RISK IN TIME: The Intertwined Nature of Risk Taking and Time Discounting*

Thomas F. Epper[1] and Helga Fehr-Duda[2]

[1]CNRS, IESEG School of Management, Univ. Lille, UMR 9221 - LEM - Lille Economie Management, F-59000 Lille, France.
[2]University of Zurich, Department of Banking and Finance, Plattenstrasse 32, CH-8032 Zurich, Switzerland.

November 5, 2021

Abstract

Standard economic models view risk taking and time discounting as two independent dimensions of decision making. However, mounting experimental evidence demonstrates striking parallels in patterns of risk taking and time discounting behavior and systematic interaction effects, which suggests that there may be common underlying forces driving these interactions. Here we show that the inherent uncertainty associated with future prospects together with individuals' proneness to probability weighting generates a unifying framework for explaining a large number of puzzling behavioral regularities: delay-dependent risk tolerance, aversion to sequential resolution of uncertainty, preferences for the timing of the resolution of uncertainty, the differential discounting of risky and certain outcomes, hyperbolic discounting, subadditive discounting, and the order dependence of prospect valuation. Furthermore, all these phenomena can be predicted simultaneously with the same set of preference parameters.

JEL D01, D81, D91

*We are thankful for many valuable comments and suggestions made by Enrico Diecidue, Sebastian Ebert, Kfir Eliaz, Ernst Fehr, Nicola Gennaioli, Michihiro Kandori, Alexander Koch, David Laibson, Antoine Nebout, Andrew Oswald, Ran Spiegler, Bob Sugden, and by seminar and conference participants at the Aarhus University, Columbia University, Copenhagen University, ETH Zurich, Frankfurt School of Finance and Management, Harvard University, Ludwig-Maximilians-University Munich, New York University, University of Basel, University of Konstanz, University of Lausanne, University of Lucerne, University of Mannheim, University of Nottingham, University of Warwick, University of Zurich, as well as the FUR conference, the ESA conference, the Summer School of the Econometric Society, the European Meeting of the Econometric Society, and the ESA conference. This work was supported by a grant from the Swiss National Science Foundation (#100018M_182128).

1 Introduction

Whatever the nature of our decisions, hardly ever can we be sure about their outcomes. In particular, the consequences of the most important decisions in our lives, such as what line of business to enter or whom to get married to, do not materialize immediately but usually take time to unfold. In other words, these important decisions involve both risk and delay. Driven by the evidence challenging expected utility theory and discounted utility theory, the past half century has seen a surge of new models of decision making for the domains of risk taking and time discounting (Starmer, 2000; Frederick, Loewenstein, and O'Donoghue, 2002; Wakker, 2010; Ericson and Laibson, 2019). A considerable body of experimental evidence suggests, however, that risk taking and time discounting are linked and interact with each other in important ways summarized in Table 1 below.

First, risk aversion has been shown to be lower for risks materializing in the more remote future than for risks materializing in the more imminent future (e.g. Shelley (1994)).[1] Lower risk aversion for remote risks may be one reason why the mobilization of public support for policies combating global warming is so difficult. Thus, economic models of climate policy may benefit from recognizing that risk aversion decreases with time delay. Asset markets constitute another area where delay-dependent risk aversion may play an important role in understanding the downward sloping structure of risk premia, i.e. the fact that risk premia decline with maturity (van Binsbergen, Brandt, and Koijen, 2012).

A second fact is based on a considerable body of evidence that impatience tends to decrease when outcomes are shifted into the more remote future - a finding on which the large literature on hyperbolic discounting is based (Loewenstein and Thaler, 1989; Laibson, 1997).

Third, the evidence indicates that many people seem to have a preference with respect to the way uncertainty resolves, i.e. sequentially or in one shot. Often, sequential evaluation of prospects renders decision makers less risk tolerant (Abdellaoui, Klibanoff, and Placido, 2015). In the domain of financial decisions, this phenomenon may underlie the large equity premia observed around the globe.

Fourth, regarding time discounting, a similar phenomenon has been observed: discount rates compounded over subperiods tend to be higher than the discount rate applied to the total period. This incidence of process dependence, labeled *subadditive discounting*, has been put forward as an alternative explanation to hyperbolic discounting to account for the observed patterns in discounting behavior (Read, 2001; Dohmen, Falk, Huffman, and Sunde, 2017).

Fifth, many people also exhibit a preference regarding the timing of the resolution of uncertainty. Usually, there is a substantial share of participants who favor delayed resolution of uncertainty in situations when money is at stake even though it should be beneficial to know the outcome of one's financial decisions as early as possible. This finding triggered a large theoretical literature following the seminal work of Kreps and Porteus (1978).

[1] In Section 5 we provide a detailed list of references regarding the empirical evidence and a discussion of extant theories that address various subsets of these facts.

A sixth regularity indicates that the presence of risk influences time discounting in an unexpected way: Sure outcomes appear to be discounted more heavily than uncertain ones, discussed in the literature under the heading *diminishing immediacy* (Keren and Roelofsma, 1995; Weber and Chapman, 2005).

Finally, people's evaluations of future risky payoffs depend on the order by which they are devalued for risk and for delay, which should not make any difference according to the standard view (Öncüler and Onay, 2009). These regularities suggest that theories that are restricted to either domain cannot easily account for the intertwined nature of risk taking and time discounting.

Table 1: Seven Stylized Facts on Risk Taking and Time Discounting

Dimension	Fact	Observed risk tolerance	Fact	Observed patience
Delay dependence	#1	increases with delay	#2	increases with delay
Process dependence	#3	higher for one-shot than for sequential valuation	#4	higher for one-shot than for sequential valuation
Timing dependence	#5	intrinsic preference for timing of uncertainty resolution		–
Risk dependence		–	#6	higher for risky payoffs than for certain ones
Order dependence	#7	depends on order of delay and risk discounting		–

The table describes seven regularities in experimental findings on risk taking and time discounting behavior with respect to delay, process, timing, risk, and order effects.

The main purpose of our paper is to provide a unifying account of all these phenomena by integrating risk taking and time discounting into one theoretical approach. Thus, our goal is to develop a formal model that is capable of explaining all the regularities on the basis of a parsimonious set of assumptions. Our approach rests on two key assumptions: First, there is risk attached to any future prospect because only immediate consequences can be totally certain. We believe that this is a plausible assumption because it is impossible to foresee all future contingencies. Accordingly, Prelec and Loewenstein (1991) claim that *"anything that is delayed is almost by definition uncertain"* (page 784). In particular, it is always possible that an event may occur that prevents the realization of a future outcome, i.e. something may go wrong before payoffs actually materialize. An unforeseen contingency may arise, such as missing one's transatlantic flight because the taxi driver was late, or realizing that one has forgotten one's passport at home. Presumably, almost everyone can readily recall such an incident. We capture the notion that something may go wrong by introducing a survival probability $0 < s < 1$ that applies also to allegedly certain future outcomes.

Second, if future prospects are perceived as inherently risky, people's risk tolerance must play a role in their valuations of future prospects. Therefore, the characteristics of (atemporal) risk preferences are crucial not only for evaluating delayed risky prospects but also for delayed (allegedly) certain ones. There is abundant evidence from the field and the laboratory that risk taking behavior depends nonlinearly on the objective probabilities (Prelec, 1998; Fehr-Duda and Epper, 2012; Barberis, 2013; O'Donoghue and Somerville, 2018). For this reason, models involving probability weighting, such as rank-dependent utility theory (RDU) (Quiggin, 1982) and cumulative prospect theory (Tversky and Kahneman, 1992) have been strong contenders of expected utility theory (EUT) (Wakker, 2010).

Our approach relies on a key characteristic of probability weighting, proneness to Allais-type *common-ratio violations*, that is one of the most widely replicated experimental regularities, found in human and animal behavior: Probabilistically mixing two lotteries with an inferior lottery frequently leads to preference reversals (Kahneman and Tversky, 1979; Gonzalez and Wu, 1999). This feature of probability weighting is called *subproportionality* and was characterized axiomatically by Prelec (1998).

Our contribution to the literature is fourfold. First and foremost, we show for general m-outcome prospects that subproportional probability weighting under rank-dependent utility (RDU) together with the assumption that (even allegedly certain) future outcomes are inherently risky provides an integrative account of all the above mentioned experimental regularities. We rely on a well-established model of risk preferences with axiomatic foundations (e.g. Quiggin (1982); Yaari (1987); Segal (1990); Wakker (1994); Chateauneuf and Wakker (1999); Abdellaoui (2002)) that we combine with the plausible assumption that something may go wrong in the future.

Second, we demonstrate that all seven stylized facts can be quantitatively predicted with the same set of preference parameters with survival probability lying in a narrow and plausible range. For this exercise, we quantitatively predict such diverse magnitudes as present certainty equivalents, discount rates, discount fractions and probability weights.

Third, addressing our theoretical contribution, we build on Halevy (2008), followed by Saito (2011) and Chakraborty, Halevy, and Saito (2020), who demonstrate that there is a tight two-way relationship between hyperbolic discounting and the property of subproportionality.[2] Our paper is inspired by their insight and extends it by showing that, beyond the two-way relationship between hyperbolic discounting and subproportionality, many additional important regularities in observed risk taking and time discounting follow from subproportional probability weighting and the assumption of an inherently uncertain future. In particular, we derive novel predictions regarding (i) the delay dependence of probability weights and (ii) the intrinsic preference for late resolution of uncertainty. Furthermore, we take advantage of Segal's work (Segal, 1987a,b, 1990) on two-stage prospect evaluation by explicitly integrating the dimension of time delay and show that (i) subproportional risk preferences imply subadditive discounting and (ii) that, under

[2]Relatedly, Chakraborty (2021) explores risk-time separability violations by adopting a weaker version of the stationarity axiom to simple risky prospects $(x, p; 0, 1-p)$.

certain circumstances, an aversion to sequential resolution of uncertainty arises and remains intact under inherent future uncertainty.

Fourth, we derive new predictions which can be put to the test by future research. We show for example, that the decrease in risk tolerance, induced by the sequential resolution of uncertainty of (atemporal) prospect risks, carries over, under certain conditions, to the sequential resolution of uncertainty for delayed future prospects. This prediction is important as many societal risks (e.g. climate risks) and asset market risks resolve sequentially over time. However, so far this prediction has not been tested.

While there is a large empirical and theoretical literature on the domain of risk taking and an equally large one on time discounting, there are, in comparison, relatively few papers dealing with an integrated view of risk and time. However, the subject has recently gained traction. For example, motivated by the similarities of anomalies in risk taking and time discounting behaviors, Prelec and Loewenstein (1991) develop psychological properties of multi-attribute prospect valuation that may be common in both decision domains. Thus, common ratio violations and decreasing impatience may be driven by the same psychological principles. The authors do not address how features of risk preferences and time preferences interact with each other, however.

Similarly, Quiggin and Horowitz (1995) analyze parallels between the theories of choice under risk and choice over time and show the usefulness of RDU for understanding the analogy between risk aversion and impatience. Baucells and Heukamp (2012) restrict their analysis to the case of simple prospects $(x, p; 0, 1-p)$ that pay x with probability p at time t and zero otherwise, and link risk taking and time discounting by making assumptions on how people trade off delays in future outcomes against reductions in the probability with which these outcomes occur.

Leland and Schneider (2017) propose a different theory that can account for many anomalies in risk taking and time discounting behavior. Their approach extends the concept of salience from outcome differences to differences in probabilities and differences in delays. This enables the authors to explain a large set of interesting facts in risk taking, time discounting and consumer behavior. However, they explicitly mention on page 20 that their theory "does not account for interaction effects between risk and time" that are exactly the object of our paper. On the other hand, our paper does not explain facts such as labeling effects, framing effects or peanut effects which are the explicanda of Leland and Schneider (2017)'s paper.

DeJarnette, Dillenberger, Gottlieb, and Ortoleva (2020) study a setting that is complementary to ours: Their *time lotteries* have fixed prices, but random payment dates. In contrast, we explicitly abstract from uncertainty with regard to the timing of outcomes. However, extending our approach to their time-lotteries setting may be an interesting direction of future research.

The remainder of the paper is organized as follows: Section 2 discusses the key assumptions of our model. Its implications for explaining the seven stylized facts are developed in Section 3. Section 4 is devoted to a quantitative assessment of our model predictions. Section 5 presents the experimental findings on the seven stylized facts in more detail and discusses other theoretical approaches that address some of these empirical regularities. Finally, Section 6 concludes.

Propositions including proofs and complementary materials are available in the appendix.

2 The Model

In the following, we will first present the general setup of our approach. Second, we justify our assumptions on the characteristics of the probability weighting function. Finally, we explain how we integrate that "something may go wrong" into the model.

2.1 Risk Preferences

In this paper, we rely on rank dependent utility theory (RDU), a generalization of expected utility theory (EUT), that allows for nonlinear weighting of the probabilities. This additional feature has proven to be an exceptionally powerful component. First, overweighting of small probabilities may counteract risk aversion embodied in the utility function. Thus, RDU can handle the empirically observed probability dependence of risk tolerance (for an early example see Preston and Baratta (1948)). Probability weighting captures the intuition that *"attention given to an outcome depends not only on the probability of the outcome but also on the favorability of the outcome in comparison to the other possible outcomes"* (Diecidue and Wakker (2001), page 284). Typically, decision makers focus on the worst and best possible outcomes and give much less attention to intermediate outcomes that will generally be underweighted even when they have the same objective probabilities as the extreme outcomes (Quiggin, 1982).

Second, RDU displays first-order risk aversion, i.e. preferences between prospects whose consequences are sufficiently close to one another do not necessarily tend to risk neutrality. Thus, the experimental evidence of pronounced risk aversion over small stakes favors RDU over many other approaches that accommodate non-EUT behavior but display only second-order risk aversion (Sugden, 2004). Additionally, RDU was axiomatized by several authors and respects completeness, transitivity, continuity, and first-order stochastic dominance, qualities that many economists are hesitant to dispense with.[3] Of course, RDU is not perfect, in the sense that it is always the best model to rationalize data. For example, it cannot handle choice-set dependent preferences, the strength of regret theory (Loomes and Sugden, 1987) and salience theory (Bordalo, Gennaioli, and Shleifer, 2012). In our view, choice sets can be tightly controlled in experimental settings, but are usually quite elusive in other contexts.

According to RDU, a decision maker's atemporal risk preferences over prospects that are played out and paid out with negligible time delay can be represented by a rank-dependent functional. Consider a prospect $P = (x_1, p_1; ...; x_m, p_m)$ over (terminal) monetary outcomes $x_1 > x_2 > ... > x_m$ with $x_i \in X \subset \mathbb{R}, p_i \in [0, 1]$ and $\Sigma p_i = 1$. The function u denotes the utility

[3] RDU can handle other behavioral phenomena as well, such as correlation aversion: When decision makers evaluate risky consumption streams they often have a preference for diversifying consumption across time, i.e. they prefer some good and some bad to all or nothing (Kihlstrom and Mirman, 1974; Richard, 1975; Epstein and Tanny, 1980; Bommier, 2007; Denuit, Eeckhoudt, and Rey, 2010). Epper and Fehr-Duda (2015) show that RDU implies correlation aversion if the decision maker is sufficiently pessimistic, which is usually borne out by the data.

of monetary amounts x, and w denotes the subjective probability weight attached to p_1, the probability of the best outcome x_1. As usual, both u and w are assumed to be monotonically increasing, w to be twice differentiable on $(0,1)$ and to satisfy $w(0) = 0$ and $w(1) = 1$. Decision weights π_i are defined as[4]

$$\pi_i = \begin{cases} w(p_1) & \text{for } i = 1 \\ w\left(\sum_{k=1}^{i} p_k\right) - w\left(\sum_{k=1}^{i-1} p_k\right) & \text{for } 1 < i \leq m \end{cases} . \tag{1}$$

Thus, the decision weight of x_i is the probability weight attached to the probability of obtaining something at least as good as x_i minus the probability weight attached to the probability of obtaining something strictly better than x_i. Consequently, decision weights sum to one. Finally, the prospect's value is represented by

$$V(P) = \sum_{i=1}^{m} u(x_i)\pi_i \,. \tag{2}$$

To keep the logic of our approach as transparent as possible we present the following steps for $m = 2$ and delegate the general case to Appendix A.1. For $m = 2$, the prospect reduces to $P = (x_1, p; x_2, 1 - p)$ and Equation 2 reads as

$$\begin{aligned} V(P) &= u(x_1)w(p) + u(x_2)\Big(1 - w(p)\Big) \\ &= \Big(u(x_1) - u(x_2)\Big)w(p) + u(x_2)\,. \end{aligned} \tag{3}$$

This representation of V clarifies that x_2 is effectively a sure thing whereas obtaining something better than x_2 is risky.

If the prospect is not played out and paid out in the present, but at some future time $t > 0$, prospect value is affected by time discounting as well. We follow the standard approach and model people's willingness to postpone gratification by a constant rate of time preference $\eta \geq 0$, yielding a discount weight of $\rho(t) = \exp(-\eta t)$.[5] A prospect to be played out and paid out at $t > 0$ is discounted for time in the following standard way:

$$V_0(P) = V(P)\rho(t)\,. \tag{4}$$

Abundant empirical evidence has demonstrated that risk taking behavior depends nonlinearly on the probabilities (Starmer, 2000; Fehr-Duda and Epper, 2012). However, in order to explain the observed interaction effects, we need to put more structure on the type of nonlinearity.

[4]Alternatively, decision weights π_i can be expressed in terms of the cumulative distribution function F of the outcomes x_i: $\pi_i = w(1 - F(x_{i+1})) - w(1 - F(x_i))$ for $1 \leq i \leq m$, where $F(x_{m+1}) := 0$.

[5]This assumption is not crucial for our results - neither a zero rate of time preference, i.e. $\rho = 1$, nor genuinely hyperbolic time preferences affect our conclusions.

2.2 Probability Weighting

Inspired by one of Allais (1953)'s famous examples, Kahneman and Tversky (1979) presented subjects with the following experiment: Subjects had to choose between 3000 Israeli pounds for sure and 4000 Israeli pounds materializing with a probability of 80%. Most people chose the sure option of 3000 pounds. When confronted with the choice between a 25%-chance of receiving 3000 pounds and a 20%-chance of receiving 4000 pounds, the majority opted for the 4000-pound alternative. Scaling down the probabilities of 100% and 80% by a common factor, in this example 1/4, induced many people to reverse their preferences. Obviously, such Allais-type behavior is inconsistent with EUT.

An intuitive explanation for common-ratio violations is fear of disappointment:[6] Losing a gamble over very likely 4000 pounds is anticipated to be much more disappointing than losing a gamble over 4000 pounds with only a small chance of materializing. On the other hand, winning 4000 pounds in the unlikely situation of a 20%-chance may trigger feelings of elation. Thus, when people are prone to disappointment and/or elation, their behavior appears to depend nonlinearly on the probabilities (Bell, 1985; Loomes and Sugden, 1986; Gul, 1991; Wu, 1999; Rottenstreich and Hsee, 2001; Walther, 2003).[7]

In RDU, common-ratio violations are mapped by subproportionality of probability weights. Formally, *subproportionality* holds if $1 \geq p > q > 0$ and $0 < \lambda < 1$ imply the inequality

$$\frac{w(p)}{w(q)} > \frac{w(\lambda p)}{w(\lambda q)} \tag{5}$$

(Prelec, 1998). Kahneman and Tversky (1979) note that this property imposes considerable constraints on the shape of w: it holds if and only if $\ln w$ *is a convex function of* $\ln p$. In other words, $\left(\frac{d\ln w}{d\ln p}\right)' > 0$, or the elasticity of w, $\varepsilon_w(p) = \frac{d\ln w}{d\ln p}$, is increasing in p.[8]

Subproportionality implies the *certainty effect*, which constitutes the special case of $p = 1$. Therefore,

$$w(\lambda q) > w(\lambda)w(q) \tag{6}$$

is satisfied for any λ, q such that $0 < \lambda, q < 1$.[9] This feature of subproportional probability

[6]It is interesting to note that Gul (1991)'s theory of disappointment aversion is observationally equivalent to RDU with a specific convex subproproportional probability weighting function if prospects have only two outcomes, see also Appendix B.4.

[7]Perceptual and procedural factors are potential drivers of probability distortions as well. The fathers of prospect theory, Kahneman and Tversky, attributed probability dependence to the psychophysics of perception according to which the sensitivity toward changes in probabilities diminishes with the distance to the natural reference points of certainty and impossibility (Tversky and Kahneman, 1992). Several other contributions focused on procedural aspects of choice (Rubinstein, 1988; Loomes, 2010). In these models, a prospect's value depends not only on the prospect's own characteristics but also on other prospects in the choice set. A recent contribution in this category is Bordalo, Gennaioli, and Shleifer (2012) who posit that probabilities are distorted in favor of payoffs that are perceived as particularly salient. Whatever its cause, we interpret probability weighting in rank-dependent models as a kind of reduced form generated by some psychological mechanism.

[8]The equivalence of subproportionality and increasing elasticity is shown in Appendix B.1.

[9]The numerical example above is a manifestation of the certainty effect, as the smaller outcome in the first decision situation, 3000 pounds, materializes with certainty.

weights has a crucial implication: It produces an aversion to compounding of probability weights (Segal, 1987a,b, 1990). We will use this insight when we discuss aspects of uncertainty resolution. There is also ample evidence for general common-ratio violations that do not involve a sure outcome (e.g. Loomes and Sugden (1987); Nebout and Dubois (2014)). However, not everyone is prone to common ratio violations. Usually, there is great heterogeneity in people's behaviors, with the aggregate often exhibiting the common ratio effect.

On average, estimated probability weighting curves overweight small probabilities and underweight large probabilities of the best outcome, which is also a common characteristic of individual estimates (Gonzalez and Wu, 1999; Bruhin, Fehr-Duda, and Epper, 2010). This feature is labeled regressiveness, i.e. the probability weighting curve cuts the diagonal from above. In the context of rank-dependent models, regressiveness of the probability weighting function generates overweighting of a prospect's tail outcomes and underweighting of its intermediate outcomes, which nicely captures the notion that more extreme outcomes within a given prospect are more salient. While the driver of our results is subproportionality, regressiveness is an independent additional characteristic that captures key features of observed behavior.[10] Aside from regressive shapes, convex weighting curves which globally underweight probabilities comprise another common category of individuals' probability weighting functions (see e.g. van de Kuilen and Wakker (2011)), which may be subproportional as well.

When inspecting the graph of $w(p)$, one cannot detect subproportionality with the naked eye. In fact, many different shapes of $w(p)$ display subproportionality, at least over some range of probabilities. Figure 1 depicts three examples of subproportional probability weighting functions that display starkly different shapes, a regressive function and two non-regressive ones.

Aside from the examples in Figure 1, many other functional specifications have been proposed in the literature (see Appendix B.4). Perhaps the most widely used representative of a globally subproportional function is Prelec (1998)'s flexible two-parameter specification of the compound invariant class, designed to map common-ratio violations. Throughout the paper, we will use this "standard" functional specification to illustrate our results,[11] defined as

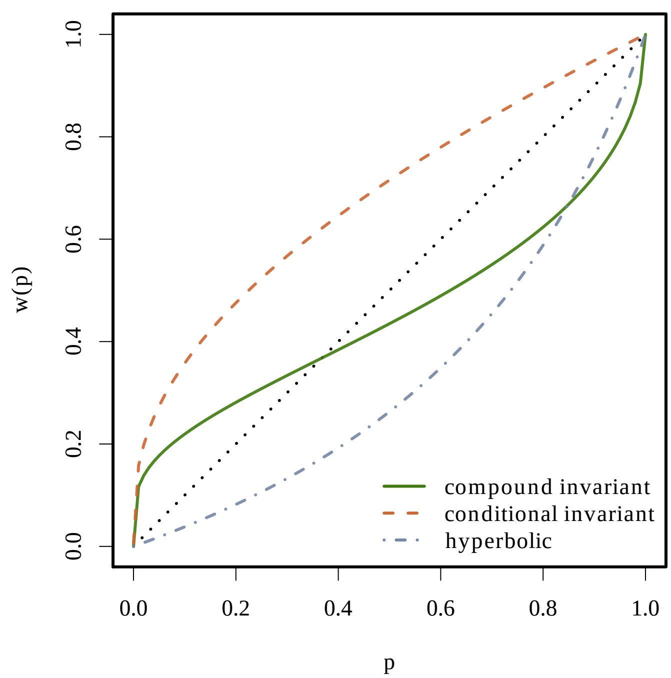
$$w(p) = \exp\left(-\beta\Big(-\ln(p)\Big)^{\alpha}\right), \tag{7}$$

where $0 < \alpha < 1$ indicates the degree of subproportionality and $0 < \beta$ governs the range of convexity. The smaller is α, the more pronounced is subproportionality and the greater is the departure from linearity. The greater is β, the greater, ceteris paribus, is the range of probabilities for which the curve is convex, i.e. underweighting p.

[10] Empirical estimates are often based on inverse-S shaped functional forms, that are concave over small probabilities and convex over large probabilities.

[11] Aydogan, Bleichrodt, and Gao (2016) provide experimental support for the compound invariant specification at the level of preference conditions.

Figure 1: Examples of Subproportional Probability Weighting Functions



Compound invariant: $w(p) = \exp\Big(-(-\ln(p))^{0.5}\Big)$ (Prelec, 1998). This function is globally subproportional and regressive. **Conditional invariant:** $w(p) = \exp\Big(-5\left(1-p^{0.1}\right)\Big)$ (Prelec, 1998). This function is globally subproportional and concave, i.e. non-regressive. **Hyperbolic:** $w(p) = \frac{p}{p+2.8(1-p)}$ (Rachlin, Raineri, and Cross, 1991). This function is globally subproportional and convex, i.e. non-regressive.

2.3 Future Uncertainty

The final building block of our model concerns the integration of "something may go wrong" due to events unrelated to the prospect under consideration. This (uninsurable) risk inherent in the future, *survival risk* for short, turns allegedly guaranteed payoffs into risky ones and introduces an additional layer of risk over and above the objective probability distributions of risky payoffs (henceforth referred to as *prospect risk*). Consequently, there are two distinct types of risk, *prospect risk* which may resolve at any time between the present and the payment date, and *survival risk* which resolves fully only at the payment date. Thus, the subjective perception of future uncertainty changes the nature of the prospect. Formally, let $0 < s < 1$ denote the constant per-period probability of prospect survival, i.e. the probability that the decision maker will actually obtain the promised rewards by the end of the period. Essentially, there are two ways of accounting for this subjective probability s. First, for a delay t, the probability s^t is transformed according to the decision maker's probability weighting function, and the resulting $w(s^t)$ affects the prospect as a whole, i.e. all outcomes equally. In this case, prospect value amounts to

$$V_0(P) = V(P)w(s^t)\rho(t)\,. \tag{8}$$

Such an approach only affects measured discount rates but cannot handle the observed interaction effects. Thus, we work with the second solution, namely that s impacts the perceived probability distribution of the prospect. Then the probability that the allegedly guaranteed payment x_m materializes at the end of period t is perceived to be s^t, and the probabilities of obtaining something better than x_m are scaled down by s^t. Therefore, the objective m-outcome prospect is subjectively perceived as an *(m+1)*-outcome prospect. Focusing on $m = 2$ again, $\tilde{P} = (P, s^t; \underline{x}, 1-s^t) = \Big(x_1, ps^t; x_2, (1-p)s^t; \underline{x}, 1-s^t\Big)$, where $\underline{x} < x_m$ captures that "something may go wrong".

Setting $u(\underline{x}) = 0$, the subjective present value of the prospect amounts to

$$\begin{aligned} V_0(\tilde{P}) = \ & \Big(\Big(u(x_1) - u(x_2)\Big)w(ps^t) + u(x_2)w(s^t)\Big)\rho(t) \\ = \ & \Big(\Big(u(x_1) - u(x_2)\Big)\tfrac{w(ps^t)}{w(s^t)} + u(x_2)\Big)w(s^t)\rho(t)\,. \end{aligned} \tag{9}$$

From the point of view of an outsider, the subjective probability distribution of prospect P is not observable. Consequently, she infers probability weights $\tilde{w}$ and discount weights $\tilde{\rho}$ from observed behavior on the presumption that the decision maker evaluates the objectively given prospect P, and estimates preference parameters according to RDU in the standard way:[12]

$$V_0(\tilde{P}) = \Big(\Big(u(x_1) - u(x_2)\Big)\tilde{w}(p) + u(x_2)\Big)\tilde{\rho}(t)\,, \tag{10}$$

[12]Note that it takes at least two non-zero outcomes to separate risk taking and time discounting.

interpreting $\tilde{w}$ as true probability weights and $\tilde{\rho}$ as true discount weights, while in fact the weights are distorted by survival risk. By comparing Equation 9 with Equation 10 we can see that the relationships between true and observed weights are given by

$$\tilde{w}(p) = \frac{w(ps^t)}{w(s^t)} \tag{11}$$

and

$$\tilde{\rho}(t) = w(s^t)\rho(t)\,. \tag{12}$$

These equations define the central relationships between observed and true underlying probability and discount weights. Concerning observed discount weights, Equation 12 resembles Halevy (2008)'s representation for consumption streams in discrete time, while the focus here is on temporal prospects in continuous time. This equation is the basis for explaining hyperbolic and subadditive discounting, while all the other facts rest on Equation 11, a novel prediction of our model.

Since $\tilde{w}(p) \neq w(p)$ and $\tilde{\rho}(t) \neq \rho(t)$, the presence of survival risk drives a wedge between true underlying preferences and observed risk taking and discounting behavior. Thus, future risk conjointly with proneness to Allais-type behavior provides the mechanism by which behavior under risk and behavior over time are intertwined. A summary of the model variables is provided in Table 2.

Table 2: Model Variables

	Variable	Description	Characteristics
Prospects	x	monetary payoff	$x \geq 0$
	p	probability of x	$0 \leq p \leq 1$
	s	probability of prospect survival	$0 < s < 1$
	$1-s$	survival risk	"
	t	length of time delay	$t \geq 0$
Preferences	$u(x)$	utility function	$u(0) = 0, u' > 0$
	$w(p)$	true probability weight	$w(0) = 0, w(1) = 1, w' > 0$
	η	rate of pure time preference	$\eta \geq 0$, constant
	$\rho(t)$	discount weight	$\rho(t) = \exp(-\eta t)$
Behavior	$\tilde{w}(p)$	observed probability weight	$\tilde{w}(p) = \frac{w(ps^t)}{w(s^t)}$
	$\tilde{\rho}(t)$	observed discount weight	$\tilde{\rho}(t) = w(s^t)\rho(t)$
	$\tilde{\eta}(t)$	observed discount rate	$\tilde{\eta}(t) = -\frac{\tilde{\rho}'(t)}{\tilde{\rho}(t)}$

3 Unifying the Experimental Evidence

In the following, we discuss the implications of our approach for the experimental phenomena listed in Table 1 and demonstrate that, qualitatively, all the Facts #1 through #7 can be explained within our framework. A quantitative assessment of the model predictions is presented in Section 4. While we retain some of the fundamental calculations in the main text, propositions and their proofs are presented in Appendix A.

3.1 Fact #1: Delay Dependence of Risk Taking Behavior

The first fact in our list considers the observation that risk tolerance for delayed prospects seems to be higher than risk tolerance for present ones. Concerning delayed risky prospects, we examine the case when prospect risk and survival risk are resolved simultaneously in one shot at time t. We have seen from Equation 11 that observed probability weights $\tilde{w}(p)$ deviate systematically from the underlying atemporal ones $w(p)$,

$$\tilde{w}(p) = \frac{w(ps^t)}{w(s^t)}\,.$$

As $w(s^t) < 1$, the denominator boosts observed probability weights, whereas the additional s^t in the argument of w in the numerator distorts them. Due to subproportionality

$$\tilde{w}(p) = \frac{w(ps^t)}{w(s^t)} > \frac{w(p)}{w(1)} = w(p)\,, \tag{13}$$

implying that $\tilde{w}$ is more elevated than w, i.e. that $\tilde{w}$ lies above w, which constitutes one of the central implications of our model. Since the probability weighting function maps the decision weight of the best possible outcome, an increase in the elevation of the probability weighting curve gets directly translated into higher revealed risk tolerance. For $m = 2$ and a given observed discount weight $\tilde{\rho}(t) = w(s^t)\rho(t)$,

$$\begin{aligned} V_0(\tilde{P}) = &\quad \Big(\Big(u(x_1) - u(x_2)\Big)\tilde{w}(p) + u(x_2)\Big)w(s^t)\rho(t) > \\ V_0(P) = &\quad \Big(\Big(u(x_1) - u(x_2)\Big)w(p) + u(x_2)\Big)w(s^t)\rho(t)\,. \end{aligned} \tag{14}$$

Thus, the presence of survival risk makes people appear more risk tolerant for delayed prospects than for present ones. Intuitively, the event of something going wrong takes on the role of the perceived sure outcome, which makes x_2 an intermediate one and, thus, less salient to the decision maker. In addition, this risk-tolerance increasing effect is particularly strong for small probabilities, i.e. positively skewed prospects are subject to more pronounced increases in risk tolerance, as $\frac{\tilde{w}(p)}{w(p)}$ declines in p (see Proposition 1 in Appendix A.2). Such a prediction would not be possible if the utility function were the carrier of delay dependence, as for instance in

Eisenbach and Schmalz (2016).

The delay dependence of observed probability weights $\tilde{w}$ is illustrated in Figure 2. The top row of the figure characterizes preferences in the atemporal case. Panel 1a shows a typical specimen of a regressive subproportional probability weighting function w for $t = 0$, underweighting large probabilities and overweighting small probabilities of the best outcome. For illustrative purposes, Panel 1b on the right side depicts the corresponding decision weights π_i for a prospect involving 21 equiprobable outcome levels, with outcome rank 1 denoting the best outcome and outcome rank 21 denoting the worst one. Their objective probabilities are represented on the horizontal gray line. As one can see, a regressive w generates strong overweighting of the extreme outcomes and underweighting of the intermediate ones relative to the objective probability distribution.

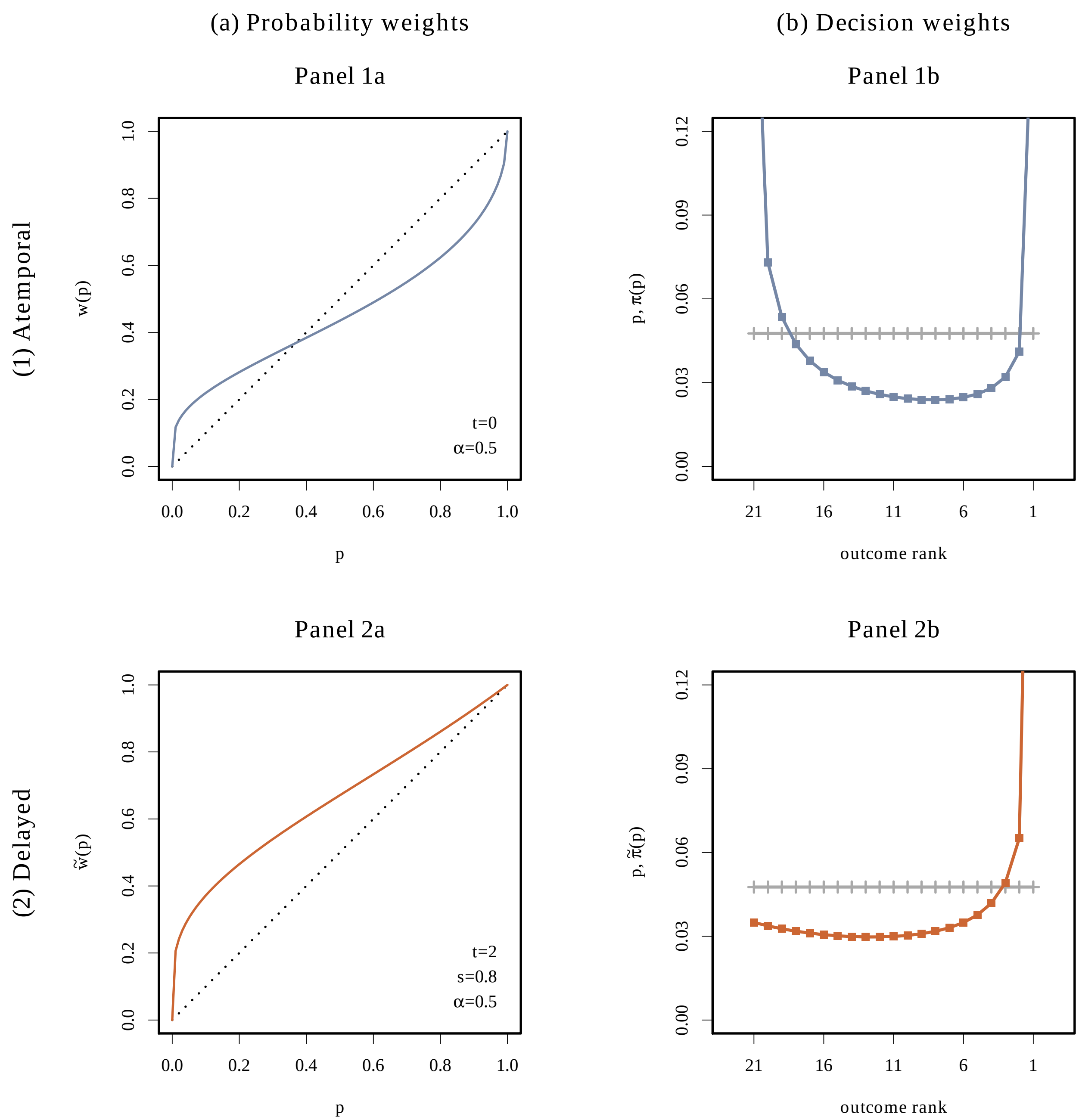
The bottom row of Figure 2 demonstrates the predictions for the case when prospects are played out and paid out simultaneously in the future, the focus of this section. Future uncertainty is captured by the parameter $s = 0.8$, i.e. the per-period prospect survival rate is perceived to be 80%. When payoffs are delayed by two periods, $t = 2$, observed probability weights $\tilde{w}$ shift upwards, as shown in Panel 2a. This shift rotates the decision weights $\tilde{\pi}_i$ counterclockwise, as depicted in Panel 2b. Now the worst outcomes are underweighted while the best ones are more strongly overweighted. For longer time delays these effects become more pronounced and may lead to a substantial underweighting of the worst outcomes. Thus, underweighting of adverse extreme events becomes more likely with longer time horizons.

3.2 Fact #2: Delay Dependence of Time Discounting Behavior

The following section is dedicated to the fact that observed discount rates decrease with the length of delay, i.e. exhibit a hyperbolic decline. Allegedly guaranteed future payoffs constitute a special case of risky ones. As is evident from Equation 12, the observed discount weight for time equals $\tilde{\rho}(t) = w(s^t)\rho(t)$. Clearly, if w is linear, $\tilde{\rho}$ declines exponentially irrespective of the magnitude of s. To see this, note that $\rho(t) = \exp(-\eta t)$ and $s^t = \exp\Big(- \big(- \ln(s)\big)t\Big)$, implying a discount rate $\tilde{\eta} = \eta - \ln(s) > \eta$ for $0 < s < 1$. In this case, uncertainty *per se* increases the absolute level of revealed impatience, but cannot account for declining discount rates. Thus, an expected utility maximizer will exhibit a constant discount rate that is higher than her underlying rate of pure time preference, but her behavior will not show any of the interaction effects addressed in this paper. If, however, w is subproportional and $0 < s < 1$, the component $w(s^t)$ distorts the discount weight in a predictable way (see details in Proposition 2 in Appendix A.3): The discount function $\tilde{\rho}(t)$ declines at a decreasing rate, i.e. in a hyperbolic way. To show this result, we set $\rho = 1$ without loss of generality. The rate $\tilde{\eta}(t)$ at which $w(s^t)$ declines is defined as

$$\tilde{\eta}(t) = -\frac{\frac{\partial w(s^t)}{\partial t}}{w(s^t)} = -\frac{w'(s^t)s^t \ln s}{w(s^t)} = -\varepsilon(s^t)\ln s\,, \qquad (15)$$

Figure 2: Fact #1. Delay Dependence of Risk Tolerance



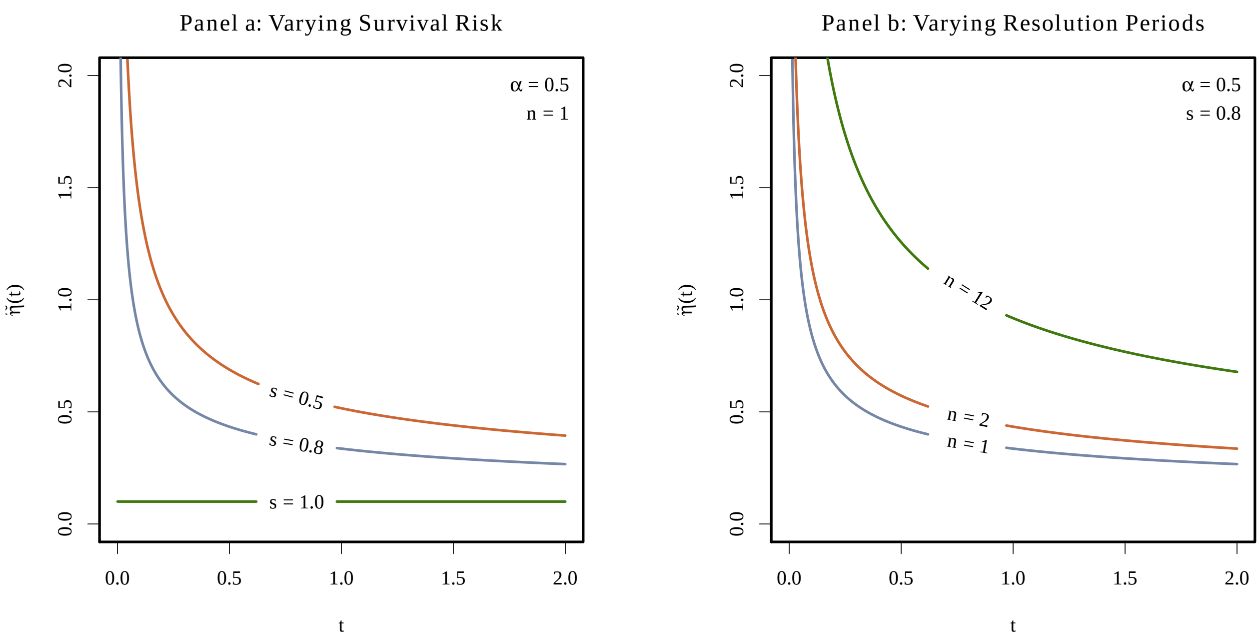
The Figure contrasts atemporal probability and decision weights with weights delayed by $t = 2$ periods. For purposes of illustration, the probability weighting curves are derived from Prelec (1998)'s two-parameter probability weighting function $w(p) = \exp\left(-\beta(-\ln(p))^{\alpha}\right)$, assuming degrees of subproportionality $\alpha = 0.5$ and of convexity $\beta = 1$. Survival risk s is set at 0.8 per period. **Top row - (1) Atemporal:** The graphs show atemporal probability weights w (Panel 1a) and their associated decision weights π (Panel 1b) for a prospect involving 21 equiprobable outcomes, with outcome rank 1 denoting the best outcome. Their objective probabilities are represented on the horizontal gray line. **Bottom row - (2) Delayed:** Panel 2a and 2b show $\tilde{w}$ and $\tilde{\pi}$ for a delay of two periods when uncertainty resolves at $t = 2$

where ε denotes the elasticity of w. Therefore,

$$\tilde{\eta}'(t) = -\varepsilon'(s^t)s^t(\ln s)^2 < 0\,, \tag{16}$$

since the elasticity of w is increasing. As Chakraborty, Halevy, and Saito (2020) have clarified, subproportionality not only predicts hyperbolic discounting, but the reverse relationship also holds in our setting.

Figure 3: Facts #2 and #4: Hyperbolic and Subadditive Discount Rates $\tilde{\eta}$



Panel a shows discount rates as they move with the length of delay t for different levels of survival risk $1 - s$, where s denotes the probability of prospect survival. When there is no survival risk, $s = 1$, the observed discount rate is constant and equals the rate of pure time preference (line labeled by $s = 1.0$). The higher is the level of risk, the lower s, the more pronounced the hyperbolic decline of discount rates over time is for decision makers with subproportional probability weights (curves labeled by $s = 0.5$ and $s = 0.8$). $\tilde{\eta}(t) := -\frac{\partial \tilde{\rho}}{\partial t}/\tilde{\rho}$. w is specified as Prelec's probability weighting function (in this example $\alpha = 0.5$ and $\beta = 1$). **Panel b** depicts discount rates for a constant level of survival probability $s = 0.8$ and varying number of resolution stages n. The more often a particular delay is divided into subintervals (of equal length in this graph), the higher is the discount rate, a manifestation of subadditive discounting.

Thus, decreasing impatience is not a manifestation of pure time preferences but a consequence of survival risk changing the subjective nature of future prospects. At the level of observed behavior, decreasing impatience is the mirror image of increasing risk tolerance if survival risk is integrated into the prospect's probability distribution. In fact, the degree of proneness to common-ratio violations, the degree of subproportionality, can be interpreted as the degree of time insensitivity. Intuitively, when the future is inherently risky, promised rewards do not materialize with certainty and, therefore, they incorporate the potential of disappointment. Because more immediate payoffs are more likely to actually materialize than more remote payoffs, this potential is perceived to decline with the passage of time and becomes almost negligible for payoffs far out in the future. Technically, since shifting a payoff into the future amounts to scaling down its probability, which constitutes an intertemporal variant of the Allais common ratio effect, a decision maker with subproportional preferences becomes progressively insensitive to a given timing difference. This insight provides a test bed for analyzing risk taking and time discounting behavior at the individual level because the characteristics of the probability weighting function feed directly into the characteristics of the observed discount function. For example, a Prelec compound invariant probability weighting function with $\alpha < 1$ generates a *Constant Relative Decreasing Impatience (CRDI)* discount function, frequently used to map hyperbolic discounting (Bleichrodt, Rohde, and Wakker, 2009).

The effects of survival risk on revealed discount rates are presented in Panel a of Figure 3, which depicts a typical decision maker's observed *discount rates* $\tilde{\eta}$ as they react to varying levels of s. The horizontal line represents the case of no survival risk, $s = 1$. In this case, the observed discount rate $\tilde{\eta}$ is constant and coincides with the true underlying rate of time preference η. When survival risk comes into play, however, discount rates decline in a hyperbolic fashion, and depart from constant discounting increasingly strongly with rising uncertainty, as shown by the curves for $s = 0.8$ and $s = 0.5$, respectively.

3.3 Fact #3: Process Dependence of Risk Taking

So far, we have considered the case when future prospects are evaluated in one single shot. In the following section we analyze the situation of uncertainty resolving in several distinct stages. In the domain of risk, sequential resolution of uncertainty frequently reduces a prospect's value relative to its one-shot counterpart, Fact #3.

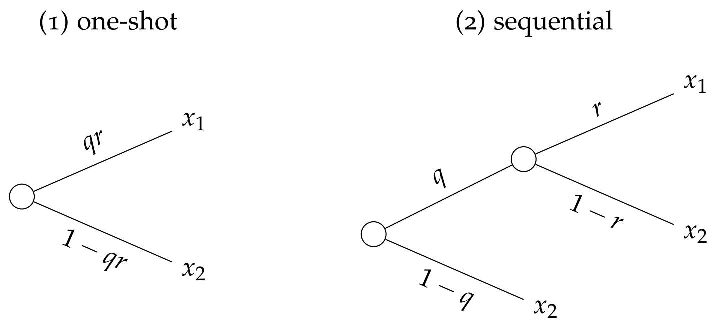
In order to derive our predictions for sequential resolution of uncertainty, we need to discuss the method by which multi-stage prospects are transformed into single-stage ones, the domain over which risk preferences are defined. In principle, there are two different transformation methods, reduction by probability calculus and *folding back*.[13] Reduction involves the calculation of the probabilities of the final outcomes and the transformation of these values by the appropriate weighting function. Folding back, on the other hand, weights the probabilities at each stage and then compounds these weights. Segal (1990) argues that folding back is particularly plausible when the stages are clearly distinct. It is well known that a naive RDU decision maker will be dynamically inconsistent if she cares only about the probabilities of the final outcomes - as the payment date draws near, she will re-evaluate the prospect and, because of the delay dependence of risk tolerance, become comparatively more risk averse. Folding back ensures dynamic consistency but has substantial consequences for prospect valuation.

Experiments on compound risks show that people frequently violate the reduction axiom

[13] Alternatively, the term *backward induction* has been used in the literature. Segal (1990) replaces the reduction axiom by an axiom of compound independence which ensures the applicability of folding back as transformation mechanism.

of EUT, i.e. the value of a prospect resolving in several stages differs from the value of the probabilistically equivalent one-stage prospect. Therefore, we assume that the decision maker applies folding back when evaluating the prospect under consideration.

Figure 4: One-Shot and Sequential Resolution of Prospect Risk



(1) one-shot: The tree depicts uncertainty resolution in one stage. **(2) sequential:** The probability tree shows the sequential resolution of uncertainty of a prospect $P = (x_1, qr; x_2)$ in two stages with partial probabilities q and r.

Figure 4 depicts the sequential resolution of a two-outcome prospect $P = (x_1, p; x_2, 1-p)$ in $n = 2$ stages with partial probabilities q and r and the corresponding one-shot resolution case. Under folding back, the prospects' values are given by

$$V_1(P) = \Big(u(x_1) - u(x_2)\Big)w(qr) + u(x_2)\,,$$

and

$$V_2(P) = \Big(u(x_1) - u(x_2)\Big)w(r)w(q) + u(x_2)\,,$$

where the subscripts of V denote the number of resolution stages. As already noted, and discussed in detail in Segal's work (Segal, 1987a,b, 1990), the certainty effect embodied in subproportional preferences generates an aversion to compounded probability weights: For $1 > p = qr > 0$ the compounding of the respective weights always leads to lower prospect values, i.e. $w(qr) > w(q)w(r)$ holds whatever are the values of q and r. Here the order of r and q, i.e. which probability resolves first, does not play a role, a feature labeled *event commutativity* (Chung, von Winterfeldt, and Luce, 1994). Furthermore, a prospect's minimum value is attained when compounding occurs over equiprobable stages, i.e. when $r = q = \sqrt{p}$. Partitions of equal length correspond to the least degenerate multi-stage prospect and can be interpreted as the comparatively most ambiguous situation, which is strongly disliked by people with subproportional preferences.[14]

[14]Because of this characteristic, Segal (1987b) proposes to model ambiguity aversion by subproportional risk preferences over two-stage lotteries. A recent paper by Dillenberger and Segal (2014) shows that such an approach has another attractive implication: It is able to solve Machina (2009, 2014)'s paradoxes which involve a number of situations where standard models of ambiguity aversion are unable to capture plausible features of ambiguity attitudes (Baillon, l'Haridon, and Placido, 2011).

In principle, uncertainty may resolve in many different ways. This raises the question how general the prediction is that subproportional decision makers undervalue prospects with sequential resolution of uncertainty relative to those with one-shot resolution. Here we focus on sequential resolution in the form of prospect survival, i.e. we only consider probability trees that, at each stage, render either x_2 or the chance that x_1 is still available at a later stage. We term trees with this structure "survival trees". For example, a third stage with partial probability v could be appended to the tree in Figure 4, such that x_1 materializes with probability qrv. For survival trees with $m = 2$, Segal's insights on two-stage prospects generalize to $n > 2$ stages, i.e. $w(\prod_{i=1}^{n} q_i) > \prod_{i=1}^{n} w(q_i)$ for $\prod_{i=1}^{n} q_i = p$, as shown in Proposition 3 in Appendix A.5.

Another type of survival tree emerges when, at each stage, either the worst possible outcome materializes or "everything is still possible" which could be any number m of probabilistic outcomes that materialize at the final stage. Thus, the survival tree has two branches at all the chance nodes before the final stage, and m branches at the terminal resolution of uncertainty. An example for $m = 3$ and $n = 3$ is discussed in Appendix A.4. Subproportionality makes clear predictions for this type of sequential resolution of uncertainty as well: The prospect's one-shot value will be greater than its folded back version. Thus, such a resolution process has the flavor of disappointment aversion since at each stage something better than x_m may turn out to be unreachable. For n resolution stages and m outcomes, the resulting probability weighting function for $\prod_{i=1}^{n} q_i = p$ is given by

$$w_n(p) = \prod_{i=1}^{n} w(q_i). \tag{17}$$

Details are set out in Proposition 3 in Appendix A.5.

The top row of Figure 5 shows the basic probabiliy weighting function and the decision weights of 21 equiprobable outcomes when uncertainty resolves in one shot. On the bottom, the probability weighting function and the corresponding decision weights are displayed that result from compounding over 12 stages of equal partial probability. As one can see, the originally regressive probability weighting function is transformed into a strongly convex one. The decision weight curve now rotates clockwise, implying substantial underweighting of the best outcomes and overweighting of the worst outcomes, as is evident in Panel 2b. Thus, compounding probability weights greatly reduces risk tolerance. Sequential valuation of this type, therefore, has a dramatic effect on the overweighting of adverse tail events. This effect may be called *myopic probability weighting* in the style of myopic loss aversion (Benartzi and Thaler, 1995) which has similar consequences on risk taking behavior when short-sighted investors are frequently exposed to the possibility of facing losses.

To sum up: If uncertainty resolves according to a survival tree, one-shot resolution is always preferred to sequential resolution of uncertainty. A preference for one-shot resolution of uncertainty does not hold generally under subproportionality in RDU (Dillenberger, 2010), however, but only applies to the class of resolution processes studied here.[15] For details see our discussion

[15]An example for which a general preference for one-shot resolution cannot be predicted is the series of experiments

in Appendix B.2.

3.4 Extension: Process Dependence of Risk Taking and the Passage of Time

The property of aversion to compound risk carries over to the case when the passage of time with its inherent uncertainty is introduced. In our view, this situation constitutes a much more interesting case than the frequently observed aversion to sequential resolution in atemporal experimental settings. However, we are not aware of any studies involving the sequential resolution of uncertainty of genuinely delayed prospects. Thus, the following insights provide the basis for novel experimental investigations.

In the following, we set $\rho = 1$ for ease of exposition. Let us first consider a two-outcome prospect $P = (x_1, p; x_2)$ resolving in $n = 2$ stages denoted by corresponding subscripts to $\tilde{w}$ and $\tilde{\rho}$, such that uncertainty is partially resolved at some future time t_1 and fully resolved at the payment date $t > t_1$, as depicted in Panel (ii) of Figure 6. Applying folding back, the resulting two-stage prospect is evaluated as

$$\begin{aligned} [V_2(\tilde{P})]_0 &= \Big(u(x_1) - u(x_2)\Big) w\left(p^{\frac{t_1}{t}} s^{t_1}\right) w\left(p^{\frac{t-t_1}{t}} s^{t-t_1}\right) + u(x_2) w\left(s^{t_1}\right) w\left(s^{t-t_1}\right) \\ &= \left(\Big(u(x_1) - u(x_2)\Big) \frac{w\left(p^{\frac{t_1}{t}} s^{t_1}\right) w\left(p^{\frac{t-t_1}{t}} s^{t-t_1}\right)}{w\left(s^{t_1}\right) w\left(s^{t-t_1}\right)} + u(x_2)\right) w\left(s^{t_1}\right) w\left(s^{t-t_1}\right) \\ &= \left(\Big(u(x_1) - u(x_2)\Big) \tilde{w}_2(p) + u(x_2)\right) \tilde{\rho}_2(t)\,, \end{aligned} \tag{18}$$

which yields the relationships

$$\tilde{w}_2(p) = \frac{w\left(p^{\frac{t_1}{t}} s^{t_1}\right) w\left(p^{\frac{t-t_1}{t}} s^{t-t_1}\right)}{w\left(s^{t_1}\right) w\left(s^{t-t_1}\right)} \tag{19}$$

and

$$\tilde{\rho}_2(t) = w\left(s^{t_1}\right) w\left(s^{t-t_1}\right)\,, \tag{20}$$

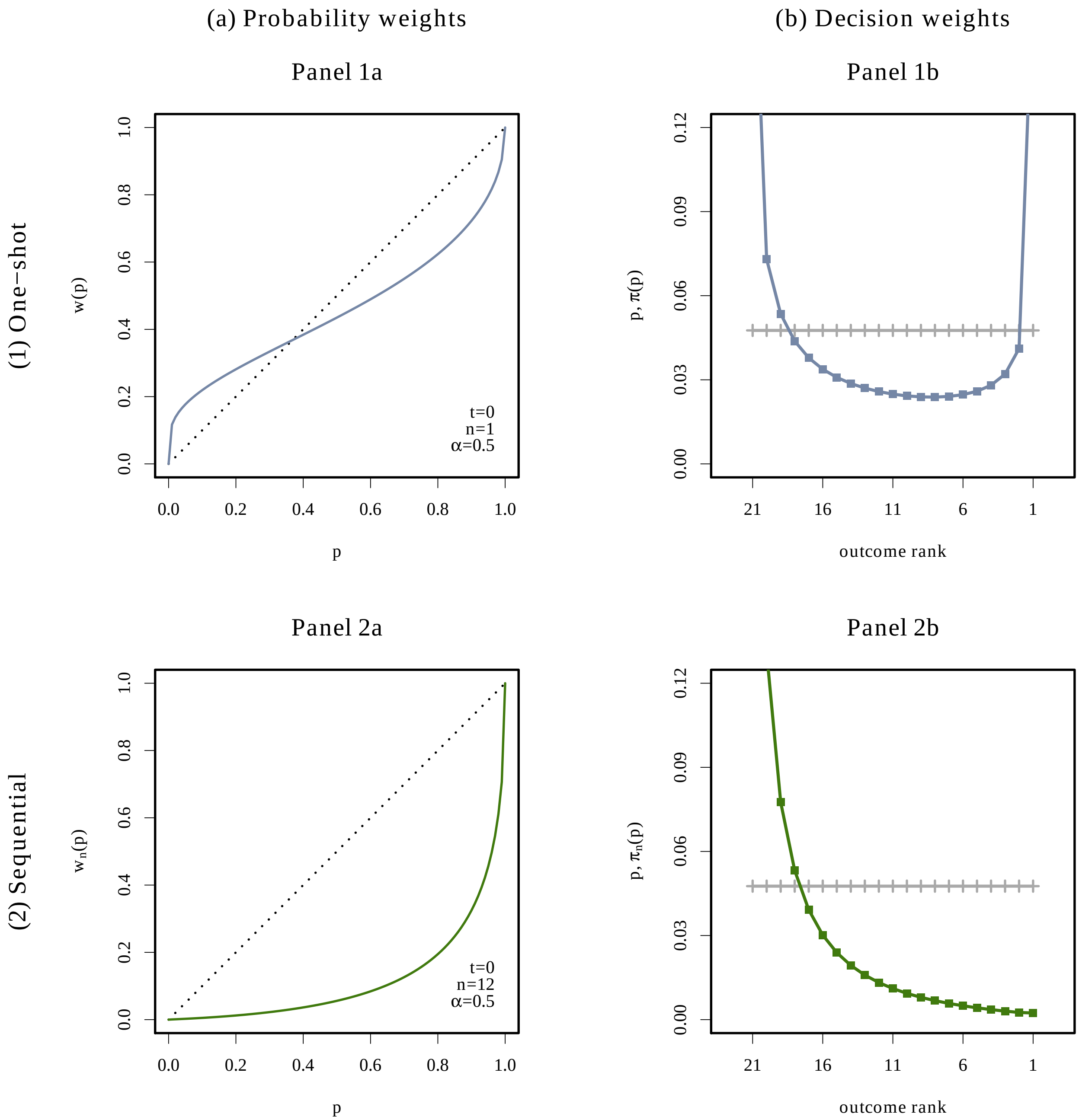
where $\tilde{\rho}_2(t)$ is interpreted as the discount weight attached to the allegedly certain outcome x_2. Subproportionality ensures that

$$\tilde{w}_2(p) = \frac{w\left(p^{\frac{t_1}{t}} s^{t_1}\right) w\left(p^{\frac{t-t_1}{t}} s^{t-t_1}\right)}{w\left(s^{t_1}\right) w\left(s^{t-t_1}\right)} < \frac{w(ps^t)}{w(s^t)} = \tilde{w}(p)\,, \tag{21}$$

i.e. under folding back observed risk tolerance is smaller than in the one-shot case, one of the main results generalized in Proposition 4 in Appendix A.6 where we provide a characterization

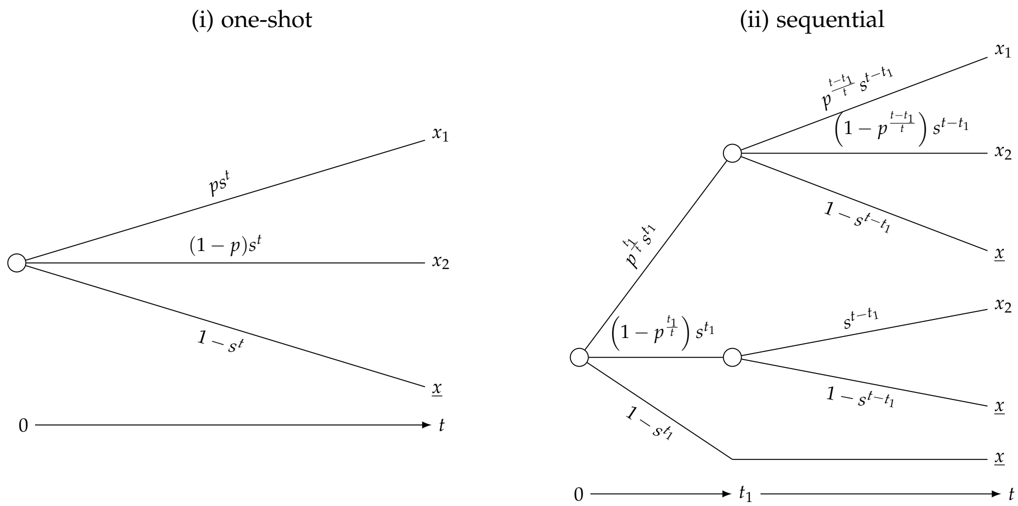
based on the investment games by Gneezy and Potters (1997); Gneezy, Kapteyn, and Potters (2003); Bellemare, Krause, Kröger, and Zhang (2005). In these experiments, participants face three independent draws of the same probability distribution and have to decide how much of their initial endowment they want to invest into the risky asset. The resulting probability tree does not belong to the class of survival trees studied in the current paper because at each stage the player may lose only one invested amount and does not face the worst case of losing the total of three invested amounts. Such a situation only arises at the third draw.

Figure 5: Fact #3: Preferences for the Resolution of Atemporal Uncertainty



The figure contrasts probability and decision weights for one-shot resolution of uncertainty with the weights for sequential resolution if the passage of time does not play a role. For purposes of illustration, the curves are derived from Prelec (1998)'s two-parameter probability weighting function $w(p) = \exp\left(-\beta(-\ln(p))^{\alpha}\right)$, assuming degrees of subproportionality $\alpha = 0.5$ and of convexity $\beta = 1$. **Top row - (1) One-shot:** The graphs show probability weights w (Panel 1a) and their associated decision weights π (Panel 1b) for a prospect involving 21 equiprobable outcomes, with outcome rank 1 denoting the best outcome when uncertainty resolves in one-shot. Their objective probabilities are represented on the horizontal gray line. **Bottom row - (2) Sequential:** Panel 2a and 2b show the compounded probability weights $w_n(p) = \prod_{i=1}^{n} w(q_i)$ and the corresponding decision weights π_n when uncertainty resolves in $n = 12$ equiprobable stages, $q_i = p^{1/12}$.

Figure 6: One-Shot and Sequential Resolution of Prospect and Survival Risk



(1) one-shot: The tree depicts uncertainty resolution of a prospect $(x_1, ps^t; x_2, (1-p)s^t; \underline{x})$ in one stage. **(2) sequential:** The probability tree shows the sequential resolution of uncertainty of the same prospect in two stages with partial probabilities $(p^{1/t}s)^{t_1}$ and $(p^{1/t}s)^{t-t_1}$.

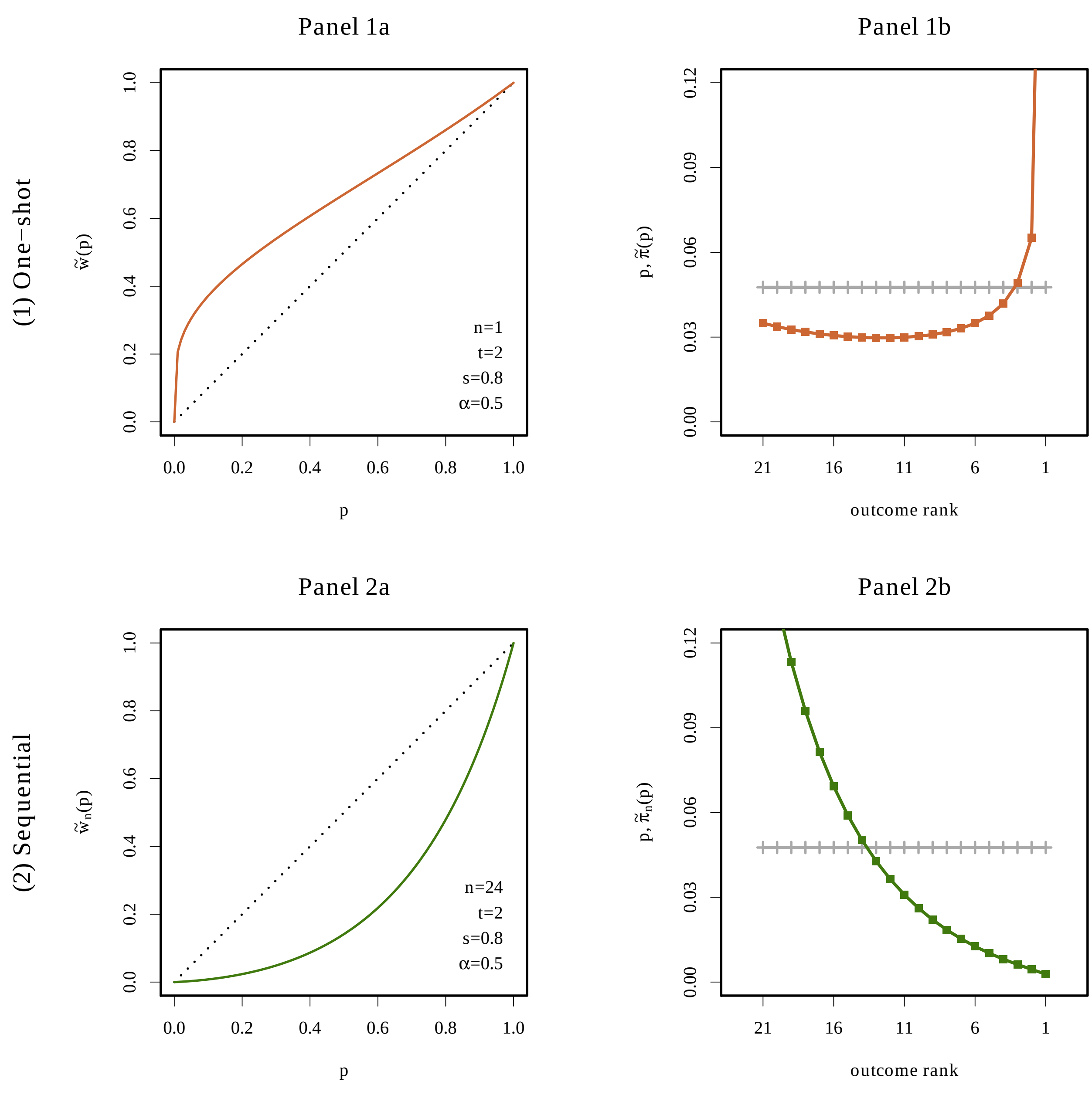
of $\tilde{w}_n$. Furthermore, total prospect value is also smaller than for one-shot resolution as both $w(ps^t)$ and $w(s^t)$ are greater than any products of probability weights of partial probabilities. Thus, the preference for one-shot resolution of uncertainty is preserved when "something may go wrong". Probability weights $\tilde{w}$ and $\tilde{w}_n$ as well as their corresponding decision weights $\tilde{\pi}$ and $\tilde{\pi}_n$ are depicted in Figure 7, which show the same patterns as for the atemporal case of Figure 5, but less pronounced because delay dependence shifts the original atemporal probability weights upwards.

3.5 Fact #4: Process Dependence of Time Discounting

Fact #4 pertains to the finding that discount rates compounded over partial periods are higher than discount rates applied to the total period under consideration, so-called subadditive discounting. As we will see shortly, we can transfer all our insights for the sequential resolution of uncertainty to discounting behavior as allegedly certain future outcomes are a special case within the class of two-outcome prospects. According to our model an allegedly certain outcome x payable at delay t is perceived as a risky future prospect $(x, s^t; \underline{x}, 1-s^t)$. Suppose now that future uncertainty resolves in two stages, first at t_1 and finally at t. Coming back to Figure 4, redefine x_2 as $\underline{x}$ and the partial probabilities as survival probabilities, $q = s^{t_1}$ and $r = s^{t-t_1}$. Subproportionaliy implies $w(s^t) > w(s^{t_1})w(s^{t-t_1})$, in other words discounting is subadditive, described as Fact #4. As before, this result holds for any number of resolution stages, and the more stages are involved the stronger the compounding effect.

Panel b of Figure 3 shows the effect of varying the number of compounding stages on ob-

Figure 7: Extension: Preferences for the Resolution of Uncertainty with Survival Risk



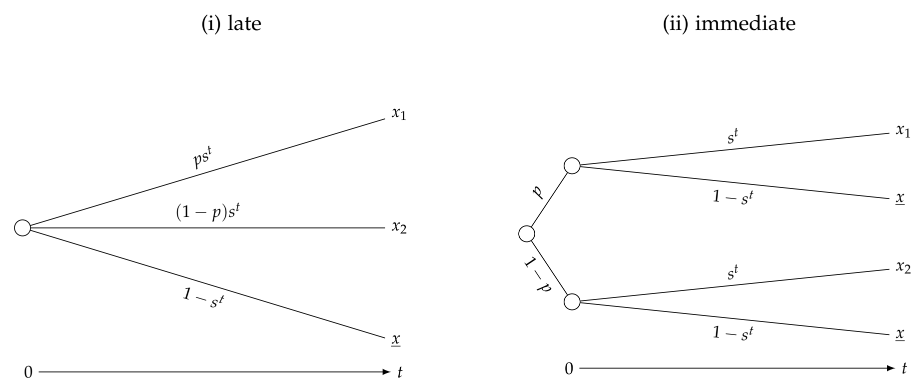
The figure shows the impact of one-shot resolution of uncertainty versus the sequential resolution of uncertainty in the presence of survival risk when the prospect under consideration is delayed by $t = 2$ periods. For purposes of illustration, the curves are derived from Prelec (1998)'s compound invariant probability weighting function $w(p) = \exp\left(-\beta(-\ln(p))^{\alpha}\right)$, assuming degrees of subproportionality $\alpha = 0.5$ and of convexity $\beta = 1$. **Top row - (1) One shot:** The graphs show delay-dependent probability weights $\tilde{w}$ (Panel 1a) and their associated decision weights $\tilde{\pi}$ (Panel 1b) for a prospect involving 21 equiprobable outcomes, with outcome rank 1 denoting the best outcome. Their objective probabilities are represented on the horizontal gray line. **Bottom row - (2) Sequential:** Panel 2a and 2b show $\tilde{w}_n(p) = \left(\frac{w((ps^t)^{1/n})}{w((s^t)^{1/n})}\right)^n$ and the corresponding decision weights π_n when uncertainty resolves in $n = 24$ equiprobable stages.

served discount rates. As predicted, discount rates increase in the number of stages. In our model, subadditive discounting is the result of decision makers' aversion to compounded probability weights and not a feature of pure time preferences themselves, as often posited in the literature.

3.6 Fact #5: Preferences for the Timing of Uncertainty Resolution

The experimental evidence has found quite a puzzling result: a substantial share of participants prefer uncertainty to be resolved at the payment date, even in circumstances when one would expect that it is advantageous to know the outcome of one's financial decisions as early as possible. In this section, we explore the consequences of subproportionality for the preferences for the timing of uncertainty resolution.

Figure 8: Late and Immediate Resolution of Prospect Risk



(1) late: The tree depicts uncertainty resolution in one stage at the payment date t. **(2) immediate:** The probability tree shows the immediate resolution of prospect risk, with survival risk resolving at t.

Figure 8 depicts two different cases of the timing of uncertainty resolution: either the prospect is played out at the payment date, corresponding to one-shot resolution and labeled "late" (Panel i), or the prospect is played out immediately after prospect valuation, labeled "immediate" (Panel ii). In the latter case, the decision maker will know the outcome right after her decision and faces only survival risk. Contrasting the resulting prospect values,

$$\begin{aligned} V_0(\tilde{P})_{late} = & \quad \left(\Big(u(x_1)-u(x_2)\Big)\frac{w(ps^t)}{w(s^t)}+u(x_2)\right)w(s^t)\rho(t) > \\ V_0(\tilde{P})_{immediate} = & \quad \left(\Big(u(x_1)-u(x_2)\Big)w(p)+u(x_2)\right)w(s^t)\rho(t)\,, \end{aligned} \tag{22}$$

shows that late resolution is always preferred as $\frac{w(ps^t)}{w(s^t)} > w(p)$ is implied by subproportionality. Thus, if no other considerations, such as being able to make better future plans, play a role, a subproportional decision maker will exhibit a preference for late resolution of uncertainty. In fact, she will prefer resolution at t to any earlier resolution time $t_1 < t$, as shown in Appendix A.7.

In our view, that subproportional risk preferences induce an intrinsic preference for late resolution of prospect risk constitutes the third important result besides delay- and process-dependence. If decision makers perceive the future as inherently risky and apply folding back, this property follows endogenously from subproportionality and does not constitute an independent preference as in the theoretical literature on resolution timing (Kreps and Porteus, 1978; Chew and Epstein, 1989; Grant, Kajii, and Polak, 2000). Moreover, our model not only predicts a general preference for late resolution of prospect risk, it also specifically addresses skewness preferences because the effect is larger for small probabilities (see Proposition 5 in Appendix A.7), which cannot be handled by utility-based explanations. Additionally, this preference for late resolution of uncertainty of positively skewed prospects increases with time delay.

3.7 Fact # 6: Risk Dependence of Patience

Researchers have been puzzled not only by delay-dependent risk tolerance and preferences with respect to resolution timing but also by other interactions between time and risk, encompassing risk-dependent discounting and diminishing immediacy: Certain outcomes tend to be discounted much more heavily than risky outcomes are (Stevenson, 1992; Ahlbrecht and Weber, 1997). As we will show below, these findings can be naturally accommodated within our framework.

Let V_0 denote the *present value* of the prospect $P = (x_1, p; x_2, 1-p)$ delayed by t periods. Hence, for $\rho = 1$,

$$V_0 = \left(\Big(u(x_1) - u(x_2) \Big) \frac{w(ps^t)}{w(s^t)} + u(x_2) \right) w(s^t) \,. \tag{23}$$

Furthermore, let V_t denote the *future value* of P as of t:

$$V_t = \Big(u(x_1) - u(x_2) \Big) w(p) + u(x_2) \,. \tag{24}$$

Discounting by $w(s^t)$ yields

$$V_t w(s^t) = \left(\Big(u(x_1) - u(x_2) \Big) w(p) + u(x_2) \right) w(s^t) \,. \tag{25}$$

According to standard discounting theory, the present value V_0 should be equal to the discounted value of V_t, namely $V_t w(s^t)$. However, because $w(p) < \frac{w(ps^t)}{w(s^t)}$, actually $V_t w(s^t) < V_0$. Therefore, it seems as if the certain value V_t is discounted more heavily than the (at t equally attractive) future prospect. The difference in the valuations is not caused by different rates of time preference for risky and certain payoffs, however, but by survival risk changing the nature of the future prospect when evaluated from the point of view of the present rather than from the point of view of the future.

The same kind of risk dependence is at work when the revealed preference for a certain smaller present payoff over an allegedly certain larger later payoff decreases substantially when both payoffs are made (objectively) probabilistic, a phenomenon termed diminishing immediacy (Keren and Roelofsma, 1995; Weber and Chapman, 2005). Because of the certainty effect, the additional layer of riskiness affects the later payoff much less than the present one because, due to survival risk, it is viewed as a risky prospect already from the outset.

3.8 Fact #7: Order Dependence of Risk Tolerance

Order dependence refers to the phenomenon that it makes a difference in which order a prospect is discounted for risk and for time. In principle, there are three different methods of establishing a decision maker's value of a prospect $P = (x_1, p; x_2, 1-p)$ delayed by t periods: the risk-first order, the time-first order, and the direct method by which both operations are performed simultaneously.

The risk-first order assesses the certainty equivalent as of time t at the first stage and its present value at the second stage. The time-first order reverses the elicitation stages and encompasses, at the first stage, the elicitation of the present risky prospect which is considered to be equivalent to the future one and, at the second stage, the elicitation of the certainty equivalent of this present risky prospect. The direct method, finally, elicits the present certainty equivalent of the delayed prospect in one single operation.

When the decision maker is required to state the prospect's value when discounting solely for risk, she ignores the dimension of time and reports V_t which gets discounted to $V_t w(s^t)$ at the second stage:

$$V_t w(s^t) = \Big((u(x_1) - u(x_2))w(p) + u(x_2)\Big)w(s^t). \tag{26}$$

Conversely, when discounting for time first, she states the present prospect which is equivalent to the delayed one. Discounting for risk at the second stage results in its value V_0, evaluated as

$$V_0 = \left(\Big(u(x_1) - u(x_2)\Big)\frac{w(ps^t)}{w(s^t)} + u(x_2)\right)w(s^t), \tag{27}$$

which is equal to the present value elicited by the direct method.

Due to subproportionality, $\frac{w(ps^t)}{w(s^t)} > w(p)$. Therefore, we predict that discounting for risk first results in a lower prospect valuation than discounting for time first. Moreover, discounting for time first is equivalent to prospect evaluation in one single operation. In their study on order dependence, Öncüler and Onay (2009) indeed found this pattern: While valuations resulting from the time-risk order and the direct method are not statistically distinguishable from each other, risk-time evaluations are significantly lower than the ones obtained from the other two methods (see also Ahlbrecht and Weber (1997)).

4 Quantitative Assessment

The previous results indicate that our model is capable of explaining all seven facts. The question remains whether the model requires vastly different parameters to explain the various facts or whether is it possible to explain them with a set of parameters within a relatively narrow and plausible range.[16] To address this question, we tie our hands and assume a fixed set of preference parameter values for (i) the utility curvature, (ii) the degrees of subproportionality and convexity of probability weights, and (iii) the rate of pure time preference, as specified in Table 3. These parameters are suggested by typical estimates in the literature (see e.g. Abdellaoui, Diecidue, and Öncüler (2011); Fehr-Duda and Epper (2012)). Together with an estimated value of the survival probability s the preference parameters generate behavioral predictions that we can quantitatively compare with the outcomes of experiments that documented the seven facts. In other words, we use only one free parameter – the value of the (annualized) survival probability s – to fit the experimental data.[17]

Ideally, for a given subject pool at a given point in time, and a given elicitation method, the estimated value of s should be rather similar across experiments because in this case the participants would have little reason to reveal different degrees of subjective uncertainty that "something may go wrong". However, the seven facts we discussed have been documented at different points in time, with different elicitation methods, and with rather different subject pools – French, Swiss, Swedish and US participants. Table 13 in the appendix summarizes the experimental studies we used for our task. Therefore, the best we can hope for is that the estimated value of s is roughly in a similar ballpark across the different experiments. In addition, because all studies have been conducted with university students in advanced Western countries with well-developed property rights, the estimated value of s should not be unreasonably low (e.g., below 0.5 or 0.6 p.a.). As we will see below, our quantitative estimates nicely confirm these expectations. The typical value of the survival probability across experiments is around 0.9 and never below 0.825. Thus, all seven facts can be quantitatively explained with a plausible and identical set of preference parameters and a narrow and plausible range of survival probabilities.

4.1 Fact #1: Delay Dependence of Probability Weights

To demonstrate the quantitative implications of our approach we proceed as follows. According to our framework the driver of risk tolerance increasing with delay are delay-dependent probability weights. Delay-dependent risk tolerance was observed in many experiments, but only a very few provide estimates of suitable probability weights. One particularly useful example is Abdellaoui, Baillon, Placido, and Wakker (2011)'s investigation of the source dependence of uncertainty attitudes. Their experiment also involved pure risk, i.e. given objective probabilities, as a special source for which uncertainty resolved at the payment date three months after the

[16]We thank an anonymous referee for proposing this calibration exercise.

[17]Note that there are so far no stylized facts regarding the perceived subjective uncertainty captured by the survival probability s. Therefore, it makes sense to estimate s.

Table 3: Global Parameter Values

Function	Specification	Parameter	Value
Probability weighting	compound invariant	α: subproportionality	0.50
		β: convexity	0.95
Utility	power	γ: curvature	0.80
Time discounting	exponential	η: rate of time preference	0.10

The functions are specified as follows. Prelec (1998)'s compound-invariant probability weighting function: $w(p) = \exp(-\beta(-\ln(p))^{\alpha})$. Power utility function: $u(x) = x^{\gamma}$. Time discount function: $\rho(t) = \exp(-\eta t)$.

experimental sessions. Abdellaoui, Baillon, Placido, and Wakker (2011) assume a Prelec (1998) compound-invariant probability weighting function and report $\hat{\alpha} = 0.67$ and $\hat{\beta} = 0.76$ for the delayed weights $w_{t=3}(p)$ (see their Figure 9 on page 713). Now, what is the level of survival probability s such that their delayed weights $w_{t=3}(p)$ can be interpreted as $\tilde{w}(p)$ based on the atemporal weights $w_{t=0}(p)$ generated by our global parameters? We estimate s by minimizing the sum of squares of the difference between $\tilde{w}$ and $w_{t=3}$. This exercise yields an optimal s^* of 0.825 p.a., which we deem a very plausible number. In other words, subjects behaved as if they thought outcomes payable in one year would actually materialize with a 82.5% chance. As one can see in Figure 9, the curve of $\tilde{w}$ predicted for this level of s^* closely matches the actual reported curve $w_{t=3}$.

4.2 Fact #2: Hyperbolic Discounting

Epper, Fehr-Duda, and Bruhin (2011) elicited both time preferences and risk preferences of a student sample.[18] Comparing the average annualized discount rates observed for a two-month delay and a four-month delay shows the usual picture: they decline from 0.368 to 0.299 when the delay increases (all these numbers can be found in the first column of their Table 2 on page 183 of the paper). Assuming that the discount rates are generated by the theoretical discount weights $w(s^t)\exp(-\eta t)$, we estimate s by minimizing the sum of squared deviations between model predictions and observations. This procedure yields $s^* = 0.947$, resulting in predicted discount rates of 0.372 for the two-month delay, and 0.293 for the four-month delay, shown in Table 5, which are very close to the observed values.

Epper, Fehr-Duda, and Bruhin (2011) provide a number of additional insights. A large body of empirical evidence documents the prevalence of common-ratio violations as well as of non-exponential discounting, at least at the level of aggregate behavior (Kahneman and Tversky, 1979; Thaler, 1981; Benzion, Rapoport, and Yagil, 1989; Starmer and Sugden, 1989; Prelec and Loewenstein, 1991). However, there is vast heterogeneity in individuals' behaviors (Hey and Orme, 1994;

[18]Based on the risk taking data, Epper, Fehr-Duda, and Bruhin (2011) estimated the mean Prelec $\alpha = 0.505$ and the mean $\beta = 0.974$, which lie very close to our global parameter values.

Figure 9: Observed versus Predicted Weighting Functions

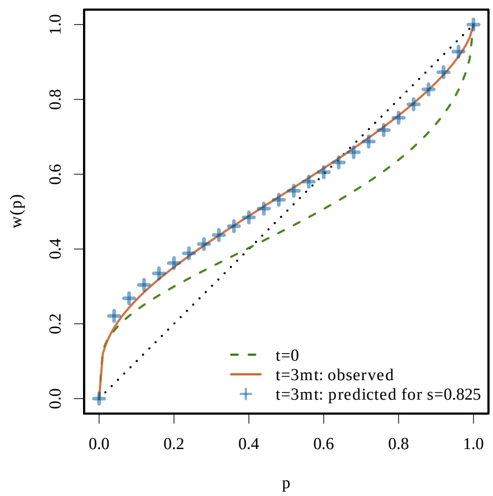


Table 4: Decision Weights

Probability	Observed	Predicted
0.125	0.289	0.308
0.250	0.388	0.395
0.500	0.552	0.544
0.750	0.719	0.710
0.875	0.821	0.821

Figure: The green curve corresponds to the atemporal probability weighting function, $w_{t=0}$, generated by our global parameter values. The red curve depicts the probability weighting function estimated by Abdellaoui, Baillon, Placido, and Wakker (2011) for uncertainty resolution in three months, $w_{t=3}$ (see Panel B in their Figure 9 on page 713). The blue cross-shaped curve is our prediction for the global parameter values and a survival probability of $s^* = 0.825$. **Table:** For each of the five probabilities in Abdellaoui, Baillon, Placido, and Wakker (2011), the table contrasts observed decision weights of the better outcome with decision weights predicted by our model with global parameter values and $s^* = 0.825$. The exact observed weights are computed using the probability weighting function estimated by Abdellaoui, Baillon, Placido, and Wakker (2011).

Chesson and Viscusi, 2000; Bruhin, Fehr-Duda, and Epper, 2010) and the question arises whether common-ratio violations and non-constant discounting are actually exhibited by the same people. To our knowledge, potential links between risk preferences and time preferences at the level of preference conditions or parameter estimates have not been explored so far. However, using the decline of discount rates as a measure of decreasing impatience, Epper, Fehr-Duda, and Bruhin (2011) provide evidence that subjects' departures from linear probability weighting are indeed highly significantly correlated with the strength of the decrease in discount rates. In fact, the only variable associated with decreasing discount rates turns out to be the degree of subproportionality of probability weights, which explains a large percentage of the variation in the extent of the decline, whereas observable individual characteristics, such as gender, age, experience with investment decisions and cognitive abilities are not significantly correlated with the degree of non-constant discounting.

4.3 Fact #3: Preference for One-Shot Resolution of Uncertainty

To the best of our knowledge, the process dependence of risk taking behavior has not been investigated experimentally in situations when there is actually a substantial time delay present.

Table 5: Hyperbolic Discounting - Observed versus Predicted Discount Rates

Delay	Observed	Predicted
2 months	0.368	0.372
4 months	0.299	0.293

The table lists observed and predicted annualized discount rates for the two different time delays in Epper, Fehr-Duda, and Bruhin (2011). The observed rates can be found in the first column of Epper, Fehr-Duda, and Bruhin (2011)'s Table 2 (page 183). The predicted rates result from our model with the global parameter values and $s^* = 0.947$.

Experimental tasks are typically based on one-stage and numerically equivalent multi-stage prospects that are resolved almost immediately. In other words, survival probability s is irrelevant in such situations. Thus, we will illustrate an atemporal version of the preference for one-shot resolution of uncertainty over sequential resolution by (i) documenting that the predicted certainty equivalents of one-shot resolved prospects are always higher than those of sequentially resolved prospects, and (ii) by comparing the actually observed certainty equivalents with the predicted certainty equivalents.

Abdellaoui, Klibanoff, and Placido (2015) report mean certainty equivalents for simple prospects $(50, p; 0, 1-p)$ (their Table 2 on page 1310) that are resolved in one stage or in two stages. Table 6 shows that one-shot certainty equivalents are always higher than sequential ones, and the predicted values, based on our global parameters, are reasonably close to the observed ones, particularly for the probabilities $1/2$ and $11/12$. For $p = 1/12$ the model overpredicts the difference between one-shot and sequential values.[19]

Table 6: Process Dependence - Observed versus Predicted Certainty Equivalents

Prospect	Condition	Observed	Predicted
$(50, 1/12; 0)$	one-shot	9.910	11.184
	sequential	9.250	7.182
$(50, 1/2; 0)$	one-shot	22.650	22.671
	sequential	20.720	19.052
$(50, 11/12; 0)$	one-shot	37.740	37.781
	sequential	34.720	35.435

The table lists certainty equivalents documented in Table 2 on page 1310 of Abdellaoui, Klibanoff, and Placido (2015) for one-shot and sequential resolution (their "CRG" condition). The predictions are obtained for our model with the global parameter values listed.

[19]Regarding the other features of our model, event commutativity and aversion to equiprobable stages, the evidence so far is mixed. For a review see Fan, Budescu, and Diecidue (2018).

4.4 Fact #4: Subadditive Discounting

To illustrate the quantitative implications of our model we again examine the discounting data of Epper, Fehr-Duda, and Bruhin (2011). In the experiment, future equivalents FEs of a fixed sooner amount of CHF 60 were elicited for various time delays. We define the observed *discount fraction* as

$$f(t_1, t_2) = \frac{60}{\text{FE}},$$

where t_1 is the payment date for the sooner amount 60 and t_2 the payment date for the later amount FE (Read, 2001). If the product $f(t_1, t_2) f(t_2, t_3)$ is smaller than the discount fraction over the total period, $f(t_1, t_3)$, then discounting is subadditive. According to our model, indifference between sooner and later payments is given by

$$u(60) w(s^{t_1}) \exp(-\eta t_1) = u(\text{FE}) w(s^{t_2 - t_1}) w(s^{t_1}) \exp(-\eta t_2).$$

Assuming power utility with parameter γ the predicted discount fraction equals to

$$f(t_1, t_2) = \frac{60}{\text{FE}} = \left(\frac{w(s^{t_2 - t_1}) \exp(-\eta t_2)}{\exp(-\eta t_1)} \right)^{\frac{1}{\gamma}}.$$

Given the optimal survival probability derived for the same data set of Fact #2, $s^* = 0.947$, the following predictions for the discount fractions result, listed in Table 7.

Table 7: Subadditive Discounting - Observed versus Predicted Discount Fractions

Discount Fraction	Observed	Predicted
$f(0,2)$	0.927	0.893
$f(2,4)$	0.941	0.893
$f(0,4)$	0.886	0.852
$f(0,2)f(2,4)$	0.872	0.797

The table lists discount fractions for various payment dates and the relevant product. Observed values are derived from the values shown in Table 2 of Epper, Fehr-Duda, and Bruhin (2011) (page 183). Predictions are derived by our model with global parameter values and $s^* = 0.947$.

Both the observed mean discount fractions and the predicted ones clearly exhibit subadditivity, with predictions fitting fairly well.

4.5 Fact #5: Preferences for the Timing of Uncertainty Resolution

Arai (1997) measured (hypothetical) strength of preference (SOP) towards resolution timing for delayed prospects that varied by outcome probability and time delay. In this case we do not have present certainty equivalents at our disposal but have to rationalize strength of preference values. We report Arai (1997)'s findings on the prospect $(5000, p; 0, 1-p)$ listed in Table 1 on page 20 of his paper. Strength of preference was measured on a scale divided into 30 equal intervals, with SOP $= 0$ denoting strong preference for immediate resolution and SOP $= 30$ denoting strong preference for late resolution. Thus, SOP $= 15$ signals indifference between immediate and late resolution of uncertainty.

Arai (1997) finds a very distinct pattern of strength of preference depending on time delay and probability: The smaller the probability and the longer the time delay, the stronger the preference for late resolution. Our task is to predict the patterns observed by Arai (1997). For this purpose we examine the wedge $W(p,t) := \frac{w(ps^t)}{w(p)w(s^t)}$ which measures the decision weight for late resolution relative to the decision weight for immediate resolution of uncertainty. We hypothesize that it is more likely to observe strength of preference SOP > 15 in favor of late resolution for greater values of the wedge $W(p,t)$. We calculate $W(p,t)$ by assuming our global parameters and survival probability $s = 0.9$, which lies in the range of optimal s^* found for the other facts (see Table 12 below.)

Table 8: Resolution Timing - $W(p,t)$ and Strength of Preference SOP

	$t = 1/4$		$t = 2$		$t = 10$	
p	$W(p,t)$	SOP	$W(p,t)$	SOP	$W(p,t)$	SOP
0.05	1.16	16.4	1.46	17.0	2.03	17.8
0.35	1.15	15.6	1.41	16.5	1.77	18.2
0.65	1.14	12.4	1.35	14.4	1.55	17.2
0.95	1.11	12.3	1.18	13.9	1.21	16.9

The table shows wedges $W(p,t) = \frac{w(ps^t)}{w(p)w(s^t)}$ predicted by our model with global parameter values and $s = 0.9$ and observed strength of preferences values reported in Arai (1997) (Table 1 on page 20).

Table 8 shows a totally consistent picture, $W(p,t)$ is predicted to decrease in p and increase in delay t, capturing the patterns in the observed strength of preference measures. The Spearman rank correlation coefficient between SOP and $W(p,t)$ amounts to 90.2%, which we deem an excellent match.

4.6 Fact #6: Risk-Dependent Discounting

In their experiments, Weber and Chapman (2005) investigated whether delaying an outcome is equivalent to making it risky. In one of these experiments participants' present certainty equivalents for delayed prospects were elicited through a series of choices using a bisection algorithm. 124 participants supplied useful responses in the immediacy task, which involved hypothetical amounts of \$100 and \$110. These amounts were payable either immediately or with various time delays, and were supposedly certain or risky materializing with a probability of $p = 0.5$.

Working with our global parameters we estimated the optimal survival probability that minimizes the sum of squares of differences between observed and predicted values. This exercise resulted in an estimate of $s^* = 0.872$, again a very reasonable number.

Table 9: Risk-Dependent Discounting - Observed versus Predicted Present Certainty Equivalents

Delay	Amount	Probability	Observed	Predicted
0	100	1.0	100.00	100.00
		0.5	38.32	37.21
4	110	1.0	70.52	81.86
		0.5	35.46	38.02
26	100	1.0	41.11	39.94
		0.5	23.34	23.40
30	110	1.0	47.85	40.16
		0.5	23.75	24.03

The table lists present certainty equivalents reported in Weber and Chapman (2005), Table 5 (page 111). The predicted present certainty equivalents are obtained using our model with global parameter values and $s^* = 0.872$.

Table 9 contrasts observed present certainty equivalents[20] with predicted ones. Generally, we are able to produce an excellent match between observed and predicted values, only the present value of 110 materializing in 4 months is overstated by the model, i.e. participants discounted 110 much more heavily than predicted. According to our model an allegedly certain outcome payable at delay t, (x, t), is evaluated as $u(x)w(s^t)\exp(-\eta t)$. Its risky counterpart is evaluated as $u(x)w(ps^t)\exp(-\eta t)$. Their corresponding non-delayed values amount to $u(x)$ and $u(x)w(p)$, respectively, implying the discount weights $w(s^t) < \frac{w(ps^t)}{w(p)}$ for the certain and risky outcomes. Comparing the entries for $p = 1$ and $p = 0.5$ for the various delays in Table 10 clearly shows a greater loss in value for allegedly certain outcomes than for risky ones.

Table 10: Diminishing Immediacy - Predicted Discount Weights

Delay t	Amount	Probability	Discount Weight
4	110	1.0	92.4%
		0.5	94.3%
26	100	1.0	59.8%
		0.5	69.0%
30	110	1.0	55.3%
		0.5	65.3%

The table list predicted discount weights for the different delayed prospects in Weber and Chapman (2005).

4.7 Fact #7: Order Dependence

In their study on order dependence, Öncüler and Onay (2009) found the following pattern: While valuations of delayed risky prospects resulting from the time-risk order ("TR", discounting for

[20] Values for present certain \$100 were not elicited.

time first and for risk thereafter) and the direct method ("D", both operations performed simultaneously) are not statistically distinguishable from each other, risk-time evaluations ("RT", discounting for risk first and for time thereafter) are significantly lower than the ones obtained from the other two methods. Here we proceeded as before, we minimized the sum of squared deviations between observation and prediction based on our global parameters which resulted in an optimal survival probability $s^* = 0.937$. We report the observed and predicted present certainty equivalents for the three elicitation methods in Table 11. Predictions match observations quite well.

Table 11: Order Dependence - Observed versus Predicted Present Certainty Equivalents

Probability	Condition	Observed	Predicted
0.5	RT	35.94	34.09
	TR	39.83	37.06
	D	39.60	37.06
0.3	RT	22.07	24.89
	TR	24.44	27.09
	D	24.14	27.09

The table shows observed present certainty equivalents reported in Öncüler and Onay (2009), Table 1 on page 285. The predictions are obtained by our model using global parameter values and $s^* = 0.937$.

4.8 Summary

The quantitative assessments conducted in this section were based on the same set of preferences parameters. We deliberately tied our hands for the quantitative predictions by assuming a plausible set of preference parameters suggested by the literature. In this way, we avoid arbitrary degrees of freedom in accommodating the data and enable a judgment to what extent our approach indeed facilitates a unifying explanation of a diverse set of facts. Table 12 shows that the objects of interest that needed to be predicted to explain the facts were quite varied – ranging from probability weights to discount rates, from discount fractions to (present) certainty equivalents. Our quantitative analysis suggests that our predictions are in general well matched with the observations. Furthermore, as Table 12 reveals, we find that observed behavior is consistent with plausible values of an annual survival probability in the range of 0.825 to 0.947. In view of the fact that the data were elicited from different subject pools in different countries and at different points in time, we deem this a remarkably narrow and plausible range of values for survival probability.

Table 12: Summary: Optimal Survival Probability s^*

Fact #	Output Variable	s^* p.a.	Remark
1	Probability weights	0.825	
2	Discount rates	0.947	
3	Certainty equivalents	-	not relevant
4	Discount fractions	0.947	same as in #2
5	Correlation with preference strength	0.900	assumed
6	Present certainty equivalents	0.872	
7	Present certainty equivalents	0.937	

The table lists estimated optimal survival probabilities for each fact (see the remarks for exceptions). Survival probabilities are estimated by minimization of the sum of square deviations between observed output variables and our model with the global parameter values.

5 Experimental Findings and Related Literature

In the following we present the experimental evidence in more detail and discuss previous explanations of the observed effects. Extant explanations usually deal with only one or a few specific regularities and do not address the entirety of the phenomena summarized in Table 1. By now, there is an extensive literature on many of these single aspects, for example on hyperbolic discounting, preferences for resolution timing and the value of information. As reviewing this literature is beyond the scope of this paper, we focus on those contributions that are more closely related to our work.

Delay dependence of risk taking behavior, Fact #1 in Table 1, has been documented by a range of papers that do not distinguish between effects of delay on utility and probability weights (Jones and Johnson, 1973; Shelley, 1994; Ahlbrecht and Weber, 1997; Sagristano, Trope, and Liberman, 2002; Noussair and Wu, 2006; Coble and Lusk, 2010). That, in fact, probability weights react to delay, rather than the utility function, was shown experimentally by Abdellaoui, Diecidue, and Öncüler (2011). They conducted a carefully designed experiment eliciting probability weights for both present and delayed prospects. Their results provide support for our approach as the probability weights of the best possible outcome, when delayed, are significantly greater than their non-delayed counterparts, both in the aggregate as well as for the majority of the individuals. In their study on ambiguity attitudes, Abdellaoui, Baillon, Placido, and Wakker (2011) show estimates of a probability weighting function derived from choices over prospects delayed by three months which we used to assess the quality of our predictions. This function is also much more elevated than typical atemporal estimates are, i.e. the curve lies above a typical atemporal one, see Figure 9.

Baucells and Heukamp (2012) restrict their analysis to the case of simple prospects $(x, p; 0, 1-p)$ that pay x with probability p at time t and zero otherwise. In this setting, the authors predict a number of effects by invoking varying additional assumptions. They model the delay dependence of risk tolerance in the following way. Aside from their fundamental axiom of a direct probability-time trade-off, mentioned in the introduction, they make two additional assumptions to predict risk premia declining with time delay. The first one is the common ratio effect, which is equivalent to subproportionality in our framework. The second crucial assumption makes the probability-time trade-off depend on outcome magnitude – the probability that renders an early prospect equally attractive as a prospect with a fixed additional delay declines with outcome magnitude. While they predict risk premia declining with delay, their time-dependent probability weights $\tilde{w}(p) = w(p \exp(-r_x t))$ clearly decrease with delay. Their approach also predicts hyperbolic discounting (Fact #2) (for this result, the common ratio effect has to hold as well as decreasing elasticity of the utility function) and risk dependence of patience (Fact #6), which is a direct consequence of the probability-time trade-off under subproportionality.

It is well known by now that delay dependence is also manifest in discounting behavior, which constitutes empirical Fact #2. There is abundant evidence that many people exhibit decreasing impatience, i.e. their discount rates are not constant but decline with the length of delay (among many others Benzion, Rapoport, and Yagil (1989); Loewenstein and Thaler (1989); Ainslie (1991); Halevy (2015)). This regularity has triggered a large literature on hyperbolic and quasi-hyperbolic time preferences (e.g. Laibson (1997), for reviews see Frederick, Loewenstein, and O'Donoghue (2002) and Ericson and Laibson (2019)). Most closely related to our approach is a string of papers following Halevy (2008). His model derives hyperbolic discounting from the same mechanism that we employ, namely a combination of future uncertainty with nonlinear probability weighting. The subsequent contributions by Saito (2011) and Chakraborty, Halevy, and Saito (2020) are concerned with establishing a two-way relationship between subproportional probability weights and hyperbolic discounting. The final paper in this series clarifies that subproportionality both implies and is implied by hyperbolic discounting in the domain of single temporal prospects in continuous time, the objects of our model. For consumption streams in discrete time, Halevy (2008)'s original topic, subproportionality still implies hyperbolic discounting, but the reverse direction requires more involved conditions, however.

Another recent contribution to modeling intertemporal choice is Kőszegi and Szeidl (2013)'s model of focusing. By explicitly taking into account attributes of the decision context, their model of attention is able to predict when people exhibit present or future bias. Our approach is able to generate future bias as well, if the decision maker is prone to a reverse common ratio effect (i.e. if the probability weighting function is supraproportional). Gabaix and Laibson (2017) propose a yet different approach to time discounting. They derive hyperbolic discounting from the assumption that that decision makers obtain unbiased but noisy simulations of future utilities. Both the source and the nature of uncertainty differ between their approach and ours: In their model, uncertainty captures the fact that the decision maker does not know the actual future utility she will experience. Simulation noise makes future utility more risky (in terms of second-order stochastic dominance). In contrast, we model the fact that "something may go wrong", which adds a downside risk to future prospects (in terms of first-order stochastic dominance). Moreover, Gabaix and Laibson (2017) do not study the interaction of risk and time preferences.

Another regularity in the data concerns the process dependence of risk taking and time discounting behavior, Facts #3 and #4. In the domain of risk, the prevalent finding is that, on average, subjects do not reduce compound probabilities according to the rules of probability calculus. For example, Aydogan, Bleichrodt, and Gao (2016) show that for their participants the reduction principle is clearly violated at the aggregate level even though 60% of subjects behave in accordance with reduction. The aggregate result is driven by a minority of participants who depart strongly from reduction - in this case in the direction of a preference for sequential resolution. The authors attribute this finding to the utility of gambling. However, there is also abundant experimental evidence that the value of a compound lottery is smaller than the value of the equivalent single-stage lottery, for example Chung, von Winterfeldt, and Luce (1994), Budescu and Fischer (2001), and Fan, Budescu, and Diecidue (2018) to name a few. It seems to be the case that the framing of the experimental tasks play a role whether one finds a preference or an aversion to compound risks (Nielsen, 2020).

One category of results concerns investment games (Gneezy and Potters, 1997; Thaler, Tversky, Kahneman, and Schwartz, 1997; Bellemare, Krause, Kröger, and Zhang, 2005; Gneezy, Kapteyn, and Potters, 2003; Haigh and List, 2005). The general finding is that people tend to invest less conservatively, i.e. they take on more risk, when they are informed about the outcomes of their decisions only infrequently. This finding is often interpreted as a manifestation of *myopic loss aversion*, a term coined by Benartzi and Thaler (1995). In this context, myopia is defined as narrow framing of decision situations which focuses on short-term consequences rather than on long-term ones. Loss aversion, one of the key constituents of prospect theory, describes people's tendency to be more sensitive to losses than to gains. According to this interpretation, if people evaluate their portfolios frequently, the probability of observing a loss is much greater than if they do so infrequently.[21] Whatever the specific experimental context, however, all these experiments share the feature that time delays were negligible. Tests of process dependence in genuinely temporal settings are still lacking.

Process dependence of risk taking was theoretically analyzed in the seminal contributions of Segal who deals with the evaluation of two-stage prospects in the domain of RDU (Segal, 1987a,b, 1990). Dillenberger (2010) provides a necessary and sufficient condition for preferences for one-shot resolution of uncertainty which holds for example in Gul (1991)'s theory of disappointment aversion, but not generally in RDU. However, we show in Appendix B.2 that this preference condition also applies to the class of resolution processes studied here.

In the domain of time discounting, a similar phenomenon of process dependence has been observed: The discounting shown over a particular delay is greater when the delay is divided into subintervals than when it is left undivided (Read, 2001; Read and Roelofsma, 2003; Ebert

[21] In these experiments subjects evaluate *sequences* of identical two-outcome lotteries over several periods where the range of potential outcomes increases with the number of periods. As we noted in Section 3.3, subproportionality does not deliver clear predictions for this class of prospects. However, Langer and Weber (2005) show that the same is true for myopic loss aversion - for specific risk profiles, myopia will not decrease but increase the attractiveness of a sequence. Blavatskyy and Pogrebna (2010) also contest the validity of the myopic loss aversion hypothesis.

and Prelec, 2007; Epper, Fehr-Duda, and Bruhin, 2009; Dohmen, Falk, Huffman, and Sunde, 2012, 2017). This regularity of subadditive discounting has usually been interpreted as a manifestation of (pure) time preferences.

Fact #5 refers to the effect of the timing of uncertainty resolution on risk taking behavior. Several experimental studies investigated people's intrinsic preferences for resolution timing. The general finding is that there are varying percentages of people with preference for early resolution, preference for late resolution and timing indifference (Chew and Ho, 1994; Ahlbrecht and Weber, 1996; Arai, 1997; Lovallo and Kahneman, 2000; Eliaz and Schotter, 2007; von Gaudecker, van Soest, and Wengström, 2011; Nielsen, 2020). Often, the percentage of people with a preference for late resolution is quite sizable (Chew and Ho, 1994; Ahlbrecht and Weber, 1996; Arai, 1997; Lovallo and Kahneman, 2000; Eliaz and Schotter, 2007; von Gaudecker, van Soest, and Wengström, 2011; Ganguly and Tasoff, 2017).[22] This finding is actually quite surprising, at least for situations when real money is at stake. Knowing early how much income to expect should always be advantageous for adapting one's consumption plans even though one might not be able to spend the money immediately.

In this context, a preference for late resolution of uncertainty can also be interpreted as an aversion to non-instrumental information. Information is non-instrumental when no further action can be taken that will change the decision maker's utility.[23] Grant, Kajii, and Polak (1998) present the following example of non-instrumental information:

"Consider, for example, the decision of whether to be tested for an incurable genetic disorder. A director of a genetic counseling program told the New York Times that there are basically two types of people. There are 'want-to-knowers' and there are 'avoiders'. There are some people who, even in the absence of being able to alter outcomes, find information of this sort beneficial. The more they know, the more their anxiety level goes down. But there are others who cope by avoiding, who would rather stay hopeful and optimistic and not have the unanswered question answered." (Grant, Kajii, and Polak (1998), page 234).

An intrinsic preference for resolution timing cannot be accommodated by EUT but is usually modeled by an additional preference parameter (Kreps and Porteus, 1978; Chew and Epstein, 1989; Grant, Kajii, and Polak, 2000). What these models cannot capture, however, is the probability dependence of timing preferences, as found by Arai (1997), for example. Epstein and Kopylov (2007)'s and Epstein (2008)'s axiomatic papers analyze resolution timing as well. According to their approach, decision makers may become more pessimistic as payoff time approaches, either due to changes in beliefs or anticipatory feelings (see also Köszegi and Rabin (2009) and Caplin and Leahy (2001)).

[22]Epstein and Zin (1991) also find a preference for late resolution of uncertainty in market data on U.S. consumption and asset returns. In line with our predictions, preference for late resolution seems to be particularly pronounced for positively skewed distributions, i.e. for prospects with small probabilities of the best outcome, and increases with time delay - a prediction that is a distinguishing feature of our model.

[23]There is a number of papers studying preference for instrumental information in non-expected utility models (see for instance Wakker (1988), Schlee (1990), Safra and Sulganik (1995). Li (2020) analyzes aversion to partial information in the context of an ambiguity averse preferences. See also the discussion of the value of information in Dillenberger (2010).

Fact #6 pertains to a number of experimental studies that report systematic effects of risk on discounting behavior: Discount rates for certain future payoffs tend to be higher than discount rates for risky future payoffs (Stevenson, 1992; Ahlbrecht and Weber, 1997; Abdellaoui, Kemel, Panin, and Vieider, 2018). Risk-dependent discounting is also evident in diminishing immediacy: People's preference for present certain outcomes over delayed ones, immediacy, weakens drastically when the outcomes become risky - they behave as if they discounted the risky reward less heavily than the original certain one (Keren and Roelofsma, 1995; Weber and Chapman, 2005; Baucells and Heukamp, 2010). This evidence motivated Halevy (2008)'s conjecture that future uncertainty might be a driver of this phenomenon.

Furthermore, the valuation of future prospects appears to be order dependent: It makes a difference whether a risky future payoff is first devalued for risk and then for delay or in the opposite order (Öncüler and Onay, 2009). When payoffs are discounted for risk first they are assigned a less favorable value than in the reverse case. Moreover, the delay-first value practically coincides with the value reported when both dimensions are accounted for in one single operation. This finding #7 can be also interpreted as a manifestation of risk dependence of discounting.

6 Concluding Remarks

We have demonstrated that our modeling approach organizes all seven stylized facts of experimental research and is also able to predict quantitatively aggregate experimental outcomes. In our view, apart from explaining the seven stylized facts uncovered by experiments, the model helps to better understand the patterns of heterogeneity in individual behaviors. Not everyone is prone to common ratio violations. In fact, almost any kind of shape of probability weighting can be found in individual estimates, and even among common-ratio violators the degree of subproportionality may vary greatly. Thus, our framework provides a host of predictions that can be investigated in future experimental research. For example, people with comparatively stronger subproportional probability weights should, *ceteris paribus*, exhibit a greater increase in risk tolerance for delayed prospects than less subproportional decision makers do. Similarly, the former group should show a greater preference for uncertainty to resolve in the future rather than in the present. Moreover, these effects are predicted to be more pronounced for positively skewed prospects - a prediction that is specific to our model. Sequential resolution of uncertainty is another area where more work needs to be done as evidence on substantially delayed prospects is still missing. Ideally, the same subjects should be exposed to the full program of experiments delineated in this paper to find out if and when our predictions materialize.

Another interesting test of the model can be based on the model's assumption that survival probability depends on time horizon according to s^t. For a given subject pool at a given point in time the preference parameters and the uncertainty perception s should not vary across time horizons. Thus, if a given subject pool faces future prospects of different delays, we should not observe a change in the estimated value of s (nor a change in estimated preference parameters) because such a change would challenge the assumption that survival probability can be represented by s^t.

The ultimate test of our model, however, is to exogenously manipulate the subjective probability that something may go wrong, s, the second crucial component of our approach aside from subproportionality. As effect sizes also depend on the perceived uncertainty of the future, such a manipulation can shed light on the question whether our model has actually identified an important causal driver of behavior.

Aside from conducting new experiments, the usefulness of our approach should be tested in the field as well. Both financial and insurance markets are fruitful areas for such an endeavor. Barberis (2013) concludes his review of 30 years of prospect theory in the following way: *"Probability weighting, [...] has drawn increasing interest in recent years. Indeed, within the risk-related areas of finance, insurance, and gambling, probability weighting plays a more central role than loss aversion and has attracted significantly more empirical support"* (page 191). Thus, our survival-risk augmented version of probability weighting could be put to the test in these fields as well. Puzzles like the maturity dependence of risk premia may appear in a new light. Another fertile application may be option prices: Polkovnichenko and Zhao (2013) show the usefulness of probability weighting for explaining option prices which could be enhanced by incorporating the maturity dimension as well.[24] Insurance markets are another domain where our approach may reconcile conflicting findings: Recognizing that risk preferences are delay dependent may help understand why people are willing to pay outrageous premiums for certain insurance contracts, such as extended warranties, and totally unwilling to take out insurance at all, such as in the health domain.

We do not claim that subproportionality plus future uncertainty are the only important drivers in the domain of risk- and time-dependent decision making. Other factors such as concave utility, intrinsically hyperbolic pure time preferences or reference dependence are also likely to play a role. However, there is accumulating evidence that risk and time preferences are intertwined and interact in systematic ways and we are just beginning to understand the factors underlying these phenomena. We have shown that subproportionality plus subjectively perceived future uncertainty provides a unifying explanation for a set of key facts – suggesting that these factors should be taken seriously in future research.

[24]We thank an anonymous referee for suggesting this application.

A Propositions and Proofs

A.1 The General m-Outcome Case

Rearranging terms in Equation 2 yields

$$\begin{aligned} V(P) &= u(x_1)w(p_1) + u(x_2)\Big(w(p_1+p_2) - w(p_1)\Big) + ... + u(x_m)\Big(1 - w(1-p_m)\Big) \\ &= \Big(u(x_1) - u(x_2)\Big)w(p_1) + ... + \Big(u(x_{m-1}) - u(x_m)\Big)w(1-p_m) + u(x_m)\,. \end{aligned} \tag{28}$$

This representation of V clarifies that x_m is effectively a sure thing whereas obtaining something better than x_m is risky.

Setting $u(\underline{x}) = 0$, the subjective present value of the prospect amounts to

$$\begin{aligned} V(\tilde{P})_0 = &\Bigg(\Big(u(x_1) - u(x_2)\Big)w(p_1 s^t) + ... \\ &... + \Big(u(x_{m-1}) - u(x_m)\Big)w\Big((1-p_m)s^t\Big) + u(x_m)w(s^t)\Bigg)\rho(t) \\ = &\Bigg(\Big(u(x_1) - u(x_2)\Big)\tfrac{w(p_1 s^t)}{w(s^t)} + ... \\ &... + \Big(u(x_{m-1}) - u(x_m)\Big)\tfrac{w\left((1-p_m)s^t\right)}{w(s^t)} + u(x_m)\Bigg)w(s^t)\rho(t)\,. \end{aligned} \tag{29}$$

From the point of view of an outsider observer, the subjective probability distribution of prospect P is not observable. Consequently, she infers probability weights $\tilde{w}$ and discount weights $\tilde{\rho}$ from observed behavior on the presumption that the decision maker evaluates the objectively given prospect P, and estimates preference parameters according to RDU in the standard way:

$$V(\tilde{P})_0 = \Bigg(\Big(u(x_1) - u(x_2)\Big)\tilde{w}(p_1) + ... + \Big(u(x_{m-1}) - u(x_m)\Big)\tilde{w}(1-p_m) + u(x_m)\Bigg)\tilde{\rho}(t)\,. \tag{30}$$

A.2 Proposition 1: Characteristics of $\tilde{w}(p)$

Given subproportionality of $w, t > 0$ and $s < 1$:

1. The function $\tilde{w}$ is a proper probability weighting function, i.e. monotonically increasing in p with $\tilde{w}(0) = 0, \tilde{w}(1) = 1$.
2. $\tilde{w}$ is subproportional.
3. $\tilde{w}$ is more elevated than w: $\tilde{w}(p) > w(p)$. The gap between $\tilde{w}(p)$ and $w(p)$ increases with
   - time delay t,
   - survival risk $1 - s$, and
   - comparatively more subproportional w.
4. The relative gap $\frac{\tilde{w}(p)}{w(p)}$ declines in p.

5. $\tilde{w}$ is less elastic than w.

6. The decision weight of the (objectively) worst possible outcome, x_m, decreases with delay t.

Proof of Proposition 1

1. Since $\tilde{w}(0) = \frac{w(0)}{w(s^t)} = 0$, $\tilde{w}(1) = \frac{w(s^t)}{w(s^t)} = 1$, and $\tilde{w}' = \frac{w'(ps^t)s^t}{w(s^t)} > 0$ hold, $\tilde{w}$ is a proper probability weighting function.

2. Subproportionality of $\tilde{w}$ follows directly from subproportionality of w as for $p > q$ and $0 < \lambda < 1$:
$$\frac{\tilde{w}(\lambda p)}{\tilde{w}(\lambda q)} = \frac{w(\lambda s^t p)}{w(\lambda s^t q)} < \frac{w(s^t p)}{w(s^t q)} = \frac{\tilde{w}(p)}{\tilde{w}(q)}\,. \tag{31}$$

3.
   - Since w is subproportional,
$$\tilde{w}(p) = \frac{w(ps^t)}{w(s^t)} > \frac{w(p)}{w(1)} = w(p) \tag{32}$$
holds for $s < 1$ and $t > 0$. Therefore, $\tilde{w}$ is more elevated than w.
   - Obviously, elevation gets progressively higher with increasing t and an equivalent effect is produced by decreasing s. Since $\tilde{w}$ increases monotonically in t and $\tilde{w} \leq 1$ for any t, elevation increases at a decreasing rate.
   - In order to show that a comparatively more subproportional probability weighting function entails a greater increase in observed risk tolerance we examine the relationship between the underlying atemporal probability weights w and observed ones $\tilde{w}$. Let w_1 and w_2 denote two probability weighting functions, with w_2 exhibiting greater subproportionality.
If $w_1(\lambda)w_1(p) = w_1(\lambda pq)$ holds for a probability $q < 1$, then $w_2(\lambda)w_2(p) < w_2(\lambda pq)$ follows as w_2 is more subproportional than w_1 (Prelec, 1998). Choose $r < 1$ such that $w_2(\lambda)w_2(p) = w_2(\lambda pqr)$. For $\lambda = s^t$, the following relationships hold:
$$\frac{\tilde{w}_1(p)}{w_1(p)} = \frac{w_1(\lambda p)}{w_1(\lambda)w_1(p)} = \frac{w_1(\lambda p)}{w_1(\lambda pq)}\,. \tag{33}$$
Applying the same logic to w_2 yields
$$\frac{\tilde{w}_2(p)}{w_2(p)} = \frac{w_2(\lambda p)}{w_2(\lambda)w_2(p)} = \frac{w_2(\lambda p)}{w_2(\lambda pqr)} > \frac{w_2(\lambda p)}{w_2(\lambda pq)}\,. \tag{34}$$
Therefore, the relative wedge $\frac{\tilde{w}_2(p)}{w_2(p)}$ caused by subproportionality is larger than the corresponding one for w_1.

4. It is straightforward to show that $\frac{\partial\left(\frac{\tilde{w}(p)}{w(p)}\right)}{\partial p} = \frac{w(ps^t)}{pw(s^t)w(p)}[\varepsilon_w(ps^t) - \varepsilon_w(p)] < 0$, as the elasticity of a subproportional w, ε_w, is increasing in p.

5. For the elasticity of $\tilde{w}$, $\varepsilon_{\tilde{w}}(p)$, the following relationship holds:

$$\varepsilon_{\tilde{w}}(p) = \frac{\tilde{w}'(p)p}{\tilde{w}(p)} = \frac{w'(ps^t)ps^t}{w(ps^t)} = \varepsilon_w(ps^t) < \varepsilon_w(p)\,, \tag{35}$$

as the elasticity ε_w increases in its argument iff w is subproportional.

6. As $\tilde{w}(p) > w(p)$ holds for any $0 < p < 1$, $\tilde{\pi}_m = 1 - \tilde{w}(1 - p_m) < 1 - w(1 - p_m) = \pi_m$ results for the decision weight of x_m. As $\tilde{w}$ increases with t, the weight of x_m declines with time delay. ■

A.3 Proposition 2: Characteristics of $\tilde{\rho}(t)$

Given subproportionality of w:

1. $\tilde{\rho}(t)$ is a proper discount function for $0 < s \leq 1$, i.e. decreasing in t, converging to zero with $t \to \infty$, and $\tilde{\rho}(0) = 1$.
2. Observed discount rates $\tilde{\eta}(t)$ are higher than the rate of pure time preference η for $s < 1$.
3. Observed discount rates decline with the length of delay for $s < 1$.
4. Greater survival risk generates a greater departure from constant discounting.
5. Comparatively more subproportional probability weighting generates a comparatively greater departure from constant discounting.

Proof of Proposition 2

1. $\tilde{\rho}(0) = w(s^0)\rho^0 = 1$. Since $w' > 0$ holds, $\frac{\partial w(s^t)}{\partial t} < 0$ and, therefore, $\tilde{\rho}' < 0$. Finally, $\lim_{t\to\infty} \tilde{\rho}(t) = 0$ (in terms of discount rates: $\lim_{t\to\infty} \tilde{\eta}(t) = \eta$).
2. Discount rates are generally defined as the rates of decline of the respective discount functions, i.e. $\eta = -\frac{\rho'(t)}{\rho(t)}$ and $\tilde{\eta}(t) = -\frac{\tilde{\rho}'(t)}{\tilde{\rho}(t)}$. Therefore,

$$\begin{aligned}
\tilde{\eta}(t) &= -\frac{\tilde{\rho}'(t)}{\tilde{\rho}(t)} \\
&= -\frac{w'(s^t)s^t \ln(s)\exp(-\eta t) - w(s^t)\exp(-\eta t)\eta}{w(s^t)\exp(-\eta t)} \\
&= -\left(\frac{w'(s^t)s^t}{w(s^t)}\ln(s) - \eta\right) \\
&= -\ln(s)\varepsilon_w(s^t) + \eta \\
&> \eta
\end{aligned} \tag{36}$$

since $\ln(s) < 0, w > 0, w' > 0$. Note that $\frac{w'(s^t)}{w(s^t)} s^t$ corresponds to the elasticity of the probability weighting function w evaluated at s^t, $\varepsilon_w(s^t)$.

3. Since the elasticity of a subproportional function is increasing in its argument, the elasticity of $w(s^t)$ is decreasing in t. Thus,

$$\tilde{\eta}'(t) = -\ln(s)\frac{\partial \varepsilon_w(s^t)}{\partial t} < 0\,. \tag{37}$$

4. In order to derive the effect of increasing survival risk, i.e. decreasing s, we examine the sensitivity of $\frac{\tilde{\rho}(t+1)}{\tilde{\rho}(t)\tilde{\rho}(1)} = \frac{w(s^{t+1})}{w(s)w(s^t)}$, which measures the departure from constant discounting between periods $t+1$ and t, with respect to changing s:

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

As $s^{t+1} < s^t < s$, $\varepsilon_w(s^{t+1}) < \varepsilon_w(s^t) < \varepsilon_w(s)$ and, hence, the sum of the elasticities in the final line of the derivation is negative. Therefore, increasing survival risk, i.e. decreasing s, entails a greater departure from constant discounting.

5. In order to examine the effect of the degree of subproportionality on decreasing impatience, 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 0 and 1:

$$\begin{aligned}
u_1(y) &= u_1(x)w_1(s)\rho \quad \text{for } 0 < y < x, \\
u_2(y') &= u_2(x')w_2(s)\rho \quad \text{for } 0 < y' < x'\,.
\end{aligned} \tag{38}$$

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

$$1 = \frac{u_1(x)w_1(s)\rho}{u_1(y)} < \frac{u_1(x)w_1(s^{t+1})\rho^{t+1}}{u_1(y)w_1(s^t)\rho^t}. \tag{39}$$

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

$$u_1(y)w_1(s^t)\rho^t = u_1(x)w_1(s^{t+1+\Delta t})\rho^{t+1}. \tag{40}$$

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.

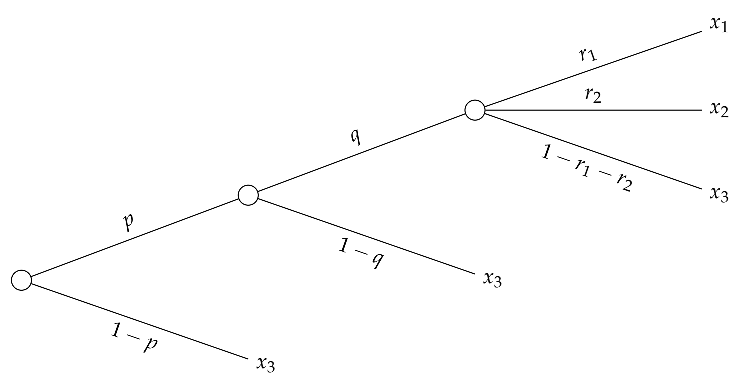
$$u_2(y')w_2(s^t)\rho^t < u_2(x')w_2(s^{t+1+\Delta t})\rho^{t+1}. \blacksquare \tag{41}$$

A.4 Folding Back of Survival Trees

In RDU, subproportional preferences are generally not sufficient to produce a preference for one-shot resolution of uncertainty. Resolution processes that can be represented by a survival tree are an exception - in this case, folding back of the tree generates compounded decision weights that are always smaller than the corresponding one-shot weights. To illustrate this result, we use an example with $n = 3$ stages and $m = 3$ outcomes, as the $n = m = 2$-case is trivial.

A survival tree is characterized by the following resolution process: At each chance node either the certain outcome materializes or the tree continues to the next stage when everything is still possible. Our example is depicted in Figure 10.

Figure 10: Survival Tree with $n = 3$ Stages and $m = 3$ Outcomes



The tree depicts the resolution of survival risk of a prospect $P = (x_1, pqr_1; x_2, pqr_2; x_3, 1 - pq(r_1 + r_2))$ in three stages.

Applying folding back, the value of the prospect is given by

$$\begin{aligned} V_3(P) =& \Big((u(x_1)-u(x_2))w(r_1)+(u(x_2)-u(x_3))w(r_1+r_2)+u(x_3)\Big)w(q)w(p) \\ &+u(x_3)(1-w(q))w(p)+1-w(p)) \\ =& \Big((u(x_1)-u(x_2))w(r_1)+(u(x_2)-u(x_3))w(r_1+r_2)+u(x_3)\Big)w(q)w(p) \\ &+u(x_3)(1-w(q)w(p)) \\ =& \Big((u(x_1)-u(x_2))w(r_1)+(u(x_2)-u(x_3))w(r_1+r_2)\Big)w(q)w(p)+u(x_3)\,. \end{aligned} \tag{42}$$

Clearly, it does not matter how many final branches the tree possesses - the formula generalizes to m outcomes in a straightforward way as the rank-dependent decision weights at the final stage get compounded with $w(p)w(q)$. The same applies if the number of stages is greater than three. If uncertainty resolves in one shot, the value of the prospect is represented by

$$V_1(P) = \Big((u(x_1)-u(x_2))w(pqr_1)+(u(x_2)-u(x_3))w(pq(r_1+r_2))+u(x_3)\,. \tag{43}$$

Subproportionality implies that $w(pqr_1) > w(q)w(p)w(r_1)$ and $w(pq(r_1+r_2)) > w(q)w(p)w(r_1+r_2)$ and, therefore, $V_1(P) > V_3(P)$. In other words, if uncertainty resolves according to a survival tree, one-shot resolution is preferred to sequential resolution.

When future uncertainty comes into play, the survival tree consists of an additional branch at each chance node, as shown in Figure 11, and the former certain outcome x_3 becomes risky as it is subjected to survival probability, here assumed to be s at each stage. The question now arises whether preference for one-shot resolution is preserved for this more complex resolution process. Recalling that $u(\underline{x}) = 0$,

$$V_3(\tilde{P}) = \Big((u(x_1)-u(x_2))w(r_1 s)+(u(x_2)-u(x_3))w((r_1+r_2)s)\Big)w(qs)w(ps)+u(x_3)(w(s))^3\,. \tag{44}$$

Its one-shot counterpart is evaluated as

$$V_1(\tilde{P}) = \Big((u(x_1)-u(x_2))w\left(pqr_1s^3\right)+(u(x_2)-u(x_3))w\left(pq(r_1+r_2)s^3\right)\Big)+u(x_3)w\left(s^3\right)\,. \tag{45}$$

Obviously, the decision weights for $V_1(\tilde{P})$ are greater than the respective ones for $V_3(\tilde{P})$. Thus, for this specific structure of uncertainty resolution, preference for one-shot resolution is preserved under subproportionality for any $n > 2$ and $m > 2$. Consequently, $\tilde{w}(pqr_1)$ is defined as

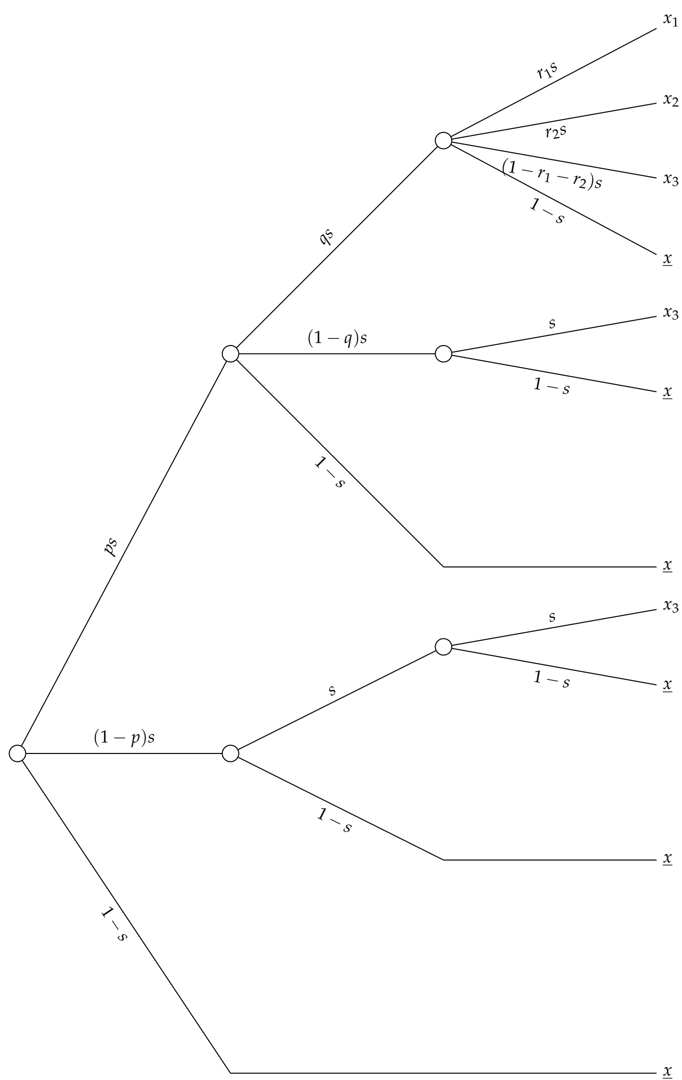
$$\tilde{w}(pqr_1) = \frac{w\left(pqr_1s^3\right)}{w\left(s^3\right)}\,, \tag{46}$$

and $\tilde{w}_3(pqr_1)$ is defined as

$$\tilde{w}_3(pqr_1) = \frac{w(ps)w(qs)w(r_1s)}{w(s)^3}\,, \tag{47}$$

which corresponds to the representation in Equation 48 where the passage of time is modeled explicitly by the partial probabilities.

Figure 11: Survival Tree with $n = 3$ Stages and $m = 3$ Outcomes with Future Uncertainty



The tree depicts the resolution of survival risk of a prospect $\tilde{P} = (x_1, pqr_1s^3; x_2, pqr_2s^3; x_3, (1-(r_1+r_2))pqs^3; \underline{x}, 1-pqs^3)$ in three stages.

These results generalize to multi-outcome prospects resolving over more than two stages if uncertainty resolves in a way analogous to the process described above: The topmost branch of the survival tree defines the path to "everything is still possible" when uncertainty resolves fully at the payment date. At each chance node along this topmost path the tree has three branches, where the two branches below the topmost one reflect the partial resolution of uncertainty of x_m contingent on its stage-by-stage prospect survival, and of $\underline{x}$, respectively. In this case, for any number of outcomes $m \geq 1$, the observed probability weights are given by

$$\tilde{w}_n(p,t) = \frac{\prod_{i=1}^{n} w\left(p^{\frac{\tau_i}{t}} s^{\tau_i}\right)}{\prod_{i=1}^{n} w\left(s^{\tau_i}\right)} = \prod_{i=1}^{n} \tilde{w}\left(p^{\frac{\tau_i}{t}}, \tau_i\right), \tag{48}$$

when the interval $[0,t]$ is partitioned into n subintervals with lengths τ_i, $i \in \{1, ..., n\}$, such that $\sum_{i=1}^{n} \tau_i = t$.

The following Proposition 3 summarizes our insights on subproportional probability weights w themselves, which drive overall prospect value, without teasing apart the separate effects on observed risk tolerance and discounting behavior. We extend these results to observed risk tolerance $\tilde{w}$ in Proposition 4. Since discount weights $\tilde{\rho}(t) = w(s^t)$ are simple probability weights themselves, Proposition 3 also speaks directly to observed discounting behavior.

Segal's work on two-stage prospects encompass the following results: For $1 > p = qr > 0$ the compounding of the respective weights always leads to lower prospect values, i.e. $w(qr) > w(q)w(r)$ holds whatever are the values of q and r. Here the order of r and q, i.e. which probability resolves first, does not play a role. Furthermore, a prospect's minimum value is attained when compounding occurs over equiprobable stages, i.e. when $r = q = \sqrt{p}$. We generalize these insights in Proposition 3.

Additionally, it can be schown that positively skewed prospects are affected more strongly by compounding of the respective probability weights: $\frac{\partial}{\partial p}\left[\frac{w(p)}{w(q)w(p/q)}\right] = \frac{w(p)}{pw(q)w(p/q)}[\varepsilon(p) - \varepsilon(p/q)] < 0$, as $p < p/q$ and the elasticity of w, ε, is increasing in p.

A.5 Proposition 3: Characteristics of $w_n(p)$

Given subproportionality of w, $s < 1$, $t > 0$, prospect risk and survival risk resolving simultaneously along a survival tree, and folding back:

1. For any number of resolution stages $n > 1$, probability weights w for one-shot resolution of uncertainty are greater than compounded probability weights for sequential resolution.
2. For a given number of resolution stages n, probability weights are smallest for evenly spaced partitions $\tau_i = \frac{t}{n} = \tau$.
3. For evenly spaced partitions, probability weights decline with the number of resolution stages n.

Proof of Proposition 3

1. Setting $q = ps^t$ or $q = s^t$, respectively, we prove by induction that $w(q) > \prod_{i=1}^{n} w(q_i)$ for probability q, $0 < q < 1$, and $q = \prod_{i=1}^{n} q_i$.
   - For $n = 2$ subproportionality implies $w(q) = w(q_1 q_2) > w(q_1)w(q_2)$.

- Assume that $w(\prod_{i=1}^{n} r_i) > \prod_{i=1}^{n} w(r_i)$ for any probabilities $0 < r_i < 1$.
- For $q = \prod_{j=1}^{n+1} q_j$ subproportionality implies

$$w(q) = w\left(q_{n+1}\prod_{i=1}^{n} q_i\right) > w(q_{n+1})w\left(\prod_{i=1}^{n} q_i\right) > w(q_{n+1})\prod_{i=1}^{n} w(q_i) = \prod_{j=1}^{n+1} w(q_j).$$

2. Without loss of generality, we reorder the sequence of subintervals such that $\tau_1 \leq \tau_2 \leq ... \leq \tau_n$. For some $i, \tau_{i-1} < \tau_i$ holds because otherwise the partition would be equally spaced right away. In this case, there exists $\varepsilon > 0$ such that $\tau_{i-1} + \varepsilon < \tau_i - \varepsilon$ is still satisfied. Due to subproportionality, the following relationship holds for $0 < q < 1$:

$$\frac{w(q^{\tau_{i-1}})}{w(q^{\tau_i - \varepsilon})} > \frac{w(q^{\tau_{i-1}}q^{\varepsilon})}{w(q^{\tau_i-\varepsilon}q^{\varepsilon})} = \frac{w(q^{\tau_{i-1}+\varepsilon})}{w(q^{\tau_i})}, \tag{49}$$

implying $w(q^{\tau_{i-1}})w(q^{\tau_i}) > w(q^{\tau_i-\varepsilon})w(q^{\tau_{i-1}+\varepsilon})$.

3. Consider two equally spaced partitions of $[0,t]$: $(\tau_i = \frac{t}{n} =: \tau)_{i=1,\dots,n}$ and $(\delta_i = \frac{t}{n-1} =: \delta)_{i=1,\dots,n-1}$. Our claim is that for $0 < p \leq 1$,

$$\prod_{i=1}^{n} w\left(p^{\frac{\tau}{t}} s^{\tau}\right) < \prod_{i=1}^{n-1} w\left(p^{\frac{\delta}{t}} s^{\delta}\right). \tag{50}$$

Setting $q = \left(p^{\frac{1}{t}} s\right)^{\frac{t}{n(n-1)}}$, we examine whether

$$\left(w\left(q^{n-1}\right)\right)^{n} < \left(w(q^{n})\right)^{n-1}. \tag{51}$$

Proceeding by induction:

- $n = 2$: Subproportionality implies $\left(w(q)\right)^2 < w\left(q^2\right)$.
- $n = 3$: Subproportionality implies $w\left(q^3\right) > \frac{\left(w(q^2)\right)^2}{w(q)}$. Thus,

$$\begin{aligned}\left(w(q^3)\right)^2 &> \frac{\left(w(q^2)\right)^2}{w(q)}\frac{\left(w(q^2)\right)^2}{w(q)} > \frac{\left(w(q^2)\right)^3 w(q^2)}{\left(w(q)\right)^2} \\ &> \frac{\left(w(q^2)\right)^3\left(w(q)\right)^2}{\left(w(q)\right)^2} = \left(w(q^2)\right)^3.\end{aligned} \tag{52}$$

- $n \to n+1$: Suppose that $\left(w(q^{n-1})\right)^{n} < \left(w(q^{n})\right)^{n-1}$ holds. Subproportionality implies

$\frac{w(q^{n-1})}{w(q^n)} > \frac{w(q^n)}{w(q^{n+1})}$. Hence,

$$\begin{aligned}\left(w(q^{n+1})\right)^n > \left(\frac{w(q^n)w(q^n)}{w(q^{n-1})}\right)^n = \frac{\left(w(q^n)\right)^{n+1}\left(w(q^n)\right)^{n-1}}{\left(w(q^{n-1})\right)^n} \\ > \frac{\left(w(q^n)\right)^{n+1}\left(w(q^{n-1})\right)^n}{\left(w(q^{n-1})\right)^n} = \left(w(q^n)\right)^{n+1}.\end{aligned} \tag{53}$$

■

Since observed risk tolerance depends on the interaction of probability weights and discount weights (subproportional probability weights themselves), it is a priori not clear whether all these characteristics carry over to observed risk tolerance. As it turns out, with one exception, the characteristics of subproportional probability weights shape observed delay-dependent risk tolerance accordingly. We enter uncharted territory with the following proposition because to our knowledge so far no experiments on the process dependence for genuinely delayed risks exist.

A.6 Proposition 4: Characteristics of $\tilde{w}_n(p)$

Given subproportionality of w, $s < 1$, $t > 0$, prospect risk and survival risk resolving simultaneously along a survival tree, and folding back:

1. For any number of resolution stages $n > 1$, risk tolerance is higher for one-shot resolution of uncertainty than for sequential resolution of uncertainty, $\tilde{w}(p,t) > \tilde{w}_n(p,t)$.
2. For a given number of resolution stages n, risk tolerance is lowest for evenly spaced partitions if the elasticity of w is concave.
3. For evenly spaced partitions, risk tolerance declines with the number of resolution stages, $\tilde{w}_n(p,t) < \tilde{w}_{n-1}(p,t)$.

Proof of Proposition 4

1. Consider Equation 48:

$$\tilde{w}_n(p,t) = \prod_{i=1}^{n} \tilde{w}\left(p^{\frac{\tau_i}{t}}, \tau_i\right).$$

Note that $\tilde{w}\left(p^{\frac{\tau_i}{t}}, \tau_i\right) = \frac{w\left(p^{\frac{\tau_i}{t}} s^{\tau_i}\right)}{w(s^{\tau_i})} < \frac{w\left(p^{\frac{\tau_i}{t}} s^{\tau_i} s^{t-\tau_i}\right)}{w\left(s^{\tau_i} s^{t-\tau_i}\right)} = \frac{w\left(p^{\frac{\tau_i}{t}} s^t\right)}{w(s^t)} = \tilde{w}\left(p^{\frac{\tau_i}{t}}, t\right).$

According to Proposition 1, $\tilde{w}\,(p,t)$ is subproportional for a fixed length of delay t and, therefore,

$$\tilde{w}_n(p,t) < \prod_{i=1}^{n} \tilde{w}\left(p^{\frac{\tau_i}{t}}, t\right) < \tilde{w}\left(\prod_{i=1}^{n} p^{\frac{\tau_i}{t}}, t\right) = \tilde{w}(p,t)\,. \tag{54}$$

2. We proceed by induction.

- Consider the case of $n = 2$ and assume that the time interval of length t is divided into two subintervals of lengths τ and $t-\tau$ with $\tau < \frac{t}{2} < t-\tau$. We compare $\tilde{w}_n$ corresponding to the evenly spaced partition $\left(\frac{t}{2}, \frac{t}{2}\right)$ with the respective $\tilde{w}_n$ for $(\tau, t-\tau)$ by examining when

$$\frac{w\left(\left(p^{\frac{1}{t}}s\right)^{\frac{t}{2}}\right) w\left(\left(p^{\frac{1}{t}}s\right)^{\frac{t}{2}}\right)}{w\left(s^{\frac{t}{2}}\right) w\left(s^{\frac{t}{2}}\right)} < \frac{w\left(\left(p^{\frac{1}{t}}s\right)^{\tau}\right) w\left(\left(p^{\frac{1}{t}}s\right)^{t-\tau}\right)}{w\left(s^{\tau}\right) w\left(s^{t-\tau}\right)}$$

  holds. Rearranging terms yields

$$\frac{w\left(\left(p^{\frac{1}{t}}s\right)^{\frac{t}{2}}\right) w\left(\left(p^{\frac{1}{t}}s\right)^{\frac{t}{2}}\right)}{w\left(\left(p^{\frac{1}{t}}s\right)^{\tau}\right) w\left(\left(p^{\frac{1}{t}}s\right)^{t-\tau}\right)} < \frac{w\left(s^{\frac{t}{2}}\right) w\left(s^{\frac{t}{2}}\right)}{w\left(s^{\tau}\right) w\left(s^{t-\tau}\right)}.$$

  Since $p^{\frac{1}{t}}s < s$ for any $0 < p < 1$, this condition amounts to requiring that $\frac{w\left(q^{\frac{t}{2}}\right)w\left(q^{\frac{t}{2}}\right)}{w(q^{\tau})w(q^{t-\tau})}$ increases in q, $0 < q < 1$. It is straightforward to show that its derivative with respect to q equals

$$\frac{\partial}{\partial q}\left(\frac{w(q^{\frac{t}{2}})w(q^{\frac{t}{2}})}{w(q^{\tau})w(q^{t-\tau})}\right) = \frac{t\left(w\left(q^{\frac{t}{2}}\right)\right)^2}{qw(q^{\tau})w(q^{t-\tau})}\left(\varepsilon_w(q^{\frac{t}{2}}) - \left(\lambda\varepsilon_w(q^{\tau}) + (1-\lambda)\varepsilon_w(q^{t-\tau})\right)\right),$$

  where $\lambda = \frac{\tau}{t}$. As $\tau < \frac{t}{2} < t-\tau$ and $\varepsilon_w\left(q^{t-\tau}\right) < \varepsilon_w\left(q^{\frac{t}{2}}\right) < \varepsilon_w\left(q^{\tau}\right)$, the term in the brackets is positive if the elasticity of w, ε_w, is a strictly concave function.

- For $n \geq 2$ the general formula for the derivative reads as

$$\frac{\left(w(q^{\frac{t}{n}})\right)^n}{q\prod_{i=1}^{n} w(q^{\tau_i})}\left(t\varepsilon_w(q^{\frac{t}{n}}) - \sum_{i=1}^{n} \tau_i\varepsilon_w(q^{\tau_i})\right),$$

  where $(\tau_i)_{i=1,\ldots,n}$ is a partition of the time interval t with $\sum_{i=1}^{n} \tau_i = t$.

- $n \to n+1$: Assume that for $t > 0$

$$t\varepsilon_w(q^{\frac{t}{n}}) - \sum_{i=1}^{n} \tau_i \varepsilon_w(q^{\tau_i}) > 0 \tag{55}$$

holds. Define a partition $(\delta_i)_{i=1,\dots,n+1}$ of t as follows:

$$\begin{aligned} \delta_i =& \frac{n\tau_i}{n+1} \quad \text{for } 1 \le i \le n \\ \delta_{n+1} =& t - \sum_{i=1}^{n} \delta_i = \frac{t}{n+1} \end{aligned}$$

Then the following relationships result:

$$\sum_{i=1}^{n+1} \delta_i \varepsilon_w \left(q^{\delta_i}\right) = \sum_{i=1}^{n} \frac{n\tau_i}{n+1} \varepsilon_w \left(q^{\frac{n\tau_i}{n+1}}\right) + \frac{t}{n+1} \varepsilon_w \left(q^{\frac{t}{n+1}}\right)$$

$$t\varepsilon_w \left(q^{\frac{t}{n+1}}\right) - \frac{t}{n+1} \varepsilon_w \left(q^{\frac{t}{n+1}}\right) = \frac{tn}{n+1} \varepsilon_w \left(q^{\frac{t}{n+1}}\right)$$

Since Equation 55 holds for any $t > 0$ and, therefore, also for $\tilde{t} = \frac{tn}{n+1}$ and $\tilde{\tau}_i = \frac{n\tau_i}{n+1}$,

$$\tilde{t}\varepsilon_w \left(q^{\frac{\tilde{t}}{n}}\right) - \sum_{i=1}^{n} \tilde{\tau}_i \varepsilon_w \left(q^{\tilde{\tau}_i}\right) > 0, \tag{56}$$

which implies

$$\frac{tn}{n+1} \varepsilon_w \left(q^{\frac{t}{n+1}}\right) - \sum_{i=1}^{n} \frac{n\tau_i}{n+1} \varepsilon_w \left(q^{\frac{n\tau_i}{n+1}}\right) > 0. \tag{57}$$

3. We examine whether $\left(\frac{w\left((p^{\frac{1}{t}}s)^{\frac{t}{n}}\right)}{w(s^{\frac{t}{n}})}\right)^n < \left(\frac{w\left((p^{\frac{1}{t}}s)^{\frac{t}{n-1}}\right)}{w(s^{\frac{t}{n-1}})}\right)^{n-1}$, which is equal to the condition that

$$\frac{\left(w\left((p^{\frac{1}{t}}s)^{\frac{t}{n}}\right)\right)^n}{\left(w\left((p^{\frac{1}{t}}s)^{\frac{t}{n-1}}\right)\right)^{n-1}} < \frac{\left(w(s^{\frac{t}{n}})\right)^n}{\left(w\left(s^{\frac{t}{n-1}}\right)\right)^{n-1}}.$$

Therefore, we examine whether the derivative of $\frac{\left(w\left(q^{\frac{t}{n}}\right)\right)^n}{\left(w\left(q^{\frac{t}{n-1}}\right)\right)^{n-1}}$ with respect to q is positive. It is straightforward to show that

$$\frac{\partial \frac{\left(w\left(q^{\frac{t}{n}}\right)\right)^n}{\left(w\left(q^{\frac{t}{n-1}}\right)\right)^{n-1}}}{\partial q} = \frac{t(w\left(q^{\frac{t}{n}}\right))^n}{q(w\left(q^{\frac{t}{n-1}}\right))^{n-1}} \left(\varepsilon_w\left(q^{\frac{t}{n}}\right) - \varepsilon_w\left(q^{\frac{t}{n-1}}\right)\right) > 0 \tag{58}$$

as the elasticity of w is increasing. ■

Contrary to the underlying probability weights w themselves, subproportionality alone does not guarantee that, for a given number of resolution stages, risk tolerance $\tilde{w}$ attains its minimum at evenly spaced partitions. The additional requirement of concavity of the elasticity of w implies that the elasticity increases more quickly for small probabilities than for large ones. While such a characteristic has not attracted any attention in the literature, there is a nice specimen of a subproportional regressive probability weighting function with concave elasticity, the so-called *neo-additive* specification

$$w(p) = \begin{cases} 0 & \text{for } p = 0 \\ \beta + \alpha p & \text{for } 0 < p < 1 \\ 1 & \text{for } p = 1 \end{cases} . \tag{59}$$

with $0 < \beta < 1, 0 < \alpha \leq 1 - \beta$. If $\beta = 0$, w is not subproportional, for $\alpha + \beta = 1$ it is not regressive. It is linear over the inner probability interval and, thus, provides an excellent approximation for the commonly used nonlinear functional forms. Since we rarely, if at all, have experimental evidence for behavior over probabilities that are extremely small or extremely large, such an approximation seems justified. This specification is also very useful for the case of ambiguity, when the probabilities are not precisely known (Chateauneuf, Eichberger, and Grant, 2007).

A.7 Proposition 5: Preferences for Resolution Timing

Given subproportionality of w, $s < 1$, $t_1 < t$, and folding back:

1. Prospects with prospect risk resolving at the time of payment t are valued more highly than prospects resolving at $t_1 < t$.
2. The wedge between late and immediate resolution, $\frac{w(ps^t)}{w(p)w(s^t)}$, declines with probability p.
3. The wedge between late and immediate resolution increases with time horizon t and survival risk $1 - s$.

Proof of Proposition 5 Without loss of generality, we set the number of outcomes $m = 2$.

1. The value of the prospect to be resolved immediately amounts to

$$\begin{aligned} &\Big(\Big(u(x_1) - u(x_2)\Big)w(p) + u(x_2)\Big)w(s^t) \\ &< \left(\Big(u(x_1) - u(x_2)\Big)\frac{w(ps^t)}{w(s^t)} + u(x_2)\right)w(s^t), \end{aligned} \tag{60}$$

as $w(ps^t) > w(p)w(s^t)$ is implied by subproportionality of w. Thus, prospects resolving at the date of payment t are valued more highly than prospects with immediate resolution.

What happens if prospect risk is not resolved immediately but rather at some later time t_1, $0 < t_1 < t$? After t_1, only survival risk remains to be resolved. In this case, the prospect's present value amounts to

$$\left(\left(u(x_1) - u(x_2)\right)\frac{w(ps^{t_1})}{w(s^{t_1})} + u(x_2)\right) w(s^{t_1})w(s^{t-t_1}) \,. \tag{61}$$

Subproportionality implies $w(p) < \frac{w(ps^{t_1})}{w(s^{t_1})} < \frac{w(ps^t)}{w(s^t)}$ and, therefore, observed risk tolerance is highest for resolution at payment time t. Moreover, the late-resolution discount weight $w(s^t) = w(s^{t_1}s^{t-t_1})$ is also greater than $w(s^{t_1})w(s^{t-t_1})$ for any earlier t_1, implying that late resolution is always preferred.

2. Examining the derivative of $\frac{w(ps^t)}{w(p)}$ with respect to p yields

$$\begin{aligned}\frac{\partial\left(\frac{w(ps^t)}{w(p)}\right)}{\partial p} &= \frac{w(ps^t)}{pw(p)}\left(\frac{w'(ps^t)ps^t}{w(ps^t)} - \frac{w'(p)p}{w(p)}\right)\\ &= \frac{w(ps^t)}{pw(p)}\left(\varepsilon_w(ps^t) - \varepsilon_w(p)\right)\\ &< 0\,,\end{aligned} \tag{62}$$

as $p > ps^t$ and the elasticity is increasing. Therefore, the wedge between late evaluation and immediate evaluation decreases with p.

3. The derivative of $\frac{w(ps^t)}{w(s^t)}$ with respect to t yields

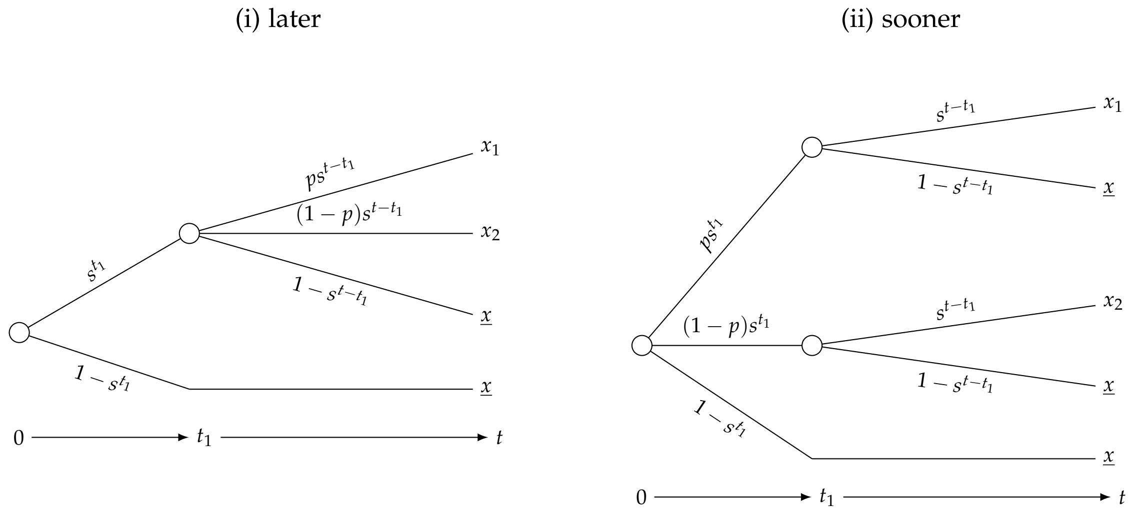
$$\begin{aligned}\frac{\partial\left(\frac{w(ps^t)}{w(s^t)}\right)}{\partial t} &= \frac{\ln(s)w(ps^t)}{w(s^t)}\left(\frac{w'(ps^t)ps^t}{w(ps^t)} - \frac{w'(s^t)s^t}{w(s^t)}\right)\\ &= \frac{\ln(s)w(ps^t)}{w(s^t)}\left(\varepsilon_w(ps^t) - \varepsilon_w(s^t)\right)\\ &> 0\,,\end{aligned} \tag{63}$$

as $\ln(s) < 0$, $s^t > ps^t$ and the elasticity is increasing. Therefore, the wedge between late and immediate evaluation increases with time horizon t and, equivalently, with survival risk $1 - s$. ■

While it is always the case that late resolution at t is preferred to any earlier resolution time t_1, we cannot ascertain that preferences for later resolution timing increase monotonically in t_1. Examining the earlier situation (Panel ii) in Figure 12, renders the prospect value (setting $\rho = 1$ again)

$$\left(u(x_1) - u(x_2)\right)w(ps^{t_1})w(s^{t-t_1}) + u(x_2)w(s^{t_1})w(s^{t-t_1})\,. \tag{64}$$

Figure 12: Later and Sooner Resolution of Prospect Risk



(1) Later: The tree depicts uncertainty resolution during the final stage. **(2) Sooner:** The probability tree shows the resolution of prospect risk after the first stage, with survival risk fully resolving at t.

We have already established that the weight of the allegedly certain outcome x_2, $w(s^{t_1})w(s^{t-t_1})$, attains its minimum value at $t_1 = t/2$. Analogously, for the risky component $ps^{t_1} = s^{t-t1}$ must hold at its minimum. Solving for t_1 yields

$$t_1^* = \frac{t}{2} - \frac{\ln(p)}{2\ln(s)}, \tag{65}$$

which lies below $\frac{t}{2}$. Regarding a simple prospect (x, p), if $t_1^* > 0$, then earlier resolution may be preferred to some later times before $\frac{t}{2}$, otherwise prospect value increases monotonically in resolution time. The latter is the case for $p \leq s^t$. For a given prospect, this condition is more likely to be met for low survival risk and/or short time horizons.

B Additional Materials

B.1 The Equivalence of Subproportionality and Increasing Elasticity

We use Prelec's (1998) definition of (strict) subproportionality: A probability weighting function $w(p)$ is subproportional if for all $1 \geq p > q > 0$ and $0 < \lambda < 1$

$$\frac{w(p)}{w(q)} > \frac{w(\lambda p)}{w(\lambda q)}. \tag{66}$$

As $p > q$, Equation 66 holds if and only if

$$\begin{aligned}\frac{w(p)}{w(\lambda p)} > \frac{w(q)}{w(\lambda q)} &\iff \frac{\partial}{\partial p}\left(\frac{w(p)}{w(\lambda p)}\right) > 0 \\ &\iff \frac{w(p)}{pw(\lambda p)}\left[\frac{w'(p)p}{w(p)} - \frac{w'(\lambda p)\lambda p}{w(\lambda p)}\right] \\ &\iff \varepsilon(p) > \varepsilon(q)\,,\end{aligned} \tag{67}$$

where ε denotes the elasticity of w, i.e. iff the elasticity of w is increasing.

B.2 A Note on Sequential Evaluation

In his Proposition 1, Dillenberger (2010) shows that, under recursivity, negative certainty independence (NCI) and a weak preference for one-shot resolution of uncertainty (PORU) are equivalent. The NCI axiom requires the following to hold: If a prospect $P = (x_1, r; x_2, 1 - r)$ is weakly preferred to a degenerate prospect $D = (y, 1)$, then mixing both with any other prospect does not result in the mixture of the degenerate prospect D being preferred to the mixture of P. This axiom is weaker than the standard independence axiom and does not put any restrictions on the reverse preference relation when a degenerate prospect is originally preferred to a non-degenerate one. The latter case characterizes the typical Allais certainty effect. NCI allows for Allais-type preference reversals but does not imply them. David Dillenberger's Proposition 3 demonstrates that NCI is generally incompatible with rank-dependent utility unless the probability weighting function is linear, i.e. unless RDU collapses to EUT. An intuitive explanation for Dillenberger's Proposition 3 is that under RDU prospect valuation is sensitive to the rank order of the outcomes and, therefore, mixtures with other prospects may affect the original rank order of outcomes in P (and D). How does Dillenberger's result relate to our claim that subproportional probability weights conjointly with recursivity imply a strong preference for one-shot resolution of uncertainty?

The crucial insight is that for the class of resolution processes studied in this paper changes in rank order do not occur and NCI is satisfied. To see this, assume that the prospect $(x_1, p; x_2, 1 - p)$, $x_1 > x_2 \geq 0$, gets resolved in two stages $\Big((x_1, r; x_2), q; (x_2, 1), 1 - q\Big)$ such that $p = qr$. In the atemporal case, when there is no additional survival risk, the two-stage prospect continues to be a strictly two-outcome one and the only relevant mixtures are those involving x_2. Suppose that $P = (x_1, r; x_2, 1 - r) \succsim (y, 1) = D$, with $x_1 > y > x_2$ and consider the following mixtures with $(x_2, 1 - \lambda)$ for some $\lambda \in (0, 1)$: $P' = (x_1, \lambda r; x_2, 1 - \lambda r)$ and $D' = (y, \lambda; x_2, 1 - \lambda)$. The following relationships hold:

$$
\begin{aligned}
P \succsim D \quad & \Rightarrow V(P) = \Big(u(x_1) - u(x_2)\Big) w(r) + u(x_2) \geq u(y) = V(D) \\
V(D') \quad & = u(y) w(\lambda) + u(x_2) \Big(1 - w(\lambda)\Big) \\
& \leq \Bigg(\Big(u(x_1) - u(x_2)\Big) w(r) + u(x_2) \Bigg) w(\lambda) + u(x_2) \Big(1 - w(\lambda)\Big) \\
& = \Big(u(x_2) - u(x_1)\Big) w(r) w(\lambda) + u(x_2) \\
& < \Big(u(x_2) - u(x_1)\Big) w(\lambda r) + u(x_2) \\
& = V(P')
\end{aligned}
\tag{68}
$$

because $w(r)w(\lambda) < w(\lambda r)$ for any $\lambda \in (0,1)$ (and hence also for $\lambda = q$) due to subproportionality of w. Consequently, for mixtures with the smaller outcome x_2, NCI, and therefore also PORU, is *strongly* satisfied. If the mixing prospect may be any arbitrary prospect, in other words if surprises are possible in the course of uncertainty resolution, this result does not hold generally. The only surprise that is still admissible is the occurrence of an outcome worse than x_2, say z. Define $P'' = \Big(x_1, \lambda r; x_2, \lambda(1-r); z, 1-\lambda\Big)$ and $D'' = (y, \lambda; z, 1-\lambda)$.

$$
\begin{aligned}
V(D'') \quad & = u(y) w(\lambda) + u(z) \Big(1 - w(\lambda)\Big) \\
& \leq \Bigg(\Big(u(x_1) - u(x_2)\Big) w(r) + u(x_2) \Bigg) w(\lambda) + u(z) \Big(1 - w(\lambda)\Big) \\
& = \Big(u(x_1) - u(x_2)\Big) w(r) w(\lambda) + \Big(u(x_2) - u(z)\Big) w(\lambda) + u(z) \\
& < \Big(u(x_1) - u(x_2)\Big) w(\lambda r) + \Big(u(x_2) - u(z)\Big) w(\lambda) + u(z) \\
& = V(P'') .
\end{aligned}
\tag{69}
$$

For $u(z) = 0$, this case is exactly the situation studied in this paper when survival risk comes into play.

B.3 Overview of Studies Used for Quantitative Assessment

Table 13: Study Overview

Fact	Study	Sample	Elicitation method	Incentives
#1	Abdellaoui, Baillon, Placido and Wakker (2011)	Study 2: 31+31 French students	certainty equivalents via bisection	31 subjects: real 31 subjects: hypothetical
#2	Epper, Fehr-Duda and Bruhin (2011)	112 Swiss students	certainty equivalents via choice lists	real
#3	Abdellaoui, Klibanoff and Placido (2015)	209 French students	certainty equivalents via choice lists	real
#4	Epper, Fehr-Duda and Bruhin (2011)	112 Swiss students	certainty equivalents via choice lists	real
#5	Arai (1997)	44 Swedish students	rating scale and choice frequencies	hypothetical
#6	Weber and Chapman (2005)	Experiment 2: 124 US students	present certainty equivalents via bisection	hypothetical
#7	Önculer and Onay (2009)	Study 1a: 39 French students	certainty/present equivalents via text field	hypothetical

B.4 Characteristics of Functional Specifications of Probability Weights

In this section we review a number of probability weighting functions that are either globally or locally subproportional. We limit our attention to functional forms with at most two parameters. Recall that subproportionality is equivalent to increasing elasticity. Consequently, if the elasticity is U-shaped, the function is supraproportional over the range of small probabilities and subproportional over large probabilities. These functions still capture the certainty effect but not necessarily general common-ratio violations. Many specifications used in the literature exhibit such a characteristic. Some experimenters found reverse common-ratio violations which require supraproportionality over the relevant probability range (see e.g. Blavatskyy (2010)). Ultimately, it is an empirical issue whether locally or globally subproportional functions fit better.

Polynomials are linear in the parameters and, thus, generally less flexible than specifications that are nonlinear in the parameters. Note that second-order polynomials demarcate the intersection of the class of quadratic utility and RDU (see also the discussion in Masatlioglu and Raymond (2016)).

Gul (1991)'s theory of disappointment aversion, for example, implies a strictly convex subproportional function in the context of RDU for two-outcome prospects. Another interesting specimen is the probability weighting function discussed in Delquié and Cillo (2006). In the context of RDU, their model of disappointment aversion generates a subproportional second-order polynomial that is equivalent to the one implied by Köszegi and Rabin (2007)'s choice-acclimating personal equilibrium, which provides an endogenous reference point (Masatlioglu and Raymond, 2016). The same polynomial also emerges in Safra and Segal (1998)'s approach to constant risk aversion. This concept captures the idea that a decision maker commits to a choice long before uncertainty is resolved, and is, therefore, particularly plausible in the context of our model. Under specific assumptions, Bordalo, Gennaioli, and Shleifer (2012) derive (discontinuous) context-dependent probability distortions from their salience theory. While their concave segment is supraproportional, the convex segment is subproportional, both of the Gul (1991) variety with $0 < \beta < 1$ and $\beta > 1$, respectively. The psychological mechanisms underlying probability weighting, therefore, often imply at least some extent of subproportionality.

An intermediate case is the constant-sensitivity specification suggested by Abdellaoui, l'Haridon, and Zank (2010) which is subproportional for large probabilities but exhibits constant elasticity for small probabilities. Thus, risk tolerance increases with delay until it hits an upper bound, staying constant afterwards. Ultimately, it is an open question whether this feature is consistent with actual behavior, which provides a fruitful avenue for future research. In particular, the associated discount function is characterized by decreasing impatience for more imminent time horizons, but constant impatience for more remote horizons. Thus, it constitutes an alternative to the quasi-hyperbolic β-δ model.

Table 14: Probability Weighting Functions

Probability weighting function w(p)	**Parameter range**	**Elasticity***	**Shape****	**Reference**
p^{α}	$\alpha > 1$	constant	convex	Luce, Mellers, and Chang (1993)
$\frac{p}{2-p}$	-	increasing	convex	Yaari (1987)
$\exp\left(-\beta(-\ln(p))^{\alpha}\right)$	$0 < \alpha < 1, \beta > 0$	increasing, concave/convex	regressive	Prelec (1998)
	$\alpha = 1, \beta > 1$	constant	convex	Prelec (1998)[1]
$\exp\left(-\frac{\beta}{\alpha}(1-p^{\alpha})\right)$	$\alpha, \beta > 0$	increasing	concave, regressive	Prelec (1998)[6]
$(1-\alpha \ln p)^{-\frac{\beta}{\alpha}}$	$\alpha, \beta > 0$	increasing	regressive	Prelec (1998)
$\frac{p^{\alpha}}{(p^{\alpha}+(1-p)^{\alpha})^{1/\alpha}}$	$0.279 < \alpha < 1$	U-shaped	regressive	Tversky and Kahneman (1992)
$\frac{\beta p^{\alpha}}{\beta p^{\alpha}+(1-p)^{\alpha}}$	$0 < \alpha < 1, \beta > 0$	U-shaped	regressive	Goldstein and Einhorn (1987)
	$0 < \alpha < 1, \beta = 1$	U-shaped	regressive	Karmarkar (1979)
	$\alpha = 1, \beta < 1$	increasing, convex	convex	Rachlin, Raineri, and Cross (1991)
		see text	see text	Bordalo, Gennaioli, and Shleifer (2012)[2]
$\frac{p+\alpha p(1-p)}{1+(\alpha+\beta)p(1-p)}$	$\alpha > 0, \beta > 0$	U-shaped	regressive	Walther (2003)
$\begin{cases} \beta^{1-\alpha}p^{\alpha} & \text{if (i) } 0 \le p \le \beta \\ 1-(1-\beta)^{1-\alpha}(1-p)^{\alpha} & \text{if (ii) } \beta < p \le 1 \end{cases}$	$0 < \alpha, \beta < 1$	(i) constant, (ii) increasing	regressive	Abdellaoui, l'Haridon, and Zank (2010)[3]
$\frac{p}{1+(1-p)\beta}$	$\beta > 1$	increasing, convex	convex	Gul (1991)
$\frac{p}{p+(1-p)\beta}$	$\beta > 1$	increasing, convex	convex	Rachlin, Raineri, and Cross (1991)
$p-\alpha p+\alpha p^2$	$0 < \alpha < 1$	increasing, concave	convex	Masatlioglu and Raymond (2016); Delquié and Cillo (2006); Safra and Segal (1998)[4]
$p+\frac{3-3\beta}{\alpha^2-\alpha+1}(\alpha p-(\alpha+1)p^2+p^3)$	$0 < \alpha, \beta < 1$	U-shaped	regressive	Rieger and Wang (2006)
$p-\alpha p(1-p)+\beta p(1-p)(1-2p)$	α depends on β	variety	variety	Blavatskyy (2014)[5]
$\begin{cases} 0 & \text{for } p = 0 \\ \beta+\alpha p & \text{for } 0 < p < 1 \\ 1 & \text{for } p = 1 \end{cases}$	$0 < \beta < 1, 0 < \alpha \le 1-\beta$	increasing, concave	regressive	Bell (1985); Cohen (1992); Chateauneuf, Eichberger, and Grant (2007)

$*$ Increasing elasticity is equivalent to *subproportionality*. $**$ An inverse-S shape means that both tails are overweighted, i.e. that the weighting function is *regressive*.

(1) Equivalent to power specification $w(p) = p^{\beta}$.

(2) The weighting function consists of a concave and a convex segment with a jump discontinuity in between (see text).

(3) For $\alpha > 1, \beta = 1$ constant elasticity, convex; for $\alpha < 1, \beta = 0$ increasing elasticity, convex.

(4) Special case of Blavatskyy (2014) with $\beta = 0$.

(5) Specific parameter constellations with $\beta > 0$ generate regressive with U-shaped elasticity.

(6) The full specification of the conditional invariant form also contains the power function (see row 1) as a special case (Prelec (1998), Proposition 4).

References

Abdellaoui, M. (2002): "A Genuine Rank-Dependent Generalization of the von Neumann-Morgenstern Expected Utility Theorem," *Econometrica*, pp. 717–736.

Abdellaoui, M., A. Baillon, L. Placido, and P. Wakker (2011): "The Rich Domain of Uncertainty: Source Functions and Their Experimental Implementation," *American Economic Review*, 101, 695–723.

Abdellaoui, M., E. Diecidue, and A. Öncüler (2011): "Risk Preferences at Different Time Periods: An Experimental Investigation," *Management Science*, 57(5), 975–987.

Abdellaoui, M., E. Kemel, A. Panin, and F. M. Vieider (2018): "Take Your Time or Take Your Chance: Time Discounting as a Distorted Probability," *Working Paper*.

Abdellaoui, M., P. Klibanoff, and L. Placido (2015): "Experiments on Compound Risk in Relation to Simple Risk and to Ambiguity," *Management Science*, 61(6), 1306–1322.

Abdellaoui, M., O. l'Haridon, and H. Zank (2010): "Separating Curvature and Elevation: A Parametric Probability Weighting Function," *Journal of Risk and Uncertainty*, 41, 39–65.

Ahlbrecht, M., and M. Weber (1996): "The Resolution of Uncertainty: An Experimental Study," *Journal of Institutional and Theoretical Economics*, 152, 593–607.

——— (1997): "An Empirical Study on Intertemporal Decision Making under Risk," *Management Science*, 43(6), 813–826.

Ainslie, G. (1991): "Derivation of 'Rational' Economic Behavior from Hyperbolic Discount Curves," *American Economic Review, Papers and Proceedings*, 81(2), 334–340.

Allais, M. (1953): "Le Comportement de L'Homme Rationnel Devant le Risque: Critique des Postulats et Axiomes de L'Ecole Americaine," *Econometrica*, 21, 503–546.

Arai, D. (1997): "Temporal Resolution of Uncertainty in Risky Choices," *Acta Psychologica*, 96, 15–26.

Aydogan, I., H. Bleichrodt, and Y. Gao (2016): "An experimental test of reduction invariance," *Journal of Mathematical Psychology*, 75, 170–182.

Baillon, A., O. l'Haridon, and L. Placido (2011): "Ambiguity Models and the Machina Paradoxes," *American Economic Review*, 101(4), 1547–1560.

Barberis, N. (2013): "Thirty Years of Prospect Theory in Economics: A Review and Assessment," *Journal of Economic Perspectives*, 27(1), 173–196.

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

——— (2012): "Probability and Time Tradeoff," *Management Science*, 58(4), 831–842.

Bell, D. (1985): "Disappointment in Decision Making under Uncertainty," *Operations Research*, 33(1), 1–27.

Bellemare, C., M. Krause, S. Kröger, and C. Zhang (2005): "Myopic Loss Aversion: Information Feedback vs. Investment Flexibility," *Economics Letters*, 87, 319–324.

Benartzi, S., and R. Thaler (1995): "Myopic Loss Aversion and the Equity Premium Puzzle," *Quarterly Journal of Economics*, 110(1), 73–92.

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

Blavatskyy, P. (2010): "Reverse Common Ratio Effect," *Journal of Risk and Uncertainty*, 40, 219–241.

Blavatskyy, P. (2014): "A Probability Weighting Function for Cumulative Prospect Theory and Mean-Gini Approach to Optimal Portfolio Investment," *mimeo*.

Blavatskyy, P., and G. Pogrebna (2010): "Reevaluating Evidence on myopic loss aversion: aggregate patterns versus individual choices," *Theory and Decision*, 68, 159–171.

Bleichrodt, H., K. I. M. Rohde, and P. Wakker (2009): "Non-Hyperbolic Time Inconsistency," *Games and Economic Behavior*, 66, 27–38.

Bommier, A. (2007): "Risk Aversion, Intertemporal Elasticity of Substitution and Correlation Aversion," *Economics Bulletin*, 4(29), 1–8.

Bordalo, P., N. Gennaioli, and A. Shleifer (2012): "Salience Theory of Choice under Risk," *Quarterly Journal of Economics*, 127(3), 1243–1285.

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

Budescu, D. V., and I. Fischer (2001): "The Same But Different: An Empirical Investigation of the Reducibility Principle," *Journal of Behavioral Decision Making*, 14, 187–206.

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

Chakraborty, A. (2021): "Present Bias," *Econometrica*, 89(4), 1921–1961.

Chakraborty, A., Y. Halevy, and K. Saito (2020): "The Relation between Behavior under Risk and over Time," *AER: Insights*, 2(1), 1–16.

Chateauneuf, A., J. Eichberger, and S. Grant (2007): "Choice Under Uncertainty With the Best and Worst in Mind: Neo-additive Capacities," *Journal of Economic Theory*, 137(1), 538–567.

Chateauneuf, A., and P. Wakker (1999): "An Axiomatization of Cumulative Prospect Theory for Decision Under Risk," *Journal of Risk and Uncertainty*.

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

Chew, S., and L. Epstein (1989): "The Structure of Preferences and Attitudes towards the Timing of the Resolution of Uncertainty," *International Economic Review*, 30(1), 103–117.

Chew, S., and J. Ho (1994): "Hope: An Empirical Study of Attitude Toward the Timing of Uncertainty Resolution," *Journal of Risk and Uncertainty*, 8, 267–288.

Chung, N., D. von Winterfeldt, and R. D. Luce (1994): "An experimental test of event commutativity in decision making under uncertainty," *Psychological Science*, 5, 394–400.

Coble, K., and J. Lusk (2010): "At the Nexus of Risk and Time Preferences: An Experimental Investigation," *Journal of Risk and Uncertainty*, 41(1), 67–79.

Cohen, M. (1992): "Security Level, Potential Level, Expected Utility: a Three-criteria Decision Model Under Risk," *Theory and Decision*, 33, 101–134.

DeJarnette, P., D. Dillenberger, D. Gottlieb, and P. Ortoleva (2020): "Time lotteries and stochastic impatience," *Econometrica*, 88(2), 619–656.

Delquié, P., and A. Cillo (2006): "Disappointment Without Prior Expectation: a Unifying Perspective on Decision Under Risk," *Journal of Risk and Uncertainty*, 33, 197–215.

Denuit, M., L. Eeckhoudt, and B. Rey (2010): "Some Consequences of Correlation Aversion in Decision Science," *Annals of Operations Research*, 176, 259–269.

Diecidue, E., and P. P. Wakker (2001): "On the Intuition of Rank-Dependent Utility," *Journal of Risk and Uncertainty*, 23(3), 281–298.

Dillenberger, D. (2010): "Preferences for One-Shot Resolution of Uncertainty," *Econometrica*, 78(6), 1973–2004.

Dillenberger, D., and U. Segal (2014): "Recursive Ambiguity and Machina's Examples," *International Economic Review*, 56, 55–61.

Dohmen, T., A. Falk, D. Huffman, and U. Sunde (2012): "Interpreting Time Horizon Effects in Inter-Temporal Choice," *IZA Discussion Paper No. 6385*.

Dohmen, T., A. Falk, D. Huffman, and U. Sunde (2017): "The Robustness and Pervasiveness of Sub-Additivity in Intertemporal Choice," *Working Paper*.

Ebert, J., and D. Prelec (2007): "The Fragility of Time: Time-Insensitivity and Valuation of the Near and Far Future," *Management Science*, 53(9), 1423–1438.

Eisenbach, T. M., and M. C. Schmalz (2016): "Anxiety in the Face of Risk," *Journal of Financial Economics*, 121, 414–426.

Eliaz, K., and A. Schotter (2007): "Experimental Testing of Intrinsic Preferences for NonInstrumental Information," *American Economic Review, Papers and Proceedings*, 97(2), 166–169.

Epper, T., and H. Fehr-Duda (2015): "Comment on "Risk Preferences Are Not Time Preferences": Balancing on a Budget Line," *American Economic Review*, 105(7), 2261–2271.

Epper, T., H. Fehr-Duda, and A. Bruhin (2009): "Uncertainty Breeds Decreasing Impatience: The Role of Risk Preferences in Time Discounting," *Working Paper, Institute for Empirical Research in Economics, University of Zurich*, 412.

——— (2011): "Viewing the Future through a Warped Lens: Why Uncertainty Generates Hyperbolic Discounting," *Journal of Risk and Uncertainty*, 43(3), 169–203.

Epstein, L., and S. Tanny (1980): "Increasing Generalized Correlation: a Definition and Some Economic Consequences," *Canadian Journal of Economics*, 13(1), 16–34.

Epstein, L., and S. Zin (1991): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis," *Journal of Political Economy*, 99(2), 263–286.

Epstein, L. G. (2008): "Living with Risk," *Review of Economic Studies*, 75, 1121–1141.

Epstein, L. G., and I. Kopylov (2007): "Cold Feet," *Theoretical Economics*, 2, 231–259.

Ericson, K. M., and D. Laibson (2019): "Intertemporal Choice," *Bernheim, D., Laibson, D., DellaVigna, S.: Handbook of Behavioral Economics - Foundations and Applications 1, Elsevier*, 2, 1–67.

Fan, Y., D. V. Budescu, and E. Diecidue (2018): "Decisions With Compound Lotteries," *Decision*, Advance online publication(http://dx.doi.org/10.1037/dec0000091), 1–25.

Fehr-Duda, H., and T. Epper (2012): "Probability and Risk: Foundations and Economic Implications of Probability-Dependent Risk Preferences," *Annual Review of Economics*, 4, 567–593.

Frederick, S., G. F. Loewenstein, and T. O'Donoghue (2002): "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature*, 40(2), 351–401.

Gabaix, X., and D. Laibson (2017): "Myopia and Discounting," *mimeo*.

Ganguly, A., and J. Tasoff (2017): "Fantasy and Dread: The Demand for Information and the Consumption Utility of the Future," *Management Science*, 63(12), 4037–4060.

Gneezy, U., A. Kapteyn, and J. Potters (2003): "Evaluation Periods and Asset Prices in a Market Experiment," *Journal of Finance*, LVIII(2), 821–837.

Gneezy, U., and J. Potters (1997): "An Experiment on Risk Taking and Evaluation Periods," *Quarterly Journal of Economics*, 112(2), 631–645.

Goldstein, W., and H. Einhorn (1987): "Expression Theory and the Preference Reversal Phenomena," *Psychological Review*, 94, 236–254.

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

Grant, S., A. Kajii, and B. Polak (1998): "Intrinsic Preference for Information," *Journal of Economic Theory*, 83, 233–259.

——— (2000): "Temporal Resolution of Uncertainty and Recursive Non-expected Utility Models," *Econometrica*, 68(2), 425–434.

Gul, F. (1991): "A Theory of Disappointment Aversion," *Econometrica*, 59(3), 667–686.

Haigh, M., and J. List (2005): "Do Professional Traders Exhibit Myopic Loss Aversion? An Experimental Analysis," *Journal of Finance*, 60(1), 523–534.

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

——— (2015): "Time Consistency: Stationarity and Time Invariance," *Econometrica*, 83(1), 335–352.

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

Jones, E. E., and C. A. Johnson (1973): "Delay of Consequences and the Riskiness of Decisions," *Journal of Personality*, 41(4), 613–637.

Kahneman, D., and A. Tversky (1979): "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47(2), 263–292.

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.

Kihlstrom, R. E., and L. J. Mirman (1974): "Risk Aversion with Many Commodities," *Journal of Economic Theory*, 8, 361–388.

Köszegi, B., and M. Rabin (2007): "Reference-Dependent Risk Attitudes," *American Economic Review*, 97(4), 1047–1073.

——— (2009): “Reference-Dependent Consumption Plans,” *American Economic Review*, 99(3), 909–936.

Kőszegi, B., and A. Szeidl (2013): “A Model of Focusing in Economic Choice,” *The Quarterly Journal of Economics*, 128(1), 53–104.

Kreps, D., and E. Porteus (1978): “Temporal Resolution of Uncertainty and Dynamic Choice Theory,” *Econometrica*, 46(1), 185–200.

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

Langer, T., and M. Weber (2005): “Myopic Prospect Theory vs. Myopic Loss Aversion: How General is the Phenomenon?,” *Journal of Economic Behavior and Organization*, 56, 25–38.

Leland, J., and M. Schneider (2017): “Risk Preference, Time Preference, and Salience Perception,” *mimeo*.

Li, J. (2020): “Preferences for Partial Information and Ambiguity,” *Theoretical Economics*, 15(3), 1059–1094.

Loewenstein, G. F., and R. Thaler (1989): “Anomalies: Intertemporal Choice,” *Journal of Economic Perspectives*, 3(4), 181–193.

Loomes, G. (2010): “Modelling Choice and Valuation in Decision Experiments,” *Psychological Review*, 117(3), 902–924.

Loomes, G., and R. Sugden (1986): “Disappointment and Dynamic Consistency in Choice under Uncertainty,” *Review of Economic Studies*, 53(2), 271–282.

——— (1987): “Testing for Regret and Disappointment in Choice Under Uncertainty,” *Economic Journal*, 97, 118–129.

Lovallo, D., and D. Kahneman (2000): “Living with Uncertainty: Attractiveness and Resolution Timing,” *Journal of Behavioral Decision Making*, 13, 179–190.

Luce, R. D., B. A. Mellers, and S.-J. Chang (1993): “Is Choice the Correct Primitive? On Using Certainty Equivalents and Reference Levels to Predict Choices Among Gambles,” *Journal of Risk and Uncertainty*, 6, 115–143.

Machina, M. (2009): “Risk, Ambiguity, and the Rank-Dependence Axioms,” *American Economic Review*, 99, 385–392.

——— (2014): “Ambiguity Aversion with Three or More Outcomes,” *American Economic Review*, 104(12), 3814–3840.

Masatlioglu, Y., and C. Raymond (2016): "A Behavioral Analysis of Stochastic Reference Dependence," *American Economic Review*, 106(9), 2760–2782.

Nebout, A., and D. Dubois (2014): "When Allais Meets Ulysses: Dynamic Axioms and the Common Ratio Effect," *Journal of Risk and Uncertainty*, 48(1), 19–49.

Nielsen, K. (2020): "Preferences for the resolution of uncertainty and the timing of information," *Journal of Economic Theory*, 189(105090), 1–39.

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

O'Donoghue, T., and J. Somerville (2018): "Modeling Risk Aversion in Economics," *Journal of Economic Perspectives*, 32(2), 91–114.

Öncüler, A., and S. Onay (2009): "How Do We Evaluate Future Gambles? Experimental Evidence on Path Dependency in Risky Intertemporal Choice," *Journal of Behavioral Decision Making*, 22, 280–300.

Polkovnichenko, V., and F. Zhao (2013): "Probability Weighting Functions Implied in Options Prices," *Journal of Financial Economics*, 107, 580–609.

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

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

Preston, M. G., and P. Baratta (1948): "An Experimental Study of the Auction-Value of an Uncertain Outcome," *American Journal of Psychology*.

Quiggin, J. (1982): "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization*, 3, 323–343.

Quiggin, J., and J. Horowitz (1995): "Time and Risk," *Journal of Risk and Uncertainty*, 10, 37–55.

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

Read, D. (2001): "Is Time-Discounting Hyperbolic or Subadditive?," *Journal of Risk and Uncertainty*, 23(1), 5–32.

Read, D., and P. Roelofsma (2003): "Subadditive versus Hyperbolic Discounting: A Comparison of Choice and Matching," *Organizational Behavior and Human Decision Processes*, 91, 140–153.

Richard, S. F. (1975): "Multivariate Risk Aversion, Utility Independence and Separable Utility Functions," *Management Science*, 22(1), 12–21.

Rieger, M., and M. Wang (2006): "Cumulative Prospect Theory and the St. Petersburg Paradox," *Economic Theory*, 28, 665–679.

Rottenstreich, Y., and C. K. Hsee (2001): "Money, Kisses and Electric Shocks: On the Affective Psychology of Risk," *Psychological Science*, 12(3), 185–190.

Rubinstein, A. (1988): "Similarity and Decision-Making Under Risk (Is There a Utility Theory Resolution to the Allais Paradox?)," *Journal of Economic Theory*, 46, 145–153.

Safra, Z., and U. Segal (1998): "Constant Risk Aversion," *Journal of Economic Theory*, 83, 19–42.

Safra, Z., and E. Sulganik (1995): "Schur Convexity, Quasi-convexity and Preference for Early Resolution of Uncertainty," *Theory and Decision*, 39, 213–218.

Sagristano, M., Y. Trope, and N. Liberman (2002): "Time-dependent Gambling: Odds Now, Money Later," *Journal of Experimental Psychology*, 131(3), 364–376.

Saito, K. (2011): "Strotz Meets Allais: Diminishing Impatience and the Certainty Effect: Comment," *American Economic Review*, 101(5), 2271–2275.

Schlee, E. (1990): "The Value of Information in Anticipated Utility," *Journal of Risk and Uncertainty*, 3, 83–92.

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

——— (1987b): "The Ellsberg Paradox and Risk Aversion: An Anticipated Utility Approach," *International Economic Review*, 28(1), 175–202.

——— (1990): "Two-Stage Lotteries Without the Reduction Axiom," *Econometrica*, 58(2), 349–377.

Shelley, M. (1994): "Gain/Loss Asymmetry in Risky Intertemporal Choice," *Organizational Behavior and Human Decision Processes*, 59, 124–159.

Starmer, C. (2000): "Developments in Non-Expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk," *Journal of Economic Literature*, 38(2), 332–382.

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.

Stevenson, M. (1992): "The Impact of Temporal Context and Risk on the Judged Value of Future Outcomes," *Organizational Behavior and Human Decision Processes*, 52, 455–491.

Sugden, R. (2004): "Alternatives to Expected Utility: Foundations," in *Handbook of Utility Theory: Volume II: Extensions*, pp. 685–755. Kluwer, Dordrecht.

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

Thaler, R. H., A. Tversky, D. Kahneman, and A. Schwartz (1997): "The Effect of Myopia and Loss Aversion on Risk Taking: An Experimental Test," *Quarterly Journal of Economics*, 112(2), 647–661.

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

van Binsbergen, J., M. Brandt, and R. Koijen (2012): "On the Timing and Pricing of Dividends," *American Economic Review*, 102(4), 1596–1618.

van de Kuilen, G., and P. Wakker (2011): "The Midweight Method to Measure Attitudes Toward Risk and Ambiguity," *Management Science*, 57(3), 582–598.

von Gaudecker, H.-M., A. van Soest, and E. Wengström (2011): "Heterogeneity in Risky Choice Behaviour in a Broad Population," *American Economic Review*, 101(2), 664–694.

Wakker, P. (1988): "Nonexpected Utility as Aversion of Information," *Journal of Behavioral Decision Making*, 1, 169–175.

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

——— (2010): *Prospect Theory for Risk and Ambiguity*. Cambridge University Press, Cambridge.

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

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.

Yaari, M. (1987): "The Dual Theory of Choice under Risk," *Econometrica*, 55(1), 95–115.