Viewing the Future through a Warped Lens: Why Uncertainty Generates Hyperbolic Discounting *

Thomas Epper Helga Fehr-Duda Adrian Bruhin

September 9, 2010

Abstract

A large body of experimental research has demonstrated that, on average, people violate the axioms of expected utility theory as well as of discounted utility theory. In particular, aggregate behavior is best characterized by probability distortions and hyperbolic discounting. But is it the same people who are prone to these behaviors? Based on an experiment with salient monetary incentives we demonstrate that there is a strong and significant relationship between greater departures from linear probability weighting and the degree of decreasing discount rates at the level of individual behavior. We argue that this relationship can be rationalized by the uncertainty inherent in any future event, linking discounting behavior directly to risk preferences. Consequently, decreasing discount rates may be generated by people's proneness to probability distortions.

Keywords: Time Preferences, Risk Preferences, Hyperbolic Discounting, Probability Weighting, Institutionally Generated Uncertainty

JEL: D01, D81, D91

*Swiss Federal Institute of Technology, Chair of Economics, Weinbergstrasse 35, CH-8092 Zurich, Switzerland. Corresponding author: Epper (epper[at]econ.gess.ethz.ch). Acknowledgements: We thank Ernst Fehr, Drew Fudenberg, David Laibson, Drazen Prelec, Peter Wakker and Herbert Walther for helpful comments. The usual disclaimer applies. This research is supported by the Swiss National Science Foundation (Grants 100012-109907 and 100014-120705). An earlier version of this paper, "Uncertainty Breeds Decreasing Impatience: The Role of Risk Preferences in Time Discounting", is available at http://www.iew.uzh.ch/wp/iewwp412.pdf.

"Future income is always subject to some uncertainty, and this uncertainty must naturally have an influence on the rate of time preference, or degree of impatience, of its possessor."

Fisher (1930)

1 Introduction

It has long been recognized by practitioners and theorists alike that the domains of choice under risk and over time are intimately related. Unlike gambling in the casino, where uncertainty is resolved almost immediately, many real world decisions, and probably the most important ones for our happiness, involve risky consequences materializing over the course of time. In the realm of economic theory, the dimensions of risk and time are treated as largely independent attributes, modeled in an equivalent way (Prelec and Loewenstein, 1991): The classical models of choice, expected utility theory (EUT) and discounted utility theory (DUT), view decision makers as maximizing a weighted sum of utilities with the weights representing either probabilities or discount factors, respectively.

A large body of empirical evidence has challenged the validity of EUT and DUT as descriptive models of choice, however. In the domain of risk, one of the best documented phenomena concerns the *Common Ratio Effect*: Often, people's preference for a smaller more probable outcome over a larger less probable one changes in favor of the larger outcome when both outcome probabilities are scaled down by a common factor. This pattern of behavior constitutes a violation of the independence axiom of EUT (Kahneman and Tversky, 1979; Starmer and Sugden, 1989).[1] The stationarity axiom of DUT, according to which preferences should depend on the absolute time interval between delivery of the objects, has met a similar fate. The *Common Difference Effect* describes the empirical regularity that preference for a smaller earlier payoff over a larger later payoff reverses when both payoffs are shifted into the more remote future, keeping the timing difference constant (Thaler, 1981; Benzion, Rapoport, and Yagil, 1989).

Researchers have reacted to these anomalies by relaxing the assumptions on the corresponding

[1] Prominent special cases are the Allais paradox (Allais, 1953) and the Bergen paradox (Hagen, 1972).

decision weights while leaving the overall structure of the models intact. Violations of independence can be captured by a suitable nonlinear transformation of the probabilities, as discussed by Quiggin (1982) and Tversky and Kahneman (1992). Violations of stationarity, on the other hand, are accounted for by allowing the discount factors to decline hyperbolically in time, i.e. at a decreasing rate (Ainslie, 1991; Laibson, 1997; Prelec, 2004). These generalizations seem to perform much better at explaining aggregate choices than do EUT and DUT (Rachlin, Raineri, and Cross, 1991; Harless and Camerer, 1994; Hey and Orme, 1994; Myerson and Green, 1995; Kirby, 1997). At the individual level, however, there is vast heterogeneity in observed behavior in both decision domains (Hey and Orme, 1994; Chesson and Viscusi, 2000; Abdellaoui, Attema, and Bleichrodt, 2010; Bruhin, Fehr-Duda, and Epper, 2010) and it is an open question whether the superior fit of the generalized models is a manifestation of common regularities of individual behavior. Clearly, endeavors of integrating risk taking and intertemporal choice in one single model, such as for example in Prelec and Loewenstein (1991), only make sense if the latter is the case. The question then arises whether violations of independence and stationarity are actually committed by the same people.

In the empirical literature, the relationship between individuals' attitudes towards risk and delay has been examined from various different angles. One strand of the literature focuses on people's risk tolerance measured independently from their degree of impatience. These studies find that more risk averse people tend to discount the future more heavily (Leigh, 1986; Anderhub, Güth, Gneezy, and Sonsino, 2001; Eckel, Johnson, and Montmarquette, 2004). Discount rates are inferred directly from choices over dated monetary amounts and, therefore, their measurement is confounded by the curvature of the utility function. Andersen, Harrison, Lau, and Rutström (2008) correct for utility curvature and still find a positive, but much reduced, correlation in their predicted degrees of risk aversion and impatience. None of the studies so far have accounted for probability weighting and, therefore, they cannot address the question of whether departures from linear probability weighting are systematically related to departures from constant discounting. Similarly, the psychological literature has dealt with comparisons of highly reduced forms of discounting functions for delay and probability, ignoring utility curvature as well as probability weighting (Myerson, Green, Hanson, Holt, and Estle, 2003). These findings also indicate moderate positive correlations between both types of discounting.

Another strand of the literature investigates people's choices when both risk and delay are present (Keren and Roelofsma, 1995; Ahlbrecht and Weber, 1997; Weber and Chapman, 2005; Noussair and Wu, 2006; Anderson and Stafford, 2009; Baucells and Heukamp, 2010; Coble and Lusk, 2010). These studies generally conclude that there are interaction effects between time and risk, such as risk tolerance increasing with delay, which are not easily justifiable within the frameworks of EUT and DUT. Again, probability weighting does not feature in any of these papers. A notable exception is the contribution by Abdellaoui, Diecidue, and Öncüler (2010) who estimate individual probability weights over varying delays, but do not elicit discount functions for guaranteed payoffs.

Finally, some recent papers examine the effects of risk in the payment date, rather than in outcome magnitude. Parallel to the findings on delayed guaranteed outcomes, Chesson and Viscusi (2000) report discount rates to decline with time horizon. Moreover, Chesson and Viscusi (2003) show that aversion to timing risk is positively related to ambiguity aversion, suggesting that uncertainty may be processed similarly in both the dimensions of time and risk. In a follow-up study Onay and Öncüler (2007) argue that the prevalence of timing risk aversion, which runs against the predictions of EUT, can be accomodated within a rank-dependent model involving probability weighting. They did not test their conjecture empirically, however.

This brief review of the literature shows that, to the best of our knowledge, there is no previous study that investigates the same individuals' probability weights and discount functions. While evidence of hyperbolic discounting is occasionally reported, simultaneous estimates of individual probability weights are usually not provided. This lack may be due to the fact that a comparatively rich data set, and for that matter also a fairly sophisticated estimation strategy, is needed to be able to disentangle utility curvature and probability weighting.

In order to close this gap, we conducted an experiment with salient monetary incentives, which exhibits a number of distinguishing features: First, the experiment generated data rich enough to be able to estimate *individual* probability weighting functions and relate them to the same subjects' revealed discount rates. Second, in contrast to many previous discounting experiments, every single subject got paid for her intertemporal choices, involving substantial payoffs, in an incentive compatible manner. Third, we kept transaction costs equal across different payment dates in order to preclude confounding effects. Finally, we controlled for utility curvature.

We present the following experimental results. First, we show that subjects' departures from linear probability weighting are highly significantly associated with the strength of decreasing discount rates. The curvature of the utility function, however, seems not to be directly related to the their decline. Second, estimation results are robust to controlling for socioeconomic characteristics, such as gender, age, experience with investment decisions and cognitive abilities. In fact, the only variable associated with decreasing discount rates turns out to be the degree of nonlinearity of probability weights, which explains a, by any standard, large percentage of the variation in the extent of the decline. In particular, cognitive abilities, as measured by the *Cognitive Reflection Test* (Frederick, 2005), cannot account for the link between proneness to probability distortions and hyperbolic discounting. Moreover, all our results are insensitive to model specification.

Our findings demonstrate that it is the same people who tend to violate the axioms of independence and stationarity. We suggest that this relationship is driven by a natural link between the domains of time and risk (Halevy, 2008; Walther, 2010): Arguably, only immediate consequences can be totally certain whereas delayed ones are uncertain by their very nature. For instance, a promised reward may, due to unforeseen circumstances, materialize later or turn out to be smaller than expected, or sudden illness or death may keep the decision maker from collecting her reward. For these reasons, future consequences are inextricably associated with uncertainty, implying that the decision maker's valuation of delayed outcomes not only depends on her *pure* time preference, i.e. her preference for immediate utility over delayed utility, but also on her perception of uncertainty and, consequently, on her risk preferences. We show theoretically that stronger departures from linear probability weighting entail more strongly declining discount rates, providing us with a theoretical underpinning of our experimental results. Figuratively speaking, hyperbolic discounting is driven by viewing the uncertain future through a warped lens, produced by systematic distortions of probabilities. This theoretical framework not only organizes our experimental findings but also accounts for previous evidence of interactions of time and risk. A number of studies detected preference reversals when either risk is added to temporal prospects (Baucells and Heukamp, 2010) or delay is added to simple risky prospects (Keren and Roelofsma, 1995; Weber and Chapman, 2005).

Our analysis suggests that institutionally generated uncertainties, such as lack of contract enforcement and weak property rights, may induce extreme short-run impatience even if people's pure rate of time preference is constant and relatively low. This insight is important because it implies that revealed behavior may be predominantly driven by environmental factors rather than by the underlying preferences themselves and, consequently, may be amenable to economic policy. While uncertainty may be an important channel through which hyperbolicity of discount rates is generated there may be other sources of hyperbolic discounting behavior as well. For instance, pure time preferences may be hyperbolic *per se*, as could be argued for addictive behavior. And when visceral motives, such as hunger or lust, come into play, uncertainty may not be the dominant dimension decision makers are concerned about. An excessive preference for the present may then be driven by factors other than potential disappearance of the object of desire.

The remainder of the paper is structured as follows: In Section 2 we describe the experimental design and procedures. Section 3 outlines our approach to estimation. Section 4 presents our results. Section 5 discusses our hypothesis on the role of risk preferences in time discounting. Section 6 concludes.

2 Experimental Design

The experiment took place at the Institute for Empirical Research in Economics (IERE), University of Zurich, in May 2006. Participants were recruited from the IERE subject pool, which consists of students from all fields offered at the University of Zurich and the Swiss Federal Institute of Technology Zurich. In total, we analyzed 112 subject's responses.[2] The experiment consisted of two main parts, one dedicated to eliciting certainty equivalents for non-delayed risky prospects,[3] the other one to eliciting future equivalents and their corresponding imputed discount rates for temporal

[2]We omitted six subjects' responses from our analysis. Four subjects reported that they would not be able to cash in their delayed payments at the respective payment dates. Three of them would have been on vacation then, the fourth person had planned a long visit abroad. Hence, their choices were not informative of their time preferences. Concerning the remaining two subjects we could not disentangle utility effects from probability weighting effects. Nonetheless, our results do not change when we include these two individuals in the data set.

[3]The risk data was also used in Bruhin, Fehr-Duda, and Epper (2010).

prospects involving guaranteed payments.[4]

We used similar procedures to elicit certainty equivalents and discount rates, in order to economize on subjects' cognitive effort. For both types of tasks we implemented choice menus containing a list of 20 varying alternatives which had to be judged against a fixed option. To familiarize subjects with the nature of the procedure, the instructions contained examples and trial problems. Besides a show up fee of CHF 10 (CHF 1 $\approx$ USD 0.8 at the time of the experiment), each subject was paid according to one of her risky choices and one of her temporal choices selected randomly at the very end of the experiment. Subjects received their compensation for the risky choices and the show-up fee in cash immediately after completion of all the tasks. The compensation for their intertemporal choices was paid out to them at the respective dates when they cashed in vouchers issued to them at the end of the experiment. Payment modalities are described in detail below. Subjects could work at their own speed. On average, it took them 1.25 hours to complete the experiment, including a socioeconomic questionnaire presented after the choice tasks.

2.1 Elicitation of Certainty Equivalents

Since the objective of the risk task was to obtain data on the basis of which individual probability weights could be estimated, a fairly large number of observations per person was needed. To elicit individual lottery evaluations, subjects were presented with 20 choice menus, each one involving a specific binary lottery $\mathcal{L} = (x_1, p; x_2)$, with $x_1 > x_2 \geq 0$, labeled *Option A* in Figure I. *Option B* in the choice menu represented the guaranteed alternatives, ranging from the higher lottery outcome x_1 to the lower outcome x_2. Every subject had to choose her preferred option in each row of the choice menu. In Figure I, a hypothetical subject prefers all guaranteed payments larger than CHF 36 to the stated lottery, and prefers the lottery in the remaining rows. This subject's valuation of the lottery, her certainty equivalent ce, is calculated as the arithmetic mean of the two amounts next to her indifference point, amounting to CHF 37 in the example here. The set of lotteries, listed in Table I, included a wide range of outcomes and probabilities. Every subject was confronted with this set of lotteries once. The choice menus appeared in an individualized random order.

[4] Instructions are available upon request. The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007).

At the end of the experiment, after the subject had completed all the tasks, one row of one choice menu was randomly selected for payment. If the subject had opted for the lottery there, her decision was played out for real. If the subject had opted for the guaranteed payoff, the respective amount was paid out to her. On average, subjects earned CHF 37.22 in cash for the risk task, including the show-up fee of CHF 10, to be paid out immediately. Cash payments for the risk task were considerably higher than the local student assistant's hourly wage.

2.2 Elicitation of Discount Rates

Using a similar format as in the risk task we elicited individual discount rates for temporal prospects $\mathcal{T} = (x, t)$, with $x > 0$, over payments x delayed by t months. The choice menus, designed as in Figure II, contained 20 binary choices each.[5] Subjects had to choose between a guaranteed payment *pe* of CHF 60 the next day (*Option A*) and a guaranteed later payment x (*Option B*), delayed by two months or four months, respectively. The varying alternatives x were sorted in descending order from the highest amount to the lowest amount, incorporating an interest payment at a simple annualized rate of $\delta_t \in [0\%, 95\%]$ over the corresponding time interval $[0, t]$. These rates were exhibited in the right-most column of the choice menu.[6] The present amount of CHF 60 and the range of interest rates were chosen to provide salient incentives, so that deferring the reward was actually worthwhile. The arithmetic mean of the two monetary amounts next to the indifference point on the choice menu provided the imputed discount rate δ_t. The hypothetical subject in Figure II, for instance, is indifferent between CHF 60 and CHF 70.50, implying a discount rate of 52.5% *per annum*.

We applied a similar random payment method in the time task as in the risk task: One of each subject's choices was paid out for real at the corresponding payment date. Average payoffs for the time task amounted to CHF 64.34. Therefore, total average payments for both risk and time tasks summed to more than four times students' opportunity costs, measured by the student assistants' hourly wages.

Since our objective was to elicit discount rates over guaranteed payments, we took special care with the payment procedure: First, every single subject got paid for one of her intertemporal choices

[5] A similar design was proposed by Coller and Williams (1999).
[6] The discounting experiment consisted of one additional task not reported here.

all of which entailed a payment of the same order of magnitude. Not paying off everyone may render the stochastic nature of the experimental earnings salient and interfere with the objective of eliciting discount rates over guaranteed amounts of money. The second issue concerns the credibility of payment. In order to control for uncertainty arising from subjects' doubts about experimenter reliability, an official voucher of the Swiss Federal Institute of Technology was issued to them. This payment method was explained in detail in the instructions, and a specimen of the voucher was included in the instruction set.

A third possibly confounding factor are transaction costs. Transaction costs should be the same regardless of the payment date in order to avoid inducing present bias resulting from immediately available cash payments. Therefore, every subject had to make a trip to the cash desk to collect her earnings for the discounting task.[7]

3 Econometric Specification

The data elicited in the experiment provide two types of main variables: certainty equivalents ce for risky prospects, and imputed discount rates δ_t for temporal prospects. We first discuss our econometric approach to risky choice and, subsequently, describe the method employed to test for a link between risk preferences and time discounting.

3.1 Behavior under Risk

Modeling decisions under risk encompasses two components, a model of behavior on the one hand, and assumptions regarding decision errors on the other hand. Risk taking behavior is modeled by rank dependent utility theory (RDU).

According to RDU, an individual values a two-outcome lottery $\mathcal{L} = (x_1, p; x_2)$, where $x_1 > x_2 \geq 0$, by $w(p)u(x_1) + (1 - w(p))u(x_2)$. The function $u(x)$, with $u(0) = 0$ and $u'(x) > 0$, describes how monetary outcomes x are subjectively valued. The function $w(p)$ assigns a subjective weight to every

[7]People entitled to payoffs the next day were issued vouchers immediately after the experiment. All the others received official certificates of indebtedness after the experiment, the vouchers themselves were sent to them by registered mail several days before they could cash them in, so they did not have to worry about forgetting encashment or misplacing their vouchers.

outcome probability p, with $w(0) = 0$, $w(1) = 1$, and $w'(p) > 0$. The decision maker's predicted certainty equivalent $\hat{ce}$ for this lottery can then be written as

$$\hat{ce} = u^{-1}\left[w(p)u(x_1) + (1 - w(p))u(x_2)\right]. \tag{1}$$

In order to make RDU operational, we have to assume specific functional forms for the utility function $u(x)$ and the probability weighting function $w(p)$. Given our objective of describing individual behavior, we choose flexible functional forms for u as well as for w. A natural candidate for utility u is a power function. In its extensive form, as discussed by Wakker (2008), u is modeled as[8]

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

A variety of parameterizations of probability weighting functions $w(p)$ have been proposed in the literature (Karmarkar, 1979; Lattimore, Baker, and Witte, 1992; Tversky and Kahneman, 1992; Prelec, 1998). Since our primary interest lies in common ratio violations we focus on a specific characteristic of the weighting function, subproportionality. Subproportionality means that for a fixed ratio of probabilities the ratio of the corresponding probability weights is closer to unity when the probabilities are low than when they are high (Kahneman and Tversky, 1979). Intuitively speaking, scaling down the original probabilities makes them less distinguishable from each other, thus favoring preference reversals. Therefore, subproportionality maps common ratio violations. Expressed formally (Prelec, 1998), *subproportionality* holds if $1 \geq p > q > 0$ and $0 < \lambda < 1$ implies the inequality

$$\frac{w(p)}{w(q)} > \frac{w(\lambda p)}{w(\lambda q)}. \tag{2}$$

As we rely on subproportionality as the crucial characteristic of the probability weighting function, to be estimated for each single individual, we adopt the flexible and empirically well-founded two-

[8] Note that $\ln x$ is not defined for $x = 0$. Therefore, estimation is carried out after shifting all outcomes by one unit of money (cmp. Wakker (2008), p.1335).

parameter specification suggested by Prelec (1998),

$$w(p) = e^{-\beta(-\ln p)^{\alpha}}. \tag{3}$$

For $\alpha < 1$, the function is subproportional everywhere, with the parameter α measuring the degree of subproportionality.[9] A smaller value of α reflects a more subproportional curve. Therefore, this specification enables us to rank individuals according to their their proneness to common ratio violations. The second parameter, $\beta > 0$, is a net index of convexity in that increasing β increases the convexity of the function without affecting subproportionality (Prelec (1998), p.505). Linear weighting is characterized by $\alpha = \beta = 1$.

With regard to error specification we have to reconsider our measurement procedure. In the course of the experiment, an individual's risk taking behavior was captured by her certainty equivalents ce_l for a set of 20 different lotteries $\mathcal{L}_l = (x_{1l}, p_l; x_{2l}),\ l \in \{1, \ldots, 20\}$. Since RDU explains *deterministic* choice, actual certainty equivalents ce_l are likely to deviate from the predicted certainty equivalents $\hat{ce}_l$ by a stochastic error ϵ_l, which has to be taken account of. Therefore, we assume that the observed certainty equivalents ce_l can be expressed as $ce_l = \hat{ce}_l + \epsilon_l$, with ϵ_l being normally distributed with zero mean.[10]

Concerning the error variance, we need to account for heteroskedasticity: For each lottery subjects had to consider 20 guaranteed outcomes, equally spaced throughout the lottery's outcome range $x_{1l} - x_{2l}$. Since the observed certainty equivalent ce_l is calculated as the arithmetic mean of the smallest guaranteed amount preferred to the lottery and the subsequent guaranteed amount, the error is proportional to the outcome range. Therefore, the standard deviation ν_l of the error term distribution has to be normalized by the outcome range, yielding $\nu_l = \nu(x_{1l} - x_{2l})$, where ν denotes an additional parameter to be estimated. In total, therefore, four parameters per subject were estimated by maximum likelihood: the curvature of the utility function η, subproportionality and convexity of the probability weighting function α and β, as well as the normalized standard deviation of the

[9] Prelec (2000) uses the term "Allais paradox index" (p.78).

[10] Since *ce* is calculated as the arithmetic mean of two neighboring amounts in the choice menu it possibly contains some measurement error. As *ce* is the dependent variable in the model a measurement error does not pose a problem other than potentially increasing noise.

decision error parameter ν.

3.2 Behavior over Time

Subjects' responses to the intertemporal choice tasks in the experiment provided us with measurements of discount rates δ_2 and δ_4, imputed from the intertemporal tradeoffs between present payments pe and payments x delayed by two and four months, respectively. However, the true underlying discount factors $D(t)$ are defined in terms of utilities, not payoffs.[11] For a temporal prospect $\mathcal{T} = (x, t)$, true discount rates are inferred from the indifference relation $u(pe) = D(t)u(x)$. Measured discount rates, therefore, deviate from the underlying true rates unless u is linear. While in our specification utility curvature affects the level of discount rates but cannot, by itself, induce induce their decline, it may have a confounding effect on the magnitude of the change in the measured discount rates $\Delta\delta = \delta_2 - \delta_4$: In the presence of nonlinear probability weighting, $\Delta\delta$ gets amplified by the concavity of the power utility function (see Appendix A.4). Specifically, the more concave u, the larger the measured difference in the discount rates $\Delta\delta$ if $w(p)$ is not linear. Therefore, we have to control for the degree of concavity η in the regression model.

3.3 Regression Model

We investigate the hypothesized relationship between probability weighting and decreasing discount rates by regressing the difference between the imputed discount rates δ_2 and δ_4, $\Delta\delta$, on a vector of regressors c. In the base model, Model 1, the vector c consists of a constant and the individuals' estimated risk preference parameters: η captures concavity of the utility function, α captures subproportionality of the probability weighting function, and β its convexity. Additionally, we estimate an extended version of the base model, Model 2, by controlling for a set of individual characteristics. In particular, these controls comprise gender (labeled *Female*), age *(Age)*, the logarithm of disposable income per month *(Log-Income)*, a binary variable indicating whether the subject is familiar with investment decisions *(Experience)* as well as the test score for the *Cognitive Reflection Test (CRT)*

[11] We assume that the utility of money is a general cardinal function which applies to risky as well as to delayed payoffs. See Wakker (1994) for a justification of this assumption.

(Frederick, 2005).[12] This three-question test measures specific aspects of cognitive ability which were found to be strongly correlated with risk taking and discounting behavior.

Unlike the exemplary choice pattern displayed in Figure II, a decision maker may have opted for the same option in all rows of the choice menu, which results in a censored observation. In particular, she may have always preferred the smaller sooner option, indicating that her discount rate may lie beyond the maximum value of 95%.[13] As a consequence, the difference between the observed discount rates δ_2 and δ_4 is affected by censoring as well. As ordinary least square (OLS) yields biased estimates in this case, we account for this issue by a censored regression model, described in detail in Appendix B. The model has the following form:

$$\Delta\delta_i^* = c_i \underbrace{(\gamma_2 - \gamma_4)}_{\Delta\gamma} + \underbrace{e_{2,i} - e_{4,i}}_{\Delta e_i}\,, \tag{4}$$

where $\Delta\delta_i^*$ specifies the true, but potentially unobserved, difference between δ_2 and δ_4 for individual i, $i \in \{1, \ldots, 112\}$. The error term Δe_i is normally distributed with mean zero and variance σ^2. The interpretation of the regression coefficients $\Delta\gamma$ is equivalent to those of OLS regression, also displayed in the regression output (Table III below).

4 Results

In the following section we analyze the raw data on risk taking behavior and time discounting, and present the estimates for subjects' probability weights. Finally, we examine the relationship between subjects' sensitivities with respect to changes in probability and delay.

4.1 Descriptive Analysis

For the domain of risk taking, Figure III summarizes observed behavior by the median relative risk premia $RRP = (ev - ce)/|ev|$, where ev denotes the expected value of a lottery's payoff and ce stands for the observed certainty equivalent. $RRP > 0$ indicates risk aversion, $RRP < 0$ risk seeking, and

[12] Summary statistics of the controls are included in Appendix E, Table E.1.

[13] A decision maker may also always prefer the later larger option. In this case, we assume a discount rate of 0%. The number of observations at the boundary of the choice menu are listed in Table E.2 of Appendix E.

$RRP = 0$ risk neutrality. The median relative risk premia, sorted by the probability p of the higher gain, show a systematic relationship between aggregated risk attitudes and lottery probabilities: Subjects' choices display the familiar pattern, i.e. they are risk averse for high-probability gains, but risk seeking for low-probability gains, supporting the existence of nonlinear probability weighting.

As far as intertemporal choices are concerned, people's average behavior exhibits decreasing discount rates, i.e. subjects discount less remote outcomes more strongly than more remote ones: The first column in Table II reveals that the discount rates imputed from subject's choices decline on average by 7 percentage points *per annum* when the time horizon is extended from two months to four months. The average data veil heterogeneity as well as the extent of decreasing discount rates, however. Whereas the majority of approximately 54% of all subjects exhibit decreasing discount rates over time, $\Delta\delta > 0$ (second column), about 29% exhibit constant discount rates (third column), and the residual group reveals increasing discount rates (fourth column). Average discount rates of subjects with decreasing discount rates amount to $\delta_2 = 47\%$ $p.a.$ and $\delta_4 = 31\%$ $p.a.$, respectively, reflecting a much greater change than do the overall averages.[14]

4.2 Risk Preference Parameters

Whereas one of our central variables, change in discount rates $\Delta\delta$, is directly observable, the other one, departure from linear probability weighting, has to be estimated from our data on certainty equivalents.

Individual risk preference parameters η, α and β were estimated on the basis of the econometric model discussed in Section 3.1. As Table II reveals, the values of the curvature parameter η of the utility function reflect slight concavity or linearity. The average subproportionality index α amounts to 0.505, indicating a pronounced departure from linear probability weighting in line with previous findings (Tversky and Kahneman, 1992; Gonzalez and Wu, 1999; Abdellaoui, 2000). The average estimates for β lie in the vicinity of one, implying that the respective curves intersect the diagonal at about $p = 1/e$.[15]

The overall picture revealed by our data is consistent with the typical empirical findings: On

[14] The distributions of the observed discount rates are shown in Appendix C.
[15] Histograms of the parameter distributions are included in Appendix D.

average, subjects systematically violate linear probability weighting and constant discounting. But the central question, namely whether the degree of subproportionality of probability weighting is associated with hyperbolicity of discounting behavior at the level of the *individual* has yet to be answered.

4.3 Relationship between Probability Weights and Hyperbolic Discounting

A first indication of a systematic relationship between probability weighting and discounting can be found in Table II. The average estimated subproportionality index α varies substantially across discounting types and exhibits a systematic pattern: α is lowest for the group with decreasing discount rates and highest for the group with increasing discount rates.

This finding is confirmed by the estimates of the regression models. Table III displays the results derived by OLS as well as by the censored regression method. Inspection of the coefficients indicates that censoring seems not to be an important problem: After omitting the 23 censored observations, OLS yields coefficients very close to the estimates of the censored regression model. Furthermore, both specifications (Models 1 and 2) lead to the same conclusion: Subproportionality of probability weighting is significantly associated with decreasing discount rates $\Delta\delta$. Table III shows the estimated coefficient of α to be approximately -0.2. All the respective estimates are significant at the 1%-level and remain robust to inclusion of additional controls. The effect is not only highly significant, it is also quite substantial: A decrease in the subproportionality index α by 0.1 is associated with an increase in $\Delta\delta$ by 2 percentage points *per annum.* In particular, the decline in discount rates is related to the degree of subproportionality, but not to the index of convexity β. We obtained the same order of magnitude for the coefficient of α when we restricted β to be equal to one. Regression coefficients and their standard errors also remain stable when either η or β are deleted from the list of regressors. Moreover, it can be shown that estimates are totally robust to alternative parameterizations of the probability weighting curve as well.[16]

The coefficients of the utility parameter η are not statistically different from zero, either.[17] This result is consistent with our hypothesis that utility curvature *per se* does not impact the extent of

[16]Results are available upon request.
[17]Nor does an interaction term $\alpha \times \eta$ contribute to explaining variation in $\Delta\delta$.

decreasing discount rates. Furthermore, none of the other individual characteristics show a significant effect.[18]

An F-test comparing the OLS Model 1 with Model 2 renders a p-value of 0.670, favoring the more parsimonious Model 1, as the controls do not substantially contribute to explaining the variance in $\Delta\delta$.[19] Furthermore, the regression models explain a rather large fraction of total variance: Model 2, for instance, yields an R-squared value of 17%.[20] These findings present conclusive evidence that comparatively more subproportional probability weighting is associated with a stronger decline in discount rates.

5 Discussion

The strong and significant correlation between subproportionality of probability weighting and extent of hyperbolic discounting begs the question of whether this relationship can be explained in causal terms. In principle, there are three pathways through which correlation could be generated. First, the tendency towards hyperbolic discounting could cause distortions in probability weights. Since, in experiments, estimates of probability weights are generally based on atemporal choices, i.e. when there is practically no time delay between choice and payment, this possibility can be effectively ruled out. Second, the direction of causality could work the other way round, with proneness to probability distortions inducing hyperbolic discounting. Finally, there could be a third factor driving both types of departures from the standard model predictions. We will discuss the latter possibility first and then turn to the second alternative.

Since the *Common Ratio Effect* and the *Common Difference Effect* pertain to diminishing sensitivity towards probability and delay, respectively, similar cognitive processes may govern the evaluation of risky and delayed outcomes. A natural candidate for a common factor driving both processes is cognitive abilities. Several papers have looked into the relationship between cognitive abilities

[18] While not significantly different from zero, coefficients exhibit the expected signs: Females have a slightly more subproportional weighting curve, consistent with previous findings (Fehr-Duda, de Gennaro, and Schubert, 2006). Both experience with investment decisions and high *CRT* scores are associated with smaller departures from linearity.

[19] The same is the case when the two censored models are compared. A likelihood ratio test of Model 2 against Model 1 renders a p-value of 0.9.

[20] When regressing $\Delta\delta$ exclusively on the socioeconomic variables, R-squared amounts to 3.9%!

and risk tolerance on the one hand, and between cognitive abilities and patience on the other hand (Frederick, 2005; Benjamin, Sebastian, and Shapiro, 2006; Dohmen, Falk, Huffman, and Sunde, 2107). Generally, they conclude that better cognitive abilities tend to be associated with higher risk tolerance as well as higher patience. For instance, Frederick (2005) finds that students with high *Cognitive Reflection Test* scores gambled significantly more often that did the low *CRT* group, and exhibited lower imputed discount rates, albeit not for choices involving longer time horizons. These previous findings seem to conflict with the insignificant coefficient of *CRT* in Table III. Since *CRT* is not correlated with α,[21] the lack of correlation between *CRT* and $\Delta\delta$ indeed suggests that *CRT* scores cannot explain the variance in the hyperbolicity of discount rates. However, we are concerned with sensitivities towards changes in probability and delay and not with measures of average risk aversion and impatience, the focus of previous research. Of course, there could be other factors than cognitive abilities, or aspects of cognitive ability not captured by *CRT* scores, that drive the correlation between subproportionality and hyperbolic discounting. Clearly, this possibility cannot be ruled out and needs further exploration.

Finally, we discuss the last one of our options, direct impact of subproportionality on hyperbolicity of discounting. Many authors have noted before that "[a]nything that is delayed is almost by definition uncertain" (Prelec and Loewenstein (1991), p.784). For instance, a promised reward may, due to unforeseen circumstances, materialize later or turn out to be smaller than expected, or death may keep the decision maker from collecting her reward at all. For these reasons, future consequences are inextricably associated with uncertainty, implying that the decision maker's valuation of temporal prospects not only depends on her *pure* time preference, i.e. her preference for immediate utility over delayed utility, but also on her perception of uncertainty and, consequently, on her risk preferences. In other words, uncertainty drives a wedge between pure time preferences and time discounting.

If this account is an accurate description of intertemporal choice it has far reaching implications for observed discounting behavior, the most obvious one being that behaviorally revealed discount rates will be higher than the rate of *pure* time preference as they include a risk premium. Not surprisingly

[21] The Pearson correlation coefficient amounts to -0.0049 (p-value 0.964).

then, uncertainty has been identified to be an important confound in the measurement of time preferences, which may, at least partly, explain the notoriously high discount rates found in empirical studies (Frederick, Loewenstein, and O'Donoghue, 2002). The story does not stop here, however. If risk preferences influence time discounting, then people's proneness to probability weighting has to be taken into account as well. Recent contributions have examined the impact of nonlinear probability weighting on discounting behavior theoretically (Halevy, 2008; Walther, 2010). Halevy, for instance, motivated by interaction effects between time and risk found by Keren and Roelofsma (1995) and Weber and Chapman (2005), is concerned with convex probability transformations that can accommodate the certainty effect inherent in the classical Allais paradox. We investigate the more general case of common ratio violations which can be modeled by subproportional probability weights. Subproportionality is not confined to convex functions but may also be present in inverse S-shaped probability transformations, which organize a large part of the empirical evidence. In the following, we show that the degree of subproportionality of probability weights indeed predicts the extent of decreasing discount rates. Furthermore, we derive comparative static results with respect to degree of uncertainty and exemplify the model predictions by graphical illustrations.

5.1 A Model of Discounting: The Warped Lens

If the future is perceived as uncertain an allegedly guaranteed delayed outcome $\mathcal{T} = (x, t)$ is effectively evaluated as a risky prospect. Suppose that any future payment is perceived to materialize with a constant per-period probability of contract survival s, $0 < s \leq 1$. Consequently, $\mathcal{T}$ is evaluated as $\mathcal{L} = (x, s^t)$, rendering x with probability s^t and zero otherwise.

As far as the rate of pure time preference is concerned, we adopt the conventional assumption: the rate of pure time preference is characterized by a constant per-period rate $r \geq 0$, resulting in a pure time discount factor ρ equal to e^{-r}.

These assumptions imply that the present equivalent pe of the future payment x, such that the decision maker is indifferent between pe and x, is defined by

$$u(pe) = w(s^t)\rho^t u(x). \tag{5}$$

The *effective* discount factor $D(t)$ at delay t equals the weight attached to $u(x)$, i.e.

$$D(t) = w(s^t)\rho^t, \tag{6}$$

which depends not only on the *pure* rate of time preference r, but also on the probability of contract survival s as well as on the shape of the probability weighting function w. Clearly, if w is linear, $D(t)$ declines exponentially irrespective of the magnitude of s. If $0 < s < 1$, the resulting discount factor is lower than for $s = 1$, implying that uncertainty *per se* increases the absolute level of discount rate, but cannot account for discount rates declining over time. In fact, due to uncertainty, discounting would be observed even for a zero rate of pure time preference. If, however, w is nonlinear and $0 < s < 1$, the component $w(s^t)$ distorts the discount factor: As shown formally in Appendix A.1, subproportionality of w generates hyperbolicity of $w(s^t)$ in t and, consequently, decreasing discount rates if the future is perceived as uncertain. Metaphorically speaking, the decision maker, when looking into the future, perceives delayed events through the warped lens of probability distortions. A natural extension of this insight is that higher degrees of subproportionality induce more strongly declining discount rates (see Appendix A.2). The effective discount factor $D(t)$ also depends on the level of uncertainty s. Higher uncertainty implies more strongly declining discount rates as well. A formal proof appears in Appendix A.3.

In order to illustrate the predictions of our model, which hold for any subproportional probability weighting function, we demonstrate the comparative static effects of subproportionality α and uncertainty s graphically, using Prelec's specification. The panels on the left-hand side of Figure IV show the comparative static effects of varying α, the right-hand side ones are dedicated to varying levels of s.

Panel *A1* of Figure IV depicts the probability weighting curves for three distinct parameter values of α, with $\beta = 1$: a medium-sized departure from linearity ($\alpha = 0.5$), as exhibited on average by our experimental subjects, a strong departure from linearity ($\alpha = 0.2$), as well as the limiting case of linear probability weighting ($\alpha = 1$). Panel *B1* of Figure IV shows, for each of the three cases of probability weighting, the effective discount factors resulting from Equation 6 as they evolve over time.[22] For a decision maker with linear probability weighting the discount function, represented by the solid gray curve, is exponential. In contrast, the dotted discount function of a typical decision maker with $\alpha = 0.5$ departs from exponentiality, exhibiting an apparently hyperbolic pattern. By comparison, the decision maker characterized by the most strongly S-shaped probability weighting curve underweights (overweights) large (small) probabilities more strongly than does the decision maker with $\alpha = 0.5$, which leads to an even more pronounced departure from exponential discounting (dashed curve).

Finally, Panel *C1* of Figure IV displays the imputed discount rates d_t inferred from $D(t) = e^{-d_t t}$. The solid gray line corresponds to linear probability weighting. Since this decision maker is not prone to probability distortions, her discount rate is independent of time delay and, consequently, constant over time. In contrast to this decision maker, the discount rates of the decision makers with nonlinear probability weights start out at very high levels and then decline sharply. As is evident from comparing the dashed curve with the dotted one, the more subproportional probability weighting function generates a larger decline in discount rates between period 2 and period 1, i.e. the difference $d_2 - d_1$ is greater for higher degrees of subproportionality α. For this prediction to hold the probability of contract survival s needs to be smaller than one. Since people vary in their perceptions of uncertainty our framework predicts a correlation between subproportionality and decreasing discount rates. This is exactly what we find in our data.

Another important insight from our approach concerns the direct impact of uncertainty on discounting behavior. Hyperbolicity of discount rates is crucially influenced by people's perceptions of uncertainty: Increasing uncertainty not only raises the level of discount rates but also exacerbates revealed short-term impatience. Panels *A2* to *C2* in Figure IV illustrate this effect for $\alpha = 0.5$, the average index of subproportionality in our data, and $r = 0.1$. When the survival probability s declines from 0.8 to 0.5, the resulting discount function departs more strongly from exponentiality as Panel *B2* shows. Hence, the decrease in discount rates associated with higher uncertainty is more pronounced as well (Panel *C2*).

[22] For illustrative purposes, in Figure IV r is fixed at 0.1 and s is assumed to be 0.8, which means that 80% of all contracts are anticipated to survive at least one period.

5.2 Perceived Uncertainty and the Pure Rate of Time Preference

The model presented in the previous section provides a theoretical underpinning of our empirical finding that the degree of subproportionality α predicts the extent of hyperbolic discounting $\Delta\delta$. Our theoretical framework implies such a relationship *ceteris paribus*, holding constant the other model parameters, specifically the subjective probability of contract survival s and the pure rate of time preference r, both of which are not observable. In our experimental setting with decisions over a short time horizon, the subjective probability of contract survival s should lie very close to unity since mortality risk is very low in our age group of subjects and we took great care to communicate experimenter reliability. One way of checking the plausibility of the theoretical model is to investigate whether, on average, actual choices are indeed consistent with this conjecture, i.e. whether values of s implied by our data lie in the vicinity of one *for a wide range of plausible values of the pure rate of time preference* r.

For this purpose, we examine the combinations of s and r that are consistent with the observed average intertemporal tradeoffs between more immediate and more remote payments pe and x. We solve for all feasible combinations of $\hat{s}$ and $\hat{r}$ that are compatible with the observed choices by inserting the estimates for subjects' average behavioral parameters η, α and β into Equation 5. As is clear from Equation 5, a higher probability of contract survival needs to be compensated by a higher pure rate of time preference, *ceteris paribus*, to keep individuals indifferent between more immediate and more remote rewards.

As Figure V shows, the feasible $(\hat{s}, \hat{r})$-combinations indeed exhibit a rising profile, with $\hat{s}$ starting out at below 97% *p.a.* and converging to 100% *p.a.*, when the pure rate of time preference increases from 0% to 15% *p.a.* and beyond. For instance, $s = 99\%$ is compatible with $r \simeq 8.5\%$ *p.a.* What this exercise shows is that the data, interpreted within our framework, are consistent with a very high subjective probability that contracts survive at least one year, in accordance with our conjecture. Furthermore, accounting for inherent uncertainty implies rates of pure time preference in a plausible range lying considerably below the observed average discount rates of more than 30% *p.a.*

This suggests that even allegedly guaranteed future outcomes are viewed as slightly uncertain, in line with direct questionnaire evidence provided in Patak and Reynolds (2007). The authors asked respondents to rate their certainties for the same rewards, delayed by 1, 2, 30, 180, and 365 days, respectively, which they had encountered during the preceding choice experiment. The respondents reported ratings that clearly declined with the length of delay. Moreover, using a similar method, Takahashi, Ikeda, Hasegawa, and Greene (2007) found that such subjective probabilities of obtaining delayed rewards decay in a hyperbolic-like manner, consistent with probability weights $w(s^t)$ declining hyperbolically with delay t.

6 Conclusion

For several decades, decision research has been dominated by the quest for better descriptive theories of behavior under risk and over time, triggered by a large body of experimental evidence challenging the classical models of choice, expected utility theory and discounted utility theory. Alternative models, accounting for nonlinear probability weighting and hyperbolic discounting, describe behavior much more accurately than do the classical models, at least at the aggregate level. In this paper we address the question whether the better fit of the generalized models is actually a consequence of the same subjects' anomalous behaviors. We present the first evidence that more pronounced systematic departures from linear probability weighting are indeed associated with more strongly declining discount rates at the level of the individual decision makers. This result is robust to inclusion of additional controls as well as model specification. In fact, the only variable explaining a substantial fraction of heterogeneity in individual discounting patterns turns out to be the degree of subproportionality of probability weights.

Several authors have proposed that the existence of matching violations of the classical axioms is not coincidental, but rather reflects the close relationship between risk and delay (e.g. Prelec and Loewenstein (1991)). Some researchers have even argued that the two attributes are virtually the same, but there is no consensus as to which one is the more fundamental of the two. We favor the view that, if there is a hierarchical relationship between them at all, risk is the more likely candidate. To bolster this view, we provide a theoretical model predicting the observed link between probability distortions and decreasing discount rates. For hyperbolic discounting to emerge two factors need to interact: probability distortions and future uncertainty.

Arguably, the future is uncertain by definition. Uncertainty may stem from different sources, either tied to the individual herself, such as lifetime expectancy, or to environmental factors. Lack of contract enforcement and weak property rights, for instance, may make people skeptical that promises will be actually kept. Therefore, institutionally generated uncertainties may induce extreme short-run impatience even if people's pure rate of time preference is low and constant. This insight is important because it implies that revealed behavior may be predominantly driven by environmental factors rather than by the underlying preferences themselves and, consequently, may be amenable to economic policy.

The channel through which uncertainty generates hyperbolic discounting is nonlinear probability weighting, a robust phenomenon in the empirical literature. If probability weighting plays such an important role in risk taking and discounting behavior, the obvious question concerning the source of these probability distortions arises. Unfortunately, little is known about the forces driving probability distortions. A number of theoretical contributions have invoked emotions to explain probability weighting (Wu, 1999; Caplin and Leahy, 2001). Walther (2003, 2010), for instance, rationalizes nonlinear probability weighting by generalizing expected utility theory: He assumes that, in addition to monetary outcomes, the decision maker cares about emotions triggered by the resolution of uncertainty. His approach predicts that, if the decision maker anticipates experiencing elation or disappointment when the actual outcome lies above or below some normal level, she will distort outcome probabilities according to an S-shaped pattern. The more emotional a person expects to be, the stronger will be her departure from linear probability weighting and, consequently, the more pronounced hyperbolic discounting will be. Of course, sensitivity to anticipated emotions is not easily observable, and we have to leave it to future research to investigate whether anticipated emotions or some other factors are the primary drivers of probability weighting.

Tables

Table I: Risky Prospects

p	x_1	x_2	p	x_1	x_2
0.1	20	10	0.25	50	20
0.5	20	10	0.5	50	20
0.9	20	10	0.75	50	20
0.05	40	10	0.95	50	20
0.25	40	10	0.05	150	50
0.5	40	10	0.5	10	0
0.75	40	10	0.5	20	0
0.95	40	10	0.05	40	0
0.05	50	20	0.95	50	0
0.1	150	0	0.25	40	0

Outcomes x_1 and x_2 are stated in CHF, p denotes the probability of x_1 materializing.

Table II: Average Discount Rates and Risk Parameters

	all	Subjects with $\Delta\delta > 0$	Subjects with $\Delta\delta = 0$	Subjects with $\Delta\delta < 0$
	100%	53.9%	29.2%	16.9%
δ_2	0.368	0.465		0.328
	(0.023)	(0.029)	0.213	(0.058)
δ_4	0.299	0.307	(0.045)	0.418
	(0.020)	(0.025)		(0.068)
$\Delta\delta$	0.070	0.157	0	-0.090
	(0.012)	(0.015)		(0.019)
η	0.873	0.808	0.948	0.953
	(0.032)	(0.046)	(0.074)	(0.072)
α	0.505	0.426	0.574	0.634
	(0.021)	(0.027)	(0.040)	(0.063)
β	0.974	0.936	1.064	0.940
	(0.026)	(0.036)	(0.068)	(0.045)
Observations	89	48	26	15

Standard errors in parentheses. Excluding 23 subjects with censored observations.

Table III: Regression Results

Dependent Variable: $\Delta\delta$ ($\Delta\delta^*$)

	OLS[a]		Censored	
	Model 1	Model 2	Model 1	Model 2
Intercept	0.226***	0.279	0.247***	0.321
	(0.063)	(0.228)	(0.057)	(0.225)
η	0.018	0.002	-0.006	-0.022
	(0.042)	(0.043)	(0.039)	(0.041)
α	-0.205***	-0.220***	-0.185***	-0.203***
	(0.066)	(0.074)	(0.062)	(0.075)
β	-0.070	-0.040	-0.074	-0.045
	(0.067)	(0.068)	(0.060)	(0.063)
Female		-0.012		-0.011
		(0.031)		(0.032)
Age		-0.001		-0.002
		(0.007)		(0.007)
Log-Income		-0.013		-0.012
		(0.024)		(0.023)
Experience		0.015		0.020
		(0.032)		(0.033)
CRT		0.021		0.021
		(0.017)		(0.017)
$\hat{\sigma}$	0.123	0.124	0.084	0.082
R^2 or (LogLik)	0.137	0.170	(48.693)	(51.123)
Observations	89	89	112	112
Parameters	4	9	9	19

a) without censored observations.

*** significant at the 1% level.

Bootstrapped standard errors in parentheses (10,000 replications). Bootstrapping accounts for the fact that the regressors α, β and η are estimated quantities.

Figures

Figure I: Choice Menu — Risk

	Option A	**Your Choice**		**Option B** *(guaranteed reward)*
1	Gain of CHF 50 with a probability of 75% *and* Gain of CHF 10 with a probability of 25%	A ○	● B	CHF 50
2		A ○	● B	CHF 48
3		A ○	● B	CHF 46
4		A ○	● B	CHF 44
5		A ○	● B	CHF 42
6		A ○	● B	CHF 40
7		A ○	● B	CHF 38
8		A ●	○ B	CHF 36
9		A ●	○ B	CHF 34
10		A ●	○ B	CHF 32
11		A ●	○ B	CHF 30
12		A ●	○ B	CHF 28
13		A ●	○ B	CHF 26
14		A ●	○ B	CHF 24
15		A ●	○ B	CHF 22
16		A ●	○ B	CHF 20
17		A ●	○ B	CHF 18
18		A ●	○ B	CHF 16
19		A ●	○ B	CHF 14
20		A ●	○ B	CHF 12

Figure II: Choice Menu — Time

	Option A *payment tomorrow*	**Your Choice**		**Option B** *payment in 4 months + 1 day*	
1		A ○	● B	CHF 79	95%
2		A ○	● B	CHF 78	90%
3		A ○	● B	CHF 77	85%
4		A ○	● B	CHF 76	80%
5		A ○	● B	CHF 75	75%
6		A ○	● B	CHF 74	70%
7		A ○	● B	CHF 73	65%
8		A ○	● B	CHF 72	60%
9		A ○	● B	CHF 71	55%
10	Amount of CHF 60	A ●	○ B	CHF 70	50%
11		A ●	○ B	CHF 69	45%
12		A ●	○ B	CHF 68	40%
13		A ●	○ B	CHF 67	35%
14		A ●	○ B	CHF 66	30%
15		A ●	○ B	CHF 65	25%
16		A ●	○ B	CHF 64	20%
17		A ●	○ B	CHF 63	15%
18		A ●	○ B	CHF 62	10%
19		A ●	○ B	CHF 61	5%
20		A ●	○ B	CHF 60	0%

Figure III: Median Relative Risk Premia *(RRP)*

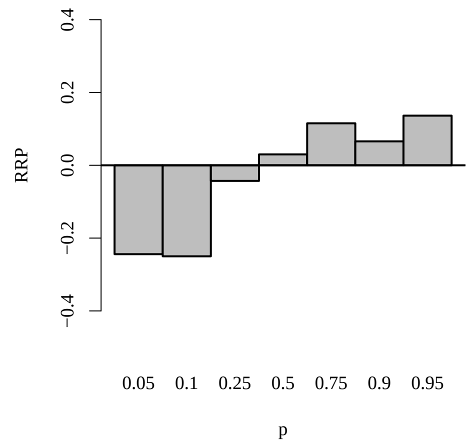


Figure IV: Probability Distortions and Discounting

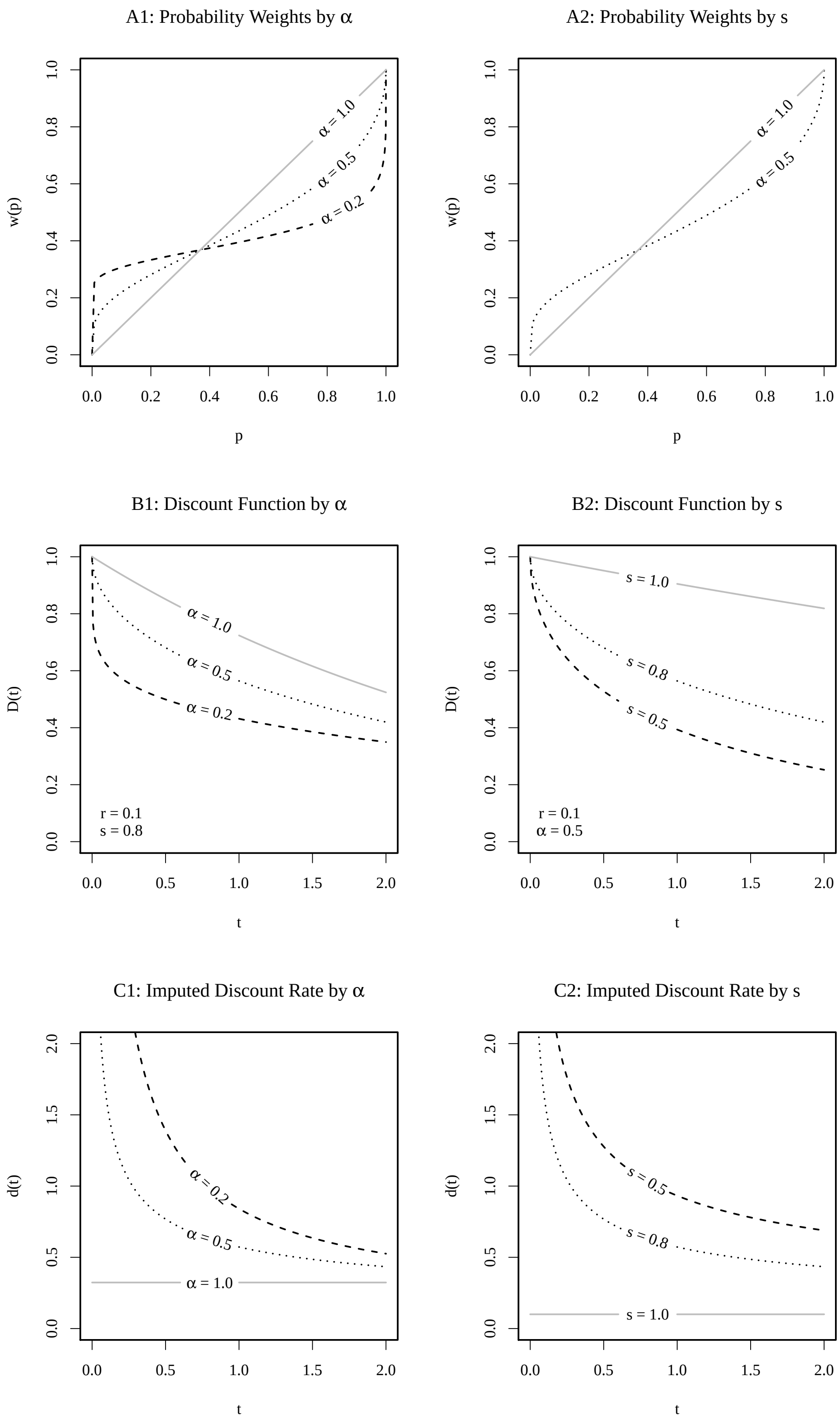
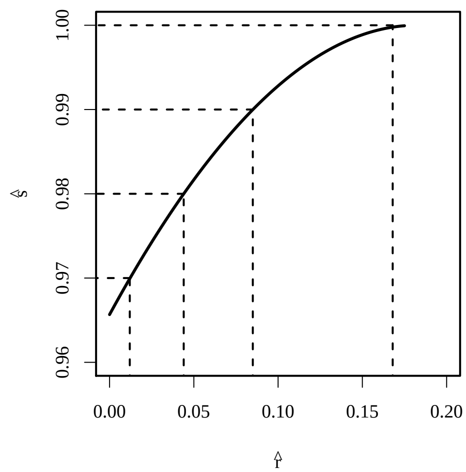


Figure V: Feasible $(\hat{s}, \hat{r})$-Combinations



A Formal Proofs

A.1 Hyperbolicity

In the framework proposed here, the discount factor $D(t)$ equals

$$D(t) = w(s^t)\rho^t, \tag{A.1}$$

with ρ defined as e^{-r}. In order to establish that subproportional probability weights are sufficient[23] for discount rates to decrease, we define decreasing impatience at t in the following way (Prelec, 2004): Let (x, t) be a temporal prospect paying off x at t with certainty. A preference relation $\succeq$ exhibits *decreasing impatience* if for any $t > 0$, $0 < x < y$, $(x, v) \sim (y, z)$ implies $(y, z+t) \succeq (x, v+t)$.

According to our framework the temporal prospects $(x, 0) \sim (y, 1)$ are evaluated as $u(x)w(s^0)\rho^0 = u(y)w(s^1)\rho^1$. As subproportionality of w implies that $w(s) < w(s^{t+1})/w(s^t)$, deferring the prospects by t periods renders

$$1 = \frac{u(y)w(s)\rho}{v(x)} < \frac{u(y)w(s^{t+1})\rho^{t+1}}{u(x)w(s^t)\rho^t} \tag{A.2}$$

and, therefore, $(y, t+1) \succ (x, t)$, meeting the requirement for decreasing impatience. ■

In the intertemporal tradeoff between the present and the subsequent period the discount factor equals $w(s)\rho$. At time t, $u(x)$ is discounted by $w(s^t)\rho^t$. Compounding by the initial one-period discount factor $w(s)\rho$ would render $w(s)w(s^t)\rho^{t+1}$ at $t+1$, but the discount factor effectively amounts to $w(s^{t+1})\rho^{t+1}$ then. Therefore, $w(s^{t+1})/(w(s)w(s^t))$, the wedge between the relative discount factors $D(0)/D(1)$ and $D(t)/D(t+1)$, provides a measure for the extent of departure from stationarity at t.

A.2 Comparative Hyperbolicity

The previous proof shows that, provided that $s < 1$, subproportionality of w engenders hyperbolic discounting. As will become clear shortly, a decision maker with a comparatively more subproportional probability weighting function will also tend to exhibit more strongly decreasing discount

[23] Note that subproportionality is not necessary.

rates:

A preference relation $\succeq_2$ exhibits *more strongly decreasing impatience* than $\succeq_1$ if for any intervals $0 \leq v < z, t, \Delta t$ and outcomes $0 < x < y$, $0 < x' < y'$, $(x, v) \sim_1 (y, z)$, $(x, v + t) \sim_1 (y, z + t + \Delta t)$, and $(x', v) \sim_2 (y', z)$ imply $(x', v + t) \preceq_2 (y', z + t + \Delta t)$ (Prelec, 2004).

In order to examine the effect of the degree of subproportionality on hyperbolicity suppose that the probability weighting function w_2 is comparatively more subproportional than w_1, as defined in Prelec (1998), and that the following indifference relations hold for two decision makers 1 and 2 at periods $v = 0$ and $z = 1$:

$$u_1(x) = u_1(y)w_1(s)\rho \text{ for } 0 < x < y,$$
$$u_2(x') = u_2(y')w_2(s)\rho \text{ for } 0 < x' < y'.$$

Due to subproportionality, the following relation holds for decision maker 1 in period t:

$$1 = \frac{u_1(y)w_1(s)\rho}{u_1(x)} < \frac{u_1(y)w_1(s^{t+1})\rho^{t+1}}{u_1(x)w_1(s^t)\rho^t}. \tag{A.3}$$

Therefore, the subjective probability of contract survival has to be reduced by compounding s over an additional time period Δt to re-establish indifference:

$$u_1(x)w_1(s^t)\rho^t = u_1(y)w_1(s^{t+1+\Delta t})\rho^{t+1}. \tag{A.4}$$

It follows from the definition of comparative subproportionality that this adjustment of the survival probability by Δt is not sufficient to re-establish indifference with respect to w_2, i.e.

$$u_2(x')w_2(s^t)\rho^t < u_2(y')w_2(s^{t+1+\Delta t})\rho^{t+1}. \tag{A.5}$$

Therefore, $(x', t) \prec (y', t + 1 + \Delta t)$. ■

A.3 Uncertainty and Hyperbolicity

In order to derive the effect of increasing uncertainty on hyperbolicity we examine the sensitivity of the departure from stationarity at t, measured by $w(s^{t+1})/(w(s)w(s^t))$, with respect to changing s:

$$
\begin{aligned}
& \frac{\partial}{\partial s}\left[\frac{w(s^{t+1})}{w(s)w(s^t)}\right] \\
= \;& \frac{1}{[w(s)w(s^t)]^2}\left[(1+t)s^t w(s)w(s^t)w'(s^{t+1}) - ts^{t-1}w(s)w(s^{t+1})w'(s^t) - w(s^t)w(s^{t+1})w'(s)\right] \\
= \;& \frac{1}{s[w(s)w(s^t)]^2}\left[(1+t)s^{t+1} w(s)w(s^t)w'(s^{t+1}) - ts^{t}w(s)w(s^{t+1})w'(s^t) - sw(s^t)w(s^{t+1})w'(s)\right] \\
= \;& \frac{w(s^{t+1})}{sw(s)w(s^t)}\left[\frac{(1+t)s^{t+1}w'(s^{t+1})}{w(s^{t+1})} - \frac{ts^t w'(s^t)}{w(s^t)} - \frac{sw'(s)}{w(s)}\right] \\
= \;& \frac{w(s^{t+1})}{sw(s)w(s^t)}\left[(1+t)\varepsilon(s^{t+1}) - t\varepsilon(s^t) - \varepsilon(s)\right] \\
< \;& 0
\end{aligned}
$$

with $\varepsilon(p) = pw'(p)/w(p)$ defined as the elasticity of the probability weighting function w. According to Segal (1987), p. 148, subproportionality holds iff $\varepsilon(p)$ is increasing. As $s^{t+1} < s^t < s$, $\varepsilon(s^{t+1}) < \varepsilon(s^t) < \varepsilon(s)$ and, hence, the sum of the elasticities in the final line of the derivation is negative. Therefore, increasing uncertainty, i.e. decreasing s, entails a greater departure from stationarity and, consequently, a higher degree of hyperbolicity. ■

A.4 Effect of Concavity

In the course of the experiment we cannot observe discount factors at delay t, $D(t)$, directly, however, but infer $\tilde{D}(t)$ from the intertemporal tradeoffs between payments at different dates, i.e. $pe = \tilde{D}(t)x_t$. According to our assumptions, utility is modeled by a power function $u(x) = x^\eta$, $\eta > 0$, which renders $\tilde{D}(t) = D(t)^{\frac{1}{\eta}}$. It follows that

$$\frac{\tilde{D}(0)/\tilde{D}(1)}{\tilde{D}(t)/\tilde{D}(t+1)} = \left(\frac{D(0)/D(1)}{D(t)/D(t+1)}\right)^{\frac{1}{\eta}} \tag{A.6}$$

and therefore the observed decrease in discount rates resulting from nonlinear probability weighting gets amplified by $\eta < 1$ and, hence, concavity has to be controlled for in the regression model.

B Censored Regression Model

This appendix discusses the way we model the difference in the censored observed discount rates, $\Delta\delta = \delta_2 - \delta_4$, and link it to individual risk preferences.

To relate time discounting to risk preferences, the model assumes the following linear relationship between the discount rate $\delta^*_{t,i}$ of individual $i \in \{1, \ldots, N\}$ over delay $t \in \{\text{two months, four months}\}$ and a vector of regressors c_i, containing a constant, the parameters of risk preferences, η_i, α_i and β_i, as well as some socioeconomic characteristics:

$$\delta^*_{t,i} = c_i \gamma_t + e_{t,i}, \tag{B.1}$$

where γ_t denotes a vector of slope parameters and $e_{t,i}$ stands for a normally distributed error term with mean zero and variance $\frac{1}{2}\sigma^2$. Under the assumption of nonnegative discounting, the choice menu, depicted in Figure II, allows us to directly elicit individual discount rates that lie between 0% and 92.5%. However, if individual i always opts for being paid out at the earlier point in time *(Option A)*, we do not necessarily observe her true discount rate $\delta^*_{t,i}$ as we only know that it amounts to at least 95%. Thus, the elicited discount rates, $\delta_{2,i}$ and/or $\delta_{4,i}$, are censored from above at $b = 0.95$. In the data we observe

$$\delta_{t,i} = \begin{cases} \delta^*_{t,i} & \text{if } \delta^*_{t,i} < b, \\ b & \text{otherwise.} \end{cases} \tag{B.2}$$

This immediately yields the difference in the discount rates over two and four months,

$$\Delta\delta^*_i = c_i \underbrace{(\gamma_2 - \gamma_4)}_{\Delta\gamma} + \underbrace{e_{2,i} - e_{4,i}}_{\Delta e_i}, \tag{B.3}$$

where Δe_i is normally distributed with mean zero and variance σ^2. Consequently, this difference $\Delta\delta^*_i$ is affected by censoring, too, and only observed if both $\delta_{2,i} < b$ and $\delta_{4,i} < b$.

In order to avoid biased estimators for γ_2, γ_4, and σ, the model needs to take the censored nature of the data into account. Therefore, its log likelihood takes on the following form:

$$
\begin{aligned}
\ln L\left(\gamma_2, \gamma_4, \sigma; c, \delta_2, \delta_4\right) &= \sum_{i:\, \delta_{2,i}=b,\, \delta_{4,i}=b} P\left(\delta_{2,i}=b,\, \delta_{4,i}=b \,|\, c, \delta_2, \delta_4\right) \\
&+ \sum_{i:\, \delta_{2,i}<b,\, \delta_{4,i}=b} P\left(\delta_{2,i}<b,\, \delta_{4,i}=b \,|\, c, \delta_2, \delta_4\right) \\
&+ \sum_{i:\, \delta_{2,i}=b,\, \delta_{4,i}<b} P\left(\delta_{2,i}=b,\, \delta_{4,i}<b \,|\, c, \delta_2, \delta_4\right) \\
&+ \sum_{i:\, \delta_{2,i}<b,\, \delta_{4,i}<b} \frac{1}{\sigma}\phi\left(\frac{\Delta\delta_i - c_i\left(\gamma_2-\gamma_4\right)}{\sigma}\right),
\end{aligned}
\tag{B.4}
$$

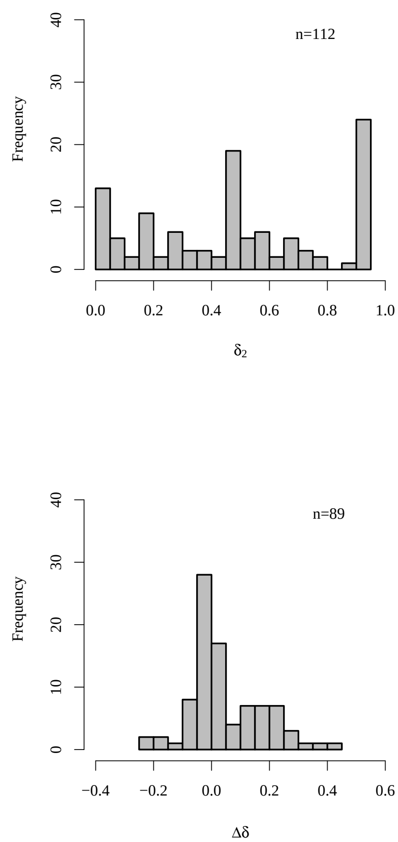
where ϕ represents the standard normal distribution's density and the probabilities P, accounting for the different ways by which an observation may be censored, are given by

$$
\begin{aligned}
P\left(\delta_{2,i}=b,\, \delta_{4,i}=b \,|\, c, \delta_2, \delta_4\right) &= \left[1-\Phi\left(\frac{b-c_i\gamma_2}{\sigma}\right)\right]\left[1-\Phi\left(\frac{b-c_i\gamma_4}{\sigma}\right)\right], \\
P\left(\delta_{2,i}<b,\, \delta_{4,i}=b \,|\, c, \delta_2, \delta_4\right) &= \Phi\left(\frac{b-c_i\gamma_2}{\sigma}\right)\left[1-\Phi\left(\frac{b-c_i\gamma_4}{\sigma}\right)\right], \\
P\left(\delta_{2,i}=b,\, \delta_{4,i}<b \,|\, c, \delta_2, \delta_4\right) &= \left[1-\Phi\left(\frac{b-c_i\gamma_2}{\sigma}\right)\right]\Phi\left(\frac{b-c_i\gamma_4}{\sigma}\right),
\end{aligned}
$$

with Φ denoting the cumulative density function of the standard normal distribution. Numerical maximization of $\ln L\left(\gamma_2, \gamma_4, \sigma; c, \delta_2, \delta_4\right)$ yields the maximum likelihood estimates of $\hat{\gamma}_2$, $\hat{\gamma}_4$, and $\hat{\sigma}$. To obtain the maximum likelihood estimate of $\Delta\hat{\gamma}$ we utilize the invariance property.

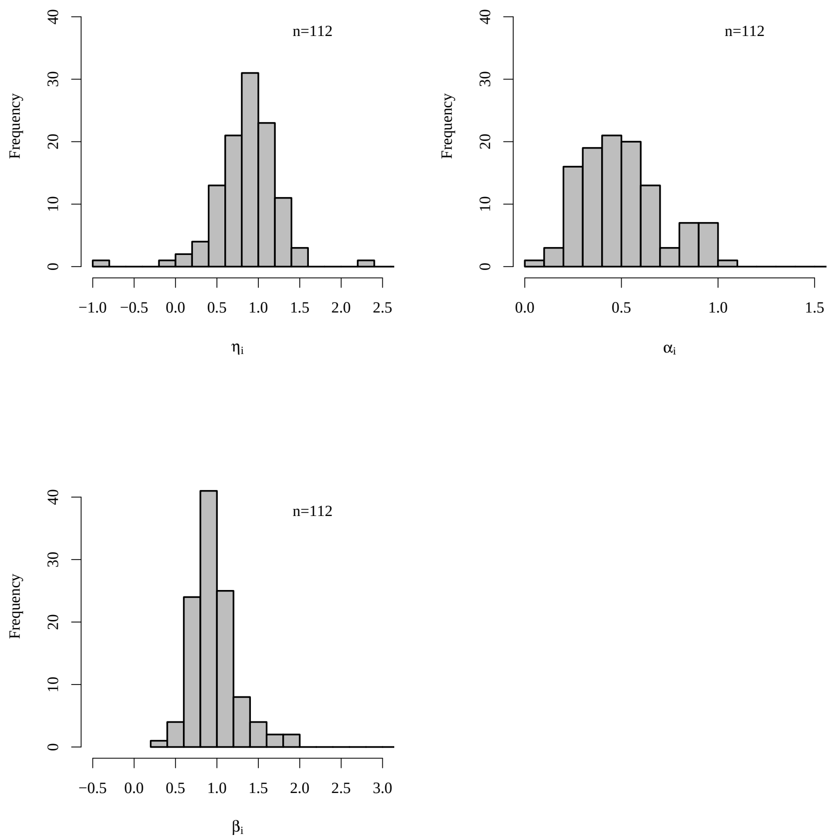
C Observed Discount Rates

Figure C.1: Discount Rates δ_2 and δ_4 and Their Change



D Estimated Risk Parameters

Figure D.1: Distribution of η, α and β



E Controls

Table E.1: Summary Statistics ($n = 112$)

	Type	**Mean**	**Std.Err.**
Female	binary	0.446	0.047
Age	numeric	22.625	0.209
Log-Income	numeric	6.380	0.067
Experience	binary	0.304	0.044
CRT	numeric	2.214	0.082

Table E.2: Number of Observations at the Bounds ($n = 112$)

	δ_2	δ_4
$\geq 95\%$	23	14
0%	2	0

References

Abdellaoui, M. (2000): “Parameter-Free Elicitation of Utility and Probability Weighting Functions,” *Management Science*, 46(11), 1497–1512.

Abdellaoui, M., A. Attema, and H. Bleichrodt (2010): “Intertemporal Tradeoffs for Gains and Losses: An Experimental Measurement of Discounted Utility,” *The Economic Journal*, 120, 845–866.

Abdellaoui, M., E. Diecidue, and A. Öncüler (2010): “Risk Preferences at Different Points in Time: An Experimental Investigation,” *mimeo*.

Ahlbrecht, M., and M. Weber (1997): “An Empirical Study on Intertemporal Decision Making under Risk,” *Management Science*, 43, 813–826.

Ainslie, G. (1991): “Derivation of ”Rational” Economic Behavior from Hyperbolic Discount Curves,” *American Economic Review*, pp. 334–340.

Allais, M. (1953): “L'Extension des Theories de l'Equilibre Economique General et du Rendement Social au Cas du Risque,” *Econometrica*, 21(2), 269–290.

Anderhub, V., W. Güth, U. Gneezy, and D. Sonsino (2001): “On the Interaction of Risk and Time Preferences: An Experimental Study,” *German Economic Review*, 2(3), 239–253.

Andersen, S., G. Harrison, M. Lau, and E. Rutström (2008): “Eliciting Risk and Time Preferences,” *Econometrica*, 76(3), 583–618.

Anderson, L., and S. Stafford (2009): “Individual Decision- Making Experiments with Risk and Intertemporal Choice,” *Journal of Risk and Uncertainty*, 38, 51–72.

Baucells, M., and F. Heukamp (2010): “Common Ratio Using Delay,” *Theory and Decision*, 68, 149–158.

Benjamin, D., A. Sebastian, and J. Shapiro (2006): “Who is 'Behavioral'? Cognitive Ability and Anomalous Preferences,” *mimeo*.

Benzion, U., A. Rapoport, and J. Yagil (1989): "Discount Rates Inferred from Decisions: An Experimental Study," *Management Science*, 35(3), 270–284.

Bruhin, A., H. Fehr-Duda, and T. Epper (2010): "Risk and Rationality: Uncovering Heterogeneity in Probability Distortion," *Econometrica*, 78(4), 1375–1412.

Caplin, A., and J. Leahy (2001): "Psychological Expected Utility Theory and Anticipatory Feelings," *Quarterly Journal of Economics*, 116(1), 55–79.

Chesson, H., and W. Viscusi (2000): "The Heterogeneity of Time-Risk Tradeoffs," *Journal of Behavioral Decision Making*, 13, 251–258.

Chesson, H., and W. Viscusi (2003): "Commonalities in Time and Ambiguity Aversion for Long-Term Risks," *Theory and Decision*, 54, 57–71.

Coble, K., and J. Lusk (2010): "At the Nexus of Risk and Time Preferences: An Experimental Investigation," *Journal of Risk and Uncertainty*, 41, 67–79.

Coller, M., and M. Williams (1999): "Eliciting Individual Discount Rates," *Experimental Economics*, 2, 107–127.

Dohmen, T., A. Falk, D. Huffman, and U. Sunde (2107): "Are Risk Aversion and Impatience Related to Cognitive Ability?," *American Economic Review*, 100(3), 1238–1260.

Eckel, C., C. Johnson, and C. Montmarquette (2004): "Saving Decisions of the Working Poor: Short- and Long-Term Horizons," in *Research in Experimental Economics Vol. 10*, pp. 219–260. Emerald Group Publishing Limited.

Fehr-Duda, H., M. de Gennaro, and R. Schubert (2006): "Gender, Financial Risk, and Probability Weights," *Theory and Decision*, 60, 283–313.

Fischbacher, U. (2007): "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments," *Experimental Economics*, 10, 171–178.

Fisher, I. (1930): "The Theory of Interest as Determined by Impatience to Spend Income and Opportunity to Invest it," *New York: Macmillan.*

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

Frederick, S., G. Loewenstein, and T. O'Donoghue (2002): "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature*, 40, 351–401.

Gonzalez, R., and G. Wu (1999): "On the Shape of the Probability Weighting Function," *Cognitive Psychology*, 38, 129–166.

Hagen, O. (1972): "Towards a Positive Theory of Preference under RIsk," *In: Allais, M., Hagen, O. (Eds.), Expected Utility and The Allais Paraox, Dordrecht, Boston*, pp. 271–302.

Halevy, Y. (2008): "Strotz Meets Allais: Diminishing Impatience and the Certainty Effect," *American Economic Review*, 98(3), 1145–1162.

Harless, D., and C. Camerer (1994): "The Predictive Utility of Generalized Expected Utility Theories," *Econometrica*, pp. 1251–1289.

Hey, J. D., and C. Orme (1994): "Investigating Generalizations of Expected Utility Theory Using Experimental Data," *Econometrica*, 62(6), 1291–1326.

Kahneman, D., and A. Tversky (1979): "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47(2), 263–291.

Karmarkar, U. S. (1979): "Subjectively Weighted Utility and the Allais Paradox," *Organizational Behavior and Human Performance*, 24, 67–72.

Keren, G., and P. Roelofsma (1995): "Immediacy and Certainty in Intertemporal Choice," *Organizational Behavior and Human Decision Processes*, 63(3), 287–297.

Kirby, K. (1997): "Bidding on the Future: Evidence Against Normative Discounting of Delayed Rewards," *Journal of Experimental Psychology*, 126(1), 54–70.

Laibson, D. (1997): "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 112(2), 443–477.

Lattimore, P., J. Baker, and A. D. Witte (1992): "The Influence of Probability on Risky Choice: A Parametric Examination," *Journal of Economic Behavior and Organization*, 17, 377–400.

Leigh, J. (1986): "Accounting for Tastes: Correlates of Risk and Time Preferences," *Journal of Post Keynesian Economics*, 9(1), 17–31.

Myerson, J., and L. Green (1995): "Discounting of Delayed Rewards: Models of Individual Choice," *Journal of the Experimental Analysis of Behavior*, 64, 263–276.

Myerson, J., L. Green, J. Hanson, D. Holt, and S. Estle (2003): "Discounting Delayed and Probabilistic Rewards: Processes and Traits," *Journal of Economic Psychology*, 24, 619–635.

Noussair, C., and P. Wu (2006): "Risk Tolerance in the Present and the Future: An Experimental Study," *Managerial and Decision Economics*, 27, 401–412.

Onay, S., and A. Öncüler (2007): "Intertemporal Choice under Timing Risk: An Experimental Approach," *Journal of Risk and Uncertainty*, 34, 99–121.

Patak, M., and B. Reynolds (2007): "Question-based Assessment of Delay Discounting: Do Respondents Spontaneously Incorporate Uncertainty into their Valuations for Delayed Rewards?," *Addictive Behaviors*, 32, 351–357.

Prelec, D. (1998): "The Probability Weighting Function," *Econometrica*, 66(3), 497–527.

Prelec, D. (2000): "Compound Invariant Weighting Functions in Prospect Theory," *"Choices, Values, and Frames", Kahneman, D. and Tversky, A. (eds.); Cambridge University Press*, pp. 67–92.

Prelec, D. (2004): "Decreasing Impatience: A Criterion for Non-stationary Time Preference and 'Hyperbolic' Discounting," *The Scandinavian Journal of Economics*, 106(3), 511–532.

Prelec, D., and G. Loewenstein (1991): "Decision Making Over Time and Under Uncertainty: A Common Approach," *Management Science*, 37(7), 770–786.

Quiggin, J. (1982): "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization.*

Rachlin, H., A. Raineri, and D. Cross (1991): "Subjective Probability and Delay," *Journal of the Experimental Analysis of Behavior*, 55(2), 233–244.

Segal, U. (1987): "Some Remarks on Quiggin's Anticipated Utility," *Journal of Economic Behavior and Organization*, 8, 145–154.

Starmer, C., and R. Sugden (1989): "Violations of the Independence Axiom in Common Ratio Problems: An Experimental Test of Some Competing Hypotheses," *Annals of Operations Research*, 19, 79–102.

Takahashi, T., K. Ikeda, T. Hasegawa, and A. Greene (2007): "A hyperbolic decay of subjective probability of obtaining delayed rewards," *Behavioral and Brain Functions*, 52(3).

Thaler, R. (1981): "Some Empirical Evidence on Dynamic Inconsistency," *Economics Letters*, 8, 201–207.

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

Wakker, P. (1994): "Separating Marginal Utility and Probabilistic Risk Aversion," *Theory and Decision*, 36, 1–44.

——— (2008): "Explaining the Characteristics of the Power (CRRA) Utility Family," *Health Economics*, 17, 1329–1344.

Walther, H. (2003): "Normal-Randomness Expected Utility, Time Preference and Emotional Distortions," *Journal of Economic Behavior and Organization*, 52, 253–266.

——— (2010): "Anomalies in Intertemporal Choice, Time-Dependent Uncertainty and Expected Utility - A Common Approach," *Journal of Economic Psychology*, 31, 114–130.

Weber, B., and G. Chapman (2005): "The combined effects of risk and time on choice: Does uncertainty eliminate the immediacy effect? Does delay eliminate the certainty effect?," *Organizational Behavior and Human Decision Processes*, 96, 104–118.

Wu, G. (1999): "Anxiety and Decision Making with Delayed Resolution of Uncertainty," *Theory and Decision*, 46, 159–198.