raw correlations. The main effect of prudence is positive and highly significant and so is the interaction term of income risk with prudence. Hence, more prudent people save more when facing higher income risk. This effect is robust to the inclusion of controls in column (2). When restricting the sample to only risk averse agents, the coefficients on the interaction term stay positive, however they are not significant at conventional levels for the interaction term, unless controls are

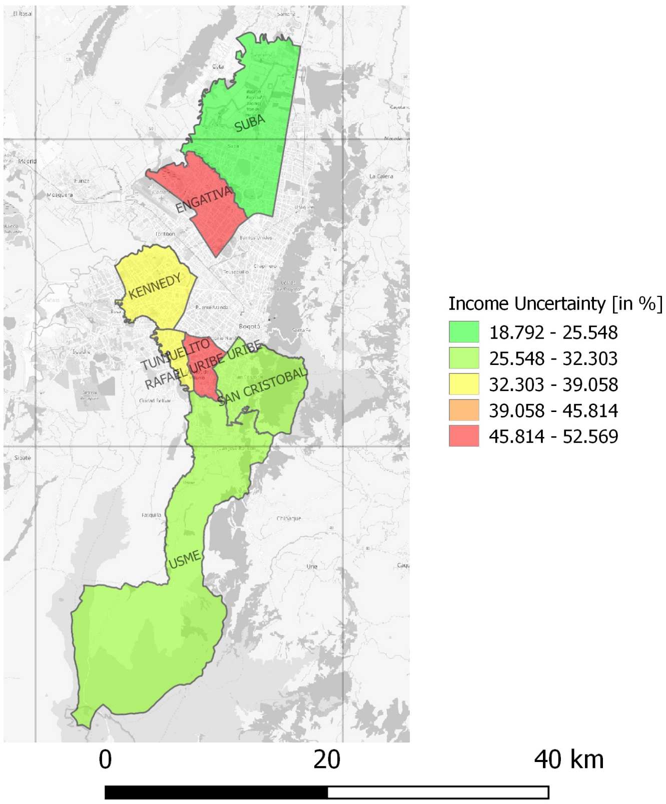


Figure 6. Income Uncertainty in Bogotá: Ratio of Closed to Existing Businesses in 2013.

Notes: Individuals are categorized based on the economic activity they perform according to the ISIC Rev. IIIa A.C. categorization as used e.g., by DANE for their household survey. At this aggregation level, official data on firm closure is available e.g., from the Knowledge Management Board of the Chamber of Commerce of Bogotá. Participants are assigned a level of income uncertainty corresponding to the ratio of firm closure in the working sector they are classified.

Table 9. Net Savings, Firm Closures and Prudence (C&E)

	Full Sample		Risk Averse		Risk Loving	
	(1)	(2)	(3)	(4)	(5)	(6)
Income risk	-1.746	-1.479	-1.863	-1.399	0.768	2.146
	(2.354)	(2.405)	(2.870)	(3.031)	(3.293)	(3.412)
Prudence (C&E)	1.240***	1.115**	2.224***	2.036***	0.298	0.199
	(0.461)	(0.471)	(0.769)	(0.739)	(0.613)	(0.660)
Prudence (C&E) $\times$ Income risk	1.015**	1.066**	0.856	1.037*	0.984*	0.925
	(0.424)	(0.416)	(0.600)	(0.566)	(0.588)	(0.566)
Risk Aversion (A&P)		4.644*		2.231		9.500*
		(2.408)		(4.460)		(5.666)
Controls	No	Yes	No	Yes	No	Yes
Observations	471	459	218	215	237	230

Notes: This table reports the results of ordinary least squares regressions on net savings. Prudence (C&E) is the Crainich-Eckhoud measure of prudence. Risk Aversion (A&P) is the Arrow-Pratt measure of risk aversion. Income risk is measured as the ratio of closed to existing businesses in 2013 in the working sector an individual was usually working in at the time of the survey. Prudence and income risk are centered. The controls are time preferences, gender, age, financial literacy, body mass index (BMI), household members (adults and children), income as measured as the average income per household member, planning horizon and education. Coefficients of controls are reported in the full regression results in Appendix E. We account for potential heteroskedasticity by robust standard errors. Results of t-tests indicated at following significance levels * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

included. For the risk-loving agents, however, we observe a positive and significant coefficient on the raw correlation of the interaction term. After including control variables, this coefficient is only significant if focusing on those that are saving or are indebted (see Table 18 in Appendix E). When using the Kimball measure of prudence we get large, but insignificant coefficients (see Table 16 in Appendix E).

6.3 Summary

Applying our method to a sample of poor households in Bogotá, Colombia we find comparable results with respect to prudence as for example Tarazona-Gomez (2004) in her experiment with students in Bogotá. The results regarding risk aversion are also in line with previous findings (Tarazona-Gomez, 2004), although we find a higher share of risk-loving individuals.

Most interestingly, however, we find strong support for the theory of precautionary saving: According to the theoretical framework going back to Leland (1968), prudent individuals react to income uncertainty by raising their saving. This should lead to higher savings, and those who are more likely to face income uncertainty should hold higher savings. In our extension of the theoretical model in Appendix A, we have shown that those who are more prudent than others with respect to different intensity measures should hold higher savings, irrespective of them being risk averse or risk loving.

This relationship can be found in the data, and it is robust—even when we pool risk-averse and risk-seeking individuals: We find prudence to be strongly linked to a higher level of net savings. While preferences like risk aversion and time preferences have failed to explain a low level of wealth empirically, prudence seems to be of importance. Our results suggests that the sample under study has a high demand for consumption smoothing and would thus profit from a suitable solution—although caution is warranted for the correlational nature of our data.

7 Conclusion

In this paper, we have developed the very first method for non-parametric elicitation of utility-based intensity measures of higher order risk preferences. We thus provide a tool to measure those utility-based measures applied in theoretical work without relying on the commonly used parametric functions that are not flexible enough for working with higher order risk preferences, for example because they cannot model risk preferences of imprudent risk averters, or intemperant risk averters, or, in case of the expo-power family not even those of imprudent agents.

Using a simulation exercise, we illustrate the favorable characteristics of our method. We show that—for the case of the Arrow-Pratt measure of risk aversion—the method performs comparable to the power utility family or the expo-power family with respect to smoothing out measurement errors. Simulating decision errors, we can also reject worries about error propagation affecting the final intensity measures as unjustified, as it leaves the correlation coefficient with the true (assumed) intensity measure *unaffected* to the third digit.

We validate our method using the only alternative to elicit proper intensity measures of higher order risk preferences—the method by Ebert and Wiesen (2014)—, which, however, does not elicit utility-based intensities as used in theoretical work, but premia. To this end, we conduct experiments both online and in the laboratory, with several hundred participants. We find the intensity measures are significantly correlated, although theoretically, the relationship is not even linear. Generally, validation results in the laboratory are stronger than those from the online sample, but also those show highly significant relations. In addition, we provide empirical evidence in favor of the common assumption (e.g., Noussair, Trautmann, and Kuilen, 2014) that choosing the prudent or temperant option more often or more consistently in the risk apportionment tasks (Eeckhoudt and Schlesinger, 2006) is related to stronger prudence and temperance intensities.

In our last chapter, we use these validiated measures for the first direct test of the precautionary saving theory due to Leland (1968) and Kimball (1990), using microdata and measures of the (original) theoretical model (and our extension). We examine the relation of precautionary savings, income risk, and utility-based prudence intensity measures resulting from our method among a sample of more than 650 poor urban households in Bogotá, Colombia, and find—although correlational—strong support for the predictions of the model. More prudent individuals save more compared to less prudent individuals, and in particular so, when facing income risk.

References

Abdellaoui, Mohammed (2000). “Parameter-free elicitation of utility and probability weighting functions”. en. *Management Science* 46 (11): 1497–512. DOI: 10.1287/mnsc.46.11.1497.12080. [11, 16, 69, 70]

Abdellaoui, Mohammed, Carolina Barrios, and Peter Wakker (2007). “Reconciling introspective utility with revealed preference: Experimental arguments based on prospect theory”. en. *Journal of Econometrics* 138 (1): 356–78. DOI: 10.1016/j.jeconom.2006.05.025. [1, 17, 61]

Abdellaoui, Mohammed, Han Bleichrodt, and Corina Paraschiv (2007). “Loss aversion under prospect theory: A parameter-free measurement”. en. *Management Science* 53 (10): 1659–74. DOI: 10.1287/mnsc.1070.0711. [11, 16, 69, 70]

Aiyagari, S. Rao (1, 1994). “Uninsured idiosyncratic risk and aggregate saving”. *The Quarterly Journal of Economics* 109 (3): 659–84. DOI: 10.2307/2118417. [33]

Aldrin, Magne (2006). “Improved predictions penalizing both slope and curvature in additive models”. en. *Computational Statistics & Data Analysis* 50 (2): 267–84. DOI: 10.1016/j.csda.2004.08.002. [14, 65, 69]

Andersen, Steffen, Glenn W. Harrison, Morten I. Lau, and E. Elisabet Rutström (2008). “Eliciting risk and time preferences”. *Econometrica* 76 (3): 583–618. http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0262.2008.00848.x/abstract (visited on 09/02/2015). [34]

Arlot, Sylvain, and Alain Celisse (2010). “A survey of cross-validation procedures for model selection”. en. *Statistics Surveys* 4: 40–79. DOI: 10.1214/09-SS054. [16, 68, 69]

Azrieli, Yaron, Christopher P. Chambers, and Paul J. Healy (2018). “Incentives in Experiments: A Theoretical Analysis”. *Journal of Political Economy* 126 (4): 1472–503. DOI: 10.1086/698136. [23]

Bollaerts, Kaatje, Paul H. C. Eilers, and Iven Mechelen (2006). “Simple and multiple P-splines regression with shape constraints”. en. *British Journal of Mathematical and Statistical Psychology* 59 (2): 451–69. DOI: 10.1348/000711005X84293. [15, 67]

Bowman, David, Deborah Minehart, and Matthew Rabin (1999). “Loss aversion in a consumptions-savings model”. *Journal of Economic Behavior & Organization* 38 (2): 155–78. http://www.sciencedirect.com/science/article/pii/S0167268199000049 (visited on 10/29/2015). [52, 55]

Cardenas, Juan Camilo, and Jeffrey Carpenter (2013). “Risk attitudes and economic well-being in Latin America”. en. *Journal of Development Economics* 103: 52–61. DOI: 10.1016/j.jdeveco.2013.01.008. [4]

Chen, Daniel L., Martin Schonger, and Chris Wickens (2016). “oTree—An open-source platform for laboratory, online, and field experiments”. *Journal of*

Behavioral and Experimental Finance 9: 88–97. DOI: 10.1016/j.jbef.2015.12.001. [26]

Crainich, David, and Louis Eeckhoudt (2008). "On the intensity of downside risk aversion". en. *Journal of Risk and Uncertainty* 36 (3): 267–76. DOI: 10.1007/s11166-008-9037-x. [3, 6–9, 17, 25, 30, 37, 50, 55, 56]

Crainich, David, Louis Eeckhoudt, and Alain Trannoy (2013). "Even (mixed) risk lovers are prudent". *The American Economic Review* 103 (4): 1529–35. [8, 36, 49, 50]

De Boor, Carl (1987). *A practical guide to splines.* Applied Mathematical Sciences. Vol. 27. New York, NY: Springer. [12, 62–64]

Deaton, Angus, and John Muellbauer (1980). *Economics and consumer behavior.* Cambridge Books. Cambridge: Cambridge University Press. [1]

Deck, Cary, and Harris Schlesinger (2010). "Exploring Higher Order Risk Effects". *The Review of Economic Studies* 77 (4): 1403–20. DOI: 10.1111/j.1467-937x.2010.00605.x. [1–4, 8, 22, 30, 31, 73]

Deck, Cary, and Harris Schlesinger (2014). "Consistency of Higher Order Risk Preferences". *Econometrica* 82 (5): 1913–43. DOI: 10.3982/ecta11396. [8, 24, 83, 84, 91]

Denuit, Michel M., and Louis Eeckhoudt (2010). "Stronger measures of higher-order risk attitudes". *Journal of Economic Theory* 145 (5): 2027–36. DOI: 10.1016/j.jet.2010.03.005. [3, 7, 8, 30]

Devaney, Sharon A., Sophia T. Anong, and Stacy E. Whirl (2007). "Household savings motives". *Journal of Consumer Affairs* 41 (1): 174–86. http://onlinelibrary.wiley.com/doi/10.1111/j.1745-6606.2006.00073.x/full (visited on 08/04/2017). [35]

Dirección de Gestión de Conocimiento, Cámara de Comercio de Bogotá (2014). "Tablero de indicadores de Bogotá y Cundinamarca 2014 No. 2". http://hdl.handle.net/11520/8554. http://bibliotecadigital.ccb.org.co/handle/11520/8554. Accessed: 2017-07-06. [33]

Dynan, Karen E. (1, 1993). "How prudent are consumers?" *Journal of Political Economy* 101 (6): 1104–13. DOI: 10.1086/261916. [2, 3, 38, 58]

Ebert, Sebastian, and Daniel Wiesen (2011). "Testing for prudence and skewness seeking". en. *Management Science* 57 (7): 1334–49. DOI: 10.1287/mnsc.1110.1354. [1, 4]

Ebert, Sebastian, and Daniel Wiesen (2014). "Joint measurement of risk aversion, prudence, and temperance". en. *Journal of Risk and Uncertainty* 48 (3): 231–52. DOI: 10.1007/s11166-014-9193-0. [1, 3, 4, 8, 22, 25, 28–30, 42, 72, 87, 92]

Eeckhoudt, Louis, and Harris Schlesinger (2006). "Putting risk in its proper place". *American Economic Review* 96 (1): 280–89. DOI: 10.1257/000282806776157777. [1, 2, 5, 6, 8, 24, 25, 28–30, 42]

Eeckhoudt, Louis, Harris Schlesinger, and Ilia Tsetlin (2009). "Apportioning of Risks Via Stochastic Dominance". *Journal of Economic Theory* 144 (3): 994–1003. DOI: 10.1016/j.jet.2008.11.005. [8, 49]

Eilers, Paul H. C., and Jelle J. Goeman (2004). "Enhancing scatterplots with smoothed densities". en. *Bioinformatics* 20 (5): 623–28. DOI: 10.1093/bioinformatics/btg454. [65, 66]

Eilers, Paul H. C., and Brian D. Marx (1996). "Flexible smoothing with B-splines and penalties". *Statistical Science* 11 (2): 89–102. http://www.jstor.org/stable/2246049 (visited on 04/18/2017). [2, 10, 13, 64, 65, 68, 70]

Eilers, Paul H. C., and Brian D. Marx (2010). "Splines, knots, and penalties". en. *Wiley Interdisciplinary Reviews: Computational Statistics* 2 (6): 637–53. DOI: 10.1002/wics.125. [13, 14, 62, 65]

Eilers, Paul H. C., Brian D. Marx, and Maria Durbán (2015). "Twenty years of P-splines". *SORT-Statistics and Operations Research Transactions* 39 (2): 149–86. http://www.raco.cat/index.php/SORT/article/view/302258 (visited on 04/18/2017). [68]

Ekern, Steinar (1980). "Increasing Nth degree risk". *Economics Letters* 6 (4): 329–33. DOI: 10.1016/0165-1765(80)90005-1. [5–7]

Etchart-Vincent, Nathalie (2004). "Is probability weighting sensitive to the magnitude of consequences? An experimental investigation on losses". *Journal of Risk and Uncertainty* 28 (3): 217–35. http://www.springerlink.com/index/HU5X52164070R336.pdf (visited on 05/11/2017). [11, 16, 69]

Fagereng, Andreas, Luigi Guiso, and Luigi Pistaferri (2017a). "Firm-related risk and precautionary saving response". *American Economic Review: Papers & Proceedings* 107 (5): 393–97. DOI: 10.1257/aer.p20171093. [3, 33, 38]

Fagereng, Andreas, Luigi Guiso, and Luigi Pistaferri (2017b). "Portfolio choices, firm shocks and uninsurable wage risk". *The Review of Economic Studies*, DOI: 10.1093/restud/rdx023. [33]

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde (2018). "Global Evidence on Economic Preferences". *The Quarterly Journal of Economics* 133 (4): 1645–92. DOI: 10.1093/qje/qjy013. eprint: /oup/backfile/content_public/journal/qje/133/4/10.1093_qje_qjy013/3/qjy013.pdf. [23]

Fennema, Hein, and Marcel Van Assen (1998). "Measuring the utility of losses by means of the tradeoff method". *Journal of Risk and Uncertainty* 17 (3): 277–96. http://www.springerlink.com/index/T754025WL88L1022.pdf (visited on 05/11/2017). [11, 16, 69]

Fisher, Patti J., and Catherine P. Montalto (2010). "Effect of saving motives and horizon on saving behaviors". en. *Journal of Economic Psychology* 31 (1): 92–105. DOI: 10.1016/j.joep.2009.11.002. [35]

Fisher, Patti J., and Catherine P. Montalto (2011). "Loss aversion and saving behavior: Evidence from the 2007 U.S. survey of consumer finances". en. *Journal* *of Family and Economic Issues* 32 (1): 4–14. DOI: 10.1007/s10834-010-9196-1. [35]

Friedman, Milton (1957). "The permanent income hypothesis". In *A theory of the consumption function*. Princeton, NJ: Princeton University Press. Chap. 3, 20–37. [54]

Fuchs-Schündeln, Nicola, and Matthias Schündeln (1, 2005). "Precautionary savings and self-selection: Evidence from the German reunification experiment". *The Quarterly Journal of Economics* 120 (3): 1085–120. DOI: 10.1093/qje/120.3.1085. [1, 33, 34]

Gilboa, Itzhak (1989). "Expectation and variation in multi-period decisions". *Econometrica: Journal of the Econometric Society* 57 (5): 1153–69. DOI: 10.2307/1913626. [51, 54]

Greiner, Ben (2015). "Subject pool recruitment procedures: organizing experiments with ORSEE". *Journal of the Economic Science Association* 1 (1): 114–25. DOI: 10.1007/s40881-015-0004-4. [26]

Guiso, Luigi, Tullio Jappelli, and Daniele Terlizzese (1, 1992). "Earnings uncertainty and precautionary saving". *Journal of Monetary Economics* 30 (2): 307–37. DOI: 10.1016/0304-3932(92)90064-9. [3, 38]

Harrison, Glenn W., Morten I. Lau, and Melonie B. Williams (2002). "Estimating individual discount rates in Denmark: A field experiment". *The American Economic Review* 92 (5): 1606–17. http://www.jstor.org/stable/3083267 (visited on 08/04/2017). [36]

Hofmann, Wilhelm, Malte Friese, and Anne Roefs (1, 2009). "Three ways to resist temptation: The independent contributions of executive attention, inhibitory control, and affect regulation to the impulse control of eating behavior". *Journal of Experimental Social Psychology* 45 (2): 431–35. DOI: 10.1016/j.jesp.2008.09.013. [35]

Holt, Charles A., and Susan K. Laury (2002). "Risk aversion and incentive effects". *American Economic Review* 92 (5): 1644–55. http://www.nber.org/~rosenbla/econ311-04/syllabus/holtlaury.pdf (visited on 02/17/2015). [1, 17, 18, 61]

Holzmeister, Felix, Juergen Huber, Michael Kirchler, Florian Lindner, Utz Weitzel, and Stefan Zeisberger (2019). "What Drives Risk Perception? A Global Survey with Financial Professionals and Lay People". DOI: 10.31219/osf.io/v6r9n. [6]

Keenan, Donald C., and Arthur Snow (2002). "Greater downside risk aversion". *Journal of Risk and Uncertainty* 24 (3): 267–77. http://www.springerlink.com/index/tmt7crbkhmcf7q4j.pdf (visited on 04/05/2017). [6, 50, 55]

Kimball, Miles S. (1990). "Precautionary saving in the small and in the large". *Econometrica* 58 (1): 53–73. DOI: 10.2307/2938334. [2, 3, 6–9, 17, 37, 42, 49–51, 55, 57, 58, 75]

Kohavi, Ron (1995). "A study of cross-validation and bootstrap for accuracy estimation and model selection". In *Proceedings of the 14th interna-*

tional joint conference on Artificial intelligence. vol. 2, 1137–43. https://pdfs.semanticscholar.org/0be0/d781305750b37acb35fa187febd8db67bfcc.pdf (visited on 05/08/2017). [68]

Kőszegi, Botond, and Matthew Rabin (2009). "Reference-dependent consumption plans". en. *American Economic Review* 99 (3): 909–36. DOI: 10.1257/aer.99.3.909. [52, 55]

Leland, Hayne E. (1968). "Saving and uncertainty: The precautionary demand for saving". *The Quarterly Journal of Economics* 82 (3): 465–73. DOI: 10.2307/1879518. [1–3, 6, 7, 41, 42, 49, 50, 52, 53, 55–57]

Loewenstein, George (1987). "Anticipation and the valuation of delayed consumption". *The Economic Journal* 97 (387): 666–84. DOI: 10.2307/2232929. [50]

Loewenstein, George F, and Dražen Prelec (1993). "Preferences for sequences of outcomes." *Psychological review* 100 (1): 91. [54]

Lusardi, Annamaria, and Olivia Mitchell (2011). "Financial Literacy and Planning: Implications for Retirement Wellbeing". Techreport. DOI: 10.3386/w17078. [34]

Maier, Johannes, and Maximilian Rüger (2011). "Experimental Evidence on Higher Order Risk Preferences with Real Monetary Losses". Munich Discussion Paper 2010-33. Volkswirtschaftliche Fakultät, Ludwig-Maximilians-Universität München. http://www.wu.ac.at/courses/other/egrie2011/programme/papers/iv.a1.rueger.pdf (visited on 01/20/2016). [1, 4]

Marx, Brian D., and Paul H. C. Eilers (2002). "Multivariate calibration stability: A comparison of methods". en. *Journal of Chemometrics* 16 (3): 129–40. DOI: 10.1002/cem.701. [14, 65]

McCord, Mark, and Richard de Neufville (1986). "Lottery Equivalents: Reduction of the certainty effect problem in utility assessment". *Management Science* 32 (1): 56–60. DOI: 10.1287/mnsc.32.1.56. [2]

Menezes, C., C. Geiss, and J. Tressler (1980). "Increasing Downside Risk". *The American Economic Review* 70 (5): 921–32. [6]

Modica, Salvatore, and Marco Scarsini (31, 2017). "A note on comparative downside risk aversion". en. *Journal of Economic Theory* 122 (2): 267–71. DOI: 10.1016/j.jet.2004.06.008. [6, 7, 50, 55]

Moffitt, Terrie E., Louise Arseneault, Daniel Belsky, Nigel Dickson, Robert J. Hancox, HonaLee Harrington, Renate Houts, Richie Poulton, Brent W. Roberts, Stephen Ross, Malcolm R. Sears, W. Murray Thomson, and Avshalom Caspi (15, 2011). "A gradient of childhood self-control predicts health, wealth, and public safety". *Proceedings of the National Academy of Sciences* 108 (7): 2693–98. DOI: 10.1073/pnas.1010076108. [35]

Noussair, Charles N., Stefan T. Trautmann, and Gijs van de Kuilen (2014). "Higher order risk attitudes, demographics, and financial decisions". en. *The Review of Economic Studies* 81 (1): 325–55. DOI: 10.1093/restud/rdt032. [1–4, 8, 9, 17, 18, 22, 24, 29–31, 33, 34, 36, 37, 42, 49, 73]

Pratt, John W. (1964). "Risk Aversion in the Small and in the Large". *Econometrica* 32 (1/2): 122. DOI: 10.2307/1913738. [3, 6, 16, 56]

Reinsch, Christian H. (1967). "Smoothing by spline functions". *Numerische Mathematik* 10 (3): 177–83. http://www.springerlink.com/index/K37198281Q833WL4.pdf (visited on 05/07/2017). [64]

Rothschild, Michael, and Joseph E Stiglitz (1970). "Increasing risk: I. A definition". *Journal of Economic Theory* 2 (3): 225–43. DOI: 10.1016/0022-0531(70)90038-4. [1, 5]

Runge, Carl (1901). "Über empirische Funktionen und die Interpolation zwischen äquidistanten Ordinaten". *Zeitschrift für Mathematik und Physik* 46: 224–43. [12]

Sandmo, A. (1970). "The effect of uncertainty on saving decisions". *The Review of Economic Studies* 37 (3): 353–60. DOI: 10.2307/2296725. [6]

Schneider, Sebastian O., and Matthias Sutter (2021). "Higher Order Risk Preferences: New Experimental Measures, Determinants and Field Behavior". Mimeo. [2]

Shao, Jun (1993). "Linear model selection by cross-validation". *Journal of the American Statistical Association* 88 (422): 486–94. DOI: 10.2307/2290328. [68]

Starmer, Chris, and Robert Sugden (1991). "Does the Random-Lottery Incentive System Elicit True Preferences? An Experimental Investigation". *The American Economic Review* 81 (4): 971–78. http://www.jstor.org/stable/2006657. [23]

Tarazona-Gomez, Marcela (2004). "Are individuals prudent? An experimental approach using lottery choices". Technical Report. EHESS Toulouse. http://www2.toulouse.inra.fr/lerna/cahiers2005/05.13.177.pdf (visited on 01/25/2016). [4, 41]

Van Rooij, Maarten CJ, Annamaria Lusardi, and Rob JM Alessie (2012). "Financial literacy, retirement planning and household wealth". *The Economic Journal* 122 (560): 449–78. http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0297.2012.02501.x/full (visited on 08/04/2017). [34]

Wakai, Katsutoshi (2008). "A model of utility smoothing". *Econometrica* 76 (1): 137–53. http://www.jstor.org/stable/4502056 (visited on 09/25/2017). [51, 54]

Wakker, Peter (2008). "Explaining the characteristics of the power (CRRA) utility family". en. *Health Economics* 17 (12): 1329–44. DOI: 10.1002/hec.1331. [59, 60]

Wakker, Peter, and Daniel Deneffe (1996). "Eliciting von Neumann-Morgenstern utilities when probabilities are distorted or unknown". *Management Science* 42 (8): 1131–50. DOI: 10.1287/mnsc.42.8.1131. [2, 10]

A A Two-Period Savings and Consumption Model

The seminal work by Leland (1968) linked a positive third derivative of the Bernoulli utility function in expected utility theory to precautionary saving. A key finding is that (precautionary) savings can be independent of risk aversion.[26] In further developments, Kimball (1990) introduced a measure for the intensity of prudence that indicates the strength of the precautionary saving motive in a two-period model of consumption and saving assuming time-separable utility, However, Kimball focuses on risk averters (characterized by a concave utility function), neglecting the existence of risk lovers (convex utility function), whereas the model of Leland refrains from imposing such restrictions.

Focusing on risk averters only, as recently been shown to be too restrictive. Building on work by Eeckhoudt, Schlesinger, and Tsetlin (2009) on "mixed risk averters" who like to combine good with bad outcomes, Crainich, Eeckhoudt, and Trannoy (2013) elegantly demonstrated the equivalence between a preference for combining good with good outcomes and a positive second, third and fourth derivative, corresponding to risk loving, prudent and intemperant behavior; this pattern they attribute to "mixed risk lovers" who agree with mixed risk averters on the sign of uneven derivatives, but disagree on even derivatives. Using a simple two period-model assuming time-separable utility, they show that prudent risk lovers devote all their income to saving in the presence of a future income risk. Empirical support for the existence of mixed risk averters and mixed risk lovers is given by Noussair, Trautmann, and Kuilen (2014), who study the prevalence of risk aversion, prudence and temperance in a student sample and the general population in the Netherlands. They find about 15 percent of their sample showing risk-loving behavior, and that "prudence is more prevalent than temperance". What is more, they find that "[t]he degree of temperance seems to be more closely related to the degree of risk aversion than prudence is". In their student sample, they find that the rank correlation between prudence and risk aversion is not significant (and negative), while the rank correlation between risk aversion and temperance is significant and positive.[27] Although they pool risk-averse and risk-loving subjects in their

[26] Noussair, Trautmann, and Kuilen (2014) confirm this conjecture empirically for a sample of Dutch individuals: While prudence is correlated with saving and wealth measures, risk aversion is not.

[27] When pooling student sample and the general population, both rank correlations are positive and significant, however, the correlation between risk aversion and prudence is still

analysis, they find a strong connection between prudence and saving and wealth, which is robust to controlling for risk aversion levels and temperance, suggesting that this connection also holds for the 15 percent showing risk-loving behavior. However, results are not presented for the risk-loving and the risk-averse sub-samples separately.

When considering risk lovers while investigating the relationship between prudence and saving, the intensity measure of prudence put forward by Kimball (1990), $-u'''/u''$ is inapplicable. The measure u'''/u', advocated for by Crainich and Eeckhoudt (2008), building on earlier work by Modica and Scarsini (2017) and Keenan and Snow (2002) is an interesting candidate, as it is defined independently of the second derivative of the utility function. We show that this measure is a good measure for the intensity of the precautionary demand for saving, building on the original and more general framework of saving and consumption by Leland (1968). In particular, one prediction emerging from our model is that also risk lovers may save a non-trivial fraction of their income proportional to the intensity measure of prudence u'''/u', which extends the finding by Crainich and Eeckhoudt (2008) in the direction of allowing risk lovers to save a non-trivial fraction of their income—proportional to the measure of prudence u'''/u'.

We first present a consumption and saving model based on the work by Leland (1968). As in the original work, we do not assume overall utility to be additively time-separable. Besides resulting in less realistic predictions regarding the savings fraction of risk lovers, the rather strong assumption of first and second period consumption utility being perfect substitutes implied by additive time-separability of overall utility has been empirically debated in the literature on time-discounting at least since G. Loewenstein (1987).[28] Double-check: Moreover, a time-separable utility function is inconsistent with empirical findings in the area of consumption and saving. For example, excess smoothness in aggregate consumption growth compared to aggregate income growth, a well-documented phenomenon, cannot be explained by a time-separable utility function (**Pagel**; **MichaelidisLudwig**). Thus, there is growing empirical evidence that this assumption might be too restrictive to explain individual and aggregate consumption and saving decisions.

---

lower than that between risk aversion and temperance. Moreover, temperance significantly increases with risk aversion, which is not the case for prudence.

[28] In fact, with time-separable utility, risk lovers either save all or none of their income (see e.g., Crainich, Eeckhoudt, and Trannoy, 2013).

Different models of inter-temporal decision-making, starting with Gilboa (1989), have been axiomatizing a utility path dependence (see also Wakai, 2008, Axiom 3 for a more recent example). In the spirit of Gilboa (1989), we assume that individuals dislike variation in consumption and experience disutility when consumption varies over time. This simple extension allows to explain a precautionary motive for saving also for risk-loving individuals—proportional to the measure of prudence u'''/u'.

A.1 A two-period model of consumption and saving

We assume that an agent lives for two periods and receives income w_t for $t \in 1, 2$. The income in the first period is deterministic, whereas in the second period, w_2, is random with known expectation and variance given by $\mathbb{E}[w_2] = \bar{w_2}$ and $\mathbb{E}[(w_2 - \bar{w_2})^2] = \sigma^2$, respectively.

The agent has access to financial markets, where the fraction of income saved in $t = 1$, $k \leq 1$, receives an interest rate $r > 0$.[29] Consumption in $t = 1$ is $c_1 = (1 - k)w_1$ and consumption in $t = 2$ is $c_2 = w_2 + (1 + r)kw_1$.

The agent's objective is to maximize the expected inter-temporal utility of consumption $\mathbb{E}[U(c_1, c_2)]$ by deciding on the fraction of income, k, that they would save in $t = 1$. The inter-temporal utility of consumption is given by:

$$U(c_1, c_2) \;=\; u_1(c_1) + u_2(c_2) + g(c_2 - c_1), \tag{3}$$

where $u_t(c_t)$ is the three times differentiable utility function of consumption for $t \in 1, 2$ and g denotes a three times differentiable reference-dependent value function that takes consumption in the first period as reference. If consumption in the second period is larger than consumption in the first period the function takes a positive value, reflecting a positive surprise. If the opposite is true, individuals suffer a utility loss that is proportional to the drop in consumption. For simplicity, we assume that the discount rate is equal to 1.

We assume that the reference dependent value function is concave. By incorporating a reference-dependent value function, we modify the assumption of additive time-separable utility previously adopted by Kimball (1990).

[29] In case of negative savings, k is restricted such that c_2 will always be non-negative.

Such an additive term relating consumption of the two periods has been applied before in the context of intertemporal consumption and saving models by Bowman, Minehart, and Rabin (1999) or Kőszegi and Rabin (2009).

We assume positive marginal utility of consumption ($u'_t > 0$) but do not make any assumption about the second derivative of the utility function and allow u''_t to be either (weakly) positive, indicating risk aversion (neutrality) or negative, indicating risk-loving behavior.

The first- and second-order condition for an interior solution imply that—in absence of uncertainty—the following conditions are satisfied for the optimal consumption bundle (c_1^*, c_2^*) resulting from the optimal saving rate k^*:

$$\frac{dU}{dk} = (1+r)\frac{\partial U}{\partial c_2} - \frac{\partial U}{\partial c_1} = 0 \tag{FOC}$$

$$\frac{d^2U}{dk^2} = \frac{d}{dk}\left[(1+r)\frac{\partial U}{\partial c_2} - \frac{\partial U}{\partial c_1}\right] < 0. \tag{SOC}$$

For an additive U with a reference-dependent value function as in (3), (SOC) writes

$$u''_1(c_1) + u''_2(c_2)(1+r)^2 + g''(c_2 - c_1)(2+r)^2 < 0.$$

Clearly, in case u_1 and u_2 are strictly concave in c_t (corresponding to a *risk-averse* individual) and when $g'' \leq 0$, the second order condition is satisfied. Moreover, if g'' is negative and the absolute value of the last term is larger than the first two in (SOC), then overall utility is concave in k. Hence, the second order condition (SOC) is satisfied, even if u_1 and u_2 are convex as in the case of a risk-loving individual.

To study how income uncertainty affects the optimal level of the saving decision, we follow Leland. Taylor series expansions of $\mathbb{E}\left[\partial U / \partial c_t\right]$ in the resulting first-order condition around the optimal consumption bundle in absence of uncertainty indicate that the effect of uncertainty on the optimal saving rate depends on the sign of

$$\left(\frac{\partial^3 U^*}{\partial c_1 \partial (c_2)^2} - (1+r)\frac{\partial^3 U^*}{\partial (c_2)^3}\right)\sigma^2, \tag{4}$$

where U^* is defined as U evaluated at the consumption bundle resulting from the optimal saving rate in absence of income uncertainty.[30] Leland (1968) infers that if (4) is negative, we may "[u]nder reasonable regularity assumptions [...] say that the optimal [saving rate] will be larger [...] when uncertainty is present, and the more uncertainty, the greater will be the optimal [saving rate]." This implies that the saving rate will be larger when income is uncertain if the following expression is negative:

$$-\left((2+r)\,g^{*'''}(c_2-c_1)+(1+r)u_2^{*'''}(c_2)\right)\sigma^2. \tag{5}$$

This corresponds to (4) for utility as given in (3). Thus, uncertainty results in a positive precautionary demand for saving for prudent individuals (those for which $u_2''' > 0$) when $g''' \geq 0$, provided that the necessary condition of a negative second derivative of overall utility with respect to the saving rate (SOC) is satisfied. We summarize this in the following proposition:

Proposition 1 *Let the overall inter-temporal utility of an agent be given by $U(c_1, c_2) = u_1(c_1) + u_2(c_2) + g(c_2 - c_1)$. Assume g such that $d^2U/dk^2 < 0$. Then a non-negative third derivative of the reference-dependent value function g and a positive third derivative of the second period utility, i.e. $g''' \geq 0$ and $u_2''' > 0$, are a sufficient (though not necessary) condition for a positive precautionary demand for saving under the assumption of the model of precautionary saving by Leland (1968). This statement holds independently of the sign of the second derivative of the second period utility.*

A.2 Exemplary reference-dependent value function

So far the 'reference-dependent value function' has been characterized as a concave function of the difference of consumption in the first and in the second period and hence as a concave function of the proportion k of income saved. We now provide an example of a particular reference-dependent value function. We assume that

$$g(x) \;:=\; -lx^2 \tag{6}$$

with $l \geq 0$; for l large enough, $d^2U/dk^2 < 0$ and the extreme of the utility function at the critical point satisfying the first-order condition is a utility maximum (for a risk-averse individual, $l = 0$ is large enough).

[30] More generally, for a function $f(k)$, we define f^* as f evaluated at the optimal saving rate k^* in absence of uncertainty.

Introducing a function $g(c_2 - c_1) = -l(c_2 - c_1)^2$ for $l > 0$ decreases overall utility with an increasing difference between consumption in the first period and consumption in the second period. This special choice of the reference-dependent value function incorporates variation aversion (for $l > 0$) as observed empirically by G. F. Loewenstein and Prelec (1993), which challenged time-separable utility. Gilboa (1989) axiomatically derived a utility function with path-dependence, where overall utility is decreased as the difference between consumption in the first and consumption in the second period grows, holding any of the two constant. More specifically, Gilboa (1989) proposed a reference-dependent value function of the form $|u(c_2) - u(c_1)|$. In a more recent work, Wakai (2008), translates the idea of variation aversion in a setting where negative variation is more unpleasant than positive variation is pleasant.

The way we incorporate variation aversion is an analytically convenient version of the utility function proposed by Gilboa (1989), which also decreases overall utility as the difference between consumption in the first and consumption in the second period grows, but our function does so independently of the absolute levels of consumption. This simplification might seem strict, as for high consumption levels, a relatively small difference might loom less than the same absolute difference for low consumption levels or vice versa. As the scaling parameter l might capture these individual wealth levels, however, we argue it is appropriate.

With respect to life-time saving, our choice of the reference-dependent value function is a way to incorporate the aim of consumption smoothing as predicted by the permanent income hypothesis (Friedman, 1957) or simply the commitment of living and consuming in the next period.

This choice of utility could, additionally to any risk-averse individual, also represent an individual that is willing to take risk in each period, but only as long as the difference between consumption in both periods does not get too large—impeding 'ruthless' over-consumption in any of both periods.

If $g(c_2 - c_1) = -l(c_2 - c_1)^2$ as in the example above, then $g''' \equiv 0$ and we rewrite (5):

$$-(1 + r)u_2^{*'''}(c_2)\sigma^2. \tag{7}$$

From this we see that just as in the case of simple additive, time-separable utility, a positive precautionary saving demand results solely from a positive third derivative of second period utility with respect to k.

We summarize these findings in a corollary:

Corollary 1 *Let the overall intertemporal utility of an agent be given by* $U(c_1, c_2) = u_1(c_1) + u_2(c_2) - l(c_2 - c_1)^2$*. Assume* l *large enough such that* $d^2U/dk^2 < 0$*. Then a positive third derivative of the second period utility, i.e.* $u_2''' > 0$*, is both a sufficient and a necessary condition for a positive precautionary demand for saving under the assumption of the model of precautionary saving by Leland (1968). This statement holds independently of the sign of the second derivative of the second period utility.*

Corollary 1 states that—given the reference-dependent value function, and thus the overall utility, is of the most simple form allowing for a regular utility maximum—a positive third derivative of the second period utility alone causes a positive precautionary demand for saving, for risk-loving and risk-averse individuals, where for the latter, $l = 0$ is large enough.

Unfortunately, the existence of a reference-dependent value function and its particular shape is not directly testable. This drawback is also inherent in the models by Kőszegi and Rabin (2009) and Bowman, Minehart, and Rabin (1999).

A.3 Measuring the Strength of the Precautionary Saving Motive

We now show, building on the model by Leland (1968), that the measure by Crainich and Eeckhoudt (2008) can also be directly interpreted as a measure of intensity of a precautionary savings demand whereas the measure by Kimball (1990) is restricted to the case of risk-averse individuals. This second measure is adequate when comparing risk-averse individuals only, but cannot generally be used in the framework of Leland's model or the extension of the model that we present here.

Crainich and Eeckhoudt measure Building on previous work on downside risk aversion,[31] Crainich and Eeckhoudt (2008) suggest measuring the degree of prudence by $\pi = u'''/u'$ (Keenan and Snow, 2002; Modica and Scarsini, 2017). The intuitive interpretation they give of this measure is the analog to

[31] Downside risk aversion is equivalent to prudence for three times differentiable utility functions.

the utility premium for compensating the pain of a zero-mean risk. When there is 'misapportionment of risk' (meaning risk added to a state of lower wealth instead of to the state of higher wealth), $\pi = u'''/u'$ is proportional to the money equivalent of pain induced by this misapportionment. One advantage of this measure is, that it is independent of the sign of the second derivative. Hence, it is closely related to the theoretical prediction of the model by Leland (1968) as it can be computed for both risk-averse and risk-loving individuals leading to a similar interpretation.

To see how the measure by Crainich and Eeckhoudt (2008) can be incorporated in the Leland (1968) framework, rewrite (4) as

$$\left(\frac{\partial^3 U^*}{\partial c_1 \partial (c_2)^2} - \frac{\partial U^*}{\partial c_1} \Big/ \frac{\partial U^*}{\partial c_2} \frac{\partial^3 U^*}{\partial (c_2)^3}\right) \sigma^2 \tag{8}$$

$$= \left(\frac{\partial^3 U^*}{\partial c_1 \partial (c_2)^2} - \frac{\partial U^*}{\partial c_1} \frac{\partial^3 U^*}{\partial (c_2)^3} \Big/ \frac{\partial U^*}{\partial c_2}\right) \sigma^2, \tag{9}$$

where we used the first-order condition (FOC) and rearranged terms.

If the utility is additively time-separable (i.e. $g \equiv l \equiv 0$), (8) can be written as

$$-u_1^{*\prime} \pi^* \sigma^2. \tag{10}$$

Hence, it is clear that the savings rate increases with the intensity of prudence $\pi^* = \frac{\partial^3 U^*}{\partial (c_2)^3} / \frac{\partial U^*}{\partial c_2} = \frac{u_2^{*\prime\prime\prime}}{u_2^{*\prime}}$.

Let us now turn to the case of a utility with an additive reference-dependent value function of the form $-l(c_2 - c_1)^2$ relating consumption in the two periods. (8) in this case equals (5) with $g''' \equiv 0$. We focus on the second term,

$$-(1 + r) u_2^{*\prime\prime\prime}(c_2) \tag{11}$$

and find that the larger $u_2^{*\prime\prime\prime}$, the larger the precautionary savings demand under the conditions derived before. Also in this case, π is a good measure of the precautionary savings demand. First, because dividing $u_2^{*\prime\prime\prime}$ by $u_2^{*\prime}$ leaves the sign unchanged. Second, following the rationale by Pratt (1964) when justifying his measure, multiplying u with a positive constant does not change behavior, but it changes u'''. The measure π is unaffected by such a transformation.

We summarize these findings:

Proposition 2 *Let the overall intertemporal utility of an agent be given by* $U(c_1, c_2) = u_1(c_1) + u_2(c_2) - l(c_2 - c_1)^2$. *Assume* l *large enough such that* $d^2U/dk^2 < 0$. *Then, all else equal,* $m^* = \frac{\partial^3 U^*}{\partial(c_2)^3} / \frac{\partial U^*}{\partial c_2}$ *indicates the strength of a precautionary demand for saving under the assumption of the model of precautionary saving by Leland (1968), including regularity assumptions. This statement holds independently of the sign of the second derivative of the second period utility.*

Note that for risk-averse individuals, Proposition 2 holds for $l = 0$, i.e. under the usual assumption of time-separable utility.

Kimball (1990) measure The first measure of the degree of intensity of prudence was proposed by Kimball (1990). Using a close analogy to the Arrow-Pratt measure of risk aversion, the intensity of prudence is defined as $-u'''/u''$. In a simple two-period model with additive time-separable utility, Kimball (1990) shows that the savings function of a globally more prudent individual at a given level of saving moves upward at a lower level of risk. This measure thus is directly related to the intensity of the precautionary saving motive.

The measure by Kimball (1990) has two shortcomings that are especially relevant when trying to apply the concept empirically, see Appendix B for an example. First, since the measure depends on the second derivative of the (per-period) utility function, it implies that precautionary savings depend on risk aversion. However, Leland (1968) shows that the precautionary demand for savings is independent of the degree of risk aversion.

Second, when focusing on precautionary savings, Kimball (1990) neglects the possibility of a convex utility function. Hence, the proposed measure is only meaningful for risk-averse individuals, as for them, it shows a positive value when prudent, but not for risk-loving individuals, for whom its value is negative if prudent. Similarly, this measure yields a positive intensity of prudence for an imprudent individual (negative third derivative) that is risk loving, leading to a contradiction.

B Dynan (1993)

Dynan (1993) assumes a concave utility, which is additive over time and establishes the following link between expected consumption growth, prudence and the variation in consumption growth:

$$\mathbb{E}_t \left[\frac{c_{i,t+1} - c_{i,t}}{c_{i,t}} \right] = \frac{1}{\rho_r} \left(\frac{r_i - \delta}{1 + r_i} \right) + \frac{\xi}{2} \mathbb{E}_t \left[\left(\frac{c_{i,t+1} - c_{i,t}}{c_{i,t}} \right)^2 \right], \tag{12}$$

where $\mathbb{E}_t$ is the expectation conditional on information available at time t, $c_{i,t}$ represents consumption, δ is the constant time preference rate, and r_i is the household's interest rate. $\rho_r = -c_{i,t}(u''/u')$ is the coefficient of relative risk aversion and $\xi = -c_{i,t}(u'''/u'')$ is the coefficient of relative prudence as defined by Kimball (1990). This equation has been derived using a Taylor series approximation for $u'(c_{i+1,t})$ around $c_{i,t}$:

$$u'(c_{i,t+1}) \approx u'(c_{i,t}) + u''(c_{i,t})(c_{i,t+1} - c_{i,t}) + \frac{u'''(c_{i,t})}{2}(c_{i,t+1} - c_{i,t})^2. \tag{13}$$

Note that this approximation does not impose any structural constraints on any of the derivatives. This result is combined with the first order condition Dynan (1993, Equation 3) resulting from solving the maximization problem as stated in Dynan (Equation 1 1993):

$$\left(\frac{1 + r_i}{1 + \delta} \right) \mathbb{E}_t \left[u'(c_{i,t+1}) \right] = u'(c_{i,t}). \tag{14}$$

What is problematic about that approach is that ξ as used in equation (12) does only correspond to Kimball's definition, when $u'''(c_{i,t})u''(c_{i,t}) < 0$. Only then, it can be interpreted as a measure of prudence with respect to a precautionary savings demand. From the description of the empirical approach applied by Dynan (1993), it does not seem like this is taken into account. In addition, the estimations of $1/\rho_r$ (the coefficients of $(r_i - \delta)/(1 + r_i)$) and its large standard errors suggest that the sign of ρ_r is not necessarily positive.

C Shortcomings of Parametric Utility Functions for the Study of Higher Order Risk Preferences

Widely used parametric functions are not suited for (empirically) studying higher order risk preferences, because they are too stylized to possess the flexibility to combine any shape of the utility function with any shape of the second, third and higher order derivatives. This is in particular true for utility functions belonging to the family of switching sign utilities (i.e. functions where $\operatorname{sgn}(u^{(n)}) = -\operatorname{sgn}(u^{(n-1)})$). According to these functions, a risk-averse individual (negative second derivative) would always be classified as prudent (positive third derivative) and temperant (negative fourth derivative), whereas risk-seeking individuals (positive second derivative) would always be classified as imprudent and intemperant (negative third and positive fourth derivative, respectively).

We illustrate this shortcoming for the power (CRRA) utility family.

Power utility family For $x > 0$, the family is defined as

$$u(x) = \begin{cases} x^b & \text{if } b > 0 \\ \ln(x) & \text{if } b = 0 \\ -x^b & \text{if } b < 0. \end{cases}$$

Note that, if u is an interval scale (meaning u is unique up to unit and level), it can be multiplied by any positive factor and any constant can be added without affecting any relevant empirical aspect (Wakker, 2008). This is, in mathematical terms, a monotonic transformation, that does not affect the maximization process. In particular, u as defined above is then equivalent to alternative formulations (with restricted domain, i.e. where $b < 1$) such as

$$u(x) = \begin{cases} \frac{x^{1-\eta}-1}{1-\eta} & \text{if } \eta > 0, \eta \neq 1 \\ \ln(x) & \text{if } \eta = 1, \end{cases}$$

where $b = \eta - 1$.

Thus, we analyze the shape of utility and its derivatives according to the power family as defined in its general definition. The second derivative is given by

$$u''(x) = \begin{cases} \frac{\partial}{\partial x} bx^{b-1} = b(b-1)x^{b-2} & \text{if } b > 0 \\ \frac{\partial}{\partial x} \frac{1}{x} = -\frac{1}{x^2} & \text{if } b = 0 \\ \frac{\partial}{\partial x} - bx^{b-1} = -b(b-1)x^{b-2} & \text{if } b < 0. \end{cases}$$

Note that as $x > 0$, the first derivative of u is positive for all $b \in \mathbb{R}$. The second derivative is negative and the utility itself concave for $b < 1$ and thus, in the expected utility framework would correspond to a risk averse individual.

The third derivative is given by

$$u'''(x) = \begin{cases} \frac{\partial}{\partial x} b(b-1)x^{b-2} = b(b-1)(b-2)x^{b-3} & \text{if } b > 0 \\ \frac{\partial}{\partial x} - \frac{1}{x^2} = \frac{1}{x^3} & \text{if } b = 0 \\ \frac{\partial}{\partial x} - b(b-1)x^{b-2} = -b(b-1)(b-2)x^{b-3} & \text{if } b < 0. \end{cases}$$

For $b \leq 0$, we see immediately that the third derivative is positive. For $0 < b < 1$ and $b > 2$, we have a positive third derivative. For $1 < b < 2$, we have a negative third derivative.

Thus, only for b in the interval $(1, 2)$, could the third derivative be negative. On that interval however, the second derivative is always positive, and thus—using this utility family—an imprudent individual can never be risk averse. Similarly, a risk averse individual can never be imprudent.

Exponential utility family We now turn to the exponential (CARA) utility family. Assuming u is unique up to unit and level, the formulation of this family again does not depend on multiplicative factors or the addition of constants (Wakker, 2008).

We define the exponential family for $x > 0$ as:

$$u(x) = \begin{cases} (1 - e^{-bx})/b & \text{if } b \neq 0 \\ x & \text{if } b = 0. \end{cases}$$

Its derivatives are given by

$$u'''(x) = \begin{cases} \frac{\partial^2}{\partial x^2} e^{-bx} = \frac{\partial}{\partial x} - be^{-bx} = (-b)^2 e^{-bx} & \text{if } b \neq 0 \\ \frac{\partial^2}{\partial x^2} 1 = \frac{\partial}{\partial x} 0 = 0 & \text{if } b = 0. \end{cases}$$

We see that the third derivative can never be negative and thus, assuming this utility family, we would never classify an individual as imprudent. Further, the fourth derivative will always have the same sign as the second derivative, and so we can never classify an individual as intemperant and risk averse relying on the utility functions of this family.

Expo-Power Family The expo-power family has been proposed by Abdellaoui, Barrios, and Wakker (2007) and was used e.g., by Holt and Laury (2002). It exhibits an increasing measure of relative risk aversion in x and a decreasing measure of absolute risk aversion in x for $0 < b < 1$ and x in the interval $[0, 1]$ (resulting from normalizations of x, see Abdellaoui, Barrios, and Wakker, 2007). On the interval $(0, 1]$, it is defined by

$$u(x) = \begin{cases} -e^{-x^b/b} & \text{if } b \neq 0 \\ -1/x & \text{if } b = 0. \end{cases}$$

We take a look at its derivatives:

$$u''(x) = \begin{cases} \frac{\partial}{\partial x} e^{-x^b/b} x^{b-1} = e^{-x^b/b} x^{b-2} \left(-x^b + b - 1\right) & \text{if } b \neq 0 \\ \frac{\partial}{\partial x} 1/x^2 = -2/x^3 & \text{if } b = 0. \end{cases}$$

For $b = 0$, the second derivative is always negative and the function itself is concave. For $b \neq 0$, the sign of the second derivative depends on the sign of $-x^b + b - 1$. It is negative for $b - 1 < x^b$. As $x \in (0, 1]$, this is the case for $b < 1$. Contrarily, for $b > 2$, the above term is always positive and so is the second derivative, implying a convex utility function.

Let's now turn to the third derivative:

$$u'''(x) = \begin{cases} \frac{\partial}{\partial x} u''(x) = e^{-x^b/b} x^{b-3} \left(b^2 + x^{2b} + 3x^b - 3b(x^b + 1) + 2\right) & \text{if } b \neq 0 \\ \frac{\partial}{\partial x} u''(x) = 6/x^4 & \text{if } b = 0. \end{cases}$$

For $b \neq 0$, the sign of the third derivative depends on the term $b^2 + x^{2b} + 3x^b - 3b(x^b + 1) + 2$. Numerically, one finds that this term is uniformly negative in $x \in (0, 1]$ for values of b between roughly 1.27 and 2. That is, for a risk-loving or a risk-averse individual (i.e. individuals to which a parameter value of $r < 1$ and $r > 2$ correspond, respectively) there is at least one point $x \in (0,1]$, for which the third derivative is positive. Therefore, neither a risk averse nor a risk loving individual will ever be classified as imprudent according to this utility family.

D Methodology: Details

D.1 An Introduction To Spline Regression

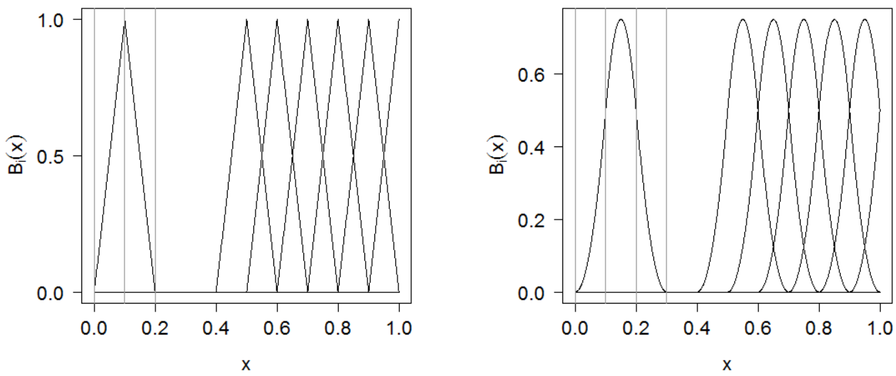
For spline regression, the domain of definition $[x_{\text{min}}, x_{\text{max}}]$ is divided into $k-1$ subintervals, where the k boundaries are called inner knots. Local basis functions, defined depending on the knots, are placed (often equidistantly, thus independent of the data) such that they cover the domain of definition.

One such choice of basis functions, so called B-splines, have proven to be numerically stable and efficient for computation.[32] A single B-spline of degree[33] p is a combination of $p+1$ polynomial pieces of degree p that are joined $p-1$ times continuously differentiable at the knots. It is different from zero only on a small subinterval of the domain of definition (spanned by $p+2$ knots) and zero otherwise. Figure 7 illustrates single B-splines of degree $p=1$ and degree $p=2$. For illustration purposes, the first $p+2$ inner knots are indicated by the gray vertical lines at $x=0.1$, 0.2 and for the B-spline of degree 2 additionally at $x=0.3$.

We use a B-spline basis consisting of $k+p-1$ equally spaced B-splines spanned by $k+2p$ knots; see Figure 2(a) for an illustration of an exemplary B-splines basis. Also in Figure 2(a) we see that at any point x in the interval $[x_{\text{min}}, x_{\text{max}}]$, a B-spline basis decomposes 1, i.e. $\forall x \in [x_{\text{min}}, x_{\text{max}}] : \sum_{j=1}^{k+p-1} B_j(x,p) = 1$, where we denote the value of the jth B-spline (the B-spline with support interval starting at knot j) at x with $B_j(x,p)$ and where p is the degree of the B-spline. For regressing the y-values of a set of N data points (x_i, y_i) on the B-spline basis $(B_j)_{j=1,\ldots,k+p-1}$,

[32] De Boor (1987) gives a recursive formula for computation of B-splines from a lower degree B-spline Since a B-spline of degree 0 is just a constant between subsequent knots, this facilitates computation. However, B-splines can also be constructed as linear combinations of truncated power functions. Eilers and Marx (2010) show that it is numerically stable. We use the latter approach for computation of our B-spline basis.

[33] Note that, in the B-spline literature, usually *order* is used instead of *degree*, where *order = degree + 1*. In the P-spline literature however, degree is preferred, as order mostly is referring to the order of differences used in the penalty.



(a) B-spline of Degree $p = 1$ (b) B-spline of Degree $p = 2$

Figure 7. Illustration of a Single B-spline and the Corresponding B-spline Basis

we evaluate the $k + p - 1$ B-splines at the given x-values, which yields the $(N \times (k + p - 1))$ design matrix $\boldsymbol{B}$. In matrix notation, the regression approach is to minimize

$$Q_B(\boldsymbol{\alpha}) = \|\boldsymbol{y} - \boldsymbol{B\alpha}\|^2, \tag{15}$$

and the result is a fitted curve $\hat{y}(x) = \sum_{j=1}^{k+p-1} \hat{a}_j B_j(x, p)$.

A particularly useful feature of B-spline regression for our problem is established by the following formula for the m-th derivative of a B-spline function $f(x)$:

$$f^{(m)}(x) = \frac{1}{h^m} \sum_{j=m+1}^{k+p-1} \triangle^m a_j B_j(x, p - m), \tag{16}$$

where h is the knot distance, and $\triangle^m a_j = \triangle(\triangle^{m-1} a_j)$ with $\triangle a_j := (a_j - a_{j-1})$. For a derivation of this result, see Appendix D.2 or De Boor (1987, Ch. 10).

Equation (16) illustrates that the derivatives of a spline function can be computed conveniently by differencing its B-spline coefficients. Once a fitted curve is established, its derivatives are automatically obtained without the need to determine them numerically.

D.2 Derivatives of a B-spline Function

As laid out in De Boor (1987, Ch. 10, equations (12) & (16)), the derivative of a B-spline function $f(x)$ is given by

$$\begin{aligned} f^{(1)}(x) &= \frac{\partial f(x)}{\partial x} = \frac{\partial}{\partial x} \sum_{j=1}^{k+p-1} a_j B_j(x,p) = \sum_{j=1}^{k+p-1} a_j \frac{\partial}{\partial x} B_j(x,p) \\ &= \sum_{j=1}^{k+p} (a_j - a_{j-1}) \frac{1}{h} B_j(x, p-1) = \frac{1}{h} \sum_{j=1}^{k+p} \triangle a_j B_j(x, p-1), \end{aligned} \tag{17}$$

where h is the knot distance, and $\triangle a_j := (a_j - a_{j-1})$, $a_0 := 0$ and $a_{k+p} := 0$. Note that, sticking to the indices, the B-splines $B_j(x, p-1)$ vanish on the interval $[x_{\min}, x_{\max}]$ for $j < 2$ and $j > k + p - 1$. Accounting for this fact and iteratively applying (17) yields:

$$\begin{aligned} f^{(m)}(x) &= \frac{1}{h^m} \sum_{j=1}^{k+p+m-1} \triangle^m a_j B_j(x, p-m) \\ &= \frac{1}{h^m} \sum_{j=m+1}^{k+p-1} \triangle^m a_j B_j(x, p-m), \end{aligned} \tag{18}$$

where $\triangle^m a_j = \triangle(\triangle^{m-1} a_j)$ is the m-order difference of the sequence (a_j).

D.3 Choice of Penalty: Jointly Smoothing Multiple Derivatives

For their exposition of P-splines, Eilers and Marx (1996) use an unspecified order d of penalization. A penalty based on the second derivative was introduced in the smoothing context by Reinsch (1967), mainly "because it leads to a very simple algorithm". Probably, a penalty of order 2 is still the most common penalty used. However, Eilers and Marx (1996) note that, besides convenient computation, there is no specific reason for this choice.

Here, we are interested in smoothing the utility function itself, and additionally the third and the fourth derivative of the utility function. In this regard, Eilers and Marx (1996) note that the "kth difference penalty is a good discrete approximation to the integrated square of the kth derivative"[34]. Further, for a penalty of order d, the fitted curve approaches a poly-

[34] They illustrate this point with a penalty on second order differences.

nomial of degree $d - 1$, as the penalty increases (Eilers and Marx, 1996). Lastly, interpolation is affected by the order of the penalty: With a penalty of order d, interpolation of utility is of degree $2d - 1$. This means that the third derivative of the interpolating curve has degree $2d - 4$ and the fourth derivative has degree $2d - 5$ (Eilers and Marx, 2010).

Considering these aspects suggests using a penalty of order 3 or 4, where the latter is to be preferred for the limiting behavior of the spline function when the penalty increases and for the degree of the fourth derivative of the interpolation curve. However, this choice leads to fluctuations in the third derivative not caused by the data, so additionally introducing a penalty on the third derivative is necessary.

The objective function then writes

$$Q_B(\boldsymbol{\alpha}) \;=\; \|\boldsymbol{y} - \boldsymbol{B\alpha}\|^2 + \omega_3 \|\boldsymbol{D_3 \alpha}\|^2 + \omega_4 \|\boldsymbol{D_4 \alpha}\|^2. \tag{19}$$

Penalization of multiple orders has been applied in other studies, remarkably with a focus on the quality of prediction (i.e. interpolation): Marx and Eilers (2002) introduced the use of multiple orders independently, including a ridge penalty (corresponding to order $d = 0$) in addition to any penalty order $d = 1, 2, 3$. Aldrin (2006) shows in a simulation experiment that the prediction performance when penalizing both slope ($d = 0$) and curvature ($d = 2$) is always at least as good as penalizing curvature only.

Our goal is to smooth the third and fourth derivative jointly corresponding to using both penalties of order $d = 3$ and $d = 4$. When choosing an optimal parameter with respect to prediction quality, the third order penalty and the fourth order penalty are to a high degree exchangeable. As the third order differences of the B-spline coefficients generally will be much higher than the fourth order differences, the third order penalty will in general dominate the fourth order penalty, causing the effect of the latter to vanish and resulting in unnecessary fluctuations of the fourth derivative. Using a penalty of order $d = 4$ alone, however, results in unnecessary fluctuations of the third derivative.

This issue has been dealt with by Eilers and Goeman (2004), who present the first approach we are aware of to jointly smooth multiple orders $d > 0$. They study how signals consisting of largely flat areas combined with a sharp pulse can be smoothed. This phenomenon resembles to a certain degree the pattern of a considerable share of our utility functions, where parts of nearly zero marginal utility follow parts of sharp increases or vice versa. Eilers and Goeman (2004) use a combination of a first and second order penalty by setting $\omega_2 = \nu\omega_1^2$, corresponding to penalty terms of order $d = 2$ and $d = 1$, where they found by trial and error that $\nu = 1/4$ works well.

We consider a couple of hundreds of utility functions, and in this case, visual inspection is clearly too time-consuming. Furthermore, the choice of ν should be independent of subjective judgment as it could affect the classification and intensities of risk preferences. Therefore, we develop a data-driven approach.

We propose a solution in which the fourth order penalty 'drives the shape' of the utility function while the third order penalty is limited to avoiding unnecessary fluctuations in the third derivative. This is achieved by using one penalty parameter, appropriately scaled for both orders.

Specifically, we choose the scaling parameter ν such that

$$\omega_4\|\boldsymbol{D_4\alpha_0}\| \;\approx\; \nu\omega_4\|\boldsymbol{D_3\alpha_0}\|,$$

i.e. such that the penalty terms have approximately equal weight, where we used $\nu = 0.001$ for a computation of $\|\boldsymbol{D_3\alpha_0}\|$. Then, we set $\omega_3 = \nu\omega_4/5$, to ensure $\omega_4\|\boldsymbol{D_4\alpha}\| > \omega_3\|\boldsymbol{D_3\alpha}\|$.

The choice of ν for a first computation of $\|\boldsymbol{D_3\alpha_0}\|$ practically neglects any third order penalty and a good fit using solely the fourth order penalty is achieved. Then, ν is set as a fifth of the ratio of the sum of absolute differences of the fourth order differences of the B-spline coefficients over the third order differences. The ratio measures how much stronger a third order penalty will affect the smoothing process as compared to a fourth order penalty and will in most cases be the maximum factor needed to weight both penalties equally. To avoid the third order penalty dominating the fourth order penalty, we divide this ratio by five, which proved to yield the desired behavior.

Thus, we elaborate an approach which is to be preferred when visual inspection is inappropriate—be it for objectivity reasons or for time constraints impeding the researcher to visually investigate a large amount of individual curves.

D.4 Incorporating a Monotonicity Constraint In P-Spline Regression

The rationale of the approach by Bollaerts, Eilers, and Mechelen (2006) to approximately incorporate a monotonicity constraint in P-Spline regression is simple: Although in P-Spline regression, a penalization term is added to the objective function, the predicted function is still a B-spline function, and thus, its derivatives are given by equations (17) and (16). From these formulae, a sufficient condition for the first derivative to be positive and the utility to be a monotone increasing function can be easily deduced: Since h, the knot distance, is positive and B non-negative for all x, p and j, all $\triangle a_j$ have to be positive.

Thus, differences of coefficients of the B-splines that are negative have to be penalized, whereas positive differences do not, which makes the penalty asymmetric. This penalization is achieved with the following penalty:

$$\sum_{j=2}^{k+p-1} w(\alpha)_j (\triangle a_j)^2,$$

where

$$w(\alpha)_j = \begin{cases} 0, & \text{if } \triangle a_j \geq 0 \\ 1, & \text{otherwise.} \end{cases}$$

The objective function (19) now writes

$$Q_B(\boldsymbol{a}) = \|\boldsymbol{y} - \boldsymbol{B\alpha}\|^2 + \nu\omega_4\|\boldsymbol{D_3\alpha}\|^2 + \omega_4\|\boldsymbol{D_4\alpha}\|^2 + \kappa\|\boldsymbol{W^{1/2}D_1\alpha}\|^2 \tag{20}$$

which—in case Q_B is convex—is minimized if

$$(\boldsymbol{B'B} + \nu\omega_4\boldsymbol{D_3'D_3} + \omega_4\boldsymbol{D_4'D_4} + \kappa\boldsymbol{D_1'WD_1})\hat{\boldsymbol{\alpha}} = \boldsymbol{B'y},$$

where $\boldsymbol{y}$, $\boldsymbol{B}$, $\boldsymbol{D_d}$ and ω_4 are defined as in (19) and ν as determined in Section D.3. $\boldsymbol{W}$ and $\boldsymbol{W^{1/2}}$ are diagonal matrices with elements $w(\alpha)_j$ and $\sqrt{w(\alpha)}$, respectively, and the impact of the constraint penalty on the solution is tuned by a sufficiently high[35] (positive) constraint parameter κ.

Bollaerts, Eilers, and Mechelen (2006) show that Q_B is convex in $\boldsymbol{\alpha}$ and propose using a Newton-Raphson procedure to find an optimal solution. We

[35] We chose $\kappa = 10^8$, Bollaerts, Eilers, and Mechelen (2006) chose $\kappa = 10^6$ in their application.

follow this suggestion, stopping the algorithm after 10 iterations, which led to a monotone increasing function in 99.5% of all cases.

D.5 Choosing the Degree of Smoothness

Eilers and Marx (1996) suggest two classical objective and data-driven selection criteria, AIC and (leave-one-out) cross validation, for the choice of the degree of smoothness of the P-spline function to be established as tuned by ω in order to balance between fidelity to the data and smoothness.

Cross validation (CV) is independent of distributional or asymptotic assumptions and "should be preferred to any model selection procedure relying on assumptions which are likely to be wrong" (Arlot and Celisse, 2010). In our case, for the computation of AIC, at least the asymptotic argument that the approximation of standard errors relies on is unlikely to hold for our moderate number of data points. We thus apply cross validation as selection criteria for choosing ω_4 as in (20). The principle of CV is simple: The model is fit with only a part of the data and the remainder is used to compute prediction errors. With leave-one-out CV, for N data-points, the model is fit N times, and each time, one data point is left out and used for computation of the prediction error. Then, the model—in our setting the penalty parameter—is chosen that minimizes the average prediction error over N predictions. For this case, exact formulae for convenient computation exist without the need to actually fit the model N times (Eilers and Marx, 1996; Eilers, Marx, and Durbán, 2015).

In literature, however, it has been noted that leave-one-out CV is not the ideal choice regarding model choice (see e.g., Kohavi, 1995, and the references therein). One argument against leave-one-out CV is that the probability of choosing the model with the best predictive quality does not converge to 1 as the number of observation increases (Shao, 1993). Further, Eilers, Marx, and Durbán (2015) warn that leave-one-out CV severely overfits the data in case of correlated observations.

We address this issue with the following strategy: As proposed in literature for the purpose of model identification (e.g., Arlot and Celisse, 2010; Shao, 1993), we increase the number of points left out for prediction, which—fortunately, with a maximum of 9 elicited points—is computationally still feasible to do exhaustively. More specifically, we perform permuted leave-k-out cross validation (Aldrin, 2006): We choose the degree of smoothness, that minimizes

$$\frac{1}{V}\sum_{v=1}^{V}\sum_{i\in I^{(v)}}\{y_i - \hat{y}_{(-v)}(x_i)\}^2, \tag{21}$$

where $V = \binom{N}{k}$ is the number of possibilities to chose k out of N points for validation, $I^{(v)}$ is the set containing the vth choice of k points for validation and $\hat{y}_{-v}(x_i)$ the prediction of y_i, obtained by estimating the model using all points but those in $I^{(v)}$.

Increasing the number of points left out for prediction results in less 'weight' for a single point; it can therefore be seen as a mean of error correction, as long as the remaining points are sufficient to establish a meaningful utility curve. As pointed out in the main text, reversal rates of one third are common in choice tasks[36]. For this reason, we perform 'leave-at-least-$1/3N$-out' cross validation. In case the maximum number of utility points was elicited for the individual under study,[37] this choice results in leave-3-out CV. This strategy also accounts for correlated observations (Arlot and Celisse, 2010, Chapter 8.1) possibly resulting from the chain structure of the experiment.

In addition, we develop and apply a data-driven minimum for the penalty parameter to rule out overfitting resulting from large distances between utility points. We compute the number of balls with radius equal to the knot distance needed to cover the elicited points. For a minimum of two balls, overfitting the data is impossible and the method has to compromise between data-fidelity and smoothness. For a maximum of nine balls, however, and for low values of ω_4, the fitted function usually perfectly predicts every data point used for estimation, and minimizes the prediction error for those points left out for validation. Thus, according to CV, the minimal ω_4 is chosen in those cases—but the fitted function is considerably overfitting the data. We impede this by setting a higher minimal penalty parameter in those cases.

[36] See e.g.,Abdellaoui (2000), Abdellaoui, Bleichrodt, and Paraschiv (2007), Etchart-Vincent (2004), and Fennema and Van Assen (1998)

[37] Note that we excluded the points $(0, 0)$ and $(1, 1)$ for computing the average prediction error.

Specifically, for individual i, the minimal smoothness parameter is calculated using the following formula:

$$\omega_{4,i}^{\min} = (b_i * (n_{\max}/n_i) - 1)^{2.5} \tag{22}$$

where b_i is the number of balls with radius equal to the knot-distance needed to cover all elicited utility points of individual i, $n_{\max}$ is the maximum number of elicited points possible for all individuals (we have $n_{\max} = 9$) and n_i is the number of elicited points for individual i.

If the maximum of 9 points are elicited, and all points have a pairwise distance above the knot-distance, then the number of balls to cover all points will be 9, and the minimal value for $\omega_{4,i}$ will be roughly 180. This value is still low enough to allow the fitted function to be a polynomial of degree $p > 3 = d - 1$, but in most cases, it is high enough to prevent overfitting. In some cases, the data is overfit, indicating that the minimal smoothness parameter is chosen conservatively. If all elicited points in $(0, 1)$ lie close together, the minimal smoothness parameter $\omega_{4,i}$ would be 1, i.e. a minimal smoothness parameter that results in hardly any penalization.

For some individuals, less than 9 utility points were elicited, since due to the implementation of the protocol followed (Abdellaoui, Bleichrodt, and Paraschiv, 2007), in some choice tasks, one option is stochastically dominated and the resulting utility point has to be erased following, (e.g., Abdellaoui, 2000). If the number of elicited utility points is less or equal to the order of penalty, we have to perform leave-one-out CV according to the formula by Eilers and Marx (1996). In those cases, overfitting is only prevented by the increased minimal ω_4 in case of sparse information per data knot as expressed by formula (22).

E Full Tables and Robustness Tests

E.1 Validation

Table 10. Distribution of risk and prudence attitude

Panel A: Online		Risk averse	Risk loving	All
Non-prudent	%	12.30	5.75	18.04
Prudent	%	62.20	19.76	81.96
All	%	74.50	25.50	100.00

Panel B: Lab		Risk averse	Risk loving	All
Non-prudent	%	16.67	4.39	21.05
Prudent	%	71.05	7.89	78.95
All	%	87.72	12.28	100.00

Table 11. Prediction of (Higher Order) Risk Premia with Utility-Based Intensity Measures

Panel (a): Laboratory

	Risk Premium		Prudence Premium (EW)		Temperance Premium (EW)	
	(1)	(2)	(1)	(2)	(1)	(2)
AP Risk ($-u''/u'$)	0.790***	0.255***				
	(0.087)	(0.035)				
CE Prudence (u'''/u')			0.807***	0.169***		
			(0.126)	(0.048)		
DE Temperance ($-u^{(iv)}/u'$)					0.695***	0.201*
					(0.118)	(0.081)
Cognitive Abilities		0.795***		0.874***		0.888***
		(0.026)		(0.043)		(0.046)
Num.Obs.	114	114	60	60	60	60
R2	0.624	0.969	0.420	0.927	0.168	0.878
R2 Adj.	0.620	0.968	0.410	0.924	0.153	0.874

Panel (b): Online

	Risk Premium		Prudence Premium (EW)		Temperance Premium (EW)	
	(1)	(2)	(1)	(2)	(1)	(2)
AP Risk ($-u''/u'$)	0.650***	0.489***				
	(0.027)	(0.023)				
CE Prudence (u'''/u')			0.518***	0.364***		
			(0.055)	(0.047)		
DE Temperance ($-u^{(iv)}/u'$)					0.190***	0.108*
					(0.043)	(0.042)
Cognitive Abilities		0.496***		0.477***		0.569***
		(0.012)		(0.028)		(0.026)
Num.Obs.	992	992	308	308	308	308
R2	0.422	0.642	0.277	0.486	0.042	0.364
R2 Adj.	0.421	0.642	0.275	0.482	0.038	0.360

Notes: The utility-based intensity measures of higher order risk preferences, i.e., the Arrow-Pratt coefficient of risk aversion, the Crainich-Eeckhoudt coefficient of prudence and the Denuit-Eeckhoudt coefficient of temperance result from our method. Prudence and temperance premia are elicited using the method by Ebert and Wiesen (2014). Risk premium is the certainty equivalent of the lottery $(0, 0.5, 140)$ subtracted from 1. All measures are expressed in standard deviations.
+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table 12. Prediction of (Higher Order) Risk Premia with Utility-Based Intensity Measures

Panel (a): Laboratory

	Prudence			Temperance		
	Pooled	N	DS	Pooled	N	DS
Coef.	0.127***	0.104***	0.155***	0.080**	0.062**	0.102*
	(0.018)	(0.017)	(0.034)	(0.024)	(0.021)	(0.048)
Num.Obs.	114	59	55	114	59	55
R2	0.290	0.389	0.270	0.082	0.144	0.075
R2 Adj.	0.283	0.379	0.256	0.074	0.130	0.058

Panel (b): Online

	Prudence			Temperance		
	Pooled	N	DS	Pooled	N	DS
Coef.	0.164***	0.178***	0.151***	0.068***	0.060***	0.077***
	(0.008)	(0.014)	(0.010)	(0.012)	(0.017)	(0.018)
Num.Obs.	684	345	339	684	345	339
R2	0.322	0.315	0.335	0.046	0.035	0.060
R2 Adj.	0.321	0.313	0.333	0.044	0.033	0.057

Notes: This table reports results from OLS regressions, where the dependent variables are the utility-based intensity measures of higher order risk preferences, i.e., the Crainich-Eeckhoudt coefficient of prudence (columns 1-3) and the Denuit-Eeckhoudt coefficient of temperance (columns 4-6), as resulting from our method (see Section 3). Explanatory variables are the number of prudent and temperant choices, respectively, in the risk apportionment tasks as implemented by Noussair, Trautmann, and Kuilen (2014) with the "dice-design" (columns 2 and 4, labelled 'N') and by Deck and Schlesinger (2010) with the "fortune-wheel design" (columns 3 and 6, labelled 'DS'). Prudence and temperance intensity measures are expressed in standard deviations.
+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

E.2 Bogotá saving experiment

Table 13. Net Savings and Prudence (C&E) Showing Coefficients on Covariates

	Full Sample			Risk Averse		Risk Loving	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Prudence (C&E)	1.140⋆⋆⋆	1.161⋆⋆⋆	1.114⋆⋆	1.707⋆⋆⋆	1.468⋆⋆⋆	0.560	0.651
	(0.422)	(0.422)	(0.431)	(0.641)	(0.559)	(0.601)	(0.724)
Risk Aversion (A&P)		3.329⋆	3.291		0.637		6.038
		(2.002)	(2.134)		(3.120)		(5.325)
Male			-7.647		-9.299		-2.863
			(6.847)		(8.889)		(10.34)
Age			0.101		0.0610		0.426⋆
			(0.152)		(0.186)		(0.218)
Financial literacy			-1.601⋆		0.390		-3.841⋆⋆
			(0.933)		(1.109)		(1.702)
BMI			-0.161		-0.846		0.783
			(0.618)		(0.809)		(1.000)
Adult HH members			-2.588		-5.937⋆		-0.357
			(1.868)		(3.196)		(2.720)
Children HH members			0.192		1.434		0.0408
			(2.243)		(2.934)		(3.564)
Income			-0.695		2.053		-2.382
			(1.392)		(1.909)		(2.085)
Impatience			0.0535		0.0442		0.170
			(0.169)		(0.240)		(0.255)
Increase in patience			0.246		0.338		0.196
			(0.243)		(0.288)		(0.396)
Planning horizon			ref.		ref.		ref.
– Next months			9.439⋆		3.429		16.65⋆
			(5.138)		(6.698)		(8.949)
– Next year			-1.060		2.360		-4.154
			(9.930)		(15.58)		(13.28)
– Next two to five years			13.02⋆		5.171		31.24⋆⋆
			(6.971)		(8.873)		(14.90)
– 5 or more years			11.20⋆		7.431		15.19
			(6.667)		(11.00)		(10.89)
Constant	-14.69⋆⋆⋆	-14.64⋆⋆⋆	-0.453	-11.78⋆⋆⋆	17.46	-17.04⋆⋆⋆	-44.73
	(2.722)	(2.705)	(22.32)	(3.443)	(30.90)	(4.321)	(29.45)
Education	No	No	Yes	No	Yes	No	Yes
Observations	567	567	554	270	267	279	271

Notes: This table reports the results of ordinary least squares regressions on net savings. Prudence (C&E) is the Crainich-Eckhoud measure of prudence. Risk Aversion (A&P) is the Arrow-Pratt measure of risk aversion. The controls are time preferences, gender, age, financial literacy, body mass index (BMI), household members (adults and children), income as measured as the average income per household member, planning horizon and education. *Income* in 100k Colombian pesos. We account for potential heteroskedasticity by robust standard errors. Results of t-tests indicated at following significance levels ⋆ $p < 0.10$; ⋆⋆ $p < 0.05$; ⋆⋆⋆ $p < 0.01$.

Table 14. Net Savings and Prudence (Kimball) Showing Coefficients on Covariates

	(1)	(2)	(3)	(4)	(5)	(6)
Prudence (Kimball)	12.73*		7.236		7.245	
	(7.428)		(4.774)		(6.190)	
Prudence (C&E)		2.191*		1.096*		1.293
		(1.206)		(0.619)		(0.795)
Risk Aversion (A&P)			12.55	13.47	10.24	10.18
			(9.618)	(10.36)	(9.797)	(9.678)
Male					6.227	5.018
					(10.17)	(9.790)
Age					0.164	0.190
					(0.333)	(0.333)
Financial literacy					0.915	0.798
					(2.237)	(2.213)
BMI					-1.093	-1.244
					(1.081)	(1.156)
Adult HH members					-4.062	-4.327
					(4.260)	(4.258)
Children HH members					2.146	2.291
					(4.330)	(4.326)
Income					4.647*	4.850*
					(2.756)	(2.808)
Impatience					0.742**	0.745**
					(0.363)	(0.360)
Increase in patience					-0.0836	-0.109
					(0.219)	(0.216)
Planning horizon					ref.	ref.
– Next months					-2.176	-0.177
					(15.66)	(15.13)
– Next year					32.58	33.40
					(20.36)	(21.08)
– Next two to five years					10.77	9.571
					(17.66)	(15.76)
Constant	-14.31***	-15.79***	-33.37*	-35.49*	-37.68	-33.41
	(5.178)	(5.937)	(19.03)	(20.54)	(42.81)	(42.04)
Education	No	No	No	No	Yes	Yes
Observations	120	120	120	120	119	119

Notes: This table reports the results of ordinary least squares regressions on net savings. The prudence measure by Kimball is only defined for subjects that are risk averse at all points of evaluation, which reduces the number of observations to 120. For comparison, columns (2), (4) and (6) show the results for the C&E measure for the same sample. The controls are time preferences, gender, age, financial literacy, body mass index (BMI), household members (adults and children), income as measured as the average income per household member, planning horizon and education. *Income* in 100k Colombian pesos. We account for potential heteroskedasticity by robust standard errors. Results of t-tests indicated at following significance levels * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table 15. Net Savings and Prudence (C&E) for Non-zero Net Savings Showing Coefficients on Covariates

	Full Sample			Risk Averse		Risk Loving	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Prudence (C&E)	2.297***	2.435***	2.749***	2.992**	2.629***	1.561	3.565**
	(0.861)	(0.863)	(0.804)	(1.168)	(0.974)	(1.358)	(1.712)
Risk Aversion (A&P)		7.030*	10.25**		-0.207		17.00
		(4.033)	(4.608)		(6.468)		(11.98)
Male			-17.84		-13.73		-13.11
			(14.30)		(17.98)		(23.27)
Age			-0.179		-0.660		0.858
			(0.380)		(0.472)		(0.575)
Financial literacy			-3.645*		2.987		-7.631**
			(2.024)		(3.019)		(2.963)
BMI			0.379		-0.615		2.338
			(1.314)		(1.700)		(2.451)
Adult HH members			-9.435**		-9.395		-9.444
			(4.334)		(5.929)		(7.890)
Children HH members			1.379		4.760		5.386
			(4.531)		(5.790)		(7.426)
Income			0.264		3.970		-1.639
			(2.272)		(3.577)		(3.549)
Impatience			-0.0205		-0.360		0.547
			(0.358)		(0.541)		(0.561)
Increase in patience			0.319		0.608		0.329
			(0.453)		(0.610)		(0.748)
Planning horizon			ref.		ref.		ref.
– Next months			14.86		0.637		31.79*
			(10.57)		(13.85)		(18.92)
– Next year			-2.107		4.974		-17.91
			(17.67)		(26.44)		(19.41)
– Next two to five years			40.30***		89.71**		71.81**
			(15.06)		(38.82)		(34.55)
– 5 or more years			27.02*				35.82
			(15.54)				(22.27)
Constant	-31.94***	-31.54***	-38.58	-23.41***	51.30	-37.80***	-202.0*
	(5.712)	(5.586)	(67.88)	(6.700)	(75.05)	(9.305)	(115.8)
Education	No	No	Yes	No	Yes	No	Yes
Observations	257	257	249	124	124	127	120

Notes: This table reports the results of ordinary least squares regressions on net savings. The sample is restricted to subjects who have non-zero net-savings. Prudence (C&E) is the Crainich-Eckhoud measure of prudence. Risk Aversion (A&P) is the Arrow-Pratt measure of risk aversion. The controls are time preferences, gender, age, financial literacy, body mass index (BMI), household members (adults and children), income as measured as the average income per household member, planning horizon and education. *Income* in 100k Colombian pesos. We account for potential heteroskedasticity by robust standard errors. Results of t-tests indicated at following significance levels * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table 16. Net Savings, Firm Closures and Prudence (C&E & Kimball) Showing Coefficients on Covariates

	Full Sample C&E		Risk Averse C&E		Strictly Risk Averse Kimball	
	(1)	(2)	(3)	(4)	(5)	(6)
Income risk	-1.746	-1.479	-1.863	-1.399	-4.985	-4.154
	(2.354)	(2.405)	(2.870)	(3.031)	(6.150)	(7.472)
Prudence	1.240***	1.115**	2.224***	2.036***	18.37*	9.263
	(0.461)	(0.471)	(0.769)	(0.739)	(9.852)	(9.574)
Prudence × Income risk	1.015**	1.066**	0.856	1.037*	9.654	12.99
	(0.424)	(0.416)	(0.600)	(0.566)	(9.694)	(9.948)
Risk Aversion (A&P)		4.644*		2.231		18.98
		(2.408)		(4.460)		(18.73)
Male		-10.50		-16.27		10.19
		(8.132)		(10.95)		(13.49)
Age		-0.00373		-0.155		-0.269
		(0.194)		(0.275)		(0.468)
Financial literacy		-1.469		1.966		3.523
		(1.208)		(1.465)		(3.197)
BMI		0.0508		-1.284		-1.391
		(0.721)		(1.146)		(1.473)
Adult HH members		-1.627		-3.763		-1.443
		(2.026)		(3.145)		(3.969)
Children HH members		-2.266		-1.202		-2.708
		(2.506)		(3.452)		(5.328)
Income		-0.889		3.040		5.866
		(1.637)		(2.577)		(4.058)
Impatience		-0.103		-0.127		0.558
		(0.181)		(0.325)		(0.406)
Increase in patience		0.381		0.431		-0.259
		(0.297)		(0.400)		(0.329)
– Next months		9.617		7.929		12.88
		(6.426)		(9.350)		(24.90)
– Next year		1.775		14.00		53.13
		(10.76)		(17.33)		(32.59)
– Next two to five years		14.13		27.74		55.94*
		(9.795)		(25.44)		(29.22)
– 5 or more years		13.87		-6.131		
		(9.759)		(11.93)		
Constant	-15.33***	-6.033	-12.58***	20.66	-16.45**	-61.69
	(2.995)	(29.66)	(4.158)	(39.63)	(6.542)	(71.05)
Controls	No	Yes	No	Yes	No	Yes
Observations	471	459	218	215	93	92

Notes: This table reports the results of ordinary least squares regressions on net savings. Prudence is the Crainich-Eckhoud measure of prudence in columns (1) to (4) and the Kimball measure in columns (5) and (6). Risk Aversion (A&P) is the Arrow-Pratt measure of risk aversion. Income risk is measured as the ratio of closed to existing businesses in 2013 in the working sector an individual was usually working in at the time of the survey. Prudence and income risk are centered. The controls are time preferences, gender, age, financial literacy, body mass index (BMI), household members (adults and children), income as measured as the average income per household member, planning horizon and education. *Income* in 100k Colombian pesos. We account for potential heteroskedasticity by robust standard errors. Results of t-tests indicated at following significance levels * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table 17. Net Savings, Firm Closures and Prudence (C&E) Showing Coefficients on Covariates

	Full Sample		Risk Averse		Risk Loving	
	(1)	(2)	(3)	(4)	(5)	(6)
Income risk	-1.746	-1.479	-1.863	-1.399	0.768	2.146
	(2.354)	(2.405)	(2.870)	(3.031)	(3.293)	(3.412)
Prudence (C&E)	1.240***	1.115**	2.224***	2.036***	0.298	0.199
	(0.461)	(0.471)	(0.769)	(0.739)	(0.613)	(0.660)
Prudence (C&E) × Income risk	1.015**	1.066**	0.856	1.037*	0.984*	0.925
	(0.424)	(0.416)	(0.600)	(0.566)	(0.588)	(0.566)
Risk Aversion (A&P)		4.644*		2.231		9.500*
		(2.408)		(4.460)		(5.666)
Male		-10.50		-16.27		-4.157
		(8.132)		(10.95)		(12.26)
Age		-0.00373		-0.155		0.391
		(0.194)		(0.275)		(0.244)
Financial literacy		-1.469		1.966		-4.361**
		(1.208)		(1.465)		(2.085)
BMI		0.0508		-1.284		1.928*
		(0.721)		(1.146)		(0.994)
Adult HH members		-1.627		-3.763		0.386
		(2.026)		(3.145)		(3.102)
Children HH members		-2.266		-1.202		-3.604
		(2.506)		(3.452)		(3.174)
Income		-0.889		3.040		-3.419
		(1.637)		(2.577)		(2.248)
Impatience		-0.103		-0.127		0.0365
		(0.181)		(0.325)		(0.237)
Increase in patience		0.381		0.431		0.270
		(0.297)		(0.400)		(0.437)
Planning horizon		ref.		ref.		ref.
– Next months		9.617		7.929		12.01
		(6.426)		(9.350)		(9.271)
– Next year		1.775		14.00		-12.33
		(10.76)		(17.33)		(14.24)
– Next two to five years		14.13		27.74		33.20*
		(9.795)		(25.44)		(18.35)
– 5 or more years		13.87		-6.131		19.77
		(9.759)		(11.93)		(15.42)
Constant	-15.33***	-6.033	-12.58***	20.66	-17.37***	-53.10
	(2.995)	(29.66)	(4.158)	(39.63)	(4.519)	(33.94)
Controls	No	Yes	No	Yes	No	Yes
Observations	471	459	218	215	237	230

Notes: This table reports the results of ordinary least squares regressions on net savings. The sample is restricted to subjects who have non-zero net-savings. Prudence (C&E) is the Crainich-Eckhoud measure of prudence. Risk Aversion (A&P) is the Arrow-Pratt measure of risk aversion. Income risk is measured as the ratio of closed to existing businesses in 2013 in the working sector an individual was usually working in at the time of the survey. Prudence and income risk are centered. The controls are time preferences, gender, age, financial literacy, body mass index (BMI), household members (adults and children), income as measured as the average income per household member, planning horizon and education. *Income* in 100k Colombian pesos. We account for potential heteroskedasticity by robust standard errors. Results of t-tests indicated at following significance levels * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table 18. Net Savings, Firm Closures and Prudence (C&E) for Non-zero Net Savings Showing Coefficients on Covariates

	Full Sample		Risk Averse		Risk Loving	
	(1)	(2)	(3)	(4)	(5)	(6)
Income risk	-0.959	-0.760	-0.562	0.374	3.103	-0.0879
	(3.959)	(4.329)	(4.299)	(6.526)	(5.844)	(6.291)
Prudence (C&E)	2.247**	2.484***	3.628***	3.478***	0.856	2.309*
	(0.871)	(0.785)	(1.289)	(1.302)	(1.302)	(1.327)
Prudence (C&E) × Income risk	1.551**	1.841***	1.433	1.702	1.574	2.398**
	(0.683)	(0.684)	(0.935)	(1.080)	(0.988)	(1.154)
Risk Aversion (A&P)		11.43**		1.979		11.00
		(4.615)		(7.762)		(10.39)
Male		-25.27		-28.75		-19.87
		(15.93)		(21.98)		(26.13)
Age		-0.116		-0.631		0.913
		(0.436)		(0.653)		(0.698)
Financial literacy		-2.727		5.485		-7.977**
		(2.272)		(3.483)		(3.374)
BMI		0.908		-0.860		4.629**
		(1.383)		(1.784)		(2.314)
Adult HH members		-9.051*		-5.144		-7.627
		(5.125)		(7.019)		(10.06)
Children HH members		-1.777		-0.121		2.373
		(4.964)		(6.405)		(6.734)
Income		-0.724		2.981		-2.307
		(2.578)		(4.009)		(3.928)
Impatience		-0.335		-1.074		0.229
		(0.373)		(0.687)		(0.509)
Increase in patience		0.697		0.880		0.801
		(0.523)		(0.700)		(0.810)
Planning horizon		ref.		ref.		ref.
– Next months		9.929		2.249		27.81
		(13.33)		(17.55)		(21.05)
– Next year		0.588		12.86		-27.17
		(17.11)		(28.80)		(24.58)
– Next two to five years		46.50**		126.6**		83.38*
		(21.63)		(51.44)		(45.08)
– 5 or more years		33.90*				40.41
		(18.62)				(25.39)
Constant	-30.33***	-56.16	-20.82***	54.77	-36.02***	-252.6**
	(5.829)	(80.54)	(7.246)	(96.24)	(9.216)	(117.6)
Controls	No	Yes	No	Yes	No	Yes
Observations	231	223	111	111	114	107

Notes: This table reports the results of ordinary least squares regressions on net savings. The sample is restricted to subjects who have non-zero net-savings. Prudence (C&E) is the Crainich-Eckhoud measure of prudence. Risk Aversion (A&P) is the Arrow-Pratt measure of risk aversion. Income risk is measured as the ratio of closed to existing businesses in 2013 in the working sector an individual was usually working in at the time of the survey. Prudence and income risk are centered. The controls are time preferences, gender, age, financial literacy, body mass index (BMI), household members (adults and children), income as measured as the average income per household member, planning horizon and education. *Income* in 100k Colombian pesos. We account for potential heteroskedasticity by robust standard errors. Results of t-tests indicated at following significance levels * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

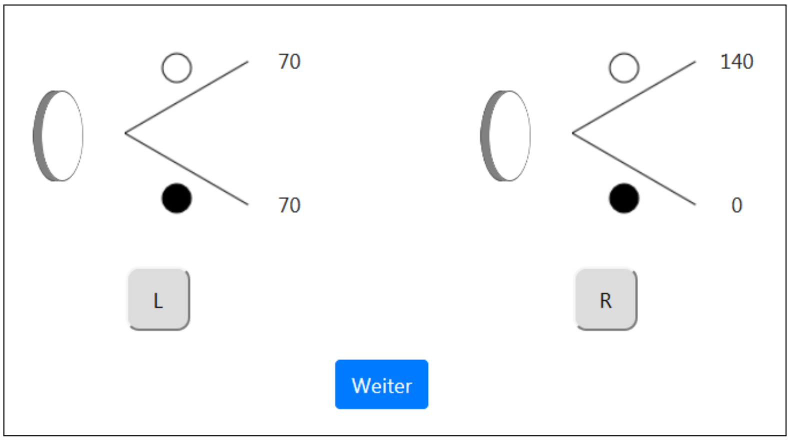
F Validation Experiment: Instructions, Decision Tasks and Exemplary Decision Screens

F.1 Risk Tasks

F.1.1 Certainty equivalents and utility mid-points.

Coin toss game

In the coin toss game, coins are tossed (resp. their toss is simulated by the computer). Thereby the payment is determined. You can choose between two options, called “Option L” (left) and “Option R” (right). An example of a choice is given below:



A coin is tossed in both options. In this example, “Option L” yields 70$, no matter if the coin lands with the white or the black side at the top. In ”Option R” there is something different in this example: Here you receive 140$, if the coin lands with the white side at the top. If the coin lands with the black side on the top, you receive 0$ - therefore nothing.

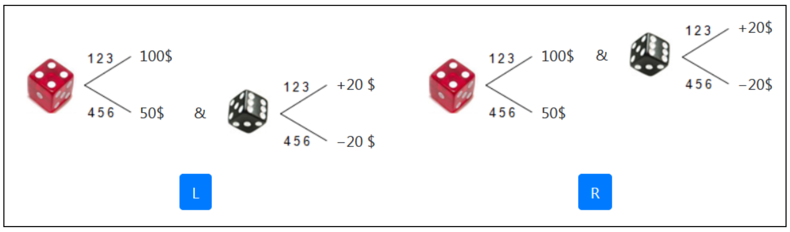
Overall, we show you 18 decision situations in the coin toss game. The amounts are different in each situation. In the end, only one choice (of the coin toss game and the following games) will be selected by the computer – this one determines the payment. You should therefore always choose the option you prefer.

For entering your choice, please click the button “L” or “R” first and then confirm your choice by clicking “Next”. If you have no more questions, you can start by clicking “Next”.

F.1.2 Risk Apportionment: Prudence (Noussair2013).

1. dice game

In the first dice game, a die is rolled (resp. a throw is simulated by the computer). This determines the payment in this game. Again, you can choose between two options, “Option L” (left) and “Option R” (right). In both options a red die is rolled first. Additionally, in both options a black die is rolled sometimes. An example of a choice is given below:



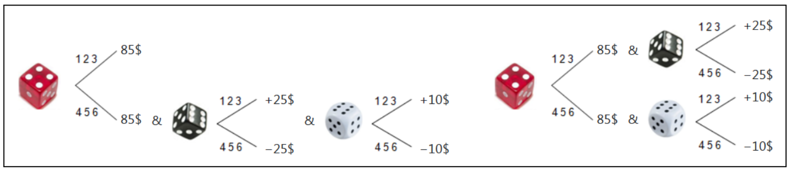
First, the red die is rolled. In this example, “Option L” and “Option R” yield 100$, if the roll of the red die is 1, 2 or 3. If the roll of the red die is 4, 5 or 6, you receive 50$.
The black die is used sometimes. In the example, you see that the black die is rolled, if the roll of the red die is 4, 5, or 6 in “Option L”, whereas in “Option R” the black die is rolled if the roll of the red is 1, 2 or 3. The black die determines, if an amount is added to the payment of the red die or if it is subtracted. In the example, 20$ are added to the payment determined by the red die, if the roll of the die is 1, 2 or 3. If the roll of the black die is 4, 5 or 6, 20$ are subtracted from the payment.

Overall, we show you 5 decision situations in the first dice game. The amounts are different in each situation. In the end, only one choice (of the first dice game and the other games) will be selected by the computer — this one determines the payment. You should therefore always choose the option you prefer.

For entering your choice, please click the button “L” or “R”. There is no confirmation ensued. If you have no more questions, you can start by clicking “Next”.

F.1.3 Risk Apportionment: Temperance (Noussair2013).

2. Dice game Additional to the red and the black die, in the second dice game a white die is rolled in both options sometimes. An example of a choice is given below:



The black and the white die are used sometimes. In the example, you can see that in "Option L" the white and the black die are rolled, if the roll of the red die is 4, 5, or 6. In "Option R" the white die is rolled, if the roll of the red die is 4, 5 or 6, and the black die is rolled, if the roll of the red die is 1, 2 or 3. The white and the black die determine whether an amount is added to or subtracted from the payment of the red die. In the example, 25$ are added to the payment determined by the red die, if the roll of the black die is 1, 2 or 3. If the roll of the black die is 4, 5 or 6, 25$ are subtracted from the payment. If the roll of the white die is 1, 2 or 3, 10$ are added to the payment determined by the red die. If the roll of the white die is 4, 5 or 6, 10$ are subtracted from the payment.

Overall, we show you 5 decision situations in the second dice game. The amounts are different in each situation. In the end, only one choice (of the second dice game and the other games) will be selected by the computer — this one determines the payment. You should therefore always choose the option you prefer.

For entering your choice, please click the button "L" or "R". There is no confirmation ensued. If you have no more questions, you can start by clicking "Next".

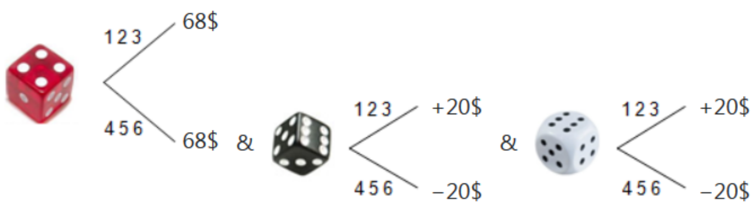
Control Questions

If you have selected the following option, which one is the smallest amount possible?

○ -20 ○ 28 ○ 68

...and which one is the largest amount possible?

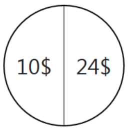
○ 68 ○ 88 ○ 108



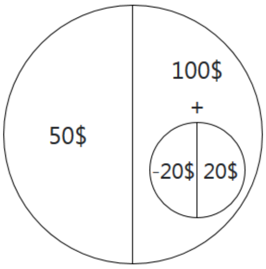
F.1.4 Risk Apportionment: Prudence (Fortune-Wheel-Design; Deck and Schlesinger, 2014).

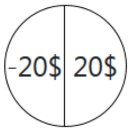
First screen

1. wheel of fortune game

In the first wheel of fortune game, a wheel of fortune is turned (resp. the turn is simulated by the computer). This determines the payment in this game. Again, you can choose between two options, “Option A” and “Option B”. Each option involves amounts of money and one or more 50-50 lotteries represented as a circle with a line through the middle. A 50-50 lottery means there is a 50For example, 

is a 50-50 lottery in which you can receive either 10$ or 24$; each with an equal chance. If the wheel of fortune stops at the right side, you would receive 24$. If it stops at the left side, you would receive 10$.

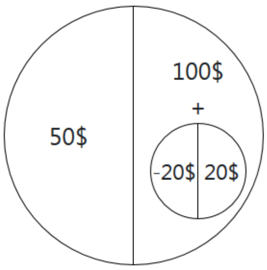
In some cases, one of the items in a 50-50 lottery may be another lottery. For example, 

is such a lottery. Here, you would receive either 50$ or 100$ plus the lottery 

.

Further explanations are given on the next page.

Second Screen

In the example 

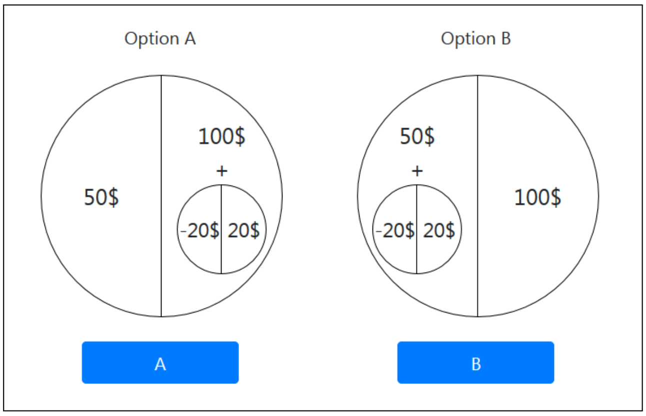
, there is a 50There is also a 50Conditional on this outcome for the big 50-50 lottery, you would then have a 50Therefore, the chance that you would end up with 100$+20$ = 120$ is

0,5 * 0,5 = 0,25 = 25%. The chance that you would end up with 100-20 = 80$ also is 0,5 * 0,5 = 0,25 = 25%.

Overall, we show you 5 decision situations in the first wheel of fortune game. The amounts are different in each situation. In the end, only one choice (of the first wheel of fortune game and the other games) will be selected by the computer — this one determines the payment. You should therefore always choose the option you prefer.

For entering your choice, please click the button "A" or "B". There is no confirmation ensued. If you have no more questions, you can start by clicking "Next".

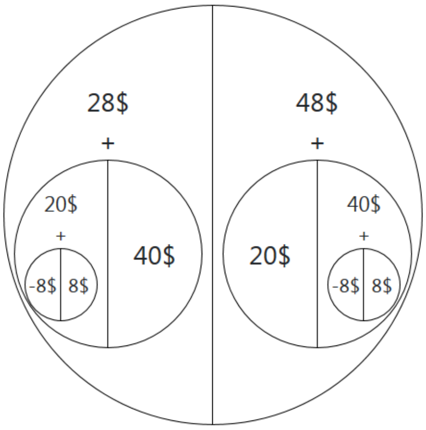
Exemplary Decision Situation

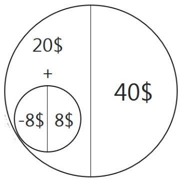


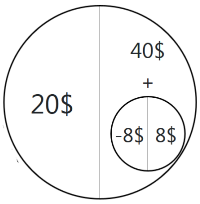
F.1.5 Risk Apportionment: Temperance (Fortune-Wheel-Design; Deck and Schlesinger, 2014).

First Screen

2. wheel of fortune game

In the second wheel of fortune game, wheels of fortune are turned, too (resp. the turns are simulated by the computer). However, further lotteries are added: 

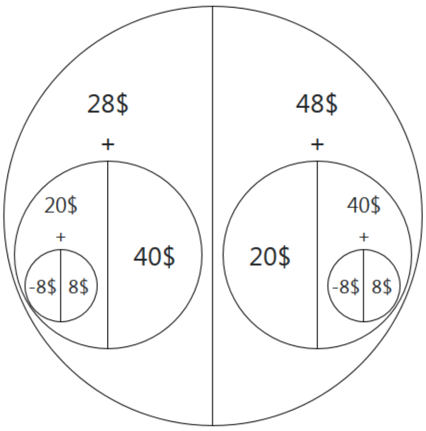
is a 50-50 lottery where you receive either 28$ plus the 50-50 lottery 

or you receive 48$ plus the 50-50 lottery 

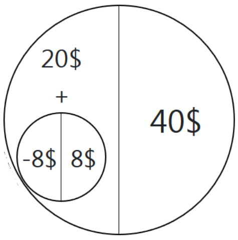
, both of which include an additional 50-50 lottery.

Further explanations are given on the next page.

Second Screen

In 

you could earn 40$ if you get 28$ +



in the big lottery and then earn 20$ in the middle lottery and -8$ in the small lottery. This occurs with a 0,5 * 0,5 * 0,5 = 12,5% chance. Whereas if you receive 8$ in the small lottery, you could earn 56$. This also occurs with a 12,5% chance. With a 50Notice that you could earn 68$ by 1) earning 28$ in the big lottery and then 40$ in the middle lottery or 2) by earning 48$ in the big lottery and 20$ in the middle lottery. Both happens with a 0,5 * 0,5 = 25% chance, therefore 25% + 25% = 50%. Besides, you could receive either 80$ or 96$ with a 0,5 * 0,5 * 0,5 = 12,5% chance for each by earning 48$ in the big lottery first, then 40$ in the middle lottery and afterwards -8$ resp. 8$ in the small lottery.

Overall, we show you 5 decision situations in the second wheel of fortune game. The amounts are different in each situation. In the end, only one choice (of the second wheel of fortune game and the other games) will be selected by the computer — this one determines the payment. You should therefore always choose the option you prefer.

For entering your choice, please click the button "A" or "B". There is no confirmation ensued. If you have no more questions, you can start by clicking "Next".

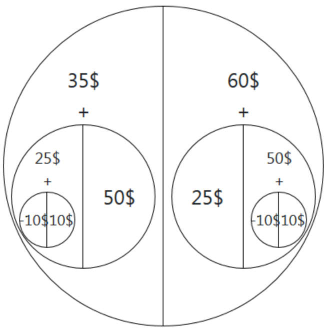
Control Questions

If you have selected the following option, which one is the smallest amount possible?

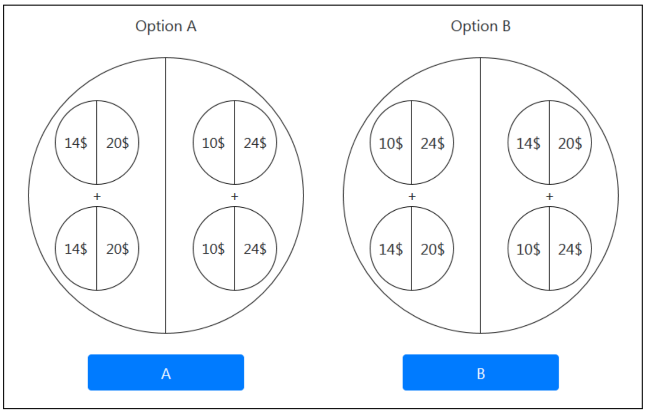
○ -10 ○ 0 ○ 50

And which one is the largest amount possible?

○ 85 ○ 110 ○ 120



Exemplary Decision Situation



F.1.6 Risk Apportionment with Compensation: Prudence (Urn-Design; Ebert and Wiesen, 2014).

First Screen

1. urn game

In the first urn game, it is drawn from different urns (resp. the draw is simulated by the computer). So these are also decision situations in which coincidence plays a part. Therefore, in these situations the outcome is uncertain again.

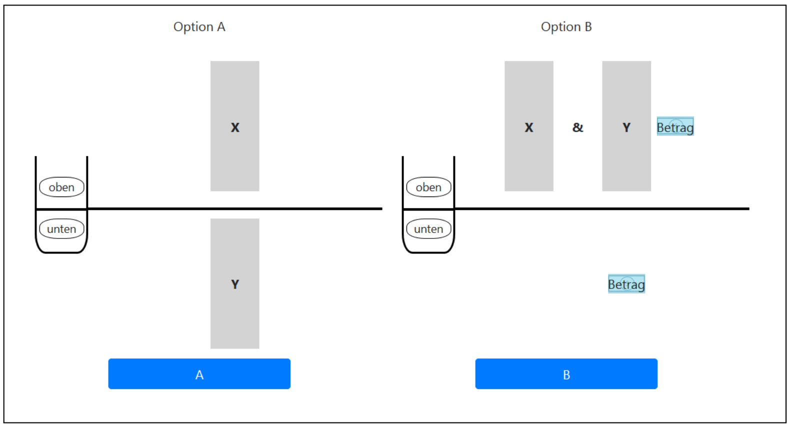
Again, you choose between two options with random outcome. The options are described and explained in detail later. Overall, there are 60 decisions in the first urn game and 40 decisions in the second urn game which approximately correspond with five decision situations in the previous games.

In the end, only one choice (of the urn game and the other games) will be selected by the computer — this one determines the payment. You should therefore always choose the option you prefer.

Further explanations are given on the next page.

Second Screen

The following figure describes the decision situations in the urn games schematically and simplified.



In each decision situation, you decide which of the two risky events, "Option A" or "Option B", you prefer. Both options, "Option A" and "Option B", comprise a random draw (RANDOM DRAW 1) that is depicted as an urn with two balls "Up" and "Down". RANDOM DRAW 1 is: With a 50% chance you are in situation "Up" and with 50% chance you are in situation "Down".

We now look at the risky event **"Option A"**. If ball "Up" is drawn, the outcome is **X**. **X** can either be a *fixed amount* or another *random draw (RANDOM DRAW **X**)*. If ball "Down" is drawn, the outcome is **Y**. Likewise, **Y** can either be a *fixed amount* or another *random draw (RANDOM DRAW **Y**)*.

In risky event **"Option B"**, both **X** and **Y** follow if ball "Up" is drawn. In addition, an *amount* (blue bank note) is added to the outcome in both situations. If ball "Down" is drawn, you receive the amount indicated on the bank note. If ball "Up" is drawn, **X** and **Y** follow and the amount (blue bank note) is added.

The *amount* on the blue bank note can take the following values:

-11,25, -10.13, -9.00, ..., -1.13, 0.00, 1.13, ..., 9.00, 10.13.

Hence, for each of these 20 amounts, one decision situation with two risky events follows. The *amount* on the blue bank note is always added to both situations ("Up" and "Down").

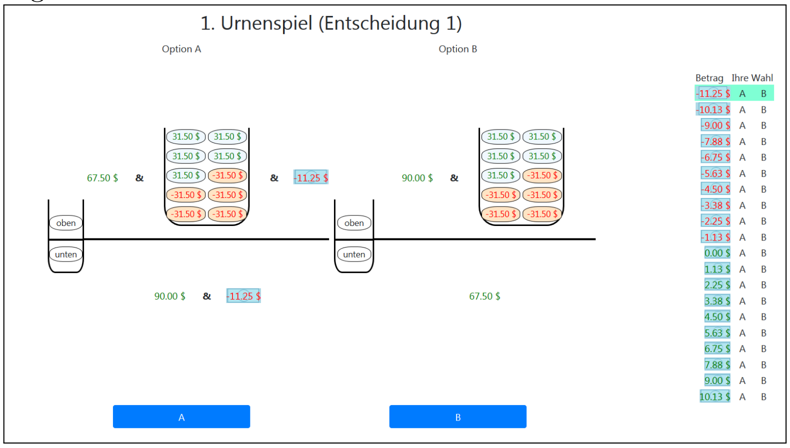
Note that on your decision screens, the risky event where the *amount* (blue bank note) is added can either be the right or the left option.

Further explanations are given on the next page.

Third Screen

In the first urn game, you make 60 decisions. These are displayed in three decision situations and each situation comprises 20 decisions between two options – “Option A” and “Option B”.

The following figure describes an example of a decision situation in the first urn game.



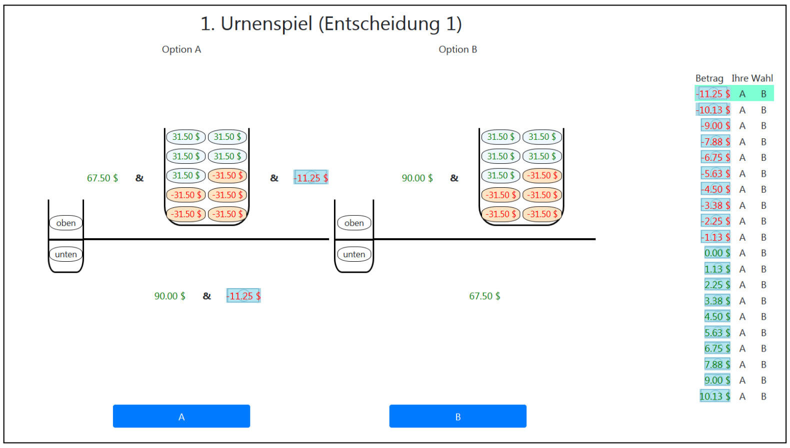
In this example, the *amount* (blue bank note) is added to Option A. The size of the added *amount* can be found in the column *amount* on the right-hand side of the screen. For each amount, you decide whether you prefer Option A or Option B by clicking on the respective blue button under the option. The green frame in the right chart helps you to overview how you have decided for which amount. Aside from the amounts on the bank notes, the decision situation is the same for all 20 amounts that are displayed on the right.

How is the payment determined in the first urn game?
For *RANDOM DRAW 1*, there are two balls in an urn – one with label ‘Up”, another with label “Down”. Both balls can be drawn with the same chance. As shown in the figure above, a second random draw (*RANDOM DRAW X*) can be necessary to determine your payoff. In *RANDOM DRAW X*, a ball is drawn from an urn containing 10 balls. The ball can either be blue or orange. Note that the composition of blue and orange balls can change in the three decision situations. The urn always contains 10 balls, and for 20 decisions the composition of blue and orange balls is identical.

Further explanations are given on the next page.

Fourth Screen

We now look at a concrete example of the recent decision situation:



If in **Option A** in *RANDOM DRAW 1* ball “Up” is drawn, the outcome is 67.5, *RANDOM DRAW X* follows and -11.25 (*amount* on the blue bank note) is added (therefore, 11.25 is subtracted).

- If in *RANDOM DRAW X* an orange ball is drawn, you lose 31.50. Overall, you receive 24.75 (=67.50 - 31.50 - 11.25) in that case.
- If in *RANDOM DRAW X* a blue ball is drawn, you win 31.50. Overall, you receive 87.75 (=67.50 + 31.50 - 11.25)in that case.

If in **Option A** in *RANDOM DRAW 1* ball “Down” is drawn, the outcome is 90.00, and -11.25 (*amount* on the blue bank note) is added (therefore, 11.25 is subtracted). Overall, you receive 87.75 (=90.00 - 11.25) in that case.

If in **Option B** in *RANDOM DRAW 1* ball “Up” is drawn, the outcome is 90.00, and *RANDOM DRAW X* follows.

- If in *RANDOM DRAW X* an orange ball is drawn, you lose 31.50. Overall, you receive 58.50 (=90 - 31.50) in that case.
- If in *RANDOM DRAW X* a blue ball is drawn, you win 31.50. Overall, you receive 121.50 (=90 + 31.50) in that case.

If in **Option B** in *RANDOM DRAW 1* ball “Down” is drawn, the outcome is 67.50.

Table 19. Risk Apportionment Tasks for Elicitation of Prudence and Temperance

Prudence: Fortune-Wheel- and Die-Design				
Original Task	Construction	Scale	Option A	Option B
D&S 11	–	x10	$[\$50+[\$-20,\$20],\$100]$	$[\$50,\$100+[\$-20,\$20]]$
D&S 13	–	x10	$[\$50+[\$-40,\$40],\$100]$	$[\$50,\$100+[\$-40,\$40]]$
D&S 14	–	x10	$[\$20+[\$10,\$-10],\$40]$	$[\$20,\$40+[\$10,\$-10]]$
D&S 16	–	x10	$[\$80+[\$20,\$-20],\$100]$	$[\$80,\$100+[\$20,\$-20]]$
D&S 17	–	x10	$[\$120+[\$10,\$-10],\$140]$	$[\$120,\$140+[\$10,\$-10]]$
Temperance: Fortune-Wheel-Design				
Original Task	Construction	Scale	Option A	Option B
D&S 18	2+2	x1	$[[\$14,\$20]+[\$14,\$20],[\$10,\$24]+[\$10,\$24]]$	$[[\$10,\$24]+[\$14,\$20],[\$14,\$20]+[\$10,\$24]]$
D&S 19	2+2	x5	$[[\$35,\$50]+[\$35,\$50],[\$25,\$60]+[\$25,\$60]]$	$[[\$25,\$60]+[\$35,\$50],[\$35,\$50]+[\$25,\$60]]$
D&S 21	2+2	x5	$[[\$5,\$80]+[\$5,\$80],[\$25,\$60]+[\$25,\$60]]$	$[[\$25,\$60]+[\$5,\$80],[\$5,\$80]+[\$25,\$60]]$
D&S 22	1+3	x2	$[\$28+[\$20+[\$-8,\$8],\$40],\$48+[\$20,\$40+[\$-8,\$8]]]$	$[\$28+[\$20,\$40+[\$-8,\$8]],\$48+[\$20+[\$-8,\$8],\$40]]$
D&S 23	1+3	x5	$[\$35+[\$25+[\$-10,\$10],\$50],\$60+[\$25,\$50+[\$-10,\$10]]]$	$[\$35+[\$25,\$50+[\$-10,\$10]],\$60+[\$25+[\$-10,\$10],\$50]]$
Temperance: Die-Design				
Original Task	Construction	Scale	Option A	Option B
NTvdK 1			$[\$85,\$85+[\$25,\$-25]+[\$25,\$-25]]$	$[\$85+[\$25,\$-25],\$85+[\$25,\$-25]]$
NTvdK 2			$[\$85,\$85+[\$25,\$-25]+[\$10,\$-10]]$	$[\$85+[\$25,\$-25],\$85+[\$10,-\$10]]$
NTvdK 3			$[\$85,\$85+[\$50,\$-50]+[\$25,\$-25]]$	$[\$85+[\$50,\$-50],\$85+[\$25,\$-25]]$
NTvdK 4			$[\$34,\$34+[\$10,\$-10]+[\$10,\$-10]]$	$[\$34+[\$10,\$-10],\$34+[\$10,\$-10]]$
NTvdK 5			$[\$68,\$68+[\$20,\$-20]+[\$20,\$-20]]$	$[\$68+[\$20,\$-20],\$68+[\$20,\$-20]]$

Notes: As in Deck and Schlesinger (2014), [x,y] denotes a 50-50 lottery with equally likely outcomes x and y. Original Task refers to the Task Number given in Deck and Schlesinger (2014) and Scale informs about the scaling parameter that we used to obtain Options A and B from the original tasks as in Deck and Schlesinger (2014).

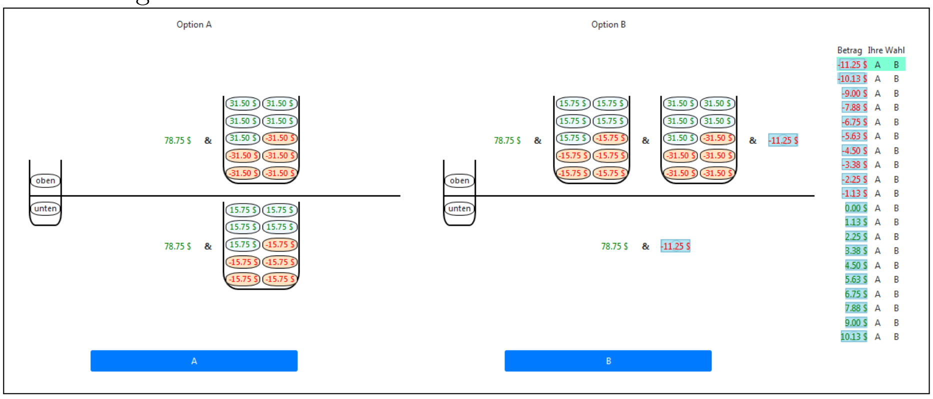
F.1.7 Risk Apportionment with Compensation: Temperance (Urn-Design; Ebert and Wiesen, 2014).

First Screen

2. urn game

In the second urn game, it is drawn from different urns, too (resp. the draw is simulated by the computer). So these are also decision situations in which coincidence plays a part and in which the outcome is uncertain. In the second urn game, you make 40 decisions. These are displayed in two decision situations and each situation comprises 20 decisions between two options – “Option A” and “Option B”.

The following figure describes an example of a decision situation in the second urn game.



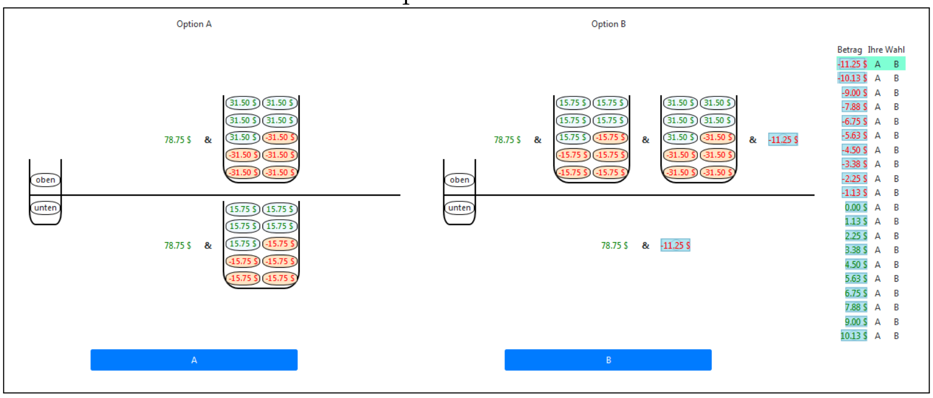
In this example, the *amount* (blue bank note) is added to Option B. The size of the added *amount* can be found in the column amount on the right-hand side of the screen. For each amount you decide whether you prefer Option A or Option B by clicking on the respective blue button under the option. The green frame in the right chart helps you to overview how you have decided for which amount. Aside from the amounts on the bank notes, the decision situation is the same for all 20 amounts that are displayed on the right.

How is the payment determined in the second urn game? For *RANDOM DRAW 1*, there are two balls in an urn – one with label “Up”, another with label “Down”. Both balls can be drawn with the same chance (analogous to the first urn game). As shown in the figure above, a second random draw (*RANDOM DRAW X*) and/or a third random draw (*RANDOM DRAW Y* can be necessary to determine your payoff. In *RANDOM DRAW X*, a ball is drawn from an urn containing 10 balls. The ball can either be blue or orange. Note that the composition of blue and orange balls can change in the decision situations. The urn always contains 10 balls, and for 20 decisions the composition of blue and orange balls is identical. This holds analogously for *RANDOM DRAW Y*. Notice that the composition of blue and orange balls across *RANDOM DRAW X* and *RANDOM DRAW Y* can differ.

Further explanations are given on the next page.

Second Screen

We now look at a concrete example of the recent decision situation:



If in **Option A** in *RANDOM DRAW 1* ball "Up" is drawn, the outcome is 78.75, and *RANDOM DRAW X* follows.

- If in *RANDOM DRAW X* an orange ball is drawn, you lose 31.50. Overall, you receive 47.25 (=78.75 - 31.50) in that case.
- If in *RANDOM DRAW X* a blue ball is drawn, you win 31.50. Overall, you receive 110.25 (=78.75 + 31.50) in that case.

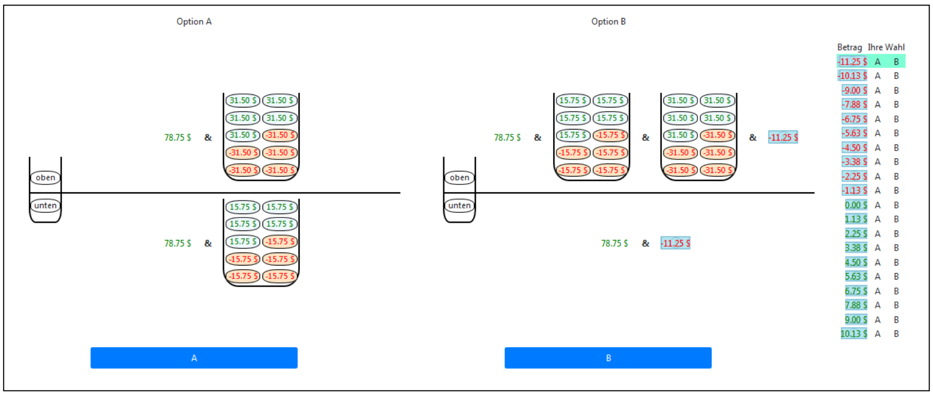
If in *RANDOM DRAW 1* ball "Down" is drawn, the outcome is 78.75 and *RANDOM DRAW Y* follows.

- If in *RANDOM DRAW Y* an orange ball is drawn, you lose 15.75. Overall, you receive 63 (=78.75 - 15.75) in that case.
- If in *RANDOM DRAW Y* a blue ball is drawn, you win 15.75. Overall, you receive 94.5 (=78.75 + 15.75) in that case.

On the next page, the calculation of outcomes for Option B follows.

Third Screen

Now to the calculation of payments in Option B in the recent example of a decision situation:



If in **Option B** in *RANDOM DRAW 1* ball “Up” is drawn, the outcome is 78.75, *RANDOM DRAW X* and *RANDOM DRAW Y* follow, and -11.25 (*amount* on the blue bank note) is added (therefore, 11.25 is subtracted).

- If in *RANDOM DRAW X* and in *RANDOM DRAW Y* an orange ball is drawn, you lose 31.50 (from RANDOM DRAW X) and 15.75 (from RANDOM DRAW Y). Overall, you receive 20.25 (=78.75 - 31.50 - 15.75 -11.25) in that case.
- If in *RANDOM DRAW X* and in *RANDOM DRAW Y* a blue ball is drawn, you win 31.50 (from RANDOM DRAW X) and 15.75 (from RANDOM DRAW Y). Overall, you receive 123.75 (=78.75 + 31.50 + 15.75 - 11.25) in that case.
- If in *RANDOM DRAW X* a blue ball and in *RANDOM DRAW Y* an orange ball is drawn, you win 31.50 (from RANDOM DRAW X) and lose 15.75 (from RANDOM DRAW Y). Overall, you receive 83.25 (=78.75 + 31.50 - 15.75 - 11.25) in that case.
- If in *RANDOM DRAW X* an orange ball and in *RANDOM DRAW Y* a blue ball is drawn, you lose 31.50 (from RANDOM DRAW X) and win 15.75 (from RANDOM DRAW Y). Overall, you receive 51.75 (=78.75 - 31.50 + 15.75 - 11.25) in that case.

If in Option B in *RANDOM DRAW 1* ball “Down” is drawn, the outcome is 78.75, and -11.25 (*amount* on the blue bank note) is added (therefore, 11.25 is subtracted). Overall, you receive 67.5 (=78.75 - 11.25) in that case.

On the next page, you are asked some comprehension questions.

Control Questions

If you have selected the following option, which one is the smallest amount possible?

○ -49,95 ○ -11,25 ○ 4,95

And which one is the largest amount possible?

○ 67,5 ○ 78,75 ○ 130,05

F.2 Real-Effort Tasks