CESifo Working Papers

www.cesifo.org/wp

Once Upon a Time Preference
How Rationality and Risk Aversion Change the Rationale for Discounting

Christian P. Traeger

CESIFO WORKING PAPER NO. 3793
CATEGORY 2: PUBLIC CHOICE
APRIL 2012

CESifo *Working Paper No. 3793*

Once Upon a Time Preference
How Rationality and Risk Aversion Change the Rationale for Discounting

Abstract

The paper develops an axiomatic framework for rational decision making. The von Neumann-Morgenstern axioms give rise to a richer risk attitude than that captured in the standard discounted expected utility model. I derive three models that permit a more comprehensive risk evaluation. These preference representations differ regarding the consistency requirements that are imposed in the evaluation of uncertain scenarios. Imposing all rationality constraints jointly eliminates pure time preference from economic evaluation. The resulting preference representation still gives reduced weight to expected future utility when uncertainty increases over time. The more we know about the future welfare consequences of our (in)actions, the more weight they receive. If uncertainty is endogenous to the decision process, the new rationale for discounting will yield quite different policy implications than the discounted expected utility model based on pure time preference.

JEL-Code: D010, D600, D810, D910, H430, Q010, Q510, Q540.

Keywords: time preference, discounting, uncertainty, expected utility, recursive utility, risk aversion, intertemporal substitutability, stationarity, certainty additivity, temporal lotteries, intertemporal risk aversion, temporal resolution of risk, discount rate.

Christian P. Traeger
University of California, Berkeley
207 Giannini Hall #3310
USA – Berkeley, CA 94720-3310
traeger@berkeley.edu

April 13, 2012
I thank Larry Karp, Chris Shannon, David Ahn, Hans Gersbach, Juergen Eichberger, Ralph Winkler, Jeffrey LaFrance, and several seminar audiences for their helpful comments. Financial support by the German Research Foundation (DFG) and the German Foreign Exchange Service (DAAD) are gratefully acknowledged.

1 Introduction

Harrod (1948) stated that discounting utility represents "a polite expression for rapacity and the conquest of reason by passion" (40). The present paper shows how Harrod's statement is literally true in an axiomatic choice framework: reason constrains choice under conditions of uncertainty, and this constraint eliminates the free parameter associated with pure time preference, also known as the utility discount rate. This paper does not impose ethical axioms. The choice-restricting axioms in this paper are all satisfied under the standard economic model. The power of these restrictions is instead derived from a more comprehensive description of risk attitude.

1.1 Motivation, Model Ingredients, and Results

The discounted expected utility model is arguably the workhorse of economic modeling. An integral part of this standard economic model is the rate of pure time preference. A positive rate of pure time preference implies that a decision maker cares intrinsically less about the future than about the present. This devaluation of the future is not a consequence of real economic effects such as consumption growth, an individual's probability of death or society's extinction. Koopmans (1960) has shown that the rate of pure time preference arises as a free parameter in an axiomatic model of stationary decision making. Koopmans' setting is a certain world with an infinite planning horizon. In this setting, the rate of pure time preference is strictly positive. The underlying argument is familiar: pure time preference helps the welfare function to converge and makes it possible to order infinite consumption streams just as we would order consumption bundles today. However, a certain future is certainly the exception rather than the rule, and no decision maker will truly claim that he can predict an infinite consumption stream with certainty.

This paper drops the assumptions of a certain future and an infinite planning horizon. Moreover, it relaxes the assumption that risk aversion is only driven by our propensity to smooth consumption over time. In an intertemporal setting, the classic von Neumann & Morgenstern (1944) axioms give rise to a more comprehensive risk attitude. The asset pricing literature has exploited these more general preferences to solve a multitude of puzzles (Vissing-Jørgensen & Attanasio 2003, Bansal & Yaron 2004, Bansal, Kiku & Yaron 2010, Nakamura, Steinsson, Barro & Ursua 2010). There is no convincing normative argument for denying a decision maker a general risk attitude. I extend the discounted expected utility model to incorporate a general risk attitude and show that, unless we are willing to abandon rationality constraints, this step eliminates the rate of pure time preference. I summarize the three main components of my analysis in the following three paragraphs.

Uncertainty can influence choice and evaluation through two different channels. First, a decision maker who prefers smooth consumption paths is averse to stochasticity because it generates wiggles. Second, a decision maker can be intrinsically averse to risk. The discounted expected utility model neglects intrinsic risk aversion. It does so by assuming that a single function captures both, attitude to risk and attitude to consumption smoothing. The usual axioms of rational choice do not imply such a restriction. Empirical evidence suggests that people are indeed intrinsically averse to risk and that the Arrow Pratt measure of relative risk aversion is greater than the aversion to intertemporal substitution (measured by the inverse of the intertemporal elasticity of substitution). The concept of intertemporal risk aversion distills this assumption into an axiom whose plausibility the reader can easily evaluate.

The rate of pure time preference is defined only in models with stationary preferences. Stationarity of preferences over certain consumption paths implies that an agent has the same ranking of different alternatives today and in the future. The axiom implies that the utility function coincides in different periods, but the weight on utility can differ across periods. For a positive rate of pure time preference, future utility receives less weight than current utility. Risk stationarity of preferences also requires that the ranking of lotteries today and in the future remains constant. In the standard economic model, the assumption of risk stationarity has no more bite than the assumption of certainty stationarity: risk aversion in the standard model is solely generated by the preference for intertemporal consumption smoothing, which is already fully determined by choice under certainty. This situation crucially changes in a model with a comprehensive characterization of risk attitude. Here, risk stationarity restricts the way in which intrinsic risk aversion changes over time.

In a setting of general risk attitude, a decision maker can exhibit an intrinsic preference for early or late resolution of uncertainty (Kreps & Porteus 1978). A decision maker with a strict preference for early resolution of uncertainty is willing to pay to receive information earlier that is of no consequential use. A decision maker with a strict preference for late resolution of uncertainty is willing to pay for delaying this information. Such an intrinsic preference for early or late resolution cannot arise in the standard economic model. I argue that, indeed, a rational decision maker should not be willing to pay for information that is of no consequential use. Hence, I eliminate the subset of preferences giving rise to such behavior.

The paper shows how these ingredients eliminate the rate of pure time preference from economic evaluation. The intuition for this finding builds on a formalization of intrinsic risk aversion. For that purpose, I use Traeger's (2010) concept of intertemporal risk aversion, for which I present a simplified axiomatic characterization that makes use of stationarity. Stationarity, a common assumption in finite-horizon models, receives an axiomatic foundation only in an infinite horizon setting. I offer an axiomatic foundation for the case of a finite planning horizon. Here, an assumption of weak future independence buried in the stationarity assumption plays a prominent role. This assumption facilitates a general comparison of scenarios that are independent of the planning horizon. I show that a discounting of expected utility occurs even without pure time preference. This discounting, however, is a function of risk aversion and uncertainty about the future rather than the consequence of an exogenous impatience parameter.

1.2 Relation to the Literature

A series of prominent advocates have argued for a zero rate of pure time preference (Ramsey 1928, Pigou 1932, Harrod 1948, Koopmans 1963, Solow 1974, Broome 1992). All of their arguments depend on ethical reasoning and not on rationality constraints imposed on intertemporal decision making. Koopmans's (1960) wide-spread axiomatization of stationary decision making expresses the idea that the mere passage of time does not change preferences. His formulation crucially depends on the assumption of an infinite planning horizon and he shows how a strictly positive rate of pure time preference emerges. I offer an axiomatization of stationarity for decision makers with a finite, and possibly rolling, planning horizon. The axiom emphasizes an element of weak future independence. The infinite horizon setting with a strictly positive rate of pure time preference always satisfies this type of weak future independence. In a finite time horizon, however, weak future independence plays a more prominent role. It permits reasonable comparisons of long-term projects without discounting away the future. This future independence aspect of the axiom is not crucial for the finding that the pure time preference must equal zero.

My setup closely relates to that of Kreps & Porteus (1978). First, I employ Kreps & Porteus' (1978) temporal lotteries, which are generalized descriptions of uncertainty trees. Kreps & Porteus (1978) show that preferences over temporal lotteries can give rise to a (intrinsic) preference for early or late resolution of uncertainty. This finding gives rise to one of my axioms. Second, I derive a modified version of their preference representation. Kreps & Porteus' (1978) representation uses expected values to aggregate uncertainty but employs a nonlinear time aggregation. I introduce a non-linear uncertainty aggregation to keep time aggregation linear and to relate directly to the discounted utility model. The linear time step implies that the utility function represents the decision maker's propensity to smooth consumption over time. The nonlinearity of the risk aggregator represents intertemporal risk aversion, a concept introduced by Traeger (2010). The present paper provides a simpler axiomatic characterization of intertemporal risk aversion making use of stationarity. Epstein & Zin (1989) and Weil (1990) similarly disentangle risk attitude from the propensity to smooth consumption over time restriction attention to isoelastic functional forms.

I briefly characterize the differences among the present model, the Epstein-Zin-Weil model, and the common intertemporally additive expected utility model. The point of departure is the observation that risk aversion can derive from two different sources. First, agents with a propensity to smooth consumption over time dislike wiggles in the consumption path resulting from stochasticity. This is the only type of risk aversion captured in the discounted expected utility model. Second, agents can be intrinsically averse to risk. These agents dislike uncertainty merely for its being uncertain. Arrow Pratt risk aversion in the Epstein-Zin-Weil model measures the overall effect of risk on evaluation, i.e., the combined aversion produced by consumption fluctuation over time and intrinsic risk aversion. In contrast, intertemporal risk aversion measures intrinsic risk aversion directly. For the current analysis, this second measure has four advantages. First, intertemporal risk aversion is a (unidimensional) measure of risk aversion in a multidimensional setting. A one-commodity setting, such as that in the Epstein-Zin-Weil model, would seriously strengthen the assumption of additive separability on certain consumption paths by excluding habit formation. Second, the axiomatic characterization of intertemporal risk attitude directly captures the type of risk attitude that rules out pure time preference. Third, the set of axioms in this paper confines the risk aggregators to a class of constant absolute intertemporal risk aversion functions; it does not confine Arrow Pratt risk aversion to a simple functional class. Fourth, the change of intertemporal risk aversion over time is the key to understanding why risk stationarity, together with an indifference to the timing of risk resolution, rules out the rate of pure time preference. The isoelastic functional forms assumed by Epstein & Zin (1991) and Weil (1990) do not generally satisfy the axioms of this paper.

I found two insights in the literature related to my main result. Epstein (1992) explores the recursive approach to disentangling Arrow Pratt risk aversion from intertemporal substitutability. A note motivates this approach by pointing out that a disentanglement in a non-recursive approach would imply a contradiction to positive discounting in his infinite horizon setting. Moreover, Bommier (2007) points out that in a non-recursive model agents with a finite life time and a zero discount rate can exhibit more general risk aversion. Both authors work in a stationary setting and observe a tension among disentanglement, discounting, and a welfare function that I show to correspond to constant absolute intertemporal risk aversion. None of the authors derives an absence of pure time preference from imposing rationality axioms, including indifference to the timing of risk resolution and intertemporal risk aversion, as I do here. Moreover, none of them derives an intuition why the tension arises. My comparing of measures of intertemporal risk aversion between the risk stationary and the timing-indifferent representations explains the tension.

The risk-sensitive and robust control literature contains prominent examples for all three preference representations derived in this paper. Jacobsen (1973) and Whittle (1981) initiated this strand of the literature with the linear exponential quadratic Gaussian (LEQG) control problem. Their undiscounted objective function is precisely the preference representation satisfying all the axioms of the current paper (assuming, in addition, quadratic utility). Later, Whittle (1990, 2002), Bouakiz & Sobel (1992), and Hansen & Sargent (1995) introduced different versions of a discounted LEQG problem. Bouakiz & Sobel's (1992) and Whittle's (2002) suggestion corresponds to a quadratic utility special case of the representation in Theorem 4. These authors find that the control rules exhibit a strange time dependence. This time dependence is simply the result of a representation that abandons risk stationarity (or time consistency). Hansen & Sargent (1995) follow a recursive approach instead and suggest a representation that corresponds to Theorem 3 with quadratic utility. These authors conserve risk stationarity, but an agent with such preferences pays positive money for non-consequential information. The present paper shows that the parameter discussed as risk-sensitivity in this literature has an axiomatically well-founded interpretation: risk-sensitivity is the coefficient of absolute intertemporal risk aversion. The representations in this strand of literature all exhibit a constant coefficient of *absolute* intertemporal risk aversion, as opposed to the Epstein-Zin-Weil model, which exhibits a constant coefficient of *relative* intertemporal risk aversion. These LEQG settings can also be interpreted as the objective function of an agent playing a zero sum game against nature (Jacobsen 1973, Hansen & Sargent 2007), which gives rise to the term "robust control". Moreover, Hansen & Sargent (2001) show that the robust control models, corresponding to the constant absolute intertemporal risk aversion representations of the present paper, can be interpreted in light of the early ambiguity model of Gilboa & Schmeidler (1989). This so-called maximin expected utility model relaxes the independence axiom and represents preferences over sets of probabilities. A decision maker selects the act that maximizes the worst expected value with respect to any distribution in the set of priors. For a particular construction of this set of priors, the model is equivalent to the robust control approach. The current paper points out that this robustness, or risk-sensitivity, or ambiguity aversion, arises even within a standard preference representation building on the usual von Neumann & Morgenstern (1944) axioms and objective risk.

I briefly relate the concept of intertemporal risk aversion to three other extensions of risk aversion to a multi-commodity or intertemporal setting. First, Kihlstrom & Mirman (1974) introduce a directional coefficient of risk aversion, implying an Arrow Pratt measure of risk aversion that depends on the commodity consumed. In contrast, intertemporal risk aversion measures stay unidimensional also in a multi-commodity world. As opposed to the usual Arrow-Pratt measure (and its Kihlstrom-Mirman extension) the intertemporal risk aversion measure is also independent of the scale employed to measure a particular commodity. Second, Kihlstrom & Mirman (1981) introduce a univariate concept of risk aversion for homothetic utility functions. These authors employ the uniqueness of a least concave representation of homothetic preferences to characterize increasing and decreasing relative risk aversion. Quantitative risk aversion measures are somewhat difficult to define in this setting. Related to this second approach, intertemporal risk aversion obtains the necessary uniqueness of a utility function (covering more general preferences) from a deterministic intertemporal choice as opposed to the least concave property of (only) homothetic preferences.[1] Finally, Richard (1975) introduces a concept of "multivariate risk aversion". In his intertemporal interpretation, the measure is multivariate only in the time dimension. The axiomatic definition assumes a naturally ordered set of unidimensional payoffs. Richard's risk comparison is between two different risky consumption paths and has been reinterpreted as correlation aversion. It is easy to show that correlation aversion is a straight-forward consequence of intertemporal risk aversion and that intertemporal risk aversion can be used to extend correlation aversion to settings with multidimensional payoffs. The axiom of intertemporal risk aversion compares risky to certain scenarios and offers a new and simpler measure of the underlying risk aversion. In contrast to the papers discussed in this paragraph, I start from the general recursive setting of Kreps & Porteus (1978), as is done in the Epstein-Zin-Weil model. The current paper shows that the (time consistent subset of) non-recursive approaches discussed in this paragraph emerge once again as the special cases corresponding to constant absolute intertemporal risk aversion.

The motivation of von Neumann & Morgenstern's (1944) framework and of the present axiomatic system slightly differs from the current main-stream decision theoretic literature, which has a strong focus on axioms that give rise to observed behavior. These papers do not usually assume a separation between the description of uncertainty and the evaluation of outcomes. In contrast, the current paper develops a decision-support framework for an agent or an agency. It therefore builds deliberately on the separation of uncertainty description from of outcome evaluation, developing a rational and consistent approach that helps the agent to rank different scenarios. It suffices to assume that decision makers use such a model when facing objectively given probabilities to eliminate pure time preference also from a more comprehensive framework.

Section 2 introduces the setup, axiomatic background, and concept of uncertainty aggregation rules. Section 3 discusses certainty stationarity and the resulting preference

[1]Both approaches only result in a uniqueness up to affine transformations. As opposed to the discounted expected utility model, affine transformations can matter for choice under risk. While Kihlstrom & Mirman (1981) normalize affine transformation away by assumption, Traeger (2010) discusses explicitly the meaning of different affine transformations of utility and normalizations.

representation. Section 4 introduces the concept of intertemporal risk aversion. Section 5 elaborates the consequences of a stationary evaluation of uncertain scenarios. Section 6 analyzes the consequences of indifference to the timing of risk resolution for certainty stationary preferences. Section 7 brings together risk stationarity and indifference to the timing of risk resolution and derives the main result of the paper and its consequences for evaluation. Section 8 gives a critical discussion of the model assumptions. Section 9 concludes. Proofs are provided in an appendix.

2 Preliminaries

Section 2 lays out the setup and the common axiomatic background for the representations in this paper. It also introduces nonlinear uncertainty aggregation rules that capture intertemporal risk aversion without abandoning a time-additive evaluation.

2.1 Setup

Let X be a connected compact metric space that characterizes factors determining welfare within a period. These factors can be consumption levels as well as more abstract descriptions of quality or the state of an ecosystem. I refer to the elements x of X as outcomes. Time is discrete with planning horizon $T \in \mathbb{N}$. The space $\mathsf{X}^t = X^{T-t+1}$ denotes the Cartesian product equipped with the product metric and characterizes the set of all certain consumption paths from period t to period T.[2] A consumption path $\mathsf{x} \in \mathsf{X}^t$ is written $\mathsf{x} = (\mathsf{x}_t, \mathsf{x}_{t+1}, ..., \mathsf{x}_T) = (x_t, x_{t+1}, ..., x_T)$. Given $\mathsf{x} \in \mathsf{X}^t$, I define $(\mathsf{x}_{-i}, x) = (\mathsf{x}_t, ..., \mathsf{x}_{i-1}, x, \mathsf{x}_{i+1}, ..., \mathsf{x}_T) \in \mathsf{X}^t$ as the consumption path that coincides with x in all but the i^{th} period, in which it yields outcome x.

Uncertain outcomes are modeled as temporal lotteries (Kreps & Porteus 1978). For any compact metric space Y, let $\Delta(Y)$ denote the set of Borel probability measures equipped with the Prohorov metric (giving rise to the topology of weak convergence).[3] I define $\tilde{X}_T = X$ and recursively $\tilde{X}_{t-1} = X \times \Delta(\tilde{X}_t)$ for all $t \in \{2, ..., T\}$, where each $\tilde{X}_t$ is equipped with the product metric. I denote by $P_t = \Delta(\tilde{X}_t)$ the space of uncertain scenarios

[2] I do not distinguish different sets of outcomes for different periods. X stands for the union of all possible outcomes perceivable in any period.

[3] I suggest an epistemic foundation of probabilities as found for example in Cox (1946,1961) or Jaynes (2003). Here probabilistic beliefs replace the notion of a binary logic. Lotteries do not (only) describe draws from an urn, but correspond to general characterizations of uncertainty with respect to possible outcomes. However, thinking of the probabilities more narrowly as representing objective lotteries is enough to imply the result of vanishing time preference.

considered in period t and refer to the elements $p_t \in P_t$ as (period t) lotteries. Observe that in every period, the decision maker has a *probability distribution over* the *outcome* in the respective period *and* the *probability distribution* over the future faced in the next period. For discrete probabilities, P_t is a general characterization of an uncertainty tree. The set of degenerate lotteries in P_t is identified with the set $\tilde{X}_t$ of (generalized or multiperiod) outcomes in the usual way. A lottery yielding outcome $\tilde{x}_t$ with probability λ and outcome $\tilde{x}'_t$ with probability $1-\lambda$ is written as $\lambda\tilde{x}_t + (1-\lambda)\tilde{x}'_t \in P_t$. A 'plus' sign between outcomes always characterizes a lottery.[4] Preferences in period t are defined on the set P_t and denoted by $\succeq_t$.[5] For convenience in presentation, I assume the existence of at least two non-indifferent consumption paths in the decision maker's choice set:

A0 (non-degeneracy) There exist $\mathsf{x}, \mathsf{x}' \in \mathsf{X}^1$ such that $\mathsf{x} \not\sim_1 \mathsf{x}'$.

Preferences $\succeq$ will abbreviate a sequence of binary relations $\succeq \equiv (\succeq_t)_{t\in\{1,...,T\}}$ on $(P_t)_{t\in\{1,...,T\}}$ that satisfy axiom A0.

The representation will employ continuous functions $u : X \rightarrow \mathbb{R}$ and $\tilde{u}_t : \tilde{X}_t \rightarrow \mathbb{R}$ for evaluating outcomes x, respectively $\tilde{x}_t$, and I denote $U = [\underline{U}, \overline{U}] = \text{range}(u)$ and $\tilde{U}_t = \text{range}(\tilde{u}_t)$. For any metric space Y, I denote by $\mathcal{C}^0(Y)$ the set of continuous functions from Y into the reals. Finally, the group of strictly positive affine transformations is denoted $\mathsf{A} = \{\mathfrak{a} : \mathbb{R} \rightarrow \mathbb{R} : \mathfrak{a}(z) = a\,z + b\,,\ a, b \in \mathbb{R}, a > 0\}$ with elements $\mathfrak{a} \in \mathsf{A}$.

2.2 Axiomatic Background

The paper assumes throughout that the von Neumann Morgenstern axioms, additive separability on certain consumption paths, and time consistency are satisfied. I briefly review these axioms below.

The first three axioms are close relatives to the axioms suggested by von Neumann & Morgenstern (1944) in an atemporal framework for choice under uncertainty.

A1 (weak order) For all $t \in \{1, ..., T\}$ preferences $\succeq_t$ are transitive and complete, i.e.,

– transitive: $\forall\, p, p', p'' \in P_t :\ p \succeq_t p'$ and $p' \succeq_t p'' \Rightarrow p \succeq_t p''$.

– complete: $\forall\, p, p' \in P_t : p \succeq_t p'$ or $p' \succeq_t p$.

[4]As X and $\tilde{X}_t$ are only assumed to be compact metric there is no immediate addition defined for their elements. In case the spaces are additionally equipped with a vector space or field structure, the vector composition will not coincide with the "+" used here. The "+" sign used here alludes to the additivity of probabilities.

[5]The relations $\succeq_t$ are required to be reflexive (axiom A1). The asymmetric part is denoted by $\succ_t$ and interpreted as strict preference. The symmetric part of the relation $\succeq_t$ is denoted by $\sim_t$ and interpreted as indifference. Non-indifference is denoted by $\not\sim_t$ and defined as $\not\sim_t \equiv P_t \times P_t \setminus \sim_t$.

Axiom A1 requires a full ranking of all lotteries (completeness). Moreover, if one lottery is preferred to a second and the second is preferred to a third, then the first lottery should also be preferred to the third (transitivity). In a behavioral context, complexity usually precludes agents from having complete preferences. However, in a normative context, we require rational choice constraints to hold on the full set of choices.

A2 (independence) For all $t \in \{1, ..., T\}$ and for all $p, p', p'' \in P_t$:
$$p \sim_t p' \quad \Rightarrow \quad \lambda\, p + (1-\lambda)\, p'' \sim_t \lambda\, p' + (1-\lambda)\, p'' \quad \forall\, \lambda \in [0,1] \quad .$$

The independence axiom states the following. Let a decision maker be indifferent between a lottery p and another lottery p'. Offer him two compound lotteries, both starting out with the toss of a biased coin (tails comes up with probability λ). If heads comes up, the decision maker enters the same lottery p'' in both lotteries. If tails comes up, the decision maker faces lottery p in the first compound lottery and lottery p' in the second. Recalling that the decision maker is indifferent between lotteries p and p', the independence axiom requires the decision maker to be indifferent between the two compound lotteries as well.[6]

A3 (continuity) For all $t \in \{1, ..., T\}$ and for all $p \in P_t$:
The sets $\{p' \in P_t : p' \succeq_t p\}$ and $\{p' \in P_t : p \succeq_t p'\}$ are closed in P_t.

Continuity A3 assures that infinitesimally small changes in the probabilities do not result in finitely large changes in the evaluation. In particular, continuity implies the slightly weaker Archimedian axiom used by von Neumann & Morgenstern (1944).

To match the widespread model of additively separable utility over time on certain consumption paths, I introduce the following axiom taken from Wakker (1988):

A4 (certainty separability)
i) For all $\mathsf{x}, \mathsf{x}' \in \mathsf{X}^1$, $x, x' \in X$ and for all $t \in \{1, ..., T\}$:
$$(\mathsf{x}_{-t}, x) \succeq_1 (\mathsf{x}'_{-t}, x) \;\Leftrightarrow\; (\mathsf{x}_{-t}, x') \succeq_1 (\mathsf{x}'_{-t}, x')\,.$$
ii) If $T = 2$ additionally: For all $x_t, x'_t, x''_t \in X$, $t \in \{1, 2\}$
$$(x_1, x_2) \sim_1 (x'_1, x''_2) \;\wedge\; (x'_1, x'_2) \sim_1 (x''_1, x_2) \;\Rightarrow\; (x_1, x'_2) \sim_1 (x''_1, x''_2)\,.$$

Wakker (1988) calls part i of the axiom coordinate independence. It requires that the choice between two consumption paths does not depend on period t consumption whenever

[6] For a discussion of descriptive limitations of the independence axiom see e.g. Starmer (2000). Note that the independence axiom can be replaced by a collection of much weaker axioms elaborated by Chew & Epstein (1989, 110) in the case of indifference to the timing of risk resolution, a setting treated in sections 6 and 7.

period t consumption coincides for both paths. Part *ii* is known as the Thomsen condition. It is required only if the model is limited to $T = 2$ periods.[7] Axiom 4 allows for a certainty additive representation of the form $\sum_{t=1}^{T} u_t(\mathrm{x}_t)$. At this stage, the axioms allow the representing utility functions u_t to change arbitrarily over time. I will later impose stationarity assumptions that relate utility functions over time.

Preferences in different periods are connected by the following consistency axiom adapted from Kreps & Porteus (1978):

A5 (time consistency) For all $t \in \{1, ..., T-1\}$:

$$(x_t, p_{t+1}) \succeq_t (x_t, p'_{t+1}) \quad \Leftrightarrow \quad p_{t+1} \succeq_{t+1} p'_{t+1} \quad \forall\, x_t \in X,\; p_{t+1}, p'_{t+1} \in P_{t+1}\ .$$

This axiom requires that a decision maker who in period t prefers a certain outcome followed by lottery p_{t+1} rather than p'_{t+1} also prefers p_{t+1} in period $t+1$. In constraining a decision support framework, time and time consistency most importantly capture how the decision maker thinks that he should connect preferences over time to obtain a rationally desirable choice model.[8]

2.3 Uncertainty Aggregation Rules

The preference representations in this paper rely on non-linear aggregation of utility over uncertainty. For any $t \in \{1, ..., T\}$ and strictly monotonic function $f_t \in \mathcal{C}^0(\mathrm{I\!R})$, I define an *uncertainty aggregation rule* as the functional $\mathcal{M}^{f_t} : \Delta(\tilde{X}_t) \times \mathcal{C}^0(\tilde{X}_t) \to \mathrm{I\!R}$ with[9]

$$\mathcal{M}^{f_t}(p_t, \tilde{u}_t) = f_t^{-1} \int_{\tilde{X}_t} f_t \circ \tilde{u}_t \; dp_t\ .$$

The uncertainty aggregation rule takes as input the decision maker's perception of uncertainty, expressed by a probability measure p_t on $\tilde{X}_t$, and a valuation of the degenerate outcomes expressed by a real valued function $\tilde{u}_t$ on $\tilde{X}_t$. The uncertainty aggregation rule weighs these utility values with the function f_t, aggregates them, and applies the inverse

[7] In the case of two periods parts *i*) and *ii*) can also be replaced by the single requirement of triple cancellation (see Wakker 1988, 427). Moreover, part *ii*) can be replaced by a slightly weaker but slightly more involved assumption known as the hexagon condition (Wakker 1989, 47 et sqq., 67).

[8] Note that explicit decision nodes can be introduced at any point in the uncertainty trees the same way as done in Kreps & Porteus (1978). Optimal choices in the framework always correspond to the best (sub-) tree and there is no explicit preference for flexibility as e.g. in Kreps (1979). Therefore, no additional insight derives from explicitly introducing decision nodes and the more complicated notation would be obstructive. The application of the preference representations in a dynamic programming framework with decision making in every period is immediate.

[9] By continuity of $f_t \circ \tilde{u}_t$ and compactness of $\tilde{X}_t$, Lesbeque's dominated convergence theorem ensures integrability (Billingsley 1995, 209).

of f_t to renormalize the resulting expression. For degenerate outcomes, an uncertainty aggregation rule returns the value of $\tilde{u}_t$ itself, i.e., $\mathcal{M}^{f_t}(\tilde{x}_t, \tilde{u}_t) = \tilde{u}_t(\tilde{x}_t)$.

The simplest example of an uncertainty aggregation rule is the expected value operator, which is obtained for $f_t = \text{id}$. An uncertainty aggregation rule that will be generated by different axioms in this paper corresponds to the parameterization $f_t(z) = \exp(-\xi z) = \big(\exp(z)\big)^{-\xi}$ and yields

$$\mathcal{M}^{\exp^{-\xi}}(p_t, \tilde{u}_t) = \frac{1}{-\xi} \ln \left[\int_{\tilde{X}_t} \exp(-\xi \tilde{u}_t) \, dp_t \right].$$

It is defined for $\xi \in \mathbb{R}$ with $\mathcal{M}^{\exp^0}(p_t, \tilde{u}_t) \equiv \lim_{\xi \to 0} \mathcal{M}^{\exp^{-\xi}}(p_t, \tilde{u}_t) = E_{p_t} \tilde{u}_t$ (see proof of Theorem 3). In the limit of ξ going to plus or minus infinity, the uncertainty aggregation rule $\mathcal{M}^{\exp^{-\xi}}$ only considers the extreme outcomes (abandoning continuity in the probabilities): $\lim_{\xi \to \infty} \mathcal{M}^{\exp^{-\xi}}(p_t, \tilde{u}_t) = \min_{\tilde{x}_t} \tilde{u}_t(\tilde{x}_t)$ and $\lim_{\xi \to -\infty} \mathcal{M}^{\exp^{-\xi}}(p_t, \tilde{u}_t) = \max_{\tilde{x}_t} \tilde{u}_t(\tilde{x}_t)$. In general, a larger ξ results in a lower value (or certainty equivalent utility) of the respective uncertainty aggregation rule.[10]

3 Certainty Stationarity

The section introduces stationarity for settings with a finite planning horizon. First, I introduce the axiomatic characterization. Then, I give a certainty stationary preference representation that admits a more general risk attitude than the standard model.

3.1 Axiomatic Description for a Finite Planning Horizon

An almost ubiquitous assumption in economic modeling is that the utility evaluation of consumption coincides in different periods (discounted utility model). Although this assumption is usually adopted independently of an agent's planning horizon, the assumption has only received an axiomatic foundation in the case of an infinite planning horizon. In the latter setting, Koopmans (1960) has shown that the rate of pure time preference must be constant and strictly positive.

Koopmans' characterization of stationarity requires that a decision maker prefer a consumption path x over another consumption path x' in the present if, and only if, he prefers a consumption path (x^0, x) over a consumption path (x^0, x') in the present (Koopmans 1960, see page 13 below for details). Such an axiomatization rests on the fact that for an infinite time horizon, adding an additional outcome does not change the

[10] The uncertainty aggregation rule $\mathcal{M}^{f_t}$ is induced by the so called quasi-arithmetic or generalized mean in a way made precise in Traeger (2010).

length of a consumption path. Both paths, x and (x^0, x), are elements of X^∞ and can be compared by the same preference relation. In contrast, with a finite planning horizon, the paths x and (x^0, x) differ in length and, thus, cannot be compared by means of the same preference relation $\succeq$. The following axiom is applicable in the setting with a finite planning horizon and, there, yields the standard discounted utility model for the evaluation of certain consumption paths.

A6 (certainty stationarity) For all $\mathsf{x}, \mathsf{x}' \in \mathsf{X}^2$ and all $x \in X$:

$$(\mathsf{x}, x) \succeq_1 (\mathsf{x}', x) \quad \Leftrightarrow \quad \mathsf{x} \succeq_2 \mathsf{x}' \,. \tag{1}$$

On the right-hand side of the equivalence, the decision maker faces a comparison between consumption paths x and x' starting in period 2. On the left-hand side of the equivalence, the decision maker faces a comparison between consumption paths x and x' starting in period 1. The additional outcome x is jointly added so that the resulting paths, (x, x) and (x', x), have the appropriate length for comparing them in period 1 by means of the preference relation $\succeq_1$ of a decision maker with time horizon T. The essence of the axiom is that a decision maker anticipates ranking (certain) consumption paths tomorrow in the same way as he does today.[11]

The best interpretation of axiom A6 adopts a rolling time horizon and splits the equivalence (1) into two separate statements. Assume that a decision maker in period 1, planning with time horizon T, prefers consumption plan (x, x) over plan (x', x). Confront him in period 2 with the exact same consumption paths, (x, x) and (x', x). For this purpose, let the decision maker plan ahead in period 2 the same number of periods as he does in period 1, implying the new time horizon $T + 1$. I denote these preferences of the decision maker in period 2 with time horizon $T + 1$ by $\succeq_{2|T+1}$. I assume that the decision maker ranks (or plans to rank) the consumption paths in both choice situations the same way. Requiring the latter for all consumption paths yields the condition

$$(\mathsf{x}, x) \succeq_{1|T} (\mathsf{x}', x) \quad \Leftrightarrow \quad (\mathsf{x}, x) \succeq_{2|T+1} (\mathsf{x}', x) \tag{2}$$

for all $\mathsf{x}, \mathsf{x}' \in \mathsf{X}^2$ and $x \in X$. Condition (2) most clearly captures the intuition of stationarity in the sense that the mere passage of time does not change the evaluation.

However, so far, assumption (2) does not imply any restriction on the set of preferences. The second step in the reasoning relates the preference relation $\succeq_{2|T+1}$ to the relation $\succeq_2$ $= \succeq_{2|T}$. Both preference relations specify how the decision maker evaluates (or anticipates evaluating) consumption plans over the future in period 2. The relation $\succeq_{2|T}$ specifies his ranking when he plans ahead until period T, and the relation $\succeq_{2|T+1}$ states his ranking

[11]Note the difference to time consistency. The latter is a condition on two consumption paths starting in the same period, which yield a common outcome in the first period. Then, the *continuation* of the paths in the next period should be ranked the same way as the complete paths in the earlier period.

when he plans ahead until period $T+1$. Accepting stationarity in the rolling time horizon sense of equation (2), axiom A6 requires the following relation between $\succeq_{\cdot|T}$ and $\succeq_{\cdot|T+1}$:

$$\mathrm{x} \succeq_{2|T} \mathrm{x}' \quad \Leftrightarrow \quad (\mathrm{x}, x) \succeq_{2|T+1} (\mathrm{x}', x) \tag{3}$$

for all $\mathrm{x}, \mathrm{x}' \in \mathrm{X}^2$ and $x \in X$. If two scenarios are evaluated with a time horizon of $T+1$ and yield the same outcome in period $T+1$, then an evaluation based only on a time horizon T yields the same ranking of the scenarios. The assumption is a weak form of future independence.

3.2 Relation to the Infinite Planning Horizon Setting

This subsection relates the axiomatic reasoning above to the situation in which agents have an infinite planning horizon. Instead of the finite consumption paths (x, x) and (x', x), the decision maker now faces the infinite paths $\mathrm{x}^\infty, \mathrm{x}'^\infty \in X^\infty$. By *time consistency*, the right-hand side of equation (2) is equivalent to $(x, \mathrm{x}^\infty) \succeq_{1|\infty+1} (x, \mathrm{x}'^\infty)$ for all $\mathrm{x}^\infty, \mathrm{x}'^\infty \in X^\infty$ and $x \in X$. However, in the infinite horizon setting $\succeq_{1|T+1} = \succeq_{1|\infty+1} = \succeq_{1|\infty} = \succeq_{1|T}$ and, thus, equation (3) becomes dispensable. Combining the two equations implies the standard axiom of stationarity for the infinite planning horizon:

$$\mathrm{x}^\infty \succeq_{1|\infty} \mathrm{x}'^\infty \Leftrightarrow (x, \mathrm{x}^\infty) \succeq_{1|\infty} (x, \mathrm{x}'^\infty) \tag{4}$$

for all $x \in X$ and all $\mathrm{x}^\infty, \mathrm{x}'^\infty \in X^\infty$, dating back to Koopmans (1960, 294)[12]. Hence, at first sight, the assumption corresponding to equation (3) does not seem necessary with an infinite planning horizon. However, this conjecture is not correct. To satisfy equation (4) in a standard representation, i.e., equation (2) with an infinite time horizon, a strictly positive rate of pure time preference is necessary. Then, the weight given to future consumption converges to zero, and the coinciding outcomes in the "last" period of the planning horizon are insignificant to the ranking. Thus, under an infinite planning horizon and a positive discount rate, equation (2) already implies equation (3). In summary, the weak future independence characterized in equation (3) is implicit not only in axiom A6 but also in the infinite time horizon stationarity axiom. The time horizon here is arbitrary and it can be pushed arbitrarily far into the future.

3.3 The Certainty Stationary Representation

The following representation holds for preferences satisfying axioms A1-A5 and a stationary evaluation of certain consumption paths in the sense of axiom A6.

[12]Koopmans (1960) formulates his postulates in terms of utility functionals, but the translation of his postulate 4 into the preference setup is immediate. Koopmans (1972) carries out this translation of his general axiomatic framework into a preference setting, stationarity still corresponding to postulate 4.

Theorem 1: The preferences $\succeq$ satisfy axioms A1-A5 and certainty stationarity A6, if, and only if, there exist a continuous function $u : X \to \mathbb{R}$, a strictly positive discount factor $\beta \in \mathbb{R}_{++}$, and strictly increasing and continuous functions $f_t : \mathbb{R} \to \mathbb{R}$ for all periods $t \in \{1, ..., T\}$ that represent preferences as follows:
Define the functions $\tilde{u}_t : \tilde{X}_t \to \mathbb{R}$ for $t \in \{1, ..., T\}$ by $\tilde{u}_T(x_T) = u(x_T)$ and recursively

$$\tilde{u}_{t-1}(x_{t-1}, p_t) = u(x_{t-1}) + \beta\, \mathcal{M}^{f_t}(p_t, \tilde{u}_t)\ , \tag{5}$$

then for all periods $t \in \{1, ..., T\}$

$$p_t \succeq_t p'_t \ \Leftrightarrow\ \mathcal{M}^{f_t}(p_t, \tilde{u}_t) \geq \mathcal{M}^{f_t}(p'_t, \tilde{u}_t) \quad \forall\, p_t, p'_t \in P_t\ . \tag{6}$$

Moreover, the functions $u, f_1, ..., f_T$ and $u', f'_1, ..., f'_T$ both represent $\succeq$ in the above sense, if and only if, there exist positive affine transformations $\mathfrak{a}_0, \mathfrak{a}_1, ..., \mathfrak{a}_T \in \mathsf{A}$ such that $u' = \mathfrak{a}_0 u$ and $f'_t|_{\tilde{U}'_t} = \mathfrak{a}_t f_t|_{\tilde{U}_t}\ {\mathfrak{a}_0}^{-1}$.

The representation theorem recursively constructs an aggregate utility function $\tilde{u}_t$ that depends on the utility in the respective period $u(x_t)$ and on the welfare derived from a particular lottery p_{t+1} over the future. Whereas aggregation of current utility and future welfare is additive in all periods, uncertainty evaluation relies on the generally nonlinear uncertainty aggregation rules $\mathcal{M}^{f_t}$. Graphically, constructing aggregate period t welfare recursively works its way through the uncertainty tree corresponding to the lottery p_t. Starting in the last period, the representation evaluates possible outcomes by u and aggregates over last subtrees of the last period by employing the uncertainty aggregation rule $\mathcal{M}^{f_T}$. The second step evaluates possible outcomes in the preceding period, again by u, and adds the values to the discounted utility of the attached subtree (yielding the $\tilde{u}_{T-1}$ values). Uncertainty evaluation and time aggregation alternate until the aggregate utility for period t is obtained. A non-recursive evaluation of the decision trees is only possible for special cases, which are discussed in sections 6 and 7.

For degenerate lotteries, $\mathcal{M}^{f_t}(\tilde{x}_t, \tilde{u}_t) = \tilde{u}_t(\tilde{x}_t)$ holds. Therefore, the evaluation of certain consumption paths simply becomes

$$\mathrm{x}^t \succeq_t \mathrm{x}'^t \quad \Leftrightarrow \quad \sum_{\tau=t}^{T} \beta^{\tau-t}\, u(\mathrm{x}^t_\tau) \geq \sum_{\tau=t}^{T} \beta^{\tau-t}\, u(\mathrm{x}'^t_\tau)\,. \tag{7}$$

Thus, for certain consumption paths, the model coincides with the standard discounted utility framework. Observe that the discount factor β in Theorem 1 can be smaller, greater, or equal to unity. For uncertain consumption, Theorem 1 gives a more general evaluation framework. This theorem avoids the implicit assumptions of risk neutrality present in the intertemporally additive discounted expected utility standard model. This risk neutrality assumption is discussed in the next section.

4 Intertemporal Risk Aversion

Traeger (2010) introduces the concept of intertemporal risk attitude in a general framework with non-stationary preferences. This section gives a slightly simplified definition for decision makers with stationary preferences. The concept of intertemporal risk attitude yields a one-dimensional measure of risk aversion for a multi-commodity world. In contrast to the Arrow Pratt measure of risk aversion, it is independent of the measurement scale used for the goods under observation. I start by giving an axiomatic characterization. Then, I characterize intertemporal risk attitude in the representation of Theorem 1.

4.1 Axiomatic Characterization of Intertemporal Risk Attitude

This section employs a special notation for constant consumption paths. The path $\bar{\mathrm{x}}^t = (\bar{x}, \bar{x}, ..., \bar{x})$ denotes the certain constant consumption path that gives outcome $\bar{x}$ from t until T. Recall that for $\mathrm{x} \in \mathrm{X}^t$ the consumption path $(\bar{\mathrm{x}}_{-i}, \mathrm{x}_i)$ denotes the consumption path where the i^{th} entry of $\bar{\mathrm{x}}$ is replaced by outcome x_i. Then, $\sum_{i=t}^{T} \frac{1}{T-t+1} (\bar{\mathrm{x}}_{-i}, \mathrm{x}_i)$ describes a lottery that delivers consumption paths with outcomes $\bar{x}$ in all but one period. The risk state i determines the period in which the outcome $\bar{x}$ is replaced by outcome x_i. The lottery draws this period i from $\{t, ..., T\}$ with equal probability. The following is a two-period example of such a lottery. An economic agent seeks employment for two years. He has the choice between two different contracts. The first contract offers him a short vacation in both years. The second contract offers him a long vacation during the first year and no vacation during the second year (terminating the contract after one year is not possible). The distinct entries of his consumption paths are $\bar{\mathrm{x}} = (\text{short vacation}, \text{short vacation})$ and $\mathrm{x} = (\text{long vacation}, \text{no vacation})$. The corresponding lottery constructed as above becomes $\sum_{i=1}^{2} \frac{1}{2} (\bar{\mathrm{x}}_{-i}, \mathrm{x}_i) = \frac{1}{2}\,(\text{long vacation}, \text{short vacation}) + \frac{1}{2}\,(\text{short vacation}, \text{no vacation})$.

Now, assume (or adjust the vacation length so that) the agent is indifferent between the two contracts. Then, the agent exhibits intertemporal risk aversion if he prefers either of the two certain paths over the lottery, in which he might end up with more vacation but might also end up with less vacation. In general, a decision maker exhibits *weak intertemporal risk aversion* in period $t < T$ if, and only if, the following axiom is satisfied:

A7$^{\mathrm{w}}$ (weak intertemporal risk aversion) For all $\bar{\mathrm{x}}, \mathrm{x} \in \mathrm{X}^t$

$$\bar{\mathrm{x}} \sim_t \mathrm{x} \quad \Rightarrow \quad \bar{\mathrm{x}} \succeq_t \sum_{i=t}^{T} \frac{1}{T-t+1} (\bar{\mathrm{x}}_{-i}, \mathrm{x}_i) \ .$$

A decision maker exhibits *strict intertemporal risk aversion* in period $t < T$ if, and only if, the following axiom is satisfied:

A7$^{\text{s}}$ (strict intertemporal risk aversion) For all $\bar{\mathrm{x}}, \mathrm{x} \in \mathrm{X}^t$

$$\bar{\mathrm{x}} \sim_t \mathrm{x} \quad \wedge \quad \exists\, \tau \in \{t, ..., T\} \text{ s.th. } (\bar{\mathrm{x}}_{-\tau}, \mathrm{x}_\tau) \not\sim_t \bar{\mathrm{x}}$$

$$\Rightarrow \quad \bar{\mathrm{x}} \ \succ_t \ \textstyle\sum_{i=t}^{T} \frac{1}{T-t+1} \left(\bar{\mathrm{x}}_{-i}, \mathrm{x}_i\right) .$$

I start with the interpretation of the strict axiom. The first part of the premise in axiom A7$^{\text{s}}$ states that a decision maker is indifferent between a constant consumption path delivering outcome $\bar{x}$ in every period and the consumption path x. The second part of the premise requires that at least some outcome x_τ is considered non-indifferent to $\bar{x}$ so that the path $(\bar{\mathrm{x}}_{-\tau}, \mathrm{x}_\tau)$ is non-indifferent to the path $\bar{\mathrm{x}}$. Together, the two assumptions in the premise imply that some of the reassembled consumption paths $(\bar{\mathrm{x}}_{-i}, \mathrm{x}_i)$ with $i \in \{t, ..., T\}$ are preferred to the constant path, while others are judged inferior. The outcomes x_t that are better than and those that are worse than $\bar{x}$ balance such that receiving them all with certainty makes the decision maker indifferent with respect to the constant consumption path. The second line of axiom A7$^{\text{s}}$ states that an intertemporal risk averse decision maker prefers the certain constant consumption path over the lottery that exchanges, with equal probability, one of the outcomes. For some draws of the period i, the decision maker will receive a consumption path that makes him better off, but for others, he receives a path that makes him worse off.

The interpretation of the *weak* axiom A7$^{\text{w}}$ is analogous, with the only difference that the consumption path x is allowed to coincide with $\bar{\mathrm{x}}$, and that the implication only requires that the lottery is not strictly preferred to the certain and constant consumption path. If axiom A7$^{\text{s}}$ [A7$^{\text{w}}$] is satisfied with $\succ_t$ [$\succeq_t$] replaced by $\prec_t$ [$\preceq_t$], the decision maker is called a strong [weak] *intertemporal risk seeker* in period t. I say that a decision maker is intertemporal risk averse (seeking) if his preferences satisfy intertemporal risk aversion (seeking) in all periods. If a decision maker satisfies weak intertemporal risk aversion as well as weak intertemporal risk seeking, the decision maker is called *intertemporal risk neutral.*

The representation of Theorem 1 invites an alternative interpretation of axiom A7$^{\text{s}}$. Here, the first part of the premise requires that for two consumption paths $\bar{\mathrm{x}}$ and x, the discounted per period utility adds up to the same aggregate utility (see equation 7). The second part of the premise requires that in at least one period, the utility gained from consumption path x differs from the utility gained from the outcome $\bar{x}$. Then, the lottery in axiom A7$^{\text{s}}$ delivers *in expectation* the same utility as the certain consumption path $\bar{\mathrm{x}}$. A decision maker is defined to be strictly intertemporal risk averse if he prefers the certain consumption path $\bar{\mathrm{x}}$ over the lottery that leaves him either worse or better off, and yields the same utility as the certain consumption path in expectation. Here, intertemporal risk aversion can be interpreted as risk aversion with respect to welfare gains and losses, where welfare is described by the certainty additive utility function.

4.2 Functional Characterization of Intertemporal Risk Attitude

The following theorem characterizes intertemporal risk aversion in terms of the preference representation of Theorem 1.

Theorem 2: Let the functions $f_1, ..., f_T$ represent preferences $\succeq$ in the sense of Theorem 1. The following assertions hold for all $t \in \{1, ..., T-1\}$:
a) A decision maker is strictly intertemporal risk averse [seeking] in period t in the sense of axiom A7$^{\text{s}}$ if, and only if, $f_t|_{\tilde{U}_t}$ is strictly concave [convex].
b) A decision maker is weakly intertemporal risk averse [seeking] in period t in the sense of axiom A7$^{\text{w}}$ if, and only if, $f_t|_{\tilde{U}_t}$ is concave [convex].
c) A decision maker is intertemporal risk neutral in period t if, and only if, $f_t|_{\tilde{U}_t}$ is linear.

Intertemporal risk attitude is described by the curvature of the functions f_t. As pointed out at the end of the preceding section, in the representation of Theorem 1, intertemporal risk aversion can be interpreted as risk aversion with respect to utility gains and losses. Thus, the standard intuition associating concavity with risk aversion applies. In contrast to measures of atemporal risk aversion, the argument of f_t is not measured in consumption but in utility units. Precisely, it is measured in *current value* certainty additive utility units: combining equations (5) and (6) shows that f_t weights utility $\tilde{u}_t(x_t, p_{t+1}) = u(x_t) + \beta \mathcal{M}^{f_t}(p_{t+1}, \tilde{u}_{t+1})$, where utility from period t consumption is undiscounted and in current value units.

Combining part *c* of Theorem 2 with Theorem 1 gives an axiomatic characterization of the discounted expected utility standard model: if intertemporal risk neutrality is required in all periods,[13] then f_t is linear in all periods. A linear parameterization makes the uncertainty aggregation rule itself linear, i.e., $\mathcal{M}^{f_t} = \text{E}$. With linear uncertainty aggregation, the lotteries no longer have to be evaluated recursively, and the expected value operator in Theorem 1 can be pulled to the first period (see also section 6).[14]

To derive a quantitative characterization of risk attitude, I define for a twice-differentiable function f_t the measures of *relative intertemporal risk aversion* as the functions RIRA^*_t :

[13] Instead of characterizing intertemporal risk neutrality by joint weak intertemporal risk aversion and seeking, it can obviously be defined by replacing $\succeq_t$ in axiom A7$^{\text{w}}$ by $\sim_t$.

[14] In the case of anticipated learning a recursive evaluation could still be preferable to enriching the state space (here X) by informational variables that permit the necessary conditioning of an expectation operator that is pulled to the first period.

$\tilde{U}_t \to \mathrm{I\!R}$ with

$$\text{RIRA}^*_t(z) = -\frac{\frac{d^2}{dz^2} f_t(z)}{\frac{d}{dz} f_t(z)}\, z\,,$$

and of *absolute intertemporal risk aversion* as the functions $\text{AIRA}^*_t : \tilde{U}_t \to \mathrm{I\!R}$ with

$$\text{AIRA}^*_t(z) = -\frac{\frac{d^2}{dz^2} f_t(z)}{\frac{d}{dz} f_t(z)}\,.$$

I define $\text{RIRA}^*_t[\tilde{x}_t] = \text{RIRA}^*_t(z)|_{z=\tilde{u}_t(\tilde{x}_t)}$ and $\text{AIRA}^*_t[\tilde{x}_t] = \text{AIRA}^*_t(z)|_{z=\tilde{u}_t(\tilde{x}_t)}$ as the intertemporal risk aversion measures at the point in consumption space $\tilde{x}_t$, where f_t is defined by Theorem 1. Because of the affine freedom $\mathfrak{a}_0$ in the representation of Theorem 1, these risk measures are not yet well-defined. The reason is discussed in detail by Traeger (2010); however, the following intuition is straight forward: the standard Arrow Pratt measures of relative and absolute risk aversion are not well-defined unless a cardinal measure scale for the consumption commodity under observation is fixed. Similarly, the measures of absolute and relative intertemporal risk aversion are not well-defined, unless a cardinal measure scale is fixed. Intertemporal risk aversion and the curvature of f_t measure risk aversion with respect to welfare gains and losses, where welfare is measured by the abstract concept of the intertemporally additive utility function u. By observing trade-offs under certainty, the preference relation fixes this measure scale up to affine transformations. Analogously to standard risk aversion measures, the measure of relative intertemporal risk aversion RIRA^*_t is well-defined only if a zero welfare level is fixed. The measure of absolute intertemporal risk aversion AIRA^*_t is well-defined only if a unit of welfare is fixed.[15]

Proposition 1: Let preferences $\succeq$ be represented in the sense of Theorem 1.

a) Choose $x^0 \in X$ and fix $u(x^0) = 0$. Then, the risk measures RIRA^*_t for a preference representation in the sense of Theorem 1 are well-defined for $t \in \{1, ..., T\}$, and the value of $\text{RIRA}^*_t[\tilde{x}_t]$ only depends on the preferences $\succeq$ and the point $\tilde{x}_t$ in consumption space.

b) Choose $x^1, x^2 \in X$ satisfying $x^1 \succ_T x^2$ and fix $u(x^1) - u(x^2) = 1$. Then, the risk measures AIRA^*_t for a preference representation in the sense of Theorem 1 are well-defined for $t \in \{1, ..., T\}$, and the value of $\text{AIRA}^*_t[\tilde{x}_t]$ only depends on the pref-

[15] The affine transformation $\mathfrak{a}_0$ in the uniqueness part of Theorem 1 points out what happens if unit and/or zero level are not fixed. For example a doubling of u would correspond to an increase in welfare range and a decrease in welfare unit. The function f_t measuring risk aversion with respect to welfare gains and losses has to offset this change in measurement by applying the inverse of $\mathfrak{a}_0$. While such an adjustment is necessary to represent the same preferences, it changes the numerical value of the intertemporal risk measures.

erences $\succeq$ and the point $\tilde{x}_t$ in consumption space.

The preferences of a decision maker are concerned with real outcomes, not with numbers in the Euclidean space. Numbers only serve as a representational tool and depend on the chosen coordinate system (i.e., measurement scale). The proposition implies that the risk measures RIRA^*_t and AIRA^*_t do not depend on how one measures the point $\tilde{x}_t$ in consumption space (see Traeger (2010) for a further elaboration). This invariance is particularly attractive when modeling risk aversion with respect to characteristics, such as quality, that are not equipped with a natural measurement scale. In such cases, whether a decision maker is risk loving or risk seeking in the sense of the Arrow Pratt measures is only a question of the applied measurement scale (Traeger 2010). In contrast, the measures of intertemporal risk aversion are invariant to changes in the measurement scale of goods, and they are well-defined and one-dimensional also in a multi-commodity setting. Note that instead of fixing both unit and zero level in Proposition 1, simply fixing the range of u implies uniqueness as well.[16]

5 Risk Stationarity

In the certainty stationary setting of Theorem 1, the functions f_t characterizing intertemporal risk aversion can vary arbitrarily over time. This section extends stationarity to the evaluation of uncertain consumption plans.

5.1 Axiomatic Description for a Finite Planning Horizon

The risk stationarity axiom extends the assumption that the passage of time does not change preferences to the ranking of uncertain consumption plans. To obtain such a stationary evaluation of general uncertain outcomes, it proves sufficient to require stationarity only for "coin toss" compositions of certain consumption paths, i.e., probability a half mixtures of type $\frac{1}{2}\mathrm{x} + \frac{1}{2}\mathrm{x}'$. As with certainty stationarity in section 3.1, I split the motivation of the axiom into two steps. The first requirement, corresponding to equation (2) becomes:

$$\tfrac{1}{2}(\mathrm{x}, x) + \tfrac{1}{2}(\mathrm{x}', x) \succeq_{t|T} (\mathrm{x}'', x) \;\Leftrightarrow\; \tfrac{1}{2}(\mathrm{x}, x) + \tfrac{1}{2}(\mathrm{x}', x) \succeq_{t+1|T+1} (\mathrm{x}'', x) \tag{8}$$

[16] I marked the measures of RIRA^*_t and AIRA^*_t with an asterix because they measure intertemporal risk aversion with respect to current value welfare changes, in contrast to the measures introduced in Traeger (2010) that measure intertemporal risk aversion in the general non-stationary setting with respect to present value welfare changes.

for all $\mathsf{x}, \mathsf{x}', \mathsf{x}'' \in \mathsf{X}^{t+1}$, $x \in X$ and $t \in \{1, ..., T\}$. With a rolling time horizon, the mere passage of time does not change the ranking among different uncertain scenarios.[17]

The second step to arrive at the axiom of risk stationarity connects the relations $\succeq_{t+1|T}$ and $\succeq_{t+1|T+1}$. Again, I require that scenarios whose outcomes coincide in the last period of a finite planning horizon $T+1$ rank the same independently of whether the decision maker applies a planning horizon of $T+1$ or T. This statement translates into the following equivalence:

$$\tfrac{1}{2}\mathsf{x} + \tfrac{1}{2}\mathsf{x}' \succeq_{t+1|T} \mathsf{x}'' \quad \Leftrightarrow \quad \tfrac{1}{2}(\mathsf{x}, x) + \tfrac{1}{2}(\mathsf{x}', x) \succeq_{t+1|T+1} (\mathsf{x}'', x) \tag{9}$$

for all $\mathsf{x}, \mathsf{x}', \mathsf{x}'' \in \mathsf{X}^{t+1}$, $x \in X$ and $t \in \{1, ..., T\}$. Together, equations (8) and (9) bring about the following axiom for stationarity of risk attitude in a setting with a finite planning horizon:

A8 (risk stationarity) For all $t \in \{1, ..., T-1\}$ and $x \in X$:
$\tfrac{1}{2}(\mathsf{x}, x) + \tfrac{1}{2}(\mathsf{x}', x) \succeq_t (\mathsf{x}'', x) \;\Leftrightarrow\; \tfrac{1}{2}\mathsf{x} + \tfrac{1}{2}\mathsf{x}' \succeq_{t+1} \mathsf{x}'' \quad \forall\, \mathsf{x}, \mathsf{x}', \mathsf{x}'' \in \mathsf{X}^{t+1}$.

The decision maker ranks lotteries of the form $\frac{1}{2}\mathsf{x} + \frac{1}{2}\mathsf{x}'$ the same when they are faced in period t as when they are faced in period t+1. When facing them in period t, the additional outcome x at the end of the planning horizon, which coincides for all consumption paths, does not change his ranking.

5.2 Relation to the Infinite Planning Horizon Setting

I briefly show how to derive the usual infinite horizon risk stationarity axioms from the assumption expressed in equation (8). Instead of the finite consumption paths (x, x) and (x', x) the decision maker faces the paths $\mathsf{x}^\infty, \mathsf{x}'^\infty \in X^\infty$. In the infinite horizon setting $\succeq_{1|T+1} = \succeq_{1|\infty} = \succeq_{1|T}$. Then, by time consistency, equation (8) for $t = 1$ is equivalent to

$$\tfrac{1}{2}\mathsf{x}^\infty + \tfrac{1}{2}\mathsf{x}'^\infty \succeq_{1|\infty} \mathsf{x}''^\infty \quad \Leftrightarrow \quad (x_1, \tfrac{1}{2}\mathsf{x}^\infty + \tfrac{1}{2}\mathsf{x}'^\infty) \succeq_{1|\infty} (x_1, \mathsf{x}''^\infty)$$

for all $\mathsf{x}^\infty, \mathsf{x}'^\infty, \mathsf{x}''^\infty \in X^\infty$ and $x_1 \in X$. Similarly for $t = 2$ equation (8) is equivalent to

$$(x_1, \tfrac{1}{2}\mathsf{x}^\infty + \tfrac{1}{2}\mathsf{x}'^\infty) \succeq_{1|\infty} (x_1, \mathsf{x}''^\infty) \quad \Leftrightarrow \quad (x_1, x_2, \tfrac{1}{2}\mathsf{x}^\infty + \tfrac{1}{2}\mathsf{x}'^\infty) \succeq_{1|\infty} (x_1, x_2, \mathsf{x}''^\infty)$$

for all $\mathsf{x}^\infty, \mathsf{x}'^\infty, \mathsf{x}''^\infty \in X^\infty$ and $x_1, x_2 \in X$. The latter statement for $t = 2$ can be transformed using the corresponding statement for $t = 1$ into the requirement:

$$\tfrac{1}{2}\mathsf{x}^\infty + \tfrac{1}{2}\mathsf{x}'^\infty \succeq_{1|\infty} \mathsf{x}''^\infty \quad \Leftrightarrow \quad (x_1, x_2, \tfrac{1}{2}\mathsf{x}^\infty + \tfrac{1}{2}\mathsf{x}'^\infty) \succeq_{1|\infty} (x_1, x_2, \mathsf{x}''^\infty)$$

[17] Observe that, in contrast to equation (2), equation (8) has to hold in all periods $t \in \{1, ..., T\}$. However, by time consistency A5 all the requirements in equation (8) can be carried over into the first two periods (see subsection 5.2). The important characteristic of the requirement is only that uncertainty resolves in all the different periods.

for all $x^\infty, x'^\infty, x''^\infty \in X^\infty$ and $x_1, x_2 \in X$. By induction, one obtains the general requirement:

$$\tfrac{1}{2}x^\infty + \tfrac{1}{2}x'^\infty \succeq_{1|\infty} x''^\infty \quad \Leftrightarrow \quad (x^t, \tfrac{1}{2}x^\infty + \tfrac{1}{2}x'^\infty) \succeq_{1|\infty} (x^t, x''^\infty) \tag{10}$$

for all $x^\infty, x'^\infty, x''^\infty \in X^\infty$, $t \in \mathbb{N}$ and $x^t \in X^t$. A corresponding[18] axiom for stationarity of risk attitude is found in Chew & Epstein (1991, 356).

5.3 The Risk Stationary Representation

Preference stationarity for the evaluation of uncertain scenarios, together with assumptions A1-A5, yield the following representation.

Theorem 3: Choose a non-degenerate closed interval $U^* \subset \mathbb{R}_+$. Preferences $\succeq$ satisfy axioms A1-A5 and risk stationarity A8 if, and only if, there exist a continuous and surjective function $u : X \to U^*$, a discount factor $\beta \in \mathbb{R}_{++}$, and $\xi \in \mathbb{R}$ that represent preferences as follows:
Define $\tilde{u}_t : \tilde{X}_t \to \mathbb{R}$ for $t \in \{1, ..., T\}$ by $\tilde{u}_T(x_T) = u(x_T)$ and recursively

$$\tilde{u}_{t-1}(x_{t-1}, p_t) = u(x_{t-1}) + \beta\, \mathcal{M}^{\exp^{-\xi}}(p_t, \tilde{u}_t) \ ,$$

then, for all periods $t \in \{1, ..., T\}$

$$p_t \succeq_t p'_t \quad \Leftrightarrow \quad \mathcal{M}^{\exp^{-\xi}}(p'_t, \tilde{u}_t) \geq \mathcal{M}^{\exp^{-\xi}}(p'_t, \tilde{u}_t) \quad \forall\, p_t, p'_t \in P_t \ .$$

Moreover, given U^* the function u is determined uniquely as are the measures of intertemporal risk aversion $\text{AIRA}^*_t = \xi$ and $\text{RIRA}^*_t = \xi\,\text{id}$, where id denotes the identity.

Risk stationarity implies that the functions f_t characterizing uncertainty aggregation in Theorem 1 coincide for different periods. Moreover, the axiom yields a constant coefficient of absolute intertemporal risk aversion ξ (implying the functional form $f_t(z) = \exp(-\xi z)$). These two implications of axiom A8 can be associated with equations (8) and (9), respectively. While the idea underlying equation (8) causes uncertainty aggregation to coincide in different periods, it is the assumption of "coinciding last outcome independence" (9) that is mostly responsible for constant absolute intertemporal risk aversion. Finally, recall from page 17 in section 4.2 that AIRA^*_t measures intertemporal risk aversion in terms of *current* welfare. The measurement with respect to the undiscounted welfare scale is crucial for making intertemporal risk aversion constant over time. Stationarity is an axiom that puts the decision maker in the position of his future self. This future self evaluates

[18] Unlike the above formulation, the authors require condition (10) for all lotteries, not just for the probability a half ('coin toss') combinations that I have used and which prove sufficient in my setting.

the outcomes of the lottery in the future period without discounting, just as the present self evaluates the outcomes of the lottery in the present without discounting. Both agents evaluate outcomes in undiscounted current welfare units and have the same ranking of lotteries. It is important for the subsequent analysis that the risk aversion coefficients in Theorem 3 are constant in current value terms. For a quadratic utility function u, the representation of Theorem 3 coincides with Hansen & Sargent's (1995) model of risk-sensitive decision making. Theorem 3 identifies their risk-sensitivity parameter as the parameter of absolute intertemporal risk aversion measured in current value units.

6 Timing Indifference

Recursive utility models, as employed in the preceding sections, generally imply an intrinsic preference for early or late resolution of uncertainty. Motivated by a normative perspective on rational decision making, I will impose intrinsic timing indifference. First, I state and discuss the assumption, and then I give the representation.

6.1 Indifference to the Timing of Risk Resolution

A rational decision maker generally prefers an early resolution of uncertainty whenever it permits refining later choices and improving outcomes (or outcome probabilities). Such a preference for an early resolution is instrumental in giving rise to better outcomes (in expectation). Kreps & Porteus (1978) show that, in addition, general recursive utility models allow the decision maker to exhibit an intrinsic preference for an early or late resolution of uncertainty. Such a decision maker prefers early (or late) resolution of uncertainty even if the obtained information does not enable him to change future outcomes (or outcome probabilities). Traeger (2007) analyzes the relation between intertemporal risk aversion and an intrinsic preference for the timing of uncertainty resolution. I argue for indifference to the timing of risk resolution from a perspective of rational decision making, or from a normative perspective of a public decision maker. A decision maker who is non-indifferent to the timing of risk resolution is willing to give up wealth for advancing or postponing information that has no effect on future behavior or payoffs. Such an action does not seem rational from an individual perspective, nor would we want a public decision maker to give up social wealth for such a purpose.[19] The formal statement of the timing indifference axiom is

[19]Note that a frequent justification of non-trivial timing preference is based on the argument that only recursive models featuring non-indifference to the timing of risk resolution make it possible to disentangle Arrow Pratt risk aversion from the elasticity of intertemporal substitutability. However, while this reasoning is true for the widespread generalized isoelastic model, it is not true in general. Any

A9 (indifference to the timing of risk resolution)
For all $t \in \{1, ..., T-1\}$, $x_t \in X$, $p_{t+1}, p'_{t+1} \in P_{t+1}$ and $\lambda \in [0,1]$:

$$\lambda(x_t, p_{t+1}) + (1-\lambda)(x_t, p'_{t+1}) \quad \sim_t \quad (x_t, \lambda p_{t+1} + (1-\lambda)p'_{t+1}).$$

A decision maker who is indifferent to the timing of risk resolution does not distinguish between a lottery in which the uncertainty about the future faced in period $t+1$ (i.e., whether he faces the future p_{t+1} or p'_{t+1}) resolves in period t (lottery on the left) or in period $t+1$ (lottery on the right). This uncertainty about the future faced in period $t+1$ is described by the probability mixture λ and $1-\lambda$. Axiom 9 is always satisfied in the intertemporally additive expected utility standard model.

6.2 The Timing-Indifferent Representation

A decision maker who is indifferent to the timing of risk resolution makes no use of the information in which period uncertainty resolves, unless this information is of consequential use. Then, preferences can be represented non-recursively. First, the *description* of uncertainty can be reduced to atemporal lotteries. As opposed to Kreps & Porteus' (1978) temporal lotteries, atemporal (or standard) lotteries abandon the recursive structure and express uncertainty as probability measures over consumption paths. I denote these standard lotteries by $p_t^{\mathsf{X}}, p_t'^{\mathsf{X}} \in \Delta(\mathsf{X}^t)$. Second, the *evaluation* of uncertainty no longer requires the recursive structure, i.e., the alternating aggregation over uncertainty and time. I therefore restrict the definition of uncertainty aggregation rules to the subdomain $\mathsf{X}^t \subset \tilde{X}_t$ yielding functionals $\mathcal{M}^{f_t} : \Delta(\mathsf{X}^t) \times \mathcal{C}^0(\mathsf{X}^t) \to \mathbb{R}$. Uncertainty aggregation rules now take as inputs lotteries p_t^{X} and evaluations $\tilde{u}_t : \mathsf{X}^t \to \mathbb{R}$ of consumption *paths*. Traeger (2007) shows how the probability measures $p_t^{\mathsf{X}} \in \Delta(\mathsf{X}^t)$ are derived from their recursive counterparts $p_t \in P_t$ by "integrating out" the information on the timing of risk resolution. This relation, however, is only needed to prove the representation within the general axiomatic setting. For an application of the representation theorem, it is sufficient to describe the uncertain future directly by the measures $p_t^{\mathsf{X}} \in \Delta(\mathsf{X}^t)$, i.e., by probability distributions over consumption paths.

Theorem 4: Choose a non-degenerate closed interval $U^* \subset \mathbb{R}_+$. Preferences $\succeq$ satisfy axioms A1-A5, certainty stationarity A6, and timing indifference A9 if, and only if, there exist a continuous and surjective function $u : X \to U^*$, a discount factor $\beta \in \mathbb{R}_{++}$, and $\xi \in \mathbb{R}$ that represent preferences as follows:

model of non-trivial intertemporal risk aversion, including the upcoming representation, allows such a disentanglement of Arrow Pratt risk aversion and intertemporal substitutability.

Define the functions $\tilde{u}_t : \tilde{X}_t \to \mathbb{R}$ for $t \in \{1, ..., T\}$ by

$$\tilde{u}_t(\mathrm{x}^t) = \sum_{\tau=t}^{T} \beta^{\tau-1} u(\mathrm{x}^t_\tau) \ ,$$

then, for all periods $t \in \{1, ..., T\}$

$$p_t \succeq_t p'_t \ \Leftrightarrow \ \mathcal{M}^{\exp^{-\xi}}(p^{\mathrm{X}}_t, \tilde{u}_t) \geq \mathcal{M}^{\exp^{-\xi}}(p'^{\mathrm{X}}_t, \tilde{u}_t) \quad \forall \, p_t, p'_t \in P_t \ . \tag{11}$$

Moreover, given U^*, the function u is determined uniquely, as are the measures of intertemporal risk aversion $\mathrm{AIRA}^*_t = \beta^{t-1}\, \xi$ and $\mathrm{RIRA}^*_t = \beta^{t-1}\, \xi \, \mathrm{id}$.

The representation evaluates outcomes in all periods with the stationary utility function u. A consumption path is evaluated by the discounted sum of per period utility. To evaluate an uncertain future, the decision maker weights the aggregate welfare of the possible consumption paths with their respective probabilities and applies the uncertainty aggregation rule $\mathcal{M}^{\exp^{-\xi}}$. For a decision maker starting directly with an atemporal (standard) lottery description, equation (11) in particular implies

$$p^{\mathrm{X}}_t \succeq_t p'^{\mathrm{X}}_t \ \Leftrightarrow \ \mathcal{M}^{\exp^{\xi}}(p^{\mathrm{X}}_t, \tilde{u}_t) \geq \mathcal{M}^{\exp^{\xi}}(p'^{\mathrm{X}}_t, \tilde{u}_t) \quad \forall \, p^{\mathrm{X}}_t, p'^{\mathrm{X}}_t \in \Delta(\mathrm{X}^t) \, .$$

The model coincides with the widespread discounted utility framework for the evaluation of individual consumption paths but generally employs nonlinear uncertainty aggregation over the welfare generated by the different paths. Note that under the above assumption of axiom A9, a result by Chew & Epstein (1989, 110) permits replacing the independence axiom A3 with a collection of weaker assumptions. Assuming a quadratic utility function u yields the risk-sensitive control model of Whittle (1990, 2002) and Bouakiz & Sobel (1992). Theorem 4 identifies the corresponding risk-sensitivity parameter as the coefficient of absolute intertemporal risk aversion in the present period.[20]

As in the risk-stationary setting, uncertainty evaluation takes the form of constant absolute intertemporal risk aversion. However, in contrast to the earlier representations, the representation in Theorem 4 aggregates over uncertainty only in the present. If the decision maker is in period t, all uncertainty evaluation relies on the single function $f_t = \exp^{-\xi}$, independent of the period in which uncertainty resolves. In contrast, the recursive representation in the earlier sections relied on period-specific uncertainty aggregation. Alternatively, the preferences characterized by Theorem 4 can be expressed in terms of the general recursive representation given in Theorem 1 (see the proof of Theorem 4), where they correspond to the case $f_t = \exp^{-\xi \beta^{t-1}}$ for $t \in \{1 \ldots T\}$. As explained on page 17 in section 4.2, these functions f_t characterize the current-value measures of

[20]This parameter ξ is also the measure of absolute intertemporal risk aversion when measuring intertemporal risk aversion with respect to present value welfare.

intertemporal risk aversion, which are stated in Theorem 4. In contrast, the function $\exp^{-\xi}$ used (for all t) in Theorem 4 characterizes intertemporal risk aversion in present-value terms.[21] Thus, measured with respect to present-value welfare units, the coefficient of absolute intertemporal risk aversion is constant over time. Keeping the measurement of intertemporal risk aversion in terms of current-value welfare as before, results in coefficients that are discounted at the same rate at which the decision maker discounts utility.

In the following, I elaborate an intuition as to why the current-value measure of intertemporal risk aversion has to be discounted. I consider a decision maker who faces a lottery over outcomes in period T and uses a recursive evaluation as in Theorem 1. First, let the lottery take place in period t with $1 < t < T$. For example, a coin is flipped in period t, determining consumption levels in period T. The decision maker assigns utility to each of the possible lottery outcomes[22] and aggregates over uncertainty according to his intertemporal risk attitude in period t. Second, let the coin toss and the decision maker's risk evaluation take place in period $t-1$. Then, with respect to their values in period t, the utilities attached to the lottery outcomes are discounted by the factor β. In consequence, the welfare *variation* corresponding to the utility at risk is also *discounted.* If the decision maker's coefficient of *absolute intertemporal risk aversion* was the same in period $t-1$ as in period t, the smaller welfare variation would imply a smaller effective risk aversion. The decision maker would exhibit an intrinsic preference for early resolution of risk. Only if $\text{AIRA}^*_{t-1} > \text{AIRA}^*_t$ can the decision maker be timing indifferent. It turns out that the coefficient of absolute intertemporal risk aversion has to be discounted precisely at the same rate of pure time preference as the utility function u. The important difference to risk stationarity is that for timing indifference, the agent does not put himself into the position of a future self but evaluates risk resolutions in different periods from the present perspective. Consequently, the coefficient of absolute intertemporal risk aversion would be constant over time only if we were to measure it with respect to present-value welfare gains and losses. In summary, the discounting of welfare variations calls for a discounting of current-value intertemporal risk aversion if the decision maker is to be indifferent to the timing of risk resolution.

Preferences in Theorem 4 are not required to satisfy risk stationarity. I close this section by drawing attention to the following consequence. When the agent puts himself in the position of his future period-t self, he still discounts utility in period t with the discount

---

[21]Traeger (2010) shows how to represent the general (non-stationary) recursive representation using uncertainty aggregators that measure intertemporal risk aversion in present-value welfare units.

[22]More precisely, he assigns utilities to each of the consumption paths from t until T, which I assume to coincide except for their period T outcomes.

factor β^{t-1}. Thus, the discount factors are fixed with respect to calendar time as opposed to the amount of periods that the evaluated outcome lies in the future (futurity). If we apply the representation of Theorem 4, making the discount factor depend on futurity rather than calendar time, we violate time consistency.

7 Implications for Discounting

This section elaborates on the consequences of the above results for long-term evaluation. First, combining the assumptions of risk stationarity and indifference to the timing of risk resolution implies a zero rate of pure time preference. Second, an intertemporal risk averse decision maker gives less weight to the future the less he knows about the consequences of his (in)actions.

7.1 The Loss of Time Preference

Risk stationarity puts the decision maker in the position of his future self and requires him to rank lotteries in the same way as he would in the present period. Therefore, his intertemporal risk aversion has to be constant when measured with respect to *current-value* welfare gains and losses. Timing indifference is a requirement for a decision maker who evaluates different lotteries from the perspective of the present. The ranking of lotteries must be independent of the time of risk resolution. Herto, intertemporal risk aversion has to be constant when measured in present-value terms, and the current-value measure AIRA^*_t picks up the discount rate. In consequence, there is only one situation in which an intrinsic devaluation of the future is compatible with both axioms: for a decision maker who is intertemporal risk neutral, neither the assumption of risk stationarity nor that of timing indifference have bite.[23] The coefficient of absolute intertemporal risk aversion is $\text{AIRA}^*_t = 0$ for all $t \in \{1, ..., T\}$ and, therefore, is independent of β. For a nontrivial model of intertemporal risk averse decision making, the following result is obtained.[24]

Theorem 5: Preferences $\succeq$ axioms A1-A5, strict intertemporal risk aversion A7^{s}, risk stationarity A8, and timing indifference A9, if and only if, there exists a representation in the sense of Theorem 4 with $\xi > 0$ and $\beta = 1$.

[23] Precisely, risk stationarity has no additional bite with respect to the assumption of certainty stationarity.

[24] I do not specify the period for which axiom A7^{s} has to hold. As a consequence of the other axioms, a decision maker who satisfies strict intertemporal risk aversion in some period also satisfies strict intertemporal risk aversion in all periods.

A decision maker who accepts the axioms presented in this paper cannot discount the future for reasons of pure time preference ($\beta = 1$). Note that the non-recursive representation in the theorem is equivalent to the recursive representation in Theorem 3, with absolute intertemporal risk aversion $\xi > 0$ and a discount factor of $\beta = 1$.

From the normative perspective or rationality point of view, Theorem 5 argues for a zero rate of pure time preference. From a theoretical perspective, this theorem identifies a tension among intertemporal risk aversion, risk stationarity, and timing indifference.[25] Together, Theorems 3 to 5 also supply a simple test to determine which of the axioms is violated if β differs from unity: constancy of the coefficient of absolute intertemporal risk aversion AIRA_t^* over time identifies the violation of timing indifference, a discounted coefficient indicates risk stationarity, and a degenerate coefficient implies intertemporal risk neutrality. Finally, note that weak future independence in the stationarity axiom, as characterized by equation (9), is innocent when it comes to eliminating pure time preference. In Theorem 3, weak future independence is responsible for the constancy of absolute intertemporal risk aversion over outcomes. However, in Theorem 5, this constancy is already implied by the timing indifference represented in axiom A9.

7.2 Discounting for Reasons of Uncertainty

A decision maker satisfying the axioms underlying Theorem 5 does not discount future utility for reasons of mere impatience. However, an intertemporal risk averse decision maker can effectively discount future utility if uncertainty increases exogenously over time. To illustrate this point, I assume that outcomes are independently distributed over time and described by the measures $\text{p}_1, \text{p}_2, ..., \text{p}_T$ with $\text{p}_t \in \Delta(X)$. Then, the distribution over consumptions paths is captured by the product measure $p_t^{\mathbf{X}} = \text{p}_t \otimes ... \otimes \text{p}_T$ for $t \in \{1, ..., T\}$. I will determine the utility discount factors that would arise if preferences satisfying the axioms of Theorem 5 were to be represented by the standard discounted expected utility model. I assume that uncertainty increases over time. Because outcomes are multidimensional, and because the dimensions affected by uncertainty can change over time, I model an increase of uncertainty over the resulting utility. To make this definition precise, I denote with p_t^u the welfare lottery induced in utility space by the uncertainty p_t over outcomes x_t (pushforward of p_t under u). Let $\text{P}_t(z) = \text{p}_t^u([0, z])$ denote the cumulative distribution functions. Without loss of generality, I assume the utility range

[25]Recall that Chew & Epstein (1989, 110) have shown that in the context of axiom A9 the independence axiom can be replaced by a collection of significantly weaker axioms, making it a less likely to be violated (on its own).

$U^* = [0,1]$. Then, p^u_t second order stochastically dominates p^u_{t+1} if

$$\int_0^z \mathrm{P}_{t+1}(z') - \mathrm{P}_t(z')\, dz' \geq 0 \tag{12}$$

for all $z \in [0,1]$. Second order stochastic dominance is not just a measure of uncertainty; it can also indicate that outcomes in period $t+1$ are certain but sufficiently undesirable. I follow Rothschild & Stiglitz (1970) in confining the definition to a better-suited notion of increasing uncertainty by requiring that both probability distributions yield the same expected utility, which is equivalent to the requirement that equation (12) holds with equality for $z = 1$. The graphical interpretation of increasing uncertainty is that the probability distribution in $t+1$, in this case over utility, has more weight in the tails. To show that the utility discount factors decrease strictly over time uncertainty has to increase strictly: I assume the existence of $z^* \in (0,1)$ such that equation (12) holds with strict inequality. Then, the distribution in period $t+1$ has strictly more weight in the tails than the distribution in period t. In summary, the sequence of probability measures $\mathrm{p}_1, \mathrm{p}_2, ...\mathrm{p}_T$ exhibits *strictly increasing uncertainty* (weight in the tails) if for all $t \in \{1,...,T-1\}$: equation (12) holds for all $z \in [0,1]$ exhibiting equality for $z = 1$, and there exists $z^* \in (0,1)$ such that equation (12) holds as a strict inequality.[26]

Proposition 2: Let preferences be represented in the sense of Theorem 5. Let the sequence $\mathrm{p}_1, \mathrm{p}_2, ...\mathrm{p}_T$ exhibit strictly increasing uncertainty as defined above. Then, the unique weights $\bar\beta_1 ... \bar\beta_T$ satisfying the equation

$$\mathcal{M}^{\exp^{-\xi}}\big(\mathrm{p}_t \otimes ... \otimes \mathrm{p}_T, \textstyle\sum_{\tau=t}^T u\big) = \mathrm{E} \textstyle\sum_{\tau=t}^T \bar\beta_t u(x_t) \text{ for all } t \in \{1,...,T\} \tag{13}$$

are falling over time: $\bar\beta_{t+1} < \bar\beta_t$ for all $t \in \{1,...,T-1\}$.

The right-hand side of equation (13) represents a discounted expected utility evaluation, whereas the left-hand side represents the evaluation based on Theorem 5. The discount factors $\bar\beta_t$ in the proposition are falling over time, though not usually at a constant rate as in the standard discounted expected utility model. How quickly they fall depends on both the agent's intertemporal risk aversion and how much uncertainty increases over time. Observe that the falling sequence of discount factors depends on the particular future the decision maker faces. A reduction in risk increases these utility weights. In contrast, a reduction of risk does not affect the exogenous utility weights in the discounted expected utility model. In consequence, with endogenous uncertainty, representation Theorem 5 favors a low-risk wealthy future compared to the discounted expected utility model.

Proposition 2 shows how utility is discounted in a scenario with exogenous uncertainty. However, economic applications such as a cost benefit analysis of a long-term project

[26] An equivalent characterization is that the riskier random variable can be obtained from the less risky random variable by adding noise (Rothschild & Stiglitz 1970).

generally depend on marginal utility rather than absolute utility. Let the outcome space be Euclidian: $X \in \mathbb{R}^n$. I define a (certain) project with a payoff sequence $(\epsilon_t)_{t\in\{1,...,T\}}$, $\epsilon_t \in \mathbb{R}^n$, that adds the certain payoff ϵ_t to the uncertain baseline consumption x_t, independent of the state of the world. The project is assumed to be small in the sense that $\epsilon_t << x_t$, i.e., every dimension of the project payoff is small against baseline consumption.[27] The standard model evaluates such a project by the welfare change $\sum_{t=1}^{T} \mathrm{E}\beta^{t-1}u'(x_t)\epsilon_t$ and recommends implementation if the sum is positive. In general, $u'(x_t)$ is a gradient and ϵ_t a vector. The welfare change is easily converted into monetary equivalents if desired.

Proposition 3: Preferences and exogenous uncertainty satisfy the same conditions and are represented as in Proposition 2. A marginal project augments consumption by the certain payoff sequence $(\epsilon_t)_{t\in\{1,...,T\}}$. Let the functions $\tilde{\beta}_t : X \to \mathbb{R}$, $t \in \{1, ..., T\}$, be such that

$$\mathrm{E}\sum_{t=1}^{T} \tilde{\beta}_t(x_t)u'(x_t)\epsilon_t$$

characterizes the first order welfare change for all projects $\epsilon \in \mathbb{R}^{n\times T}$. Then,

$$\tilde{\beta}_{t+1}(x) < \tilde{\beta}_t(x) \text{ for all } x \in X,\ t \in \{1, ..., T\} \text{ and}$$

$$\tilde{\beta}_t(\underline{x}) > \tilde{\beta}_t(\overline{x}) \Leftrightarrow u(\underline{x}) < u(\overline{x}) \text{ for all } t \in \{1, ..., T\}\ .$$

For a given consumption (or utility) level, the weights on marginal utility also decrease over time. Once again, exogenous uncertainty does not usually decrease these effective discount factors $\tilde{\beta}_t(x)$ at a constant rate. Marginal utility (or project payoff) receives a higher weight in those states of the world leading to lower utility levels. Thus, compared to the standard model, preferences based on Theorem 5 give relatively more weight to payoffs in bad states of the world.

7.3 Consequences for Long-Term Evaluation

I discussed two different rationales for discounting, pure time preference and increasing uncertainty under intertemporal risk aversion. Although both rationales reduce the effective weight given to future expected welfare, there are three important differences. First, a positive rate of pure time preference implies exponential discounting (assuming time consistency). Under intertemporal risk aversion, the form of discounting will generally depend on risk aversion and the way uncertainty increases over time. Second, marginal utility and payoffs receive a higher weight, i.e., are discounted relatively less, for bad

[27]The proposition analyzes first order welfare changes for all projects. These are useful approximations to the actual welfare changes only for small projects.

states of the world. Third, once uncertainty becomes endogenous to decision making, an intertemporal risk averse decision model will gain relatively higher welfare from low-risk scenarios. Before I illustrate these differences in a conjecture on climate change evaluation, I discuss a conceptual concern regarding the choice of the time horizon in intertemporal evaluation.

Propositions 2 and 3 do not generally yield exponential discounting and convergence of the welfare function in the limit of an infinite time horizon. For economic agents, the assumption of a finite planning horizon is most likely more reasonable than that of an infinite horizon. The time span until our sun turns into a red giant is negligible as opposed to infinity. The interesting question is not whether the time horizon is finite or infinite. The interesting question is how sensitive the decision model is to the time horizon. A zero rate of pure time preference evokes the impression that the model would be highly sensitive to the cut-off of the planning horizon. This intuition is generally wrong. The model ranks different future scenarios resulting from different courses of action. A meaningful time horizon should include those periods for which we foresee (probabilistically) how our actions change the future.[28] If we adhere to this rule, the model is insensitive to the time horizon. This insensitivity is precisely the consequence of the weak future independence assumption accompanying stationarity and playing a more prominent role in the finite horizon setting: if we drop a common future when comparing two scenarios, the resulting ranking does not change.[29]

There might be actions so daunting that we do not know how long the changes persist. Decisions on war and climate change may fall into this category. With a positive rate of pure time preference, we can simply adopt an infinite time horizon, at least for a bounded utility function or a sufficiently slow growth rate in the economy. In the present undiscounted model, the situation becomes more delicate. It is useful to step back and interpret the assumption of an infinite planning horizon. Nobody truly argues that we can plan for an infinite horizon. The point is that we do not know when to cut the horizon and that we want to avoid an arbitrary cutoff because an evaluation might be sensitive to such arbitrariness. We can think of the rate of pure time preference as a continuous chipping away at the planning horizon. Rather than making a decision on

[28]I assume for the discussion that the decision maker wants to take all long-run consequences of his actions into account, just like in an infinite horizon discounted expected utility model. If a decision maker purposely only cares for e.g. his finite life time horizon the discussion is not essential, even though I still would consider a model explicitly assigning utility to the outcome 'dead' as superior (in which case the discussion applies).

[29]It follows from the risk stationary representation in Theorem 3 that weak future independence in the risk stationarity axiom A8 implies also that a common *uncertain* future does not change the ranking of lotteries.

what the planning horizon should be, we decide on a rate that chips away at it. We know that optimal policies with long-term effects are highly sensitive to this rate, just like they would be to an arbitrary cutoff of the time horizon in a deterministic world. A prominent example is the debate on climate change evaluation (Nordhaus 2007).

The present framework does not permit the continuous chipping away at the time horizon. If we deliberately decide to take the consequences of current actions into account, we have to give the resulting welfare changes full weight. This requirement is unproblematic as long as we acknowledge that predicting policy implications from now into an infinite future does not take uncertainty sufficiently seriously. Any consequences of actions we take today blur out over a sufficiently long time horizon. Even when initiating war or climate change, we cannot truly tell how these actions will affect probabilities over the welfare distribution of a world in a million years. I conclude that rather than introducing arbitrariness, the model eliminates arbitrariness by eliminating the rate of pure time preference, which chips away at the time horizon at an arbitrary[30] rate. An intertemporal risk averse decision maker gives less weight to the future the less he knows about the consequences of his (in)actions.

I use climate change evaluation to illustration how a long-term evaluation based on preferences satisfying Theorem 5 differs from an evaluation in the standard model. Following a report on climate change carried out by Sir Nicholas Stern (2007), the discount rate for long-term evaluation has been one of the most hotly debated parameters in economics. The Stern review employs a close-to-zero rate of pure time preference based on ethical arguments. Nordhaus (2007) shows that the discounting assumptions in the Stern review drive the result that the review obtains a social cost of carbon that is ten times higher than Nordhaus's (2008) estimates (for the current commitment period under optimal policy). Theorem 5 derives a zero rate of pure time preference solely from axioms on rational decision making under uncertainty rather than ethical arguments. Apart from the absence of pure time preference, preferences satisfying Theorem 5 also imply a different risk evaluation. An adaptation project that focuses on payoffs in bad states of the world will increase welfare relatively more, as compared to other projects and standard preferences. Finally, greenhouse gas mitigation projects affect the overall uncertainty regarding future production and consumption possibilities. If we fixed greenhouse gas stocks at the current (or preindustrial) levels, uncertainty over climate factors during the next centuries would be fairly low. If we double or triple the greenhouse gas stocks, climate factors will change significantly, and our knowledge regarding these changes is limited. In

[30]Arbitrary in the sense that the free parameter in Koopmans (1960) is any strictly positive number. Of course, individuals can pick it deliberately within the discounted expected utility framework. Normative reasonings usually only speak to an elimination of the rate.

consequence, high-mitigation scenarios receive an additional welfare bonus with respect to low-mitigation scenarios. Thus, replacing pure time preference with intertemporal risk aversion could yield an even higher abatement recommendation than the one put forth in the Stern (2007) review.[31]

8 Discussion of the Assumptions

The first factor in composing the model presented here is a general risk attitude. Empirical evidence shows that people are more averse to risk than they are averse to intertemporal substitution (see the end of this section). Such evidence could be meaningless if there were a normative reason to override how people behave in this context. However, there does not seem to be any such argument. The concept of intertemporal risk aversion distills the essence of the general risk attitude assumption. I have discussed this axiom in several dozen audiences of trained economists, and I have never received more consent on an axiom. I think the axiom passes introspection for a large part of decision makers. Even a decision maker who is intertemporal risk neutral might want to judge the decision model of his choice based on its robustness with respect to extensions of the preference domain. For this decision maker a positive pure time preference is knife-edge.

Opening up to intertemporal risk attitude already largely demolishes the discounted expected utility model: if a decision maker signs off on the remaining assumptions underlying the discounted expected utility model, this workhorse is lost. The two remaining ingredients relevant to my conclusion only prevent the decision maker from reconstructing the demolished workhorse. One way of getting the pure rate of time preference back into the game is by introducing an intrinsic preference for early or late resolution of risk. The price for this reconstruction effort is a willingness to pay for receiving information that is of no consequential use earlier or later. If we are fully devoted to pure time preference, this might be the least bitter pill to swallow. It is the pill that is swallowed in the Epstein-Zin-Weil model. I think it is fair to say that a good number of researchers employing this

[31] A relevant factor in the trade-off will be how much of the overall uncertainty is driven by climate change and how much of it is exogenous to climate change policy (or to what degree strong climate policy can induce other uncertainties in the economic system). Note that Nordhaus (2007) argues that pure time preference does not matter much for climate change evaluation if the elasticity of intertemporal substitution is reduced to keep the consumption discount rate the same, matching market interest. The empirical models discussed at the end of section 8 disentangle Arrow Pratt risk aversion from intertemporal substitutability and explain market interest rates as well as including the risk free rate puzzle and the equity premium puzzle. Their estimates of the elasticity of intertemporal substituion go exactly the opposite direction of what Nordhaus suggests: Despite using rates of pure time preference close to zero, they find an intertemporal elasticity of 1.5, which is even higher then the values of 0.5 used by Nordhaus and of 1 used by Stern.

model are not aware that they are even swallowing this pill. Others swallow the pill consciously but in the belief that it is the only way to disentangle Arrow Pratt risk aversion from the willingness to smooth consumption over time, i.e., to incorporate a non-trivial intertemporal risk attitude. Rationality clearly objects to paying for information without consequential use and even more to paying to receive such information earlier or later.

The third ingredient is risk stationarity. This stationarity assumption is always met in the discounted expected utility model. The question is whether we are willing to give up risk stationarity to save our workhorse. Stationarity implies that we rank different consumption bundles independently of calendar time. If this assumption is challenged, it is usually due to the exclusion of habit formation. In the one-commodity model, this criticism is warranted. However, the choice space employed in the current analysis easily absorbs state variables that capture an explicit model of habit formation. What stationarity rules out is an explicit dependence of the welfare function on calendar time. If we believe that preferences change over time, we have to construct a meaningful model of preference change. This model should be independent of calendar time. Otherwise, we introduce into the model an arbitrariness that defeats its purpose of supporting the agent in making deliberate decisions. Moreover, if we were to promote an alternative model, it must exclude the subdomain of history-independent decision making and of decision making with explicit habit formation independent of calendar time. Including these preferences in the domain is sufficient to eliminate the pure rate of time preference.

The preceding paragraph makes a strong case for stationarity of preferences in general. The paper shows that we can save the discounted expected utility model by keeping stationarity for certain choices and giving up only the stationary ranking of uncertain scenarios. It is worthwhile to briefly comment on both stationarity assumptions separately. Certainty stationarity lies at the very foundation of the discounted expected utility model. Without certainty stationarity, there is no well-defined pure rate of time preference. Risk stationarity implies that the risk aversion functions do not depend on calendar time. If we were to make time consistent choices without the assumption of risk stationary choices, we would be forced to use a different decision model in every period. If we were to give up time consistency instead, we could use the same decision model in every period, but we would revise those plans of ours that involve uncertain outcomes from one period to the next, even without any changes in information. Both options are objectionable.

Finally, I briefly examine the other axioms of the paper. Transitivity and completeness are reasonable assumptions in order to develop a decision framework that, in principle, should be able to rank all relevant scenarios. Independence is an axiom that has been relaxed on behavioral accounts, particularly in the context of ambiguity where uncertainty is not captured comprehensively by unique probability distributions. In situations of risk, the independence axiom is still the benchmark for rational choice. The requirement that the present decision support framework satisfy independence in the case of objective probabilities is suffcient to eliminate the pure rate of time preference. The assumption of certainty separability is a convenient way to relate the general framework to the discounted expected utility model given that the discounted utility framework is intertemporally additive. It proves sufficient to extend risk attitude within the certainty additive preference domain to eliminate pure time preference. The relevant question is whether we want our model to exhibit the discussed characteristics, at least in this restricted domain. If we are to take the discounted expected utility model seriously in the first place, the answer ought to be "yes". Moreover, habit formation can be absorbed even into this restricted preference domain as long as there exists an explicit calendar-time-independent model of habit formation. Finally, I assume time consistency. It is probably the least contestable axiom in a model of rational decision making. If we were to drop time consistency, the decision maker would constantly revise his plan, even without changes in information. A decision support system suggesting that a decision maker would permanently change his mind without the arrival of new information seems neither rational nor helpful.

I have discussed the assumptions entering a decision support framework designed to help one make rational decisions rather than merely describe behavior. Adhering to a normatively desirable decision framework can imply significant deviations from actual behavior. This final paragraph points out that these deviations are likely smaller than one would initially assume. The Epstein-Zin-Weil model for disentangling Arrow Pratt risk aversion from aversion to intertemporal substitution is a model of intertemporal risk aversion. It implicitly assumes constant relative intertemporal risk aversion as opposed to the constancy of the absolute measure of intertemporal risk aversion that follows from the axioms of this paper. When Epstein & Zin (1991) estimate the model, they find a rate of pure time preference around zero. Moreover, they find a utility function that is close to the logarithmic form, which is the only overlap between the models of constant absolute and constant relative risk aversion. Later estimates of the Epstein-Zin-Weil model follow Campbell's (1996) approach of log-linearizing the Euler equations, which cancels pure time preference in the estimated equations. However, Bansal & Yaron (2004) solve a variety of asset pricing puzzles by calibrating a sophisticated asset pricing model with Epstein-Zin-Weil preferences, and they calibrate pure time preference to 0.1%. The same calibration is chosen by Nakamura et al. (2010) who develop a model for analyzing consumption disasters using Epstein-Zin-Weil preferences.[32] As do other estimations like Vissing-Jørgensen & Attanasio's (2003), these papers also find that Arrow Pratt risk aver-

[32]Similarly, the well-known preference for increasing income streams (Loewenstein & Sicherman 1991) or health (Chapman 2000) is an indicator that agents are less impatient than we generally assume when it comes to far reaching decisions .

sion is much stronger than aversion to intertemporal substitution, implying intertemporal risk aversion.

9 Conclusions

The discounted expected utility model is the workhorse of economics. It lives on a long tether and grazes on all fields of our discipline. Over the years, behavioral economics has cast into doubt its performance in capturing the actual behavior of economic agents. In contrast, its axiomatic foundation has still been going strong and the model clearly holds the pole position when it comes to rational decision making. The current paper lays out why the discounted expected utility model does not win this normative rationality race either. The discipline uses the model frequently and insightfully to analyze trade-offs, economic behavior, and strategic interactions, even in behavioral economics. However, I analyzed rational, deliberate decision making and showed that the defining characteristic of the discounted expected utility model, the utility discount rate, is no longer convincing if we extend the preference domain to take risk attitude seriously.

One way to save the discounted expected utility model is by eliminating intrinsic risk aversion. A decision maker following this path must be aware that regaining pure time preference is knife edge. A different way to regain pure time preference is to accept a framework that incurs costs in order to obtain information of no consequential use at an earlier or later point in time. Theorem 3 states the decision support framework for decision makers who are willing to swallow this pill. Decision makers not prepared to do so can give up the stationary evaluation of risk. Theorem 4 states the decision support framework for these individuals or agencies. The risk aggregation in this framework depends on calendar time. Thus, the decision maker effectively switches his decision support model every period to remain time consistent. Alternatively, the decision maker can give up time consistency and employ the representation in Theorem 3 with discounting and risk aggregation depending on futurity rather than on calendar time. Then, reevaluating scenarios tomorrow would result in revisions of the (formerly) optimal plan.

Different parts of the profession will swallow different pills. I swallow the opening remark by Harrod (1948) and it will leave a much bitterer taste on my tongue than it used to. If I were to use the discounted expected utility model again as a decision support framework, I will be aware that literally my passion and impatience cheat my rationality. The decision maker seeking a rational model without bitter taste who wants to account for risk aversion has to employ the representation of Theorem 5: an undiscounted model. His weights on future welfare reflect how much he knows about the consequences of his (in)actions. Such a decision maker follows the normatively most desirable path.

Normative desirability, however, often comes at a cost. It is fair to ask how much an evaluation in this framework differs from an evaluation underlying observed choices. I argued that the difference between the present model and observed choice in a context of deliberate decision making might be smaller than we would initially think. Individuals or agencies do not foresee the future. They face serious uncertainties regarding income, funding, health, changes in the social or business network, technological innovations, natural disasters, terrorist attacks, and many more. I showed that such uncertainties can act as if agents discount future welfare and noted empirical evidence for such discounting in the asset pricing literature. However, once the agents can influence state-specific payoffs or the uncertainty itself, the two models imply very different results.

A special feature of the present framework is its finite planning horizon. It avoids Kreps & Porteus' (1978) finding that ordering certain consumption streams over an infinite time horizon requires a strictly positive time preference. I discussed whether the model might be unduly sensitive to the choice of the time horizon. If we are to account for the full consequences of an action or policy, we should include those periods for which we foresee (probabilistically) how our actions will change the future. At some finite point in time, the differences caused by our actions blur. Then, the weak future independence assumption accompanying stationarity kicks in: if we drop a common future when comparing two scenarios, the resulting ranking does not change. Thus, the present framework requires us to evaluate precisely what changes we can realistically foresee in comparing different scenarios. Then it is insensitive to the time horizon. In contrast, the discounted expected utility model chips away continuously at the time horizon by means of pure time preference. We know that the evaluation of long-term policy scenarios is highly sensitive to this utility discount rate. At the same time, this continuous chipping away at the time horizon is convenient. For example, in the climate change context, with a 3% rate of pure time preference, we do not have to worry about our model's quality in predicting global warming and sea level rise beyond some 150 years, even if these changes persist over several thousand years. Passion and impatience have always been easier to follow than deliberate decision making.

References

Aczél, J. (1966), *Lectures on Functional Equations and their Applications*, Academic Press, New York.

Bansal, R., Kiku, D. & Yaron, A. (2010), 'Long run risks, the macroeconomy, and asset prices', *American Economic Review: Papers & Proceedings* **100**, 542–546.

Bansal, R. & Yaron, A. (2004), 'Risks for the long run: A potential resolution of asset pricing puzzles', *The Journal of Finance* **59**(4), 1481–509.

Billingsley, P. (1995), *Probability and Measure*, John Wiley & Sons, New York.

Bommier, A. (2007), 'Risk aversion, intertemporal elasticity of substitution and correlation aversion', *Economic Bulletin* **4**(29), 1–8.

Bouakiz, M. & Sobel, M. J. (1992), 'Inventory control with an exponential utility criterion', *Operations Research* **40**(3), 603–608.

Broome, J. (1992), *Counting the Cost of Global Warming*, White Horse Press, Cambridge.

Campbell, J. Y. (1996), 'Understanding risk and return', *The Journal of Political Economy* **104**(2), 298–345.

Chapman, G. B. (2000), 'Preferences for improving and declining sequences of health outcomes', *Journal of Behavioral Decision Making* **13**, 203–218.

Chew, S. H. & Epstein, L. G. (1989), 'The structure of preferences and attitudes towards the timing of the resolution of uncertainty', *International Economic Review* **30**(1), 103–17.

Chew, S. H. & Epstein, L. G. (1991), Recursive utility under uncertainty, *in* M. A. Khan & N. C. Yannelis, eds, 'Equilibrium Theory in Infinite Dimensional Spaces', Springer, Heidelberg, pp. 352–369.

Cox, R. T. (1946), 'Probability, frequency and reasonable expectation', *American Journal of Physics* **14**(1), 1–13.

Cox, R. T. (1961), *The Algebra of Probable Inference*, Johns Hopkins University Press, Baltimore.

Epstein, L. (1992), Behavior under risk: Recent developments in theory and applications, *in* J.-J. Laffont, ed., 'Advances in Economic Theory', Vol. 2 of *Econometric Society Monographs*, Cambridge University Press, Cambridge, chapter 1, pp. 1–63.

Epstein, L. G. & Zin, S. E. (1989), 'Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework', *Econometrica* **57**(4), 937–69.

Epstein, L. G. & Zin, S. E. (1991), 'Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis', *Journal of Political Economy* **99**(2), 263–86.

Gilboa, I. & Schmeidler, D. (1989), 'Maxmin expected utility with non-unique prior', *Journal of Mathematical Economics* **18**(2), 141–53.

Hansen, L. P. & Sargent, T. J. (1995), 'Discounted linear exponential quadratic gaussian control', *IEEE Transactions on Automatic Control* **40**(5), 968–71.

Hansen, L. P. & Sargent, T. J. (2001), 'Robust control and model uncertainty', *American Economic Review* **91**(2), 60–66.

Hansen, L. P. & Sargent, T. J. (2007), *Robustness*, Princeton University Press.

Harrod, S. R. F. (1948), *Towards a dynamic economics*, Macmillan, London.

Jacobsen, D. H. (1973), 'Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games', *IEEE Transactions on Automatic Control* **18**(2), 124–131.

Jaynes, E. T. (2003), *Probability Theory: The Logic of Science*, Cambridge University Press, Cambridge.

Kihlstrom, R. E. & Mirman, L. J. (1974), 'Risk aversion with many commodities', *Journal of Economic Theory* **8**(3), 361–88.

Kihlstrom, R. E. & Mirman, L. J. (1981), 'Constant, increasing and decreasing risk aversion with many commodities', *Review of Economic Studies* **48**(2), 271–80.

Koopmans, T. C. (1960), 'Stationary ordinal utility and impatience', *Econometrica* **28**(2), 287–309.

Koopmans, T. C. (1963), On the concept of optimal economic growth, Cowles Foundation Discussion Papers 163, Cowles Foundation, Yale University.

Koopmans, T. C. (1972), 'Representation of preference orderings over time', *Cowles Foundation Paper* **366**(b), 79–100.

Kreps, D. M. (1979), 'A representation theorem for preference for flexibility', *Econometrica* **47**, 565–577.

Kreps, D. M. & Porteus, E. L. (1978), 'Temporal resolution of uncertainty and dynamic choice theory', *Econometrica* **46**(1), 185–200.

Loewenstein, G. F. & Sicherman, N. (1991), 'Do workers prefer increasing wage profiles?', *Journal of Labor Economics* **9**(1), 67–84.

Nakamura, E., Steinsson, J., Barro, R. & Ursua, J. (2010), 'Crises and recoveries in an empirical model of consumption disasters', *NBER* **15920**.

Nordhaus, W. (2008), *A Question of Balance: Economic Modeling of Global Warming*, Yale University Press, New Haven. Online preprint: A Question of Balance: Weighing the Options on Global Warming Policies.

Nordhaus, W. D. (2007), 'A review of the Stern review on the economics of climate change', *Journal of Economic Literature* **45**(3), 686–702.

Pigou, A. C. (1932), *The economics of welfare*, 4th edn, Macmillan, London.

Ramsey, F. P. (1928), 'A mathematical theory of saving', *The Economic Journal* **38**(152), 543–559.

Richard (1975), 'Multivariate risk aversion, utility independence and separable utility functions', *Management Science* **22**(1), 12–21.

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

Solow, R. M. (1974), 'The economics of resources of the resources of economics', *American Economic Review* **64**(2), 1–14.

Starmer, C. (2000), 'Developments in non-expected utlity theory - the hunt for a descriptive theory of choice unter risk', *Journal of Economic Literature* **38**(2), 332–382.

Stern, N., ed. (2007), *The Economics of Climate Change: The Stern Review*, Cambridge University Press, Cambridge.

Traeger, C. (2007), Theoretical aspects of long-term evaluation in environmental economics, PhD thesis, University of Heidelberg. Published online at http://www.ub.uni-heidelberg.de/archiv/7049.

Traeger, C. (2010), 'Intertemporal risk aversion - or - wouldn't it be nice to know whether Robinson is risk averse?', *CUDARE Working Paper No 1102* .

Vissing-Jørgensen, A. & Attanasio, O. P. (2003), 'Stock-market participation, intertemporal substitution, and risk-aversion', *The American Economic Review* **93**(2), 383–391.

von Neumann, J. & Morgenstern, O. (1944), *Theory of Games and Economic Behaviour*, Princeton University Press, Princeton.

Wakker, P. (1988), 'The algebraic versus the topological approach to additive representations', *Journal of Mathematical Psychology* **32**, 421–435.

Wakker, P. P. (1989), *Additive Representations of Preference : A new foundation of decision analysis*, Kluwer Academic Press, Dordrecht.

Weil, P. (1990), 'Nonexpected utility in macroeconomics', *The Quarterly Journal of Economics* **105**(1), 29–42.

Whittle, P. (1981), 'Risk-sensitive linear/quadratic/gaussian control', *Advances in Applied Probability* **13**(4), 764–777.

Whittle, P. (1990), *Risk Sensitive Optimal Control*, Wiley, New York.

Whittle, P. (2002), 'Risk sensitivity, a strangely pervasive concept', *Macroeconomic Dynamics* **6**, 5–18.

Appendix

A Proofs for Section 3

Proof of Theorem 1: The proof builds on Theorem 2 in Traeger (2010) delivering a representation for non-stationary preferences. I show that axiom A6 makes it possible to pick coinciding utility function for evaluating outcomes in different periods and to translate the recursive description of aggregate utility in Traeger (2010) into the 'current value' form of the representation stated here.

Sufficiency: For a given $a \in \mathrm{I\!R}_{++}$, the proof makes use of the notation $\mathsf{A}^a = \{\mathfrak{a}^a \in \mathsf{A} : \mathfrak{a}^a(z) = a\,z + b\,,\, b \in \mathrm{I\!R}\}$. I define for a given sequence of continuous functions $u_t : X \to \mathrm{I\!R}$ with range U_t, $t \in \{1, ..., T\}$, the normalization constants ϑ_t by setting $\vartheta_T = 0$ and for $t < T$

$$\vartheta_t = \frac{\overline{U}_{t+1}\underline{U}_t - \underline{U}_{t+1}\overline{U}_t}{\overline{U}_t - \underline{U}_t}\,.$$

By axioms A1-A5, Corollary 3 in Traeger (2010) implies the existence of the following preference representation for $\succeq$. There exist continuous functions $u_t : X \to \mathrm{I\!R}$ with range U_t and strictly increasing and continuous functions $f_t^* : \mathrm{I\!R} \to \mathrm{I\!R}$ such that defining recursively the aggregate welfare functions $\tilde{u}_t^* : \tilde{X}_t \to \mathrm{I\!R}$ by $\tilde{u}_T^*(x_T) = u(x_T)$ and

$$\tilde{u}_{t-1}^*(x_{t-1}, p_t) = u_{t-1}(x_{t-1}) + \mathcal{M}^{f_t^*}(p_t, \tilde{u}_t^*) + \frac{\vartheta_{t-1}}{\theta_t}$$

the expression $\mathcal{M}^{f_t^*}(p_t, \tilde{u}_t^*)$ represents the lotteries p_t in the sense of equation (6). Moreover, the sequences of tuples $(u_t, f_t^*)_{t\in\{1,...,T\}}$ and $(u_t', f_t^{*\prime})_{t\in\{1,...,T\}}$ both represent $\succeq$, if and only if, for some $a \in \mathrm{I\!R}_{++}$ there exist affine transformations $\mathfrak{a}_t \in \mathsf{A}$ and $\mathfrak{a}_t^a \in \mathsf{A}^a$ for all $t \in \{1, ..., T\}$ such that $(u_t', {f^*}_t|_{\tilde{U}_t}) = (\mathfrak{a}_t^a\, u_t\,,\, \mathfrak{a}_t f_t^{*\prime}|_{\tilde{U}_t'}\, \mathfrak{a}_t^{a\,-1}\,)$.

To translate axiom A6 into the representation stated above, recall that $\mathcal{M}^{f_\tau^*}(\tilde{x}_\tau, \tilde{u}_\tau^*) = \tilde{u}_\tau^*(\tilde{x}_\tau)$. Recursive application yields for $\mathrm{x} \in \mathsf{X}^t$ that $\mathcal{M}^{f_t^*}(\mathrm{x}, \tilde{u}_t^*) = \sum_{\tau=t}^T u_\tau(\mathrm{x}_\tau) + c_t$ for some constants $c_t \in \mathrm{I\!R}$. Therefore, axiom A6 with $\mathrm{x} = (\mathrm{x}_2, ..., \mathrm{x}_T)$ translates into the requirement

$$\begin{aligned} &\sum_{\tau=2}^T u_{\tau-1}(\mathrm{x}_\tau) + \cancel{u_T(x)} \geq \sum_{\tau=2}^T u_{\tau-1}(\mathrm{x}'_\tau) + \cancel{u_T(x)} \\ \Leftrightarrow &\sum_{\tau=2}^T u_\tau(\mathrm{x}_\tau) \qquad\qquad\quad \geq \sum_{\tau=2}^T u_\tau(\mathrm{x}'_\tau)\,. \end{aligned}$$

for all $\mathrm{x}, \mathrm{x}' \in \mathsf{X}^2$. The above equivalence implies that $\sum_{\tau=2}^T u_\tau(\mathrm{x}_\tau)$ and $\sum_{\tau=2}^T u_{\tau-1}(\mathrm{x}_\tau)$ are both representations for $\succeq_2|_{\mathsf{X}^2}$. In consequence, by the uniqueness result stated above, there exist $a \in \mathrm{I\!R}_{++}$ and $b_t \in \mathrm{I\!R}, t \in \{1, ..., T-1\}$, such that $u_t = a\,u_{t+1} + b_t$ for all $t \in \{1, ..., T-1\}$.[33] I use this freedom in $(u_t)_{t\in\{1,...,T\}}$ to eliminate the affine displacement

[33] Here it is $u_t' = u_{t+1}$. Coincidence of the representations (only) on the certain outcome paths is enough to assure the uniqueness result for $(u_t)_{t\in\{1,...,T\}}$.

parameter b_t by defining $u'_t = u_t - \sum_{\tau=t}^{T-1} a^{\tau-t} b_\tau$ for $t \in \{1, ..., T-1\}$. For u'_t find that $u'_t = u_t - \sum_{\tau=t}^{T-1} a^{\tau-t} b_\tau = a u_{t+1} + b_t - b_t - a \sum_{\tau=t+1}^{T-1} a^{\tau-t} b_\tau = a u'_{t+1}$. Letting $u = u'_1$ and $\beta = a^{-1}$ yields $u_t = \beta^{t-1} u$. The change of the representing utility functions from u_t to u'_t corresponds to the affine transformations $u'_t = \mathfrak{a}^1_t\, u_t$ with $\mathfrak{a}^1_t(z) = z - \sum_{\tau=t}^{T-1} a^{\tau-t} b_\tau$ and requires an according transformation of the representing functions f^*_t to $f^{*\prime}_t = f^*_t\, \mathfrak{a}^{1\,-1}$. Then, the sequence of tuples $(\beta^{t-1}u, f^{*\prime}_t)_{t\in\{1,...,T\}}$ represents $\succeq$.

The normalization constants ϑ_t for the representing tuples $(\beta^{t-1}u, f^{*\prime}_t)_{t\in\{1,...,T\}}$ are

$$\vartheta_t = \frac{\overline{U}_{t+1}\underline{U}_t - \underline{U}_{t+1}\overline{U}_t}{\overline{U}_t - \underline{U}_t} = \frac{\beta^t \overline{U}\,\beta^{t-1}\underline{U} - \beta^t \underline{U}\,\beta^{t-1}\overline{U}}{\beta^{t-1}(\overline{U} - \underline{U})} = 0\,.$$

Using the tuples $(\beta^{t-1}u, f^{*\prime}_t)_{t\in\{1,...,T\}}$ in the representation therefore yields the following recursive definition of aggregate utility

$$\tilde{u}^*_t(\cdot,\cdot) = u'_t(\cdot) + \mathcal{M}^{f^{*\prime}_{t+1}}(\cdot, \tilde{u}^*_{t+1}) + 0 = \beta^{t-1}u(\cdot) + \mathcal{M}^{f^{*\prime}_{t+1}}(\cdot, \tilde{u}^*_{t+1})\,.$$

which is equivalent to

$$\beta^{1-t}\,\tilde{u}^*_t(\cdot, p_{t+1}) = u(\cdot) + \beta\ \beta^{1-(t+1)} f^{*\,\prime\,-1}_{t+1} \left(\mathrm{E}_{p_{t+1}} f^{*\,\prime}_{t+1} \left(\beta^{(t+1)-1}\beta^{1-(t+1)}\ \tilde{u}^*_{t+1}\right)\right).$$

Defining $\tilde{u}_t = \beta^{1-t}\tilde{u}^*_t$ and $f_t(z) = f^{*\prime}_t(\beta^{t-1}z)$ for all t yields the representation

$$\tilde{u}_t(\cdot, p_{t+1}) = u(\cdot) + \beta\, f^{-1}_{t+1}\Big(\mathrm{E}_{p_{t+1}} f_{t+1} \circ \tilde{u}_{t+1}\Big)$$

stated in the theorem.

Necessity: Axioms A1-A5 follow immediately from the necessity part in Theorem 2 in Traeger (2010). To see that axiom A6 is always satisfied in the representation observe that $\mathcal{M}^{f_\tau}(\tilde{x}_\tau, \tilde{u}_\tau) = \tilde{u}_\tau(\tilde{x}_\tau)$ and recursively $\mathcal{M}^{f_t}(\mathsf{x}, \tilde{u}_t) = \sum_{\tau=t}^{T} \beta^{\tau-t}\, u(\mathsf{x}_\tau)$. Then, axiom A6 is seen to hold by verifying the following equivalence

$$\begin{aligned} &\textstyle\sum_{\tau=1}^{T-1} \beta^{\tau-1}\, u(\mathsf{x}_{\tau+1}) + \cancel{\beta^{T-1}u(x)} \geq \sum_{\tau=1}^{T-1} \beta^{\tau-1}\, u(\mathsf{x}'_{\tau+1}) + \cancel{\beta^{T-1}u(x)} \\ \Leftrightarrow &\textstyle\sum_{\tau=2}^{T} \beta^{\tau-2}\, u(\mathsf{x}_\tau) \geq \sum_{\tau=2}^{T} \beta^{\tau-2}\, u(\mathsf{x}'_\tau)\,. \end{aligned}$$

Uniqueness: The representation is a special case of the representation in the sense of Corollary 3 in Traeger (2010), stated at the beginning of this proof. Here, $u_t = \beta^{t-1}\, u$ and $f^*_t = f_t \circ \mathfrak{a}$ for some $\mathfrak{a} \in \mathsf{A}$. Thus, the uniqueness result follows immediately from Corollary 3 in Traeger (2010). □

B Proofs for Section 4

Proof of Theorem 2: The proof translates axiom A7$^{\mathrm{s}}$ into the representation of Theorem 1. Using the premise, I translate the second line of the axiom into a concavity condition for f_t. The resulting equation (17) would be a straight forward concavity condition, if the equation had to hold for all convex combinations in the argument of f_t. However, the premise only requires equation (17) to hold for a limited set of convex combinations. The proof that the limited requirement still yields concavity of the functions f_t works, with one exception, analogous to the according proof for the non-stationary analysis carried out in Traeger (2010). I point out the step that differs and, hereafter, refer the reader to the proof of Theorem 3 in Traeger (2010).

The proof is divided into four parts. The first part derives the (restricted) concavity condition. The second part relates the rest of the proof to the one carried out in Traeger (2010). The third part deals with statements b) and c) of the theorem. The fourth part takes care of the necessity of the axioms for weak/strict intertemporal risk aversion/seeking.

Sufficiency: Part I: This part of the proof translates axiom A7$^{\text{s}}$ into the representation of Theorem 1. I start with the first line, i.e the premise:

$$\bar{\mathrm{x}} \quad \sim_t \quad \mathrm{x}$$

$$\Rightarrow \sum_{\tau=t}^{T} \beta^{\tau-t}\, u(\bar{x}) = \sum_{\tau=t}^{T} \beta^{\tau-t}\, u(\mathrm{x}_\tau)\,. \tag{14}$$

The existence of $\tau \in \{t,...,T\}$ such that $(\bar{\mathrm{x}}_{-\tau}, \mathrm{x}_\tau) \not\sim_t \bar{\mathrm{x}}$ translates into

$$u(\mathrm{x}_\tau) \neq u(\bar{x}) \text{ for some } \tau \in \{t,...,T\}. \tag{15}$$

The second line of axiom A7$^{\text{s}}$ yields

$$\bar{\mathrm{x}} \quad \succ_T \quad \textstyle\sum_{i=t}^{T} \frac{1}{T-t+1}\, (\bar{\mathrm{x}}_{-i}, \mathrm{x}_i)$$

$$\Rightarrow \quad \textstyle\sum_{\tau=t}^{T} \beta^{\tau-t} u(\bar{x}) \quad > \quad f_t^{-1}\Big[\sum_{i=t}^{T} \frac{1}{T-t+1} f_t \Big[\sum_{\tau=t}^{T} \beta^{\tau-t} u\left((\bar{\mathrm{x}}_{-i}, \mathrm{x}_i)_\tau\right) \Big]\Big]$$

$$\Rightarrow f_t\Big[\textstyle\sum_{\tau=t}^{T} \beta^{\tau-t} u(\bar{x})\Big] \quad > \quad \sum_{i=t}^{T} \frac{1}{T-t+1} f_t \Big[\sum_{\tau=t}^{T} \beta^{\tau-t} u\left((\bar{\mathrm{x}}_{-i}, \mathrm{x}_i)_\tau\right) \Big].$$

Using equation (14), the left hand side can be transformed as follows:

$$\begin{aligned} f_t\Big[\textstyle\sum_{\tau=t}^{T} \beta^{\tau-t} u(\bar{x})\Big] &= f_t\left[\frac{T-t}{T-t+1} \left[\textstyle\sum_{\tau=t}^{T} \beta^{\tau-t} u(\bar{x})\right] + \frac{1}{T-t+1} \left[\textstyle\sum_{\tau=t}^{T} \beta^{\tau-t} u(\mathrm{x}_\tau)\right]\right] \\ &= f_t\left[\frac{1}{T-t+1} \left[\textstyle\sum_{i=t}^{T} \sum_{\tau=t}^{T} \beta^{\tau-t} u\left((\bar{\mathrm{x}}_{-i}, \mathrm{x}_i)_\tau\right)\right]\right] \\ &= f_t\left[\textstyle\sum_{i=t}^{T} \frac{1}{T-t+1} \left[\textstyle\sum_{\tau=t}^{T} \beta^{\tau-t} u\left((\bar{\mathrm{x}}_{-i}, \mathrm{x}_i)_\tau\right)\right]\right]\,, \end{aligned}$$

yielding the inequality

$$f_t\left[\sum_{i=t}^T \frac{1}{T-t+1}\left[\sum_{\tau=t}^T \beta^{\tau-t} u\left((\bar{\mathsf{x}}_{-i}, \mathsf{x}_i)_\tau\right)\right]\right]$$

$$> \sum_{i=t}^T \frac{1}{T-t+1} f_t\Big[\sum_{\tau=t}^T \beta^{\tau-t} u\left((\bar{\mathsf{x}}_{-i}, \mathsf{x}_i)_\tau\right)\Big]\,. \tag{16}$$

Define the function $\tilde{z} : \mathsf{X}^t \to \mathrm{I\!R}$ by $\tilde{z}(\mathsf{x}) = \sum_{\tau=t}^T \beta^{\tau-t} u\left(\mathsf{x}_\tau\right)$. Restricting the domain to those consumption paths that satisfy condition (15), the function is onto the interior of the interval $\tilde{U}_t$ which I denote by $\Gamma = \left(\sum_{\tau=t}^T \beta^{\tau-t}\underline{U}\,,\, \sum_{\tau=t}^T \beta^{\tau-t}\overline{U}\right)$. For given consumption paths $\bar{\mathsf{x}}, \mathsf{x} \in \mathsf{X}^t$ I define $z_i = \tilde{z}\big((\bar{\mathsf{x}}_{-i}, \mathsf{x}_i)\big)$. In this notation equation (16) becomes

$$f_t\left(\sum_{i=t}^T \frac{1}{T-t+1}\, z_i\right) \quad > \quad \sum_{i=t}^T \frac{1}{T-t+1} f_t(z_i). \tag{17}$$

If equation (17) had to hold for all $z_i \in \Gamma$ it would be a straight forward condition for strict convexity of f_t. However axiom A7$^{\text{s}}$ does not immediately imply that the equation has to be met for every sequence $(z_i)_{i\in\{t,\ldots,T\}}$ with $z_i \in \Gamma$. Equation (17) only has to hold for sequences $(z_i)_{i\in\{t,\ldots,T\}}$ which are generated by consumption paths $\bar{\mathsf{x}}, \mathsf{x} \in \mathsf{X}^t$ that satisfy the premise of axiom A7$^{\text{s}}$.

Part II: It is left to show that equation (17) implies strict concavity of f_t also if it only has to hold for sequences $(z_i)_{i\in\{t,\ldots,T\}}$ which are generated by consumption paths $\bar{\mathsf{x}}, \mathsf{x} \in \mathsf{X}^t$ that satisfy the premise of axiom A7$^{\text{s}}$. This part of the proof is mostly analogous to the proof of Theorem 3 in Traeger (2010), which first shows that the condition implies local strict concavity of f_t, then weak concavity on the entire set Γ, and finally strict concavity on Γ.[34] The only difference in the proof is that the stationarity assumption in this paper permits to generate every point $z^o \in \Gamma$ from a constant consumption path, i.e. for all $z^o \in \Gamma$ exists $\bar{\mathsf{x}} \in \mathsf{X}$ such that $z^o = \tilde{z}(\bar{\mathsf{x}})$. This fact is immediate from observing that $\tilde{z}(\bar{\mathsf{x}}) = u\left(\bar{x}\right) \sum_{\tau=t}^T \beta^{\tau-t}$ is continuous as a function of $\bar{x}$ and onto Γ. Therefore, for every $z^o \in \Gamma$ a set of local perturbations x around the constant consumption $\bar{\mathsf{x}}$ generating z^o, which satisfy the premise of axiom A7$^{\text{s}}$, can be used to derive local convexity of f_t. In the non-stationary setting, not all $z^o \in \Gamma$ can be generated by a constant consumption path. That fact slightly complicates the respective analysis and, there, requires a stronger axiom for the definition of intertemporal risk aversion (enlarging the domain on which the axiom of intertemporal risk aversion restricts preferences).

Part III: Assertion b) is obtained by replacing A7$^{\text{s}}$ by A7$^{\text{w}}$and the strict inequaties by their weak counterparts.[35] A decision maker is intertemporal risk neutral if his preferences satisfy weak risk seeking as well as weak risk aversion. Therefore, assertion b) implies

[34]The step proving local strict concavity corresponds to part two of the respective proof. The fact that the functions f_t are not assumed to be differentiable slightly complicates the step from local to global concavity which is shown in part three of the proof of Theorem 3 in Traeger (2010).

[35]In this case the second step in part three becomes redundant.

that the function $f_t \circ g_t^{-1}$ has to be concave and convex at the same time and, thus, linear.
Necessity: Part IV: Necessity is implied by Theorem 3 in Traeger (2010) for $\mathsf{x} = \bar{\mathsf{x}}$. □

Proof of proposition 1: The proposition is an immediate consequence of proposition 4 in Traeger (2010). For part a) observe that fixing the zero level for $u(x^0)$ in the representation of theorem 1 fixes the zero level of u_t for all periods. □

C Proofs for Section 5

Proof of Theorem 3: The proof is divided into four parts. In the first part, I translate axiom A8 into the representation of Theorem 1 and derive a functional equation for the functions f_t parameterizing uncertainty aggregation. I translate the solution back into the representation in part two of the proof. Part three shows necessity of the axioms and part four shows uniqueness and calculates the measures of intertemporal risk aversion.
Sufficiency: Part I: First, note that axiom A8 implies axiom A6 by choosing $\mathsf{x} = \mathsf{x}'$. Therefore, a representation in terms of Theorem 1 exists. In the following, I translate axiom A8 for $t \in \{1, ..., T-1\}$ into the latter representation. Note that, by definition of x as an element of X^{t+1}, the period τ entry of the consumption path $(\mathsf{x}, x) \in \mathsf{X}^t$ corresponds to $(\mathsf{x}, x)_\tau = \mathsf{x}_{\tau+1}$ for $\tau \in \{t, ..., T-1\}$. The left hand side of the equivalence in axiom A8 translates into

$$\tfrac{1}{2}(\mathsf{x}, x^0) + \tfrac{1}{2}(\mathsf{x}', x^0) \qquad \succeq_t (\mathsf{x}'', x^0)$$
$$\Leftrightarrow f_t^{-1}\Big\{\tfrac{1}{2} f_t\Big[\textstyle\sum_{\tau=t}^{T-1} \beta^{\tau-t} u(\mathsf{x}_{\tau+1}) + \beta^{T-t} u(x)\Big]$$
$$+\tfrac{1}{2} f_t\Big[\textstyle\sum_{\tau=t}^{T-1} \beta^{\tau-t} u(\mathsf{x}'_{\tau+1}) + \beta^{T-t} u(x)\Big]\Big\} \geq \Big[\textstyle\sum_{\tau=t}^{T-1} \beta^{\tau-t} u(\mathsf{x}''_{\tau+1}) + \beta^{T-t} u(x)\Big].$$

For given consumption $\mathsf{x}, \mathsf{x}', \mathsf{x}'' \in \mathsf{X}$ and $x \in X$ I define $S = \sum_{\tau=t+1}^{T} \beta^{\tau-(t+1)} u(\mathsf{x}_\tau)$, $S' = \sum_{\tau=t+1}^{T} \beta^{\tau-(t+1)} u(\mathsf{x}'_\tau)$, $S'' = \sum_{\tau=t+1}^{T} \beta^{\tau-(t+1)} u(\mathsf{x}''_\tau)$ and $A = \beta^{T-t} u(x)$. Varying the consumption paths x, x' and x'' in X^{t+1} implies a variation of S, S' and S'' over the interval $[\frac{1-\beta^{T-t}}{1-\beta} \underline{U}, \frac{1-\beta^{T-t}}{1-\beta} \overline{U}]$. Similarly, as x is varied in X the value A takes on any number in the interval $[\beta^{T-t} \underline{U}, \beta^{T-t} \overline{U}]$. The introduced notation translates the above inequality, corresponding to the left hand side of the equivalence in axiom A8, into

$$f_t^{-1}\Big\{\tfrac{1}{2} f_t\big[S + A\big] + \tfrac{1}{2} f_t\big[S' + A\big]\Big\} - A \geq S''. \tag{18}$$

In the same notation the right hand side of the equivalence in axiom A8 translates into

$$f_{t+1}^{-1}\Big\{\tfrac{1}{2}f_{t+1}\big[S\big]+\tfrac{1}{2}f_{t+1}\big[S'\big]\Big\}\geq S''. \tag{19}$$

For every lottery $p_{t+1}\in P_{t+1}$ exists a certainty equivalent that is a certain consumption path.[36] In consequence, for any choice of $\mathsf{x},\mathsf{x}'\in \mathsf{X}^{t+1}$ exists a certainty equivalent $\mathsf{x}''\in \mathsf{X}^{t+1}$ to the lottery $\frac{1}{2}\mathsf{x}+\frac{1}{2}\mathsf{x}'\in P_{t+1}$ such that equation (19) holds with equality. Then, by axiom A8 also equation (18) has to hold with equality. Combining these two equalities by replacing S'' yields the requirement

$$f_t^{-1}\Big\{\tfrac{1}{2}f_t\big[S+A\big]+\tfrac{1}{2}f_t\big[S'+A\big]\Big\}-A=f_{t+1}^{-1}\Big\{\tfrac{1}{2}f_{t+1}\big[S\big]+\tfrac{1}{2}f_{t+1}\big[S'\big]\Big\}. \tag{20}$$

for all $S,S'\in[\frac{1-\beta^{T-t}}{1-\beta}\,\underline{U},\frac{1-\beta^{T-t}}{1-\beta}\,\overline{U}]$ and $A\in[\beta^{T-t}\,\underline{U},\beta^{T-t}\,\overline{U}]$.

Part II: Observe that the right hand side of equation (20) is independent of A. Therefore, the left hand side has to be constant in A. The condition corresponds to a functional equation solved by Aczél (1966, 153). The only solutions for a continuous function f_t satisfying equation (20) are $f_t(z)=a_t\exp(-\xi_t z)+b_t$ and $f_t(z)=a_t z+b_t$ with $a_t,b_t\in\mathbb{R}$ and $\xi_t\in\mathbb{R},\xi_t\neq 0$ for $t\in\{1,...,T-1\}$.

Acknowledging that the left hand side of equation (20) is independent of A, both sides of equation (20) characterize an uncertainty aggregation rule with respect to a variation in S and S'. The parameterizing functions of uncertainty aggregation rules are unique up to affine transformations. As both sides have to coincide, f_{t+1} has to be an affine transformation of f_t. But affine transformations of the parameterizing function of an uncertainty aggregation rule do not affect the representation. Therefore, I can choose $f_t(z)=\exp(-\xi z)$ for all $t\in\{1,...,T\}$ in case of the first class of solutions. Note that Theorem 1 is stated in term of increasing functions f_t. Because $f_t(z)=\exp(-\xi z)$ can decrease in z I should more precisely pick $f_t=\text{sgn}(-\xi)\exp(-\xi z)$. However, the multiplication with a negative number does not change the resulting uncertainty aggregation rule. In the case that f_t is linear for one and thus for all t, I choose $f_t=\text{id}$ for all $t\in\{1,...,T\}$. The latter case is identified with $\xi=0$. In fact, observe that defining $\mathcal{M}^{\exp^0}=\lim_{\xi\to 0}\mathcal{M}^{\exp^{-\xi}}$ yields by l'Hospital's rule:

$$\begin{aligned}\mathcal{M}^{\exp^0}(p_t,\tilde{u}_t)&\equiv\lim_{\xi\to 0}\mathcal{M}^{\exp^{-\xi}}(p_t,\tilde{u}_t)=\lim_{\xi\to 0}\frac{\ln\big[\int dp_t\exp(-\xi\tilde{u}_t)\big]}{-\xi}\\&=\lim_{\xi\to 0}\frac{\frac{\partial}{\partial\xi}\ln\big[\int dp_t\exp(-\xi\tilde{u}_t)\big]}{\frac{\partial}{\partial\xi}-\xi}=\lim_{\xi\to 0}\frac{\int dp_t\tilde{u}_t\exp(-\xi\tilde{u}_t)}{\int dp_t\exp(-\xi\tilde{u}_t)}\\&=\frac{\int dp_t\tilde{u}_t}{1}=E_{p_t}\tilde{u}_t\,.\end{aligned}$$

[36] See proof of Theorem 2 in Traeger (2010), induction hypothesis 2.

Substituting the restricted uncertainty aggregation rules into the representation of Theorem 1 yields the representation given in Theorem 3.

Necessity: Part III: The representation is a special case of Theorem 1. Therefore, axioms A1-A5 follow immediately from the necessity part of Theorem 1. The following calculation shows that axiom A8 is satisfied for all $t \in \{1, ..., T-1\}$, $x \in X$ and $\mathfrak{x}, \mathfrak{x}', \mathfrak{x}'' \in \mathfrak{X}^{t+1}$ in the case $\xi \neq 0$:

$$\tfrac{1}{2}(\mathfrak{x}, x) + \tfrac{1}{2}(\mathfrak{x}', x) \succeq_t (\mathfrak{x}'', x)$$

$$\Leftrightarrow \tfrac{1}{-\xi} \ln \Big(\tfrac{1}{2} \exp \big[-\xi \textstyle\sum_{\tau=t}^{T-1} \beta^{\tau-t} u(\mathfrak{x}_{\tau+1}) \big] \exp \big[-\xi \beta^T u(x) \big]$$

$$+ \tfrac{1}{2} \exp \big[-\xi \textstyle\sum_{\tau=t}^{T-1} \beta^{\tau-t} u(\mathfrak{x}'_{\tau+1}) \big] \exp \big[-\xi \beta^T u(x) \big] \Big) \geq \displaystyle\sum_{\tau=t}^{T-1} \beta^{\tau-t} u(\mathfrak{x}''_{\tau+1}) + \beta^T u(x)$$

$$\Leftrightarrow \tfrac{1}{-\xi} \ln \Big(\tfrac{1}{2} \exp \Big[-\xi \textstyle\sum_{\tau=t+1}^{T} \beta^{\tau-(t+1)} u(\mathfrak{x}_\tau) \Big] + \tfrac{1}{2} \exp \Big[-\xi \textstyle\sum_{\tau=t+1}^{T} \beta^{\tau-(t+1)} u(\mathfrak{x}'_\tau) \Big] \Big)$$

$$\geq \textstyle\sum_{\tau=t+1}^{T} \beta^{\tau-(t+1)} u(\mathfrak{x}''_\tau)$$

$$\Leftrightarrow \tfrac{1}{-\xi} \ln \Big(\tfrac{1}{2} \exp \Big[-\xi \tilde{u}_{t+1}(\mathfrak{x}) \Big] + \tfrac{1}{2} \exp \Big[-\xi \tilde{u}_{t+1}(\mathfrak{x}') \Big] \Big) \geq \tilde{u}_{t+1}(\mathfrak{x}'')$$

$$\Leftrightarrow \tfrac{1}{2}\mathfrak{x} + \tfrac{1}{2}\mathfrak{x}' \succeq_{t+1} \mathfrak{x}'' .$$

In the case $\xi = 0$ both sides of the above inequalities are linear. Thus, the term $\beta^T u(x)$ cancels as well and axiom A8 is satisfied.

Uniqueness: Part IV: The restriction that u is onto a given U^* fixes u uniquely, eliminating the affine freedom $\mathfrak{a}_0$ in the representation of Theorem 1. Therefore, by the uniqueness result of Theorem 1, the functions f_t are determined up to ('outer') affine transformations yielding unique measures

$$\text{AIRA}^*_t(z) = -\frac{\frac{d^2}{dz^2} f_t(z)}{\frac{d}{dz} f_t(z)} = -\frac{\frac{d^2}{dz^2} a_t \exp(-\xi z) + b_t}{\frac{d}{dz} a_t \exp(-\xi z) + b_t} = \xi$$

and accordingly

$$\text{RIRA}^*_t(z) = -\frac{\frac{d^2}{dz^2} f_t(z)}{\frac{d}{dz} f_t(z)} z = \xi z .$$

□

D Proofs for Section 6

Proof of Theorem 4: The proof builds on a representation theorem derived in Traeger (2007) for timing-indifferent preferences in a non-stationary setting.
Sufficiency: By axioms A1-A5 and timing indifference axiom A9 Theorem 4 in Traeger (2007) gives the following representation. For all $t \in \{1,...,T\}$ exist strictly increasing continuous functions $u_t : X \to \mathrm{I\!R}$, such that defining the functions $\tilde{u}_t : \mathrm{X}^t \to \mathrm{I\!R}$ by

$$\tilde{u}_t(\mathrm{x}^t) = \sum_{\tau=t}^T u_\tau(\mathrm{x}_\tau^t)$$

the expression $\mathcal{M}^{f_t}(p_t^{\mathrm{X}_t}, \tilde{u}_t)$ represents preferences in the sense of equation (11). For a derivation of the lotteries $p^{\mathrm{X}_t} \in \Delta(\mathrm{X}^t)$ from their respective counterparts $p_t \in P_t = \Delta(\tilde{X}_t)$ by 'integrating out' the information on the timing of risk resolution I refer to Traeger (2007). The identical reasoning as for the representation in Theorem 1 shows that certainty stationarity in the sense of axiom A6 ensures the existence of a discount factor $\beta \in \mathrm{I\!R}_{++}$ and a continuous function $u : X \to \mathrm{I\!R}$ such that the functions u_t in the above representation can be chosen as $u_t = \beta^{t-1}u$.

As shown in the proof of Theorem 4 in Traeger (2007) the functions f_t in the recursive representation differ from the parameterizing function of the uncertainty aggregation rules in the non-recursive representation by a normalization factor which, in the stationary setting, coincides with the discount factor. Therefore, the representation in Theorem 4 corresponds to a representation in the sense of Theorem 1 with the choice $f_t(z) = \exp(-\xi\beta^{t-1}z)$ as is easily verified by the following calculation:

$$\begin{aligned}
\mathcal{M}^{f_t}(p_t, \tilde{u}_t) &= f_t^{-1}\Big(\mathrm{E}_{p_t} f_t\Big[u(x_t) + \beta\, f_{t+1}^{-1}\Big(\mathrm{E}_{p_{t+1}} f_{t+1} \circ \tilde{u}_{t+1}\Big)\Big]\Big) \\
&= \tfrac{1}{-\beta^{t-1}\xi} \ln\Big(\mathrm{E}_{p_t} \exp\Big[-\beta^{t-1}\xi \Big\{u(x_t) + \beta\, \tfrac{1}{-\beta^t\xi} \ln\Big(\mathrm{E}_{p_{t+1}} \exp\big[-\beta^t\,\xi\, \tilde{u}_{t+1}\big]\Big)\Big\}\Big]\Big) \\
&= \tfrac{1}{-\beta^{t-1}\xi} \ln\Big(\mathrm{E}_{p_t} \exp\Big[-\beta^{t-1}\,\xi\, u(x_t)\Big] \mathrm{E}_{p_{t+1}} \exp\big[-\beta^t\,\xi\, \big(u(x_t) + \\
&\qquad\qquad \beta\, \mathcal{M}^{f_{t+2}}(p_{t+2}, \tilde{u}_{t+2})\big)\big]\Big) \\
&= \tfrac{1}{-\beta^{t-1}\xi} \ln\Big(\mathrm{E}_{p_t} \exp\big[-\beta^{t-1}\,\xi\, u(x_t)\big]\, \mathrm{E}_{p_{t+1}} \exp\big[-\beta^t\,\xi\, u(x_t)\big]\, \mathrm{E}_{p_{t+1}}...\Big)\ ,
\end{aligned}$$

which, omitting the details on integrating out the information on the timing of risk resolution, becomes

$$\begin{aligned}
\mathcal{M}^{f_t}(p_t, \tilde{u}_t) &= \tfrac{1}{-\beta^{t-1}\xi} \ln\Big(\mathrm{E}_{p_t^{\mathrm{X}}} \exp\big[-\beta^{t-1}\,\xi\, u(x_t)\big]\, \exp\big[-\beta^t\,\xi\, u(x_t)\big]...\Big) \\
&= \tfrac{1}{-\beta^{t-1}\xi} \ln\Big(\mathrm{E}_{p_t^{\mathrm{X}}} \exp\big[\xi \textstyle\sum_{\tau=t}^T -\beta^{\tau-1}\, u(x_\tau)\big]\Big)\ ,
\end{aligned}$$

which is a strictly increasing transformation of the expression in the representing equa-

tion (11). Therefore, the measures of intertemporal risk aversion are immediately calculated from the relation $f_t(z) = \exp(-\xi\beta^{t-1}z)$, yielding the results stated in the theorem.
Necessity: Necessity of the axioms follows from their necessity in the representations of 1 above and Theorem 4 in Traeger (2007).
Uniqueness: Uniqueness of u and the measures of intertemporal risk aversion is an immediate consequence of Proposition 1.

□

E Proofs for Section 7

Proof of Theorem 5: The assertion follows from a comparison of the time dependence of intertemporal risk aversion in the representation oft Theorem 3 and Theorem 4.
Sufficiency: The representation has to be a special case of the representations in Theorem 1 satisfying risk stationarity A8 and in Theorem 4 satisfying timing indifference A9. The proof of Theorem 3 shows that risk stationarity requires the representing functions f_t, characterizing uncertainty aggregation in Theorem 1, to coincide for different periods (up to possibly time dependent affine transformations that do not affect uncertainty aggregation). The proof of Theorem 4 shows that timing indifference requires these functions f_t in Theorem 1 to be of the form $f_t(z) = \text{sgn}(-\xi)\exp(-\xi\beta^{t-1}z)$ for all $t \in \{1, ..., T\}$.[37] Theorem 2 a) a strictly intertemporal risk averse decision maker implies $\xi > 0$. Thus, $\beta \neq 1$ implies by Theorem 4 a change of f_t over time that contradicts the constancy required resulting from risk stationarity (there is no affine transformation taking f_t into f'_t for $t \neq t'$ if it is of the given form). The only representation that satisfies both axioms has to have a discount factor $\beta = 1$.
Necessity: Necessity follows from necessity in Theorems 3 and 4 and the fact that both representation coincide for $\beta = 1$. □

Proof of proposition 2: The cumulative probability distribution of welfare in period t is defined as $\mathrm{P}_t(z) = \mathrm{p}_t^u([0, z])$ on the interval $[0, 1]$ and continuous from the right. Strictly increasing uncertainty implies by definition that equation (12) holds for all $z \in [0, 1]$ exhibiting equality for $z = 1$. Then Theorem 2 in Rothschild & Stiglitz (1970, 237) shows that for any strictly increasing concave function $f \in \mathcal{C}^0([0, 1])$

$$\mathrm{E}_{\mathrm{p}_t^u} f \geq \mathrm{E}_{\mathrm{p}_{t+1}^u} f \Rightarrow f^{-1}\big(\mathrm{E}_{\mathrm{p}_t^u} f\big) \geq f^{-1}\big(\mathrm{E}_{\mathrm{p}_{t+1}^u} f\big)\,.$$

[37]Because $f_t^*(z) = \exp(\xi\beta^{t-1}z)$ is decreasing in z the theorem has to be applied to the function $f_t = \text{sgn}(\xi)\exp(\xi\beta^{t-1}z)$ which increases and characterizes the same uncertainty aggregation rule.

The uncertainty aggregation rule $\mathcal{M}^{\exp^{-\xi}}(\mathrm{p}_t, u)$ is characterized by the strictly increasing and concave function $f(z) = -\exp(-\xi\, z)$ (see footnote 37). It follows from $f^{-1}\big(\mathrm{E}\mathrm{p}_t^u\, f\big) = f^{-1}\big(\mathrm{E}\mathrm{p}_t\, f \circ u\big) = \mathcal{M}^f(\mathrm{p}_t, u)$ that

$$\mathcal{M}^{\exp^{-\xi}}\big(\mathrm{p}_{t+1}, u\big) \leq \mathcal{M}^{\exp^{-\xi}}\big(\mathrm{p}_t, u\big) \text{ for all } t \in \{1, ..., T-1\} \ .$$

Strictness of this inequality follows from the existence of $z^* \in (0,1)$ such that equation (12) holds as a strict inequality and the fact that the distribution functions P_t are continuous from the right. In contrast, the fact that equation (12) holds with equality for $z = 1$ implies the existence of $\bar{u} \in \mathbb{R}$ such that $\mathrm{E}\ u(x_t) = \bar{u}$ for all $t \in \{1, ..., T\}$. Observing

$$\mathcal{M}^{\exp^{-\xi}}\big(p_1^{\mathbf{x}}, \tilde{u}_t\big) = \mathcal{M}^{\exp^{-\xi}}\big(\mathrm{p}_1, u\big) + \mathcal{M}^{\exp^{-\xi}}\big(\mathrm{p}_2, u\big) + ... + \mathcal{M}^{\exp^{-\xi}}\big(\mathrm{p}_T, u\big).$$

I can solve equation (13) recursively for the weights $\bar{\beta}_t$ finding

$$\bar{\beta}_t = \frac{\mathcal{M}^{\exp^{-\xi}}(\mathrm{p}_t, u(x_t))}{\mathrm{E}\ u(\mathrm{x}_t)} = \frac{\mathcal{M}^{\exp^{-\xi}}(\mathrm{p}_t, u(x_t))}{\bar{u}} < \frac{\mathcal{M}^{\exp^{-\xi}}(\mathrm{p}_{t-1}, u(x_{t-1}))}{\bar{u}} = \bar{\beta}_{t-1}$$

for all $t \in \{1, ..., T-1\}$. □

Proof of proposition 3: The first order welfare change gained by adding a marginal certain consumption vector ϵ_t in all risk states is

$$\mathcal{M}^{\exp^{-\xi}}(\mathrm{p}_t, u(x_t + \epsilon_t) \ = \mathrm{E}\mathrm{p}_t \frac{\exp(-\xi u(x_t))}{\mathrm{E}\mathrm{p}_t \exp(-\xi u)} u'(x_t)\epsilon_t \ \Rightarrow \ \tilde{\beta}_t(x_t) = \frac{\exp(-\xi u(x_t))}{\mathrm{E}\mathrm{p}_t \exp(-\xi u)} \ ,$$

where ϵ_t and $u'(x_t)$ are vectors forming an inner product and $\tilde{\beta}_t : X \to \mathbb{R}$ is the discount function defined in Proposition 3. I find

$$\tilde{\beta}_{t+1}(x) < \tilde{\beta}_t(x) \Leftrightarrow \frac{\exp(-\xi u(x))}{\mathrm{E}\mathrm{p}_{t+1} \exp(-\xi u)} < \frac{\exp(-\xi u(x_t))}{\mathrm{E}\mathrm{p}_t \exp(-\xi u)}$$

$$\Leftrightarrow \frac{\exp(-\xi u(x))}{\exp(-\xi \mathcal{M}^{\exp^{-\xi}}(\mathrm{p}_{t+1}, u)} < \frac{\exp(-\xi u(x_t))}{\exp(-\xi \mathcal{M}^{\exp^{-\xi}}(\mathrm{p}_t, u)}$$

$$\Leftrightarrow \exp\Big(-\xi\big[u(x) - \mathcal{M}^{\exp^{-\xi}}(\mathrm{p}_{t+1}, u)\big]\Big) < \exp\Big(-\xi\big[u(x) - \mathcal{M}^{\exp^{-\xi}}(\mathrm{p}_t, u)\big]\Big)$$

$$\Leftrightarrow \mathcal{M}^{\exp^{-\xi}}(\mathrm{p}_{t+1}, u) < \mathcal{M}^{\exp^{-\xi}}(\mathrm{p}_t, u) \ ,$$

which was shown to hold in the proof of Proposition 2. The second statement in the proposition follows immediately from the relation

$$\tilde{\beta}_t(\underline{x}) > \tilde{\beta}_t(\overline{x}) \ \Leftrightarrow \ \frac{\exp(-\xi u(\underline{x}))}{\mathrm{E}\mathrm{p}_t \exp(-\xi u)} > \frac{\exp(-\xi u(\overline{x}))}{\mathrm{E}\mathrm{p}_t \exp(-\xi u)} \ \Leftrightarrow \ u(\underline{x}) < u(\overline{x}) \ .$$

□