A TALE OF TWO TAILS: On the Coexistence of Overweighting and Underweighting of Rare Extreme Events

Thomas Epper* Helga Fehr-Duda†

October 17, 2016

Abstract

Almost all important decisions in people's lives entail risky consequences. In many situations people display considerable risk aversion, apparently overweighting rare extreme events such as airplane and stock market crashes. However, in other situations, concerning for example natural hazards, the opposite is the case. So far, no satisfactory preference-based explanation of the coexistence of over- and underweighting of rare extreme events has emerged. Here we argue that the timing of the consequences and of uncertainty resolution are crucial for understanding these phenomena. We show that future uncertainty conjointly with people's proneness to probability distortions generates a unifying framework for explaining the coexistence of over- and underweighting of rare extreme events.

JEL Classification: D01, D81, D91

*University of St. Gallen, School of Economics and Political Science, Varnbüelstrasse 19, 9000 St. Gallen, Switzerland. Email: thomas.epper@unisg.ch.

†University of Zurich, Department of Banking and Finance, Plattenstrasse 32, 8032 Zurich, Switzerland. Email: helga.fehr@bf.uzh.ch.

1 Introduction

Whatever the nature of our decisions, may they concern health, wealth, love or education, hardly ever can we be sure of their outcomes. Thus, it is an important task for economists to understand, model and predict decisions under risk. However, it seems difficult to paint a coherent picture of people's risk preferences because in some situations their behaviors appear to be extremely risk averse while in others the opposite is the case. For example, many consumers purchase extended warranties for household appliances at exorbitant prices, i.e. they display extreme risk aversion in situations that involve comparatively low stakes (Cicchetti and Dubin, 1994; Huysentruyt and Read, 2010). According to the standard workhorse of economics, expected utility theory, consumers should be approximately risk neutral in this case (Loomes and Segal, 1994). On the other hand, many are reluctant to buy adequate life insurance thereby exposing their loved ones to considerable poverty risk (Bernheim, Forni, Gokhale, and Kotlikoff, 2003; Cutler, Finkelstein, and McGarry, 2008). Similarly, stock market participation is very low in many countries around the globe (Giannetti and Koskinen, 2010), whereas inhabitants of disaster-prone areas are often not willing to take out highly subsidized insurance even though not only their wealth but also their lives are at stake (Kunreuther, 1984; Viscusi, 2010).

These disparities can be understood in terms of how tail events, i.e. rare extreme events, are evaluated. For example, paying a multiple of expected losses for extended warranties is consistent with the overweighting of an improbable appliance breakdown. Analogously, overweighting the rare event of a stock market crash makes people shy away from investing in stocks. In both cases, therefore, overweighting of rare extreme events seems to govern behavior. Underinsurance, as apparent in life and disaster insurance choices, is consistent with underweighting of rare extreme events, which raises the question how these opposite tendencies can be rationalized.

Explanations of the overweighting of tail events center on rank-dependent models, such as Rank Dependent Utility Theory (RDU; Quiggin (1982)) and Cumulative Prospect Theory (CPT; Tversky and Kahneman (1992))[1], which feature decision weights that depend on the rank of the possible outcomes. As explained in detail below, these decision weights are constructed from a probability weighting function by a cumulative procedure. On average, relative to the objective probabilities, decision weights tend to overweight the best and the worst outcomes, whereas intermediate outcomes tend to be underweighted (Fehr-Duda and Epper, 2012). This common pattern of overweighting of both tails has an intuitive interpretation: The decision makers' attention is drawn primarily to the extreme possible outcomes, termed by Lopes (1987) "the psychology of hope and fear".

However, underweighting of rare extreme events, which seems to govern disaster and life insurance choices, seems to contradict such a decision-weight based explanation. In their original paper on Prospect Theory, Kahneman and Tversky (1979) surmise that highly unlikely events are

[1] For recent reviews of the usefulness of probability weighting see Fehr-Duda and Epper (2012) and Barberis (2013b)).

either overweighted or simply ignored because people are limited in their ability to comprehend and evaluate extreme probabilities.[2] Of course, one could also argue that rare extreme events are underweighted because people are not aware of their existence. But many insufficiently insured people live in disaster-prone areas, even in so-called red zones (Barnes, 2011). Recently, for example, an earthquake in Amatrice, located in a notoriously earthquake-prone area of central Italy, caused 300 deaths and made many more homeless. In 2009, a similar disaster occurred in the same region only 50 km from Amatrice. Unawareness of the possibility of another earthquake in the region is a highly unlikely explanation. So why was Amatrice hit so unprepared? Therefore, as Barberis (2013a) has recently noted, we need a better understanding of *why* people over- and underweight tail events in their decision making.

In this paper we argue that the coexistence of overweighting and underweighting of tail events can be explained by recognizing that risk taking behavior is largely driven by the timing and the path of uncertainty resolution. Compare, for example, two different life insurance products, a regular life insurance policy (either a term policy or a permanent one), and a flight insurance policy that covers the same event but expires immediately after the flight. Concerning regular insurance policies, there is no doubt about the fact that a substantial percentage of the U.S. population is underinsured, which can be rationalized with the underweighting of the event of premature death. However, flight insurance policies used to be extremely popular in the 1950's and '60s when they were sold at vending machines at airports. Many passengers were willing to pay outrageous premiums, obviously overweighting the rare event of an airplane crash (Fehr-Duda and Fehr, 2016). The important difference between these two products is their maturity. A regular policy extends, in principle, over a long time horizon, whereas flight insurance is very short term. Therefore, the time when uncertainty is perceived to resolve seems to play a crucial role in the decision to take out life insurance.

Taking this observation as our starting point, our approach relies on two basic insights. First, it seems indisputable that there is uncertainty attached to any future prospect as only immediate consequences can be totally certain. Something unrelated to the prospect under consideration may go wrong before outcomes materialize and reduce the chances of actually obtaining the expected outcomes. For example, important documents may go lost or appointments may not be kept because of illness. If the probability that something goes wrong is perceived to increase with the length of delay, people's risk tolerance will be affected by the length of delay as well. Second, in which way this delay dependence manifests itself is contingent on people's risk preferences, in particular on the specific characteristics of probability weighting. The representative probability weighting curve has been shown to display two key features, regressiveness and subproportionality. Regressiveness entails that small probabilities of the best possible outcome are overweighted

[2] Kahneman and Tversky (1979) go on to argue that, consequently, the probability weighting function is not well-behaved near the end-points. This argument does not appear in the cumulative version of Prospect Theory in Tversky and Kahneman (1992) any more.

and large probabilities are underweighted. As already noted above, this feature of probability weighting implies that both tails of the outcome distribution are overweighted and, hence,it can explain why people simultaneously engage in gambling and insuring, why they favor positively skewed distributions and dislike negatively skewed ones. Skewness preferences play an important role in finance, e.g. in rationalizing the cross-section of asset prices, the underdiversification of households, and account for other phenomena such as betting on long shots (Barberis and Huang, 2008; Snowberg and Wolfers, 2011; De Giorgi and Legg, 2012; Polkovnichenko and Zhao, 2013). The second important characteristic of probability weighting is subproportionality, which accommodates the famous Allais common-ratio paradox (Allais, 1953). Subproportionality of the probability weighting function is equivalent to its elasticity increasing with probability: Loosely speaking, reducing the chances of the best possible outcome hurts more when the chances are high than when they are low. This feature encompasses the certainty effect, people's tendency to overreact to the loss of certainty.

Regressiveness and subproportionality can be interpreted as people's reactions to anticipated emotions when uncertainty will resolve. Regressiveness maps emotions of elation and disappointment: Elation arises when the best possible outcome materializes in spite of an ex-ante low probability. Disappointment is anticipated to set in when the best possible outcome fails to materialize in the case of an ex-ante high probability. Subproportionality measures the strength of these emotions: The higher the degree of subproportionality, the more pronounced is the departure from linear weighting and, consequently, the reactions to elation and disappointment.

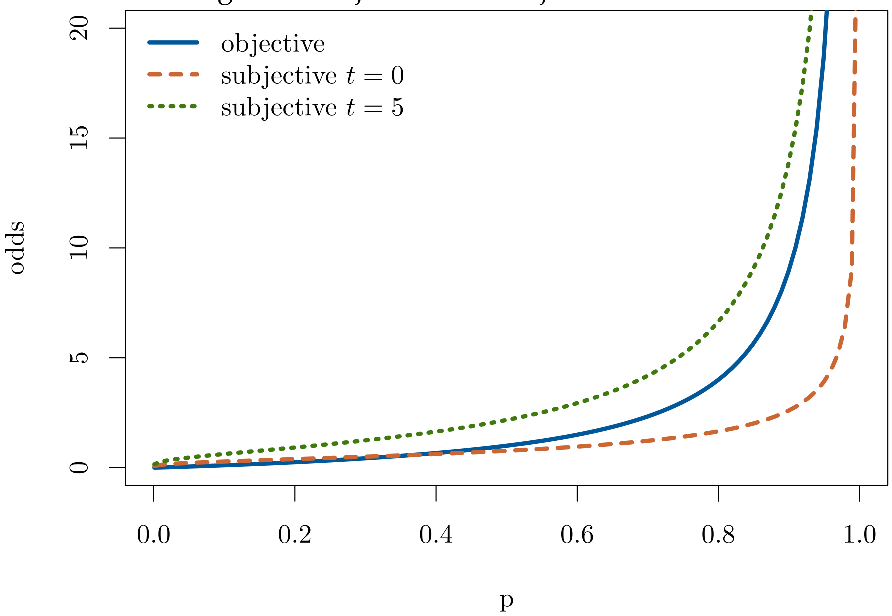
Together with future uncertainty, regressiveness and subproportionality have crucial consequences for the evaluation of tail events. We demonstrate that risk tolerance *increases with the length of delay* until outcomes materialize. To understand why, consider an insurance problem described by a two-outcome prospect $(x_1, p; x_2, 1-p)$, where $x_1 > x_2$ and x_2 is an adverse event which materializes with a small probability $1-p$. In other words, the outcome distribution has a left tail. When uncertainty resolves with a negligible time delay, the regressiveness of probability weighting implies that the adverse event is overweighted and the favorable event is underweighted, which makes it likely that the decision maker takes out insurance. However, when uncertainty resolves in the more remote future, the decision maker does not take the probabilities of x_1 and x_2 at face value because in his mind something may go wrong which decreases the chances of the prospect materializing. Regressiveness of the probability weighting function now has the following effect: As something even worse may happen, x_2 loses its extreme quality and turns into an intermediate outcome, which generally is underweighted relative to its objective probability. Consequently, the weight of x_2 will decline dramatically, which considerably decreases the attractiveness of buying insurance.

When uncertainty resolves in the remote future, the effects of subproportionality come into play as well. As the probability that the prospect will actually materialize declines with delay, the decision weight of the favorable - originally underweighted - outcome x_1 declines, too, but reacts progressively less strongly to this decline in probability since the elasticity of the probability weighting function decreases with decreasing probabilities. Hence, the ratio between the decision weights of the best outcome and the worst outcome changes continuously in favor of the best outcome. Interpreted within the context of emotions, people fear imminent disappointment much more than disappointment occurring in the remote future. Consequently, insurance contracts with short maturities are much more attractive than insurance contracts with long maturities.

Figure 1 illustrates the effects of future uncertainty by comparing the objective odds of the best outcome x_1 versus x_2, $p : (1-p)$, with the subjective odds $\pi_1(t) : \pi_2(t)$, where $\pi_1(t)$ and $\pi_2(t)$ are the corresponding decision weights, derived from a standard probability weighting function, and t denotes the length of delay. The curve in the middle corresponds to the objective odds, the curve below depicts the subjective odds when there only a negligible time delay. This curve reflects the strong propensity to take out insurance as the odds for the favorable outcome are underweighted at high levels of p (the situation we are focusing on). However, when uncertainty resolves in the future, the subjective odds are much more optimistic than the objective ones, which explains why people may be reluctant to buy insurance when uncertainty resolves in the future.

Our analysis begs the question why people shy away from investing in equities. Isn't the risk of a stock market crash a rare extreme event comparable with a natural disaster? And should we not expect high risk tolerance then? In our view, there is a crucial difference between these two types of events. Take, for example, natural disasters such as earthquakes and tsunamis. Rarely can their timing be predicted long before their actual occurrence. They literally appear out of the blue. In these cases, uncertainty resolves in one shot at some unknown time in the future, which leads to a considerable underweighting of these events. The opposite process is at work in the stock market. Information on asset prices is readily available, for many assets even in real time. Therefore, notwithstanding the longterm nature of many investments, uncertainty is perceived to resolve gradually over the course of time. One can watch price bubbles building up but unfortunately not tectonic plates shifting. Therefore, the time horizon relevant for risk taking behavior is much shorter for investments in the stock market than for disaster insurance. Moreover, we show that, due to the compounding of subproportional decision weights, prospect valuation tends to become extremely pessimistic when uncertainty resolves gradually. Our finding is reminiscent of *myopic loss aversion* (Benartzi and Thaler, 1995; Barberis, Huang, and Thaler, 2006), which makes people pronouncedly risk averse for short time horizons. Contrary to myopic loss aversion, *myopic probability weighting* is a general phenomenon that emerges independently of the location of the reference point. Overall, therefore, we predict that risk tolerance is high for prospects with long delays when uncertainty resolves in one shot. Gradual resolution of uncer-

Figure 1: Objective vs. Subjective Odds



The figure compares objective to subjective odds for two different time delays. The **solid curve** depicts the objective odds $p : (1-p)$. When the time delay is negligible, $t = 0$, the subjective odds $\pi_1(t) : \pi_2(t)$ lie below the objective odds for medium and large probabilities **(dashed curve)**. However, when uncertainty resolves in the future, here at $t = 5$, the subjective odds lie above the objective odds for the entire range of probabilities **(dotted curve)**. For construction of the figure, we set $s = 0.9$ and $w(p) = \exp(-(-\ln(p))^{0.5})$ (Prelec, 1998).

tainty, on the other hand, counteracts the otherwise risk-tolerance increasing effect of long time delays and may lead to pronounced risk aversion.

To the best of our knowledge, we are the first to present a preference-based explanation of the coexistence of overweighting and underweighting of rare extreme events, thereby providing a rationale for the observed variations in real-world risk tolerance.[3] We are not the first to acknowledge that *"[a]nything that is delayed is almost by definition uncertain"* (Prelec and Loewenstein (1991), p.784). Consequently, the focus of previous research has been on the implications of future uncertainty for discounting behavior (Sozou, 1998; Dasgupta and Maskin, 2005; Bommier, 2006; Halevy, 2008; Walther, 2010; Pennesi, 2015). Many contributions investigate risk taking in atemporal settings some of which are related to our work: Quiggin (2003) studies the consequences of background risk for generalized expected utility models, Segal (1987a,b, 1990) deals with the relationship between subproportionality and two-stage lotteries, and Dillenberger (2010) analyzes

[3]Moreover, our model delivers a unifying perspective on seemingly unrelated phenomena discovered by experimental research, such as the preference for late resolution of uncertainty, hyperbolic and subadditive discounting, the differential discounting of certain and risky prospects, and the order-dependence of prospect valuation. It can also reconcile the magnitude effect in discounting with the magnitude effect in risk taking. Refer to Epper and Fehr-Duda (2015).

preferences for one-shot resolution of uncertainty. Concerning interactions of time and risk, Baucells and Heukamp (2012) present an axiomatic model for the domain of simple prospects with only one non-zero outcome. However, none of these contributions can address the coexistence of underweighting and overweighting of tail events.

The remainder of the paper is organized as follows: The key assumptions of our model and their implications for general multi-outcome prospects are discussed in Sections 2 and 3. Model predictions are presented in Section 4. Finally, Section 5 concludes. Supplementary materials are available in the appendix where we also show that our results developed for decision under risk are portable to situations when decision makers do not know the probabilities with precision.

2 Key Assumptions

Our model builds on two basic ideas: First, there is risk attached to any future prospect. Second, people are prone to probability distortions. The risk inherent in the future, *survival risk* for short, may stem from different sources. At the personal level, it refers to a general feeling of "something may go wrong" due to unexpected contingencies, such as a check getting lost in the mail or the involvement in an accident. Another important channel through which survival risk may manifest itself is the institutional environment. Environments where property rights are only weakly protected or institutions of contract enforcement are not reliable, as is the case in many developing countries, are characterized by high survival risk. This risk turns allegedly guaranteed payoffs into risky ones and introduces an additional layer of risk over and above the objective atemporal probability distributions of risky payoffs (henceforth referred to as *base risk*). Consequently, there are two distinct types of risk, time-independent *base risk* and time-dependent *survival risk*. We model the probability of prospect survival by a constant per-period rate s. Thus, survival risk at time delay t amounts to $1 - s^t$.

The second pillar of our model concerns the characteristics of risk preferences. Abundant empirical evidence has demonstrated that risk taking behavior depends nonlinearly on the probabilities. Since our main concern is the overweighting and underweighting of tails events, we make use of the characteristics of rank-dependent models.[4] The starting point of our approach is Rank Dependent Utility Theory (RDU).[5] We assume that a decision maker's atemporal risk preferences over prospects that are played out and paid out with negligible time delay can be

[4]For an insightful discussion on the intuition of rank-dependent models see Diecidue and Wakker (2001).

[5]RDU is a generalization of expected utility theory and, thus, tacitly also assumes asset integration. While reference dependence, modeled e.g. by CPT, may be an important additional feature of risk taking behavior, it does not play a role in explaining over- and underweighting of rare extreme events. RDU has several attractive features. First, RDU respects completeness, transitivity, continuity, and first-order stochastic dominance. Moreover, RDU displays first-order attitudes toward risk, i.e. preferences between prospects the consequences of which are sufficiently close to one another do not necessarily tend to risk neutrality. In this sense, experimental evidence favors rank-dependent utility theory over many other non-expected utility approaches that only permit second-order risk aversion (Sugden, 2004). RDU is also able to accommodate correlation aversion (Fehr-Duda and Epper, 2012).

represented by a rank-dependent functional. Consider a prospect $P = (x_1, p_1; ...; x_m, p_m)$ over (terminal) monetary outcomes $x_1 > x_2 > ... > x_m \geq 0$ with $\Sigma p_i = 1$. u measures the utility 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 and to satisfy $w(0) = 0$ and $w(1) = 1$. Decision weights π_i are defined as[6]

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

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. Finally, the prospect's value is represented by

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

On average, empirical probability weighting curves are regressive, overweighting small probabilities and underweighting large probabilities (Bruhin, Fehr-Duda, and Epper, 2010), which is also a common pattern in individual data (Gonzalez and Wu, 1999):[7] A probability weighting function $w(p)$ is regressive if there exists a probability $p^* \in (0, 1)$, such that

$$\begin{array}{ll} w(p) > p & \text{for } p < p^* \\ w(p) = p^* & \text{for } p = p^* \\ w(p) < p & \text{for } p > p^* \end{array} . \tag{3}$$

In the context of rank-dependent models, regressiveness of the probability weighting function generates overweighting of a prospect's extreme outcomes and underweighting of its intermediate outcomes, which nicely captures the notion that more extreme outcomes within a given prospect are more salient (see Figure 2 in Section 4.1). Specifications of functional forms for w typically show a combination of concavity over small probabilities and convexity over large probabilities, i.e. an inverse S-shape, which is a slightly stronger requirement than regressiveness.

To see why a regressive probability weighting function generates overweighting of the tails, consider Equations 1 and 3. Suppose that the best and the worst outcomes, x_1 and x_m, materialize with small probabilities (i.e. $m > 2$). The decision weight of the right tail, π_1, equals $w(p_1)$. As p_1

[6] 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 $F(x_{m+1}) := 0$.

[7] 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)).

is small, x_1 is overweighted by w. The decision weight of the left tail, π_m, equals $1 - w(1 - p_m)$. As p_m is small, $1 - p_m$ is large and, hence, underweighted by w. Consequently, x_m is overweighted.

Another pervasive feature of risk preferences concerns proneness to Allais-type *common-ratio violations* that constitute one of the most widely replicated experimental regularities in human and animal behavior: Mixing a pair of prospects with common aversive outcomes frequently leads to preference reversals (Allais, 1953; Hagen, 1972; Kahneman and Tversky, 1979; MacCrimmon and Larsson, 1979; Battalio, Kagel, and MacDonald, 1985; Loomes and Sugden, 1987; Kagel, MacDonald, and Battalio, 1990; Nebout and Dubois, 2014; Chark, Chew, and Zhong, 2016).

Inspired by one of Allais (1953)'s famous examples, Kahneman and Tversky (1979) presented subjects with the decision situation summarized in Table 1. In the first decision situation, involving a certain option and a risky one, most people chose the certain option of 3000 dollars. When confronted with the choice between a 25%-chance of receiving 3000 dollars and a 20%-chance of receiving 4000 dollars, the majority opted for the 4000-dollar alternative, however. Multiplying the probabilities of 100% and 80% by a common factor $\lambda \in (0, 1)$, in this example by $\lambda = 1/4$, induced many people to reverse their preferences, a regularity termed *common ratio effect*.

Table 1: Allais-Type Common Ratio Pairs

First pair of options:		
\$ 3000 for sure	*or*	\$ 4000 with a probability of 80%
Second pair of options:		
\$ 3000 with a probability of 25%	*or*	\$ 4000 with a probability of 20%

Common-ratio violations are parsimoniously characterized by subproportionality of the probability weighting function w. Formally, subproportionality of w holds for probabilities p and q, if $1 \geq p > q > 0$, and $0 < \lambda < 1$ imply the inequality

$$\frac{w(p)}{w(q)} > \frac{w(\lambda p)}{w(\lambda q)} \tag{4}$$

(Prelec, 1998). Intuitively, subproportionality decreases the decision maker's sensitivity to disappointment for scaled-down probabilities, i.e. outcomes with high ex-ante probabilities of materializing carry higher disappointment potential. In this sense, the loss of certainty hurts more than the scaling down of a probability bounded away from one does. Therefore, subproportionality implies the *certainty effect*, which constitutes the special case of $p = 1$: $w(\lambda q) > w(\lambda)w(q)$ is satisfied for any λ, q such that $0 < \lambda, q < 1$. Many functional specifications proposed in the literature exhibit subproportionality over some probability range under appropriate parameter restrictions (see Appendix C). Perhaps the most prominent representative of a globally subproportional function with a regressive shape is Prelec (1998)'s flexible two-parameter specification. Throughout the paper, we will use this functional specification to illustrate our results graphically.

3 The Model

Our approach is applicable to an arbitrary number of outcomes provided that survival risk does not change the rank order of the prospects, i.e. "something may go wrong" is encoded as an outcome $\underline{x}$ no better than the prospects' minimum outcome $x_m \geq \underline{x}$. Rearranging terms in Equation 2 yields

$$\begin{aligned} V(P) &= u(x_1)w(p_1) + u(x_2)\Big(w(p_1+p_2) - w(p_1)\Big) + ... + u(x_m)\Big(1 - w(1-p_m)\Big) \\ &= \Big(u(x_1) - u(x_2)\Big)w(p_1) + ... + \Big(u(x_{m-1}) - u(x_m)\Big)w(1-p_m) + u(x_m)\,. \end{aligned} \tag{5}$$

This presentation of $V(P)$ clarifies that x_m is effectively a sure thing whereas obtaining something better than x_m is risky.

If the prospect is not played out and paid out in the present, but at some future time $t > 0$, two additional factors become important. First, 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 $\rho(t) = \exp(-\eta t)$. 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. A prospect to be played out and paid out at $t > 0$ is discounted for time in the standard way:

$$[V(P)]_0 = V(P)\rho(t)\,. \tag{6}$$

Second, and most importantly, survival risk changes the nature of the prospect. Let $0 < s \leq 1$ denote the constant per-period probability of prospect survival, i.e. the probability that the decision maker will actually obtain the promised outcomes by the end of the period.[8] 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 $\tilde{P} = \Big(x_1, p_1 s^t; x_2, p_2 s^t; ...; x_m, p_m s^t; \underline{x}, 1 - s^t\Big)$, where $\underline{x}$ captures that "something may go wrong". With the passage of time, the probability of prospect survival gets progressively scaled down.

[8]For similar approaches see Halevy (2008) and Walther (2010) who study hyperbolic discounting in the context of probability-weighting models.

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

$$
\begin{aligned}
[V(\tilde{P})]_0 = & \Big(\Big(u(x_1) - u(x_2)\Big) w(p_1 s^t) + ... \\
& ... + \Big(u(x_{m-1}) - u(x_m)\Big) w((1-p_m)s^t) + u(x_m) w(s^t)\Big) \rho(t) \\
= & \Big(\Big(u(x_1) - u(x_2)\Big) \tfrac{w(p_1 s^t)}{w(s^t)} + ... \\
& ... + \Big(u(x_{m-1}) - u(x_m)\Big) \tfrac{w((1-p_m)s^t)}{w(s^t)} + u(x_m)\Big) w(s^t) \rho(t) \, .
\end{aligned} \tag{7}
$$

Now suppose that the observer assumes that there is no survival risk, i.e. that $s = 1$, while in fact $s < 1$. 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. However, in the eye of the decision maker the prospect involves an additional layer of risk. If the observer neglects $s < 1$, she infers preference parameters from:

$$
[V(\tilde{P})]_0 = \Big(\Big(u(x_1) - u(x_2)\Big) \tilde{w}(p_1) + ... + \Big(u(x_{m-1}) - u(x_m)\Big) \tilde{w}(1-p_m) + u(x_m)\Big) \tilde{\rho}(t) \, , \tag{8}
$$

interpreting $\tilde{w}$ as true probability weights and $\tilde{\rho}$ as true discount weights, while in fact the weights are distorted by survival risk. Obviously, the measured weights differ from the underlying ones if $s < 1$. By comparing Equation 7 with Equation 8 we can see that the relationship between underlying and observed risk preference parameters is given by

$$
\tilde{w}(p) = \tilde{w}(p, t) = \frac{w(ps^t)}{w(s^t)} \, , \tag{9}
$$

as $\tilde{\rho}(t) = w(s^t)\rho(t)$ is interpreted as the discount weight attached to the allegedly certain outcome x_m.[9] Equation 9 defines the central relationship between observed and underlying probability weights. Because $\tilde{w}(p, t) \neq w(p)$ for subproportional preferences, survival risk drives a wedge between atemporal risk preferences and risk taking behavior with respect to delayed prospects. A summary of the model variables is provided in Table 2.

4 Model Predictions

In the following, we present our model predictions rationalizing the over- and underweighting of tail events. As discussed above, both the timing and the process of uncertainty resolution are crucial features of prospect valuation. We distinguish two cases: First, the prospect is played out

[9] Time discounting of a certain outcome constitutes the special case of $p = 1$. Concerning the discount weights $\tilde{\rho}(t)$, an equivalent representation was derived by Halevy (2008) for Yaari (1987)'s dual theory with a convex probability weighting function. If $\tilde{w}$ is subproportional, $\tilde{\rho}$ declines hyperbolically (see also Epper, Fehr-Duda, and Bruhin (2011)).

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	$0 < s \leq 1$
	$1 - s$	survival risk	
	t	length of time delay	$t \geq 0$
Preferences	$u(x)$	utility function	$u(0) = 0, u' > 0$
	$w(p)$	atemporal 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	$\tilde{w}(p,t)$	observed probability weight	$\tilde{w}(p,t) = \frac{w(ps^t)}{w(s^t)}$
	$\tilde{\rho}(t)$	observed discount weight	$\tilde{\rho}(t) = w(s^t)\rho(t)$

and paid out at some time in the future. This situation of one-shot resolution of uncertainty is represented by Theorem 1. Theorem 2 covers the case when uncertainty is resolved sequentially over the course of time.

4.1 One-Shot Resolution of Uncertainty

Turning to the one-shot resolution of base risk and survival risk, we see from Equation 9 that observed probability weights $\tilde{w}(p,t)$ deviate from the underlying atemporal ones $w(p)$ in two respects: First, $w(s^t) < 1$ in the denominator boosts observed weights. Second, $w(ps^t)$ in the numerator distorts observed probability weights. In the following, we suppress delay t in the notation whenever there is no ambiguity about the length of delay. The assumption of subproportional probability weights w generates clear predictions for $\tilde{w}$:

THEOREM 1:

Given subproportionality of w 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: $\tilde{w}(p) > w(p)$. Elevation increases with

   - time delay t,
   - survival risk $1 - s$, and
   - degree of subproportionality.

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

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

Proof of Theorem 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 $0 < \lambda < 1$:
$$\frac{\tilde{w}(\lambda p)}{\tilde{w}(\lambda q)} = \frac{w(\lambda s^t p)}{w(\lambda s^t q)} < \frac{w(s^t p)}{w(s^t q)} = \frac{\tilde{w}(p)}{\tilde{w}(q)}\,. \tag{10}$$

3. Since w is subproportional,
$$\tilde{w}(p) = \frac{w(ps^t)}{w(s^t)} > \frac{w(ps)}{w(s)} > \frac{w(p)}{w(1)} = w(p) \tag{11}$$
holds for $s < 1$ and $t > 1$. 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:
$$\frac{\tilde{w}_1(p)}{w_1(p)} = \frac{w_1(\lambda p)}{w_1(\lambda)w_1(p)} = \frac{w_1(\lambda p)}{w_1(\lambda)w_1(p)}\frac{w_1(\lambda)w_1(p)}{w_1(\lambda pq)} = \frac{w_1(\lambda p)}{w_1(\lambda pq)}\,. \tag{12}$$

   Applying the same logic to w_2 yields
$$\frac{\tilde{w}_2(p)}{w_2(p)} = \frac{w_2(\lambda p)}{w_2(\lambda)w_2(p)} = \frac{w_2(\lambda p)}{w_2(\lambda pqr)} > \frac{w_2(\lambda p)}{w_2(\lambda pq)}\,. \tag{13}$$

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

$$\varepsilon_{\tilde{w}}(p) = \frac{\tilde{w}'(p)p}{\tilde{w}(p)} = \frac{w'(ps^t)ps^t}{w(ps^t)} = \varepsilon_w(ps^t) < \varepsilon_w(p)\,, \tag{14}$$

as the elasticity ε_w is increasing in its argument iff w is subproportional (Segal, 1987a).

5. 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.

□

That $\tilde{w}$ is more elevated than w constitutes the central implication of our model. Due to subproportionality

$$\frac{w(p_1 s^t)}{w(p_1)} > \frac{w(s^t)}{w(1)}$$

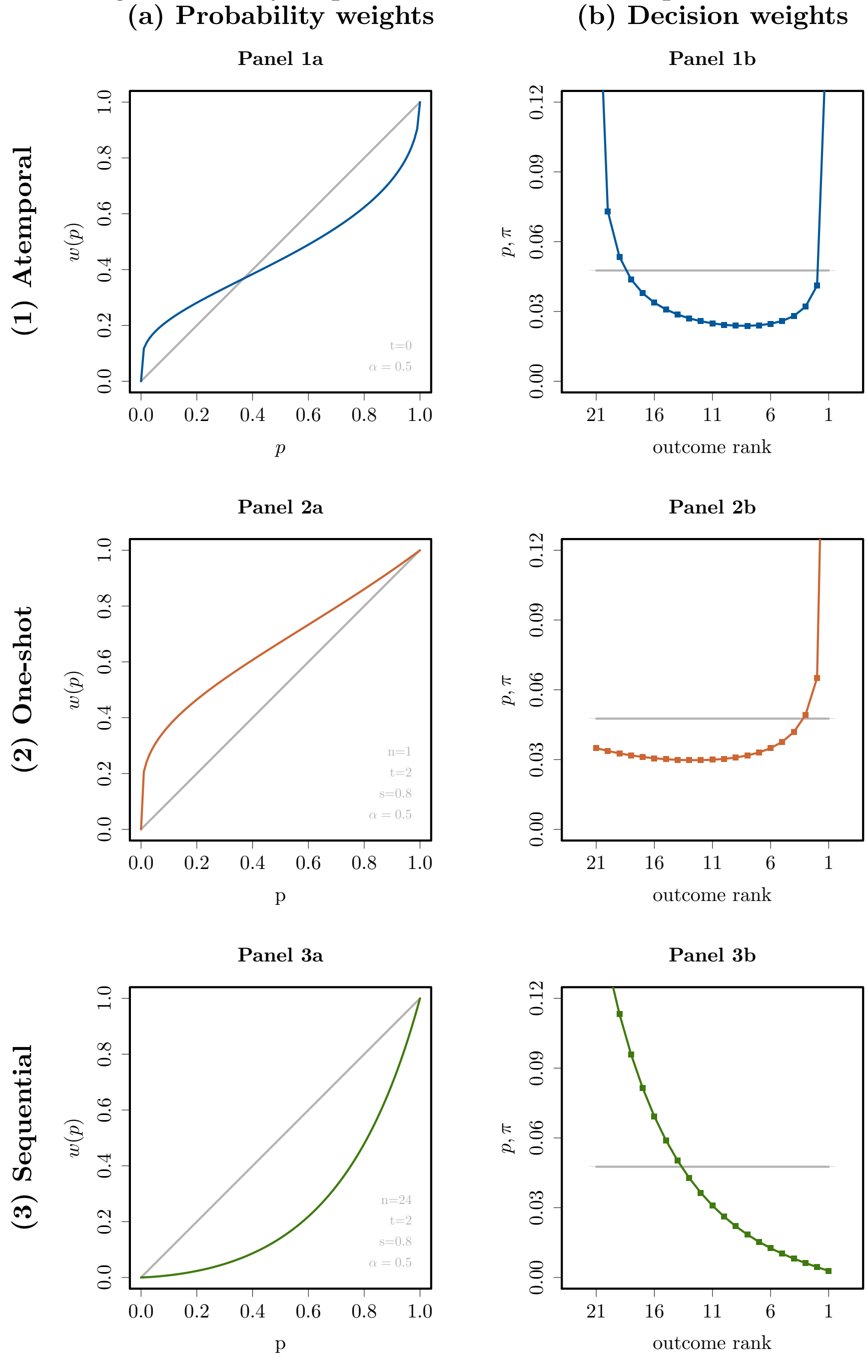
holds, i.e. comparing the delayed case with the atemporal one, the weight of the best possible outcome is devalued less than the weight of the sure component. In other words, x_m suffers more strongly from delay than does x_1. Thus, the presence of survival risk makes people appear more risk tolerant for delayed prospects than for present ones.[10] The consequences for a regressive w are clear. Figuratively speaking, the decision weight curve rotates counterclockwise: The right tail of the outcome distribution gains more weight, whereas the left tail loses weight with delay. If t is sufficiently large, the left tail may even be underweighted, as illustrated in Figure 2.

The top row of Figure 2 characterizes preferences in the atemporal case. Panel 1a shows a typical specimen of a regressive probability weighting function for delay $t = 0$, 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 for a prospect involving 21 equiprobable outcome levels, with outcome rank 1 denoting the best outcome and outcome rank 21 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 middle row of Figure 2 demonstrates the predictions for one-shot resolution of uncertainty, i.e. when prospects are played out and paid out simultaneously in the future. Future

[10] In the domain of simple prospects (x, p), Baucells and Heukamp (2012) derive a time-dependent probability weighting function $\tilde{w}(p) = w(p \exp(-r_x t))$, which obviously decreases with t. A crucial element of their model is r_x, the probability discount rate that is assumed to decrease with outcome magnitude. This assumption drives their result that risk premia decline with time delay.

Figure 2: Delay Dependence and Process Dependence



For purposes of illustration, the curves are derived from Prelec's two-parameter probability weighting function $w(p) = \exp\left(-\beta(-\ln(p))^{\alpha}\right)$ (Prelec, 1998), assuming a degree of subproportionality $\alpha = 0.5$ and convexity $\beta = 1$. Survival risk s is set at 0.8 per period. n denotes the number of (equally spaced) stages in the case of sequential evaluation. **Top row (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. **Middle row (one-shot)**: Panel 2a and 2b show $\tilde{w}$ and $\tilde{\pi}$ for a delay of two periods, $t = 2$, when uncertainty resolves in one shot $n = 1$. **Bottom row (sequential)**: Panel 3a and 3b show $\tilde{w}$ and $\tilde{\pi}$, respectively, for a delay of two periods when uncertainty resolves sequentially in $n = 24$ equally spaced stages, $\tilde{w}(p) = \left(\frac{w((ps^t)^{1/n})}{w((s^t)^{1/n})}\right)^n$.

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, $t = 2$, and uncertainty resolves in one shot ($n = 1$) observed probability weights $\tilde{w}$ shift upwards, as shown in Panel 2a. This shift transforms the decision weights 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 and, hence, underinsuring becomes more likely with longer time horizons. The delay dependence of risk tolerance, therefore, provides a rationale for the underweighting of adverse tail events.

Numerous experimental studies have found that risk tolerance is indeed higher for payoffs materializing in the future than for payoffs materializing in the present (Jones and Johnson, 1973; Shelley, 1994; Ahlbrecht and Weber, 1997; Sagristano, Trope, and Liberman, 2002; Noussair and Wu, 2006; Coble and Lusk, 2010). More specifically, Abdellaoui, Diecidue, and Öncüler (2011) conducted a carefully designed experiment eliciting probability weights for both present and delayed prospects, i.e. in our notation $w(p)$ and $\tilde{w}(p)$. Their results provide persuasive direct support for our approach. They find four distinctive characteristics of delay-dependent prospect valuation. First, the utility for money u does not react to time delay. Second, $\tilde{w}$ is significantly more elevated than w in the aggregate as well as for the majority of the individuals. Third, an additional six-month delay affects elevation less strongly than the first six-month delay. Moreover, $\tilde{w}$ appears to be less strongly curved than w.[11] Another important finding of Abdellaoui, Diecidue, and Öncüler (2011) concerns behavior under timing uncertainty. When their experimental subjects did not know the exact timing of the payoffs, they acted as if the prospects' delays were midway between the present and the longest delay in the experiment, 12 months. This finding suggests that delay dependence is also present in situations when payoff dates, and hence the resolution of uncertainty, are indeterminate.

Aside from delay-dependent risk tolerance, the model produces other interesting effects. For one, $\tilde{w}$ is less elastic than w, implying less sensitivity to anticipated disappointment with respect to delayed prospects. This prediction is in line with Trope and Liberman (2003)'s theory of tem-

[11] In their study on ambiguity, Abdellaoui, Baillon, Placido, and Wakker (2011) show estimates of a probability weighting curve derived from choices over prospects delayed by three months. This curve is also much more elevated than typical atemporal estimates are (see for example Bruhin, Fehr-Duda, and Epper (2010)).

poral construal, that posits that temporal distance changes the way people mentally represent those events. The greater the temporal distance, the more likely are events to be represented in terms of a few abstract features. Another insight concerns the impact of the degree of subproportionality on the valuation of delayed prospects. Stronger subproportionality implies a more pronounced reaction to future uncertainty, which is a plausible implication when subproportionality is interpreted as measure of emotionality. This result speaks not only to individual heterogeneity but also to situations that may trigger more or less fear. For example, in times of economic crisis people may react much more strongly to anticipated emotions (Cohn, Engelmann, Fehr, and Marechal, 2015), i.e. they may display a higher degree of subproportionality than in times of economic stability. Thus, in times of crises, they will react more strongly to imminent risks but much less strongly to risks resolving in the remote future. Furthermore, the wedge between $\tilde{w}$ and w also increases with the degree of survival risk, implying, somewhat paradoxically, that observed risk tolerance increases with subjective uncertainty.[12]

4.2 Sequential Resolution of Uncertainty

So far, we have considered the case of uncertainty resolving in one shot, the domain over which atemporal risk preferences are defined.[13] If uncertainty does not resolve in one shot but rather sequentially over the course of time, future prospects lose their single-stage quality and turn into multi-stage ones. In this case the question arises in which way multi-stage prospects are transformed into single-stage ones. Essentially, there are two different transformation methods, reduction by probability calculus and folding back (Sarin and Wakker, 1994). In the case of reduction by probability calculus, the probabilities of reaching the final outcomes are compounded and probability weights are applied only to the resulting compounded probabilities. Folding back means that a multi-stage prospect is evaluated recursively by replacing the n^{th}-stage prospect with its certainty equivalent and inserting the utility of the certainty equivalent into the $(n-1)^{th}$-stage valuation formula and so forth. Thus, decision weights get compounded.

[12] This finding mirrors Quiggin (2003)'s result of atemporal risk tolerance increasing with background risk.

[13] The ramifications of sequential prospect valuation have previously been analyzed for a different class of atemporal risk preferences. Palacios-Huerta (1999)'s contribution focuses on process dependence in the context of Gul (1991)'s model of disappointment aversion. He shows that a disappointment averse decision maker exhibits much larger risk aversion when she evaluates a prospect sequentially rather than in one shot. Dillenberger (2010) provides an axiomatic underpinning for this result and an insightful discussion of the consequences of a preference for one-shot resolution of uncertainty on the value of information. See also Cerreia-Vioglio, Dillenberger, and Ortoleva (2015).

Several authors made a case against reduction as an appropriate mechanism of transforming multi-stage prospects into single-stage ones (Segal (1990); Dekel, Safra, and Segal (1991); Grant, Kajii, and Polak (1998) among others). Segal (1990) argues that even if the decision maker accepts the basic laws of probability theory she may have a preference over the number of lotteries she participates in, which invalidates reduction by probability calculus.

However, subproportionality of risk preferences raises the issue of dynamic consistency. Dynamic consistency requires that choices made at, or plans formed at, different times conform with one another (Sugden, 2004). As Loomes and Sugden (1986) explain, any theory that accommodates the common-ratio effect must dispense either with dynamic consistency or with reduction by the probability calculus. Therefore, if the decision maker cares only about the total probabilities of the final outcomes she will be dynamically inconsistent unless she precommits herself to stick to her original plans.[14] Folding back, on the other hand, ensures dynamic consistency but, as Theorem 2 will show, has substantial consequences for revealed risk taking behavior (see also Sarin and Wakker (1992)). 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 two stages, $n = 2$, 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 Figure 3. Applying folding back, the resulting two-stage prospect is evaluated as

$$\begin{aligned}
[V_2(\tilde{P})]_0 &= \Big(u(x_1) - u(x_2)\Big) w\left(p^{\frac{t_1}{t}} s^{t_1}\right) w\left(p^{\frac{t-t_1}{t}} s^{t-t_1}\right) + u(x_2) w\left(s^{t_1}\right) w\left(s^{t-t_1}\right) \\
&= \left(\Big(u(x_1) - u(x_2)\Big) \frac{w\left(p^{\frac{t_1}{t}} s^{t_1}\right) w\left(p^{\frac{t-t_1}{t}} s^{t-t_1}\right)}{w\left(s^{t_1}\right) w\left(s^{t-t_1}\right)} + u(x_2)\right) w\left(s^{t_1}\right) w\left(s^{t-t_1}\right) \\
&= \left(\Big(u(x_1) - u(x_2)\Big) \tilde{w}_2(p) + u(x_2)\right) \tilde{\rho}_2(t)\,,
\end{aligned} \tag{15}$$

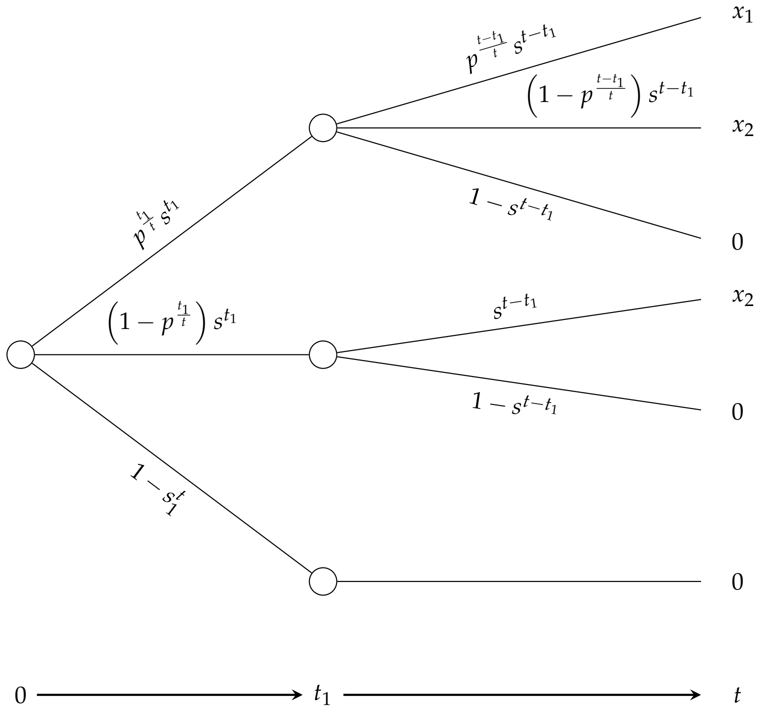
which yields the relationship

$$\tilde{w}_2(p) = \frac{w\left(p^{\frac{t_1}{t}} s^{t_1}\right) w\left(p^{\frac{t-t_1}{t}} s^{t-t_1}\right)}{w\left(s^{t_1}\right) w\left(s^{t-t_1}\right)} \tag{16}$$

as $\tilde{\rho}_2(t) = w\left(s^{t_1}\right) w\left(s^{t-t_1}\right)$ is interpreted as the discount weight attached to the allegedly

[14] A time-inconsistent decision maker will become progressively less risk tolerant as the payment date draws nearer.

Figure 3: Sequential Resolution of Uncertainty



certain outcome x_2. Subproportionality ensures that

$$\tilde{w}_2(p) = \frac{w\left(p^{\frac{t_1}{t}} s^{t_1}\right) w\left(p^{\frac{t-t_1}{t}} s^{t-t_1}\right)}{w\left(s^{t_1}\right) w\left(s^{t-t_1}\right)} < \frac{w(ps^t)}{w(s^t)} = \tilde{w}(p)\,, \tag{17}$$

the main result of Theorem 2. Now suppose that the interval $[0, t]$ is partitioned into n subintervals with lengths τ_i, $i \in \{1, ..., n\}$, such that $\sum_{i=1}^{n} \tau_i = t$. In this case, it is straightforward to show for any number of outcomes $m \geq 1$ that the observed probability weights are given by

$$\tilde{w}_n(p, t) = \frac{\prod_{i=1}^{n} w\left(p^{\frac{\tau_i}{t}} s^{\tau_i}\right)}{\prod_{i=1}^{n} w\left(s^{\tau_i}\right)} = \prod_{i=1}^{n} \tilde{w}\left(p^{\frac{\tau_i}{t}}, \tau_i\right)\,. \tag{18}$$

THEOREM 2:

Given subproportionality of w, $s \leq 1$ and folding back:

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 evaluation stages n, prospect valuation is lowest for equally spaced subintervals $\tau_i = \frac{t}{n} = \tau$.[15]

3. For equally spaced subintervals, prospect valuation declines with the number of evaluation stages: $[\tilde{V}_n]_0 < [\tilde{V}_{n-1}]_0$.

Proof of Theorem 2.

1. Consider Equation 18:

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

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

$$\tilde{w}_n(p,t) < \prod_{i=1}^{n} \tilde{w}\left(p^{\frac{\tau_i}{t}}, t\right) < \tilde{w}\left(\prod_{i=1}^{n} p^{\frac{\tau_i}{t}}, t\right) = \tilde{w}(p,t) . \tag{19}$$

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 relationships hold for $0 < q < 1$:

$$\frac{w(q^{\tau_{i-1}})}{w(q^{\tau_i - \varepsilon})} > \frac{w(q^{\tau_{i-1}} q^{\varepsilon})}{w(q^{\tau_i - \varepsilon} q^{\varepsilon})} = \frac{w(q^{\tau_{i-1}+\varepsilon})}{w(q^{\tau_i})} , \tag{20}$$

implying $w(q^{\tau_{i-1}})w(q^{\tau_i}) > w(q^{\tau_i - \varepsilon})w(q^{\tau_{i-1}+\varepsilon})$, in particular for probabilities $q = p^{1/t}s$ and $q = s$. Therefore, compounding probability weights and decision weights over a more evenly spaced partition generates a smaller prospect value.

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

$$\prod_{i=1}^{n} w\left(p^{\frac{\tau}{t}} s^{\tau}\right) < \prod_{i=1}^{n-1} w\left(p^{\frac{\delta}{t}} s^{\delta}\right) . \tag{21}$$

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

$$\left(w\left(q^{n-1}\right)\right)^{n} < \left(w(q^{n})\right)^{n-1} . \tag{22}$$

[15]For $\tilde{w}_n$ itself rather than total prospect value to be smallest for equally spaced partitions an additional condition is required: the elasticity of w has to be convex.

Proceeding by complete induction:

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

$$\Big(w(q^3)\Big)^2 > \frac{\Big(w(q^2)\Big)^2}{w(q)}\frac{\Big(w(q^2)\Big)^2}{w(q)} > \frac{\Big(w(q^2)\Big)^3 w(q^2)}{\Big(w(q)\Big)^2} > \frac{\Big(w(q^2)\Big)^3\Big(w(q)\Big)^2}{\Big(w(q)\Big)^2} = \Big(w(q^2)\Big)^3 \tag{23}$$

- $n \to n+1$: Suppose that $\Big(w(q^{n-1})\Big)^n < \Big(w(q^n)\Big)^{n-1}$ holds. Subproportionality implies $\frac{w(q^{n-1})}{w(q^n)} > \frac{w(q^n)}{w(q^{n+1})}$. Hence,

$$\begin{aligned}\Big(w(q^{n+1})\Big)^n > \left(\frac{w(q^n)w(q^n)}{w(q^{n-1})}\right)^n &= \frac{\Big(w(q^n)\Big)^{n+1}\Big(w(q^n)\Big)^{n-1}}{\Big(w(q^{n-1})\Big)^n} \\ > \frac{\Big(w(q^n)\Big)^{n+1}\Big(w(q^{n-1})\Big)^n}{\Big(w(q^{n-1})\Big)^n} &= \Big(w(q^n)\Big)^{n+1}\end{aligned} \tag{24}$$

$\square$

Theorem 2 shows that a decision maker with subproportional preferences prefers uncertainty to be resolved in one shot at the payment date t rather than sequentially over the course of time.[16] Note that this result does not hold generally under subproportionality in RDU but only applies to the class of prospects studied here, i.e. prospects that are devalued by survival risk without effects on the rank order of the outcomes (see Dillenberger (2010)'s necessary and sufficient criterion for preferences for one-shot resolution and our discussion in Appendix B).

Preference for one-shot resolution of uncertainty is embodied in the characteristics of atemporal risk preferences and, therefore, all the insights of Segal (1990), who analyzes two-stage prospects in an atemporal setting, still apply. However, risk tolerance is additionally influenced by its delay dependence. Consider a prospect with a long time horizon t. If its total uncertainty is resolved in one single stage, all the decision weights attain their maximum values. If uncertainty

[16] A special case is the valuation of allegedly certain future payoffs, which constitute simple prospects in our framework. A myopic decision maker, applying folding back, will exhibit a discount weight of $w\left(s^{t_1}\right) w\left(s^{t-t_1}\right) < w\left(s^t\right)$, an incident of *subadditive discounting*, which has found experimental support (Read, 2001; Read and Roelofsma, 2003; Ebert and Prelec, 2007; Epper, Fehr-Duda, and Bruhin, 2009; Dohmen, Falk, Huffman, and Sunde, 2012).

resolves sequentially, both probability and discount weights are smaller than in the one-shot case. The effect gets more pronounced the finer is the partition of delay t into subintervals. Therefore, anticipating to watch uncertainty resolve over time considerably dampens the effect of long time horizons on observed risk tolerance, because the decision maker is frequently exposed to the possibility of a disappointing outcome.

Our model predicts that equally spaced partitions of the time interval will be valued particularly unfavorably. 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 (Segal, 1987b). Because of this characteristic, Segal (1987b) proposes to model ambiguity aversion by subproportional risk preferences over two-stage lotteries.[17]

The consequences of sequential valuation for the tails of the outcome distribution are straight forward if w is regressive. The weight of the right tail $\tilde{\pi}_n(x_1) = \tilde{w}_n(p_1) < \tilde{w}(p_1)$. Therefore, overweighting of x_1 declines relative to the one-shot situation. The weight of the left tail increases as $\tilde{\pi}_n(x_m) = 1 - \tilde{w}_n(1 - p_m) > 1 - \tilde{w}(1 - p_m)$ holds. Depending on the number of subperiods over which probability weights are compounded, this increase may lead to a considerable overweighting of the worst outcomes and, consequently, to pronounced risk aversion.

The bottom row of Figure 2 demonstrates the effect of sequential valuation on probability weights and decision weights for a delay of $t = 2$. If a prospect is evaluated in 24 equally spaced time intervals, $n = 24$, the probability weighting curve takes on a convex form, which implies strong risk aversion. The associated decision weights for our reference prospect involving 21 equiprobable outcomes are depicted in Panel 3b. The decision weight curve now rotates clockwise: The worst outcomes are strongly overweighted while the best outcomes are considerably underweighted. Sequential valuation, 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 incurring losses.

[17]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).

5 Discussion

Most economically important decisions, may they concern health, wealth, love or education involve a significant interval between the time that the decision is made and the time that all uncertainty is completely resolved. Our contribution provides a novel view on perplexing real-world behaviors. We show that if people view the future as inherently uncertain and are susceptible to probability weighting, their risk tolerance varies greatly depending on the length of delay and their perception of uncertainty resolution. When the passage of time does not play a significant role, a typical decision maker overweights both tails of an outcome distribution. This feature of risk preferences explains people's skewness preferences, favoring positively skewed distributions and disfavoring negatively skewed ones (Lovallo and Kahneman, 2000; Barberis, 2013b). If uncertainty resolves in the future, however, adverse tail events receive progressively less weight and, for long time horizons, may even be substantially underweighted, thereby greatly reducing people's willingness to buy insurance.

Delay- and process-dependent risk tolerance not only affects individuals' welfare but also society at large. People's reluctance to take out insurance for floods and earthquakes, for example, poses serious problems when disaster actually strikes. It is practically impossible for the public authorities to deny assistance once there are identified victims and their stories are publicized in the news (Viscusi, 2010). In the context of climate policy, it takes decades or even centuries until the stock of pollutants will be sufficiently reduced to see any gaugeable effect of society's abatement endeavors. If there is both great uncertainty about the effectiveness of abatement policies and lack of feedback, the risk tolerance of a large percentage of the population may be extremely high and, therefore, it is likely that they are opposed to supporting abatement measures. It remains to be seen whether endeavors to combat global warming will be met with more support once its effects become more visible.

Stock market investors' time horizons may also be long-term in principle but, contrary to natural disasters, information on portfolio performance is easily accessible and, due to its omnipresence in the news, hard to ignore. Thus, uncertainty resolves practically in real time, which substantially counteracts the otherwise risk-tolerance increasing effect of long investment horizons. Recently, the term structure of market risk premia has attracted considerable attention (Andries, Eisenbach, and Schmalz, 2015; Eisenbach and Schmalz, 2016). The empirical evidence points to a downward sloping curve, i.e. assets with short maturities seem to earn much higher risk premia than assets with long maturities (van Binsbergen, Brandt, and Koijen, 2012), which contradicts the predictions of standard asset-pricing models. Consequently, new models of asset pricing work with the assumption of horizon-dependent risk tolerance (Khapko, 2015). Our model provides a rational for both, high risk premia, because of the short-term resolution of uncertainty, and risk premia declining with maturity, because of the delay-dependence of risk tolerance.

Referring to experimental evidence in atemporal settings, Hertwig, Barron, Weber, and Erev (2004) suggest an alternative explanation for the underweighting of tail events. They argue that overweighting occurs in situations when risks are described in abstract terms. However, when people decide on the basis of their own experience by sampling the distributions, they tend to underweight tail events. This claim has triggered a heated debate on the so-called description-experience gap (Barberis, 2013a; de Palma, Abdellaoui, Attanasi, Ben-Akiva, Erev, Fehr-Duda, Fok, Fox, Hertwig, Picard, Wakker, Walker, and Weber, 2014). Recently, Abdellaoui, L'Haridon, and Paraschiv (2011) show that having to find out themselves about outcomes and probabilities by experience sampling makes people considerably more pessimistic than in the case of fully described risks, which manifests itself in a less elevated probability weighting curve. In other words, ambiguity about distributions shifts the probability weighting curve downwards, which may explain the underweighting of rare extreme events observed in experiments. Many empirical facts in finance, insurance and gambling are consistent with the overweighting of tail events, however. According to Hertwig, Barron, Weber, and Erev (2004)'s claim all these phenomena would have to be based on described risks. In our view, it seems implausible that in many real-world situations people's decisions are based solely on abstract descriptions rather than on their own or somebody else's experience. Turning back to our example in the introduction: Why should a regular life insurance policy be driven by experience and a flight insurance by description?

Models of probability weighting have proven to be quite successful in organizing the results of countless experiments. Recently, it has been recognized that they are useful for explaining field data as well. Here we show that extending the realm of probability weighting from timeless decisions to intertemporal ones helps rationalize the coexistence of over- and underweighting of tail events, a puzzle unsolved so far. Whether the mechanism we suggest is actually driving behavior needs to be assessed by future work. The model presented in this paper provides a host of novel testable predictions which, we hope, will encourage researchers to conduct experiments to gauge the extent of its applicability.

Appendix A The Case of Ambiguity

In real-world settings probabilities are rarely known to the decision maker. With the exception of some games of chance, such as tossing a coin or playing roulette, the decision maker has to assess the likelihoods of ambiguous events. Our model is cast in terms of objectively given probabilities, however. Thus, the question arises whether our results are portable to the domain of ambiguity.

In this domain, the following framework is usually applied: S is a set of exhaustive and mutually exclusive states of nature. One of these states $s \in S$ will obtain, but the decision maker is unsure which one it will be. Subsets of S are called events and denoted by A. Prospects, often termed *acts*, are described as $P = (x_1, A_1; ...; x_m, A_m)$, which yield the monetary outcome if the event A_i contains the true state of nature. Outcomes are rank ordered x_i, $x_1 > x_2 > ... > x_m$ and $(A_1, A_2, ..., A_m)$ is a partition of the state space. To accommodate ambiguity, RDU is generalized to Choquet Expected Utility Theory (Schmeidler, 1989), which features a weighting function $W(A)$. W is a capacity satisfying $W(\oslash) = 0$, $W(S) = 1$, and monotonicity with respect to set inclusion, i.e. $A \subset B \Longrightarrow W(A) \leq W(B)$. Decision weights π_i are constructed analogously to the case of risk:

$$\pi_i = \begin{cases} W(A_1) & \text{for } i = 1, \\ W\left(\bigcup_{k=1}^{i} A_k\right) - W\left(\bigcup_{k=1}^{i-1} A_k\right) & \text{for } 1 < i \leq m. \end{cases} \tag{25}$$

As before, the prospect's value is represented by

$$V(P) = \sum_{i}^{m} u(x_i)\pi_i \, . \tag{26}$$

There is a large literature in the psychology of judgment which suggests that, generally, people tend to overweight the likelihood of rare events and to underweight the likelihood of probable events. A prominent example are the frequency estimates for causes of death reported in Tversky and Koehler (1994). The same pattern of behavior has been found in experimental research on decisions under ambiguity (Tversky and Wakker, 1995; Gonzalez and Wu, 1999; Kilka and Weber, 2001; Abdellaoui, Vossmann, and Weber, 2005). In the literature, this pattern of overweighting and underweighting is discussed under the heading of *subadditivity*, a consequence of diminishing sensitivity towards probabilities when moving away from certainty and impossibility (Einhorn and Hogarth, 1985; Fox and See, 1993; Wakker, 2004).[18] Formally, subadditivity comprises two conditions, *lower subadditivity* $\underline{SA}$ and *upper subadditivity* $\overline{SA}$:

$$\underline{\mathrm{SA}}: \quad W(A) \geq W(A \cup B) - W(B), \tag{27}$$

$$\overline{\mathrm{SA}}: \quad 1 - W(S - A) \geq W(A \cup B) - W(B), \tag{28}$$

provided that $A \cap B = \oslash$ and $W(A \cup B)$ and $W(B)$ are bounded away from 1 and 0, respectively.[19]

Experimental studies suggest that subadditivity is more pronounced under ambiguity than under risk, which induced Tversky and Fox (1995) to suggest a two-stage model, formalized in Wakker (2004): Consider an ambiguous prospect (x, A) that pays x in the event that A occurs and zero otherwise. Furthermore, assume that its value can be represented by $V((x, A)) = u(x)W(A)$. Elicit the matching probability $\hat{p}(A)$ such that the decision maker is indifferent between the risky prospect $(x, \hat{p})$ and the ambiguous prospect $(x; A)$. Then $W(A)$ can be decomposed as

$$W(A) = w(\hat{p}(A)), \tag{29}$$

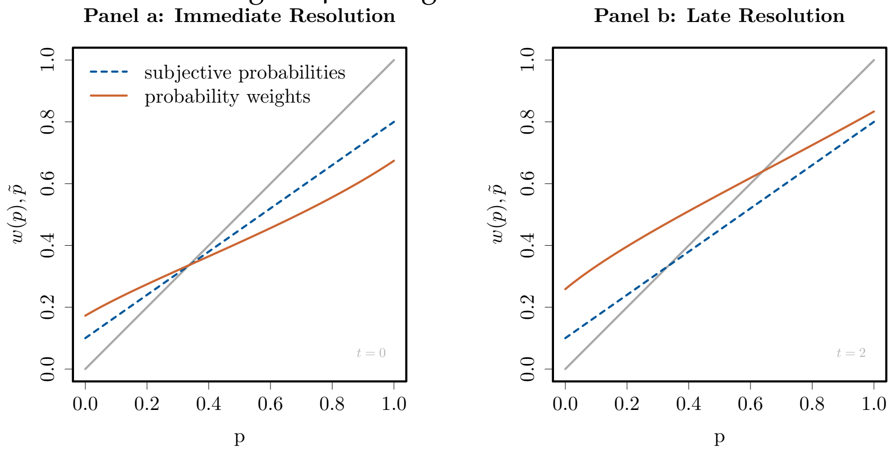
where w is the probability weighting function for risk. This decomposition has been used in a number of experimental studies (Abdellaoui, Vossmann, and Weber, 2005; Baillon, 2008; Baillon, Huang, Selim, and Wakker, 2016). The probability weighting function w for decisions under ambiguity has been found to differ from the pure risk case in that it is more strongly subadditive (Abdellaoui, Baillon, Placido, and Wakker, 2011), with the degree of departure from the risk case depending on the *source* of uncertainty, i.e. the concrete decision context (i.e. whether ambiguity concerns the composition of Ellsberg urns, the temperature in a specific city the following day, the movement of a specific stock index, etc.).

The crucial link to our analysis is that subadditivity is implied by strong regressiveness of the probability weighting function (for the proof see Prelec (1998), footnote 10). Therefore, all our predictions also apply to the case of ambiguity. We use these insights to develop a graphical representation of such a two-stage model which serves as basis for illustrating the effects of delay on behavior under ambiguity.

[18]Subadditivity also drives the famous Allais common consequence effect (Wu and Gonzalez, 1998).

[19]An example involving decision weights akin to our approach can be found in (Chateauneuf, Eichberger, and Grant, 2007).

Figure 4: Ambiguous Probabilities



Panel a: The dashed (blue) curve corresponds to subjective probabilities $\hat{p}$ assumed to follow the regularity $\hat{p} = 0.1 + 0.7p$, where p denotes empirical frequency. $w(p)$ is Prelec's functional specification with $\alpha = 0.64$ and $\beta = 1.03$ applied to $\hat{p}$ (red solid curve). **Panel b:** The red curve depicts $\tilde{w}(\hat{p})$ constructed according to Equation 9 with $s = 0.8$ and $t = 2$.

The curves in Figure 4 are constructed in the following way: In these graphs, probabilities p are in principle objectively given but the decision maker does not know them with precision, for example empirically observed frequencies of specific events. The decision maker judges the likelihoods $\hat{p}$ of these events (or reports choice-based probabilities). These probabilities are represented by the dashed lines in Figure 4, which mimic typical findings on the relationship between subjective judgments $\hat{p}$ and observed frequencies p (e.g. Fox and Tversky (1998), Figure 5). Applying w with parameters found in the literature (Abdellaoui, Baillon, Placido, and Wakker, 2011) to $\hat{p}$ renders the probability weighting curve $w(p)$ as a function of observed frequencies (solid curves in the figure). Panel a shows atemporal preferences, whereas Panel b depicts the case of delaying payoffs by 2 periods. As one can see, the resulting probability weighting curves for ambiguity are much less strongly curved than in the risky situation, but qualitatively the same implications for behavior over delayed prospects arise.

Appendix B 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 $P = (x_1, r; x_2)$ 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 nondegenerate one. The latter case characterizes the typical Allais certainty effect. NCI allows for Allais-type preference reversals but does not imply them. 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 prospects studied in this paper changes in rank order do not occur and, hence, NCI is satisfied. To see this, assume that the prospect $(x_1, p; x_2)$, $x_1 > x_2 \geq 0$, gets resolved in two stages $\Big((x_1, r; x_2), q; (x_2, 1)\Big)$ 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 $P = (x_1, r; x_2) \succsim (y, 1) = D$, with $x_1 > y > x_2$ and consider the following mixtures with $(x_2, 1 - \lambda)$ for some $\lambda \in (0, 1)$: $P' = (x_1, \lambda r; x_2)$ and $D' = (y, \lambda; x_2)$. The following relationships hold:

$$\begin{aligned}
P \succsim D \quad &\Rightarrow V(P) = \Big(u(x_1) - u(x_2)\Big)w(r) + u(x_2) \geq u(y) \\
V(D') \quad &= u(y)w(\lambda) + u(x_2)\Big(1 - w(\lambda)\Big) \\
&\leq \bigg(\Big(u(x_1) - u(x_2)\Big)w(r) + u(x_2)\bigg)w(\lambda) + u(x_2)\Big(1 - w(\lambda)\Big) \\
&= \Big(u(x_2) - u(x_1)\Big)w(r)w(\lambda) + u(x_2) \\
&< \Big(u(x_2) - u(x_1)\Big)w(\lambda r) + u(x_2) \\
&= V(P')
\end{aligned} \tag{30}$$

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 $P'' = \Big(x_1, \lambda r; x_2, \lambda(1-r); z\Big)$ and $D'' = (y, \lambda; z)$.

$$\begin{aligned}
V(D'') &= u(y)w(\lambda) + u(z)\Big(1 - w(\lambda)\Big) \\
&\leq \bigg(\Big(u(x_1) - u(x_2)\Big)w(r) + u(x_2)\bigg)w(\lambda) + u(z)\Big(1 - w(\lambda)\Big) \\
&= \Big(u(x_2) - u(x_1)\Big)w(r)w(\lambda) + \Big(u(x_2) - u(z)\Big)w(\lambda) + u(z) \\
&< \Big(u(x_2) - u(x_1)\Big)w(\lambda r) + \Big(u(x_2) - u(z)\Big)w(\lambda) + u(z) \\
&= V(P'')
\end{aligned} \tag{31}$$

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

Appendix C Subproportionality

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 superproportional over the range of small probabilities and subproportional over large probabilities. These functions 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 superproportionality 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. Bordalo, Gennaioli, and Shleifer (2012) derive (discontinuous) context-dependent probability distortions from their salience theory. While their concave segment is superproportional, the convex segment is equivalent to a subproportional probability weighting function of the Rachlin, Raineri, and Cross (1991) variety. The psychological mechanisms underlying probability weighting, therefore, often imply subproportionality.

Table 3: Probability Weighting Functions

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

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

(2) The weighting function consists of two parts and a jump in between.

(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 inverse S with U-shaped elasticity.

References

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., O. L'HARIDON, AND C. PARASCHIV (2011): "Experienced vs. Described Uncertainty: Do We Need Two Prospect Theory Specifications?," *Management Science*, 57(10), 1879–1895.

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.

ABDELLAOUI, M., F. VOSSMANN, AND M. WEBER (2005): "Choice-Based Elicitation and Decomposition of Decision Weights for Gains and Losses Under Uncertainty," *Management Science*, 51(9), 1384–1399.

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

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.

ANDRIES, M., T. M. EISENBACH, AND M. C. SCHMALZ (2015): "Asset Pricing with Horizon-Dependent Risk Aversion," *Federal Reserve Bank of New York Staff Reports*, 703.

BAILLON, A. (2008): "Eliciting Subjective Probabilities Through Exchangeable Events: An Advantage and a Limitation," *Decision Analysis*, 5(2), 76–87.

BAILLON, A., Z. HUANG, A. SELIM, AND P. WAKKER (2016): "Measuring Ambiguity Attitudes for All (Natural) Event," *mimeo*.

BAILLON, A., O. L'HARIDON, AND L. PLACIDO (2011): "Ambiguity Models and the Machina Paradoxes," *American Economic Review*, 101(4), 1547–1560.

Barberis, N. (2013a): "The Psychology of Tail Events: Progress and Challenges," *American Economic Review*, 103(3), 611–616.

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

Barberis, N., and M. Huang (2008): "Stocks as Lotteries: The Implications of Probability Weighting for Security Prices," *American Economic Review*, 98(5), 2066–2100.

Barberis, N., M. Huang, and R. H. Thaler (2006): "Individual Preferences, Monetary Gambles, and Stock Market Participation: A Case for Narrow Framing," *American Economic Review*, 96(4), 1069–1090.

Barnes, K. (2011): "Europe's Ticking Time Bomb," *Nature*, 473, 140–141.

Battalio, R., J. Kagel, and D. MacDonald (1985): "Animals' Choices over Uncertain Outcomes: Some Initial Experimental Results," *American Economic Review*, 75(4), 597–613.

Baucells, M., and F. H. Heukamp (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.

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

Bernheim, B. D., L. Forni, J. Gokhale, and L. J. Kotlikoff (2003): "The Mismatch Between Life Insurance Holdings and Financial Vulnerabilities: Evidence from the Health and Retirement Study," *American Economic Review*, 93(1), 354–365.

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*.

Bommier, A. (2006): "Uncertain Lifetime and Intertemporal Choice: Risk Aversion as a Rational for Time Discounting," *International Economic Review*, 47(4), 1223–1246.

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.

Cerreia-Vioglio, S., D. Dillenberger, and P. Ortoleva (2015): "Cautious Expected Utility and the Certainty Effect," *Econometrica*, 83(2), 693–728.

Chark, R., S. H. Chew, and S. Zhong (2016): "Individual Longshot Preferences," *mimeo*.

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.

Cicchetti, C., and J. Dubin (1994): "A Microeconometric Analysis of Risk Aversion and the Decision to Self-insure," *Journal of Political Economy*, 102(1), 169–186.

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.

Cohn, A., J. Engelmann, E. Fehr, and M. A. Marechal (2015): "Evidence for Countercyclical Risk Aversion: An Experiment with Financial Professionals," *American Economic Review*, 105(2), 860–885.

Cutler, D. M., A. Finkelstein, and K. McGarry (2008): "Preference Heterogeneity and Insurance Markets: Explaining a Puzzle of Insurance," *American Economic Review*, 98(2), 157–162.

Dasgupta, P., and E. Maskin (2005): "Uncertainty and Hyperbolic Discounting," *American Economic Review*, 95(4), 1290–1299.

De Giorgi, E., and s. Legg (2012): "Dynamic Portfolio Choice and Asset Pricing with Narrow Framing and Probability Weighting," *Journal of Economic Dynamics and Control*, 36(7), 951–972.

de Palma, A., M. Abdellaoui, G. Attanasi, M. Ben-Akiva, I. Erev, H. Fehr-Duda, D. Fok, C. R. Fox, R. Hertwig, N. Picard, P. P. Wakker, J. L. Walker, and M. Weber (2014): "Beware of Black Swans: Taking Stock of the Description-experience gap in Decision Under Uncertainty," *Marketing Letters*, 25, 269–280.

Dekel, E., Z. Safra, and U. Segal (1991): "Existence and Dynamic Consistency of Nash Equilibrium with Non-expected Utility Preferences," *Journal of Economic Theory*, 55, 229–246.

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.

Diecidue, E., and P. P. Wakker (2001): "On the Intuition of Rank-Dependent Utility," *The 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*.

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.

Einhorn, H. J., and R. M. Hogarth (1985): "Ambiguity and Uncertainty in Probabilistic Inference," *Psychological Review*, 92(4), 433–461.

Eisenbach, T. M., and M. C. Schmalz (2016): "Anxiety in the Face of Risk," *Journal of Financial Economics*, 121(2), 414–426.

Epper, T., and H. Fehr-Duda (2015): "The Missing Link: Unifying Risk Taking and Time Discounting," *Working Paper*.

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.

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.

Fehr-Duda, H., and E. Fehr (2016): "Game Human Nature," *Nature*, 530, 413–415.

Fox, C. R., and K. E. See (1993): "Belief and Preference in Decision Under Uncertainty," *Thinking: Psychological Perspectives on Reasoning, Judgment and Decision Making. Edited by David Hardman and Laura Macchi, John Wiley & Sons*, pp. 273–314.

Fox, C. R., and A. Tversky (1998): "A Belief-Based Account of Decision Under Uncertainty," *Management Science*, 44(7), 879–895.

Giannetti, M., and Y. Koskinen (2010): "Investor Protection, Equity Returns, and Financial Globalization," *Journal of Financial and Quantitative Analysis*, 45(1), 135–168.

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.

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

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

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

Hertwig, R., G. Barron, E. Weber, and I. Erev (2004): "Decisions from Experience and the Effect of Rare Events in Risky Choice," *Psychological Science*, 15(8), 534–539.

Huysentruyt, M., and D. Read (2010): "How Do People Value Extended Warranties? Evidence From Two Field Surveys," *Journal of Risk and Uncertainty*, 40, 197–218.

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

Kagel, J., D. MacDonald, and R. Battalio (1990): "Tests of 'Fanning Out' of Indifference Curves: Results from Animal and Human Experiments," *American Economic Review*, 80(4), 912–921.

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.

Khapko, M. (2015): "Asset Pricing with Dynamically Inconsistent Agents," *Working Paper, Stockholm School of Economics*.

Kilka, M., and M. Weber (2001): "What Determines the Shape of the Probability Weighting Function Under Uncertainty," *Management Science*, 47, 1712–1726.

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

Kunreuther, H. (1984): "Causes of Underinsurance against Natural Disasters," *The Geneva Papers on Risk and Insurance*, 9(31), 206–220.

Loomes, G., and U. Segal (1994): "Observing Different Orders of Risk Aversion," *Journal of Risk and Uncertainty*, 9, 239–256.

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.

Lopes, L. L. (1987): "Between Hope and Fear: The Psychology of Risk," *Advances in Experimental Psychology*, 20, 255–295.

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.

MacCrimmon, K., and S. Larsson (1979): "Utility Theory: Axioms versus Paradoxes," *In: Allais, M., Hagen, O. (Eds.), Expected Utility and The Allais Paradox, Dordrecht, Boston*, pp. 333–409.

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.

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

Palacios-Huerta, I. (1999): "The Aversion to the Sequential Resolution of Uncertainty," *Journal of Risk and Uncertainty*, 18, 249–269.

Pennesi, D. (2015): "Uncertain Discount and Hyperbolic Preferences," *THEMA Working Paper*, 2015-02.

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.

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

——— (2003): "Background Risk in Generalized Expected Utility Theory," *Economic Theory*, 22, 607–611.

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.

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

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

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

Sarin, R., and P. Wakker (1992): "A Simple Axiomatization of Nonadditive Expected Utility," *Econometrica*, 60(6), 1255–1272.

Sarin, R. K., and P. P. Wakker (1994): "Folding Back in Decision Tree Analysis," *Management Science*, 40, 625–628.

Schmeidler, D. (1989): "Subjective Probability and Expected Utility without Additivity," *Econometrica*, 57(3), 571–587.

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.

Snowberg, E., and J. Wolfers (2011): "Explaining the Favorite-Longshot Bias: Is it Risk-Love or Misperceptions?," *Journal of Political Economy*, 118(4), 723–746.

Sozou, P. (1998): "On Hyperbolic Discounting and Uncertain Hazard Rates," *Proceedings: Biological Sciences*, 265(1409), 2015–2020.

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

Trope, Y., and N. Liberman (2003): "Temporal Construal," *Psychological Review*, 110(3), 403–421.

Tversky, A., and C. R. Fox (1995): "Weighting Risk and Uncertainty," *Psychological Review*, 102(2), 269–283.

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

Tversky, A., and D. J. Koehler (1994): "Support Theory: A Nonextensional Representation of Subjective Probability," *Psychological Review*, 101(4), 547–567.

Tversky, A., and P. Wakker (1995): "Risk Attitudes and Decision Weights," *Econometrica*, 63(6), 1255–1280.

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.

Viscusi, W. K. (2010): "The Hold-up Problem. why it is Urgent to Rethink the Economics of Disaster Insurance Protection," *in: Erwann, M.-K. and Slovic, P. (Eds.), The Irrational Economist: Making Decisions in a Dangerous World*, pp. 142–148.

Wakker, P. (2004): "On the Composition of Risk Preference and Belief," *Psychological Review*, 111(1), 236–241.

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

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

Wu, G., and R. Gonzalez (1998): "Common Consequence Conditions in Decision Making under Risk," *Journal of Risk and Uncertainty*, 16, 115–139.

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