Intertemporal Risk Aversion, Stationarity and the Rate of Discount

Preliminary Version, January 2007

Christian P. TRAEGER
Department of Agricultural & Resource Economics, UC Berkeley
Department of Economics, UC Berkeley

Abstract: Two central concerns in the sustainability debate are the weight given to the needs of future generations and the application of the precautionary principle. In an economic model, these aspects of intertemporal evaluation can be mapped into time preference and intertemporal risk aversion. The current analysis points out a close link between these different aspects of intertemporal evaluation from the perspective of time consistent decision making under uncertainty. For this purpose, the paper analyzes the consequences of a stationary evaluation of the future for decision makers exhibiting a non-trivial intertemporal risk attitude. The modeling framework is a general recursive utility setting in a multi commodity world. I translate the standard stationarity assumption into a finite time horizon, in order to avoid implicit assumptions on the rate of pure time preference. To this end, the corresponding axiom is decomposed into two primitive statements. Stationary evaluation of risky outcomes is then combined with the assumption that the timing of uncertainty resolution only affects evaluation, when the derived information can be used to alter outcomes or outcome probabilities. I show that under these assumptions, a time consistent and intertemporal risk averse decision maker has to exhibit a pure rate of time preference of zero. Such a decision maker is no longer free to devalue the future for reasons of pure impatience, but gives reduced weight to welfare that is uncertain. When uncertainty increases over time, this fact resembles discounting. However, the more a decision maker can know about the future, the more weight it will carry.

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

Correspondence:
Department of Agricultural & Resource Economics
207 Giannini Hall #3310
University of California
Berkeley, CA 94720-3310
E-mail: traeger@are.berkeley.edu

1 Introduction

The concept of intertemporal risk aversion measures the higher willingness to undergo preventive action in order to avoid a threat of harm. Formally it is a measure for the difference between the decision maker's preference to smooth consumption over time and over risk. While the intertemporally additive expected utility model implicitly assumes intertemporal risk neutrality, general recursive models like the one developed by Kreps & Porteus (1978) or the generalized isoelastic model (Epstein & Zin 1989, Weil 1990) characterize as well decision makers with a nontrivial intertemporal risk attitude. In this paper I show that the assumption of risk stationarity has significantly stronger implications for intertemporally risk averse (or seeking) decision makers, then for maximizers of intertemporally additive expected utility. At the same time the paper elaborates axioms that simplify the general recursive model used to introduce the concept of intertemporal risk aversion in Traeger (2007*a*).

Kreps & Porteus (1978) extend von Neumann & Morgenstern's (1944) famous axiomatic approach to choice under uncertainty to a multiperiod framework. For this purpose, the authors develop the concept of temporal lotteries, where uncertainty is no longer expressed immediately over the set of consumption paths, but formulated recursively. As they elaborate, the framework allows the decision maker to exhibit a preference for the timing of uncertainty resolution, even if the information obtained by an early resolution of uncertainty cannot be used to alter future outcomes. Epstein & Zin (1989) show for a one commodity setting, that this generalized framework allows to disentangle intertemporal substitutability and risk aversion. Such a disentanglement is not possible in the intertemporally additive expected utility model, where the Arrow-Pratt measure of relative risk aversion is confined to the inverse of the elasticity of intertemporal substitution. In Traeger (2007*a*) I elaborate that a particular measure for the difference between risk aversion and intertemporal substitutability has itself an interpretation of risk aversion. I endow the concept with an axiomatic foundation, show that it naturally extends into the multi-commodity setup and give it the name intertemporal risk aversion.

In this paper I derive different simplifications of the general recursive representation by introducing stationarity into the evaluation of future consumption paths. In order to avoid implicit assumptions on the pure rate of time preference, the model features a finite time horizon. In consequence, the standard axiom characterizing stationarity has to be translated into a finite time horizon. To this purpose, I decompose the axiom into two parts. The well known part of the assumption says that the mere passage of time does not affect preferences. The other assumption, which is satisfied implicitly in the infinite horizon model with a strictly positive discount rate, assumes that the ranking of two consumption plans does not depend on a common outcome in the last period. While a certainty stationary evaluation mainly makes the model coincide with the standard discount utility model on certain consumption paths, the risk stationary evaluation moreover restricts the functional form of intertemporal risk aversion. Moreover I present an additional axiom relating the evaluation of risk in different periods. While it does not correspond to a stationarity assumption in the sense that the mere passage of time does not change preferences, it makes uncertainty evaluation constant over time, as for example assumed in the generalized isoelastic model.

Finally, I combine the stationarity axioms with the assumption that the decision maker does not exhibit an intrinsic preference for an early or late resolution of uncertainty. I show that, for a time consistent decision maker, a combination of the latter assumption with risk stationarity, forces the *pure rate of time preference* to zero. A result that so far, to the best of my knowledge, has only been argued for on the basis of moral considerations. Observe that the demands to place more weight on the needs of future generations, i.e. a low rate of pure time preference, and the application of the precautionary principle, i.e. an intertemporally risk averse evaluation, are two central concerns of a sustainable development. The analysis points out that these different aspects of evaluation are closely linked from the theoretical point of view of time consistent decision making under uncertainty.

In addition to Traeger (2007*a*), which introduces the concept of intertemporal risk aversion for the general non-stationary setting, there are two other accompanying papers. In Traeger (2007*d*) I set out the immediate correspondence between the concept of intertemporal risk aversion and the willingness to undergo preventive measures in order to avoid a threat of harm. I therefore argue that the concept formalizes an important aspect of the precautionary principle, a principle gaining increasing significance in international treaties and declarations. In Traeger (2007*c*) I elaborate the general relation between Kreps & Porteus' (1978) (intrinsic) preference for the timing of uncertainty resolution and the concept of intertemporal risk aversion. Moreover, I elaborate how, under the assumption of indifference to the timing of uncertainty resolution, it is possible to disentangle atemporal (or standard) risk aversion from intertemporal substitutability in a non-recursive framework.

The paper is structured as follows. First, section 2 reviews the general preference representation derived in Traeger (2007*a*). Section 2.1 introduces uncertainty aggregation rules and briefly reviews the von Neumann-Morgenstern axioms. Section 2.2 lays out the intertemporal structure of the model and discusses the related axioms. And section 2.3 states the general representation theorem. Then, section 3 introduces different stationarity assumptions. First, section 3.1 restricts choice over certain consumption paths to the standard discount utility model by requiring an axiom of certainty stationarity. Subsequently, section 3.2 introduces an axiom that makes uncertainty aggregation invariant over time. In particular, the resulting representation contains the model of generalized isoelastic preferences, usually employed to disentangle (atemporal) risk aversion from intertemporal substitutability. Third, section 3.3 works out an alternative stationarity assumption for the evaluation of uncertainty. Analogously to the axiom for certain consumption paths it builds on the assumption that the mere passage of time does not change preferences and that the ranking of two lotteries does not depend on a common certain outcome in the last period. Section 4 characterizes intertemporal risk aversion in the derived settings. First, section 4.1 works out a simplified axiomatic characterization for stationary preferences. Then, section 4.2 analyzes the measures of intertemporal risk aversion in the respective settings. Finally, section 5 adds the assumption of indifference with respect to the timing of uncertainty evaluation. Section 5.1 states the resulting non-recursive representation for certainty stationary decision making. Section 5.2 derives the implications of risk stationarity for the weight given to future welfare. Section 6 concludes. All proofs are found in the appendix. On page 38 the reader finds an overview over the notation employed in this paper.

2 Preliminaries

2.1 Uncertainty Aggregation Rules and the von Neumann-Morgenstern Axioms

Section 3 briefly reviews the representation and the concept of intertemporal risk aversion derived in Traeger (2007*a*), which is the basis for later analysis. This section starts out by introducing uncertainty aggregation rules and briefly reviewing the von Neumann-Morgenstern axioms. Let X be a connected compact metric space. The elements x of X are called consumption levels or, more general, outcomes. They may contain quantifications in terms of real numbers as well as more abstract characterizations, for example, of current climate or the state of an ecosystem before and after an invasive species has been introduced. The space of all continuous functions from outcomes into the reals is denoted by $\mathcal{C}^0(X)$. More generally, the space of all continuous functions from some metric space Y into the reals is denoted by $\mathcal{C}^0(Y)$. An element $u \in \mathcal{C}^0(X), u : X \to \mathbb{R}$, is called a Bernoulli utility function.[1] Define $\underline{U} = \min_{x\in X} u(x)$, $\overline{U} = \max_{x\in X} u(x)$ and $U = [\underline{U}, \overline{U}]$ so that the range of u is given by U.[2] The set of all Borel probability measures on X is denoted by $P = \Delta(X)$ and equipped with the Prohorov metric which gives rise to the topology of weak convergence. The elements $p \in P$ are called lotteries. I have in mind an epistemic foundation of probabilities as worked out for example in Cox (1946,1961) or Jaynes (2003).

[1]A refined definition of Bernoulli utility in relation to the representation of preference relations is given in the section (compare page 9).

[2]Note that compactness of X and continuity of u assure that the minimum and the maximum are attained.

Here probabilistic beliefs replace the notion of a binary logic. Then lotteries do not only describe draws from an urn, but correspond to general characterizations of uncertainty with respect to possible outcomes. For a closer discussion of this point see e.g. Traeger (2007d). The degenerate lotteries giving weight 1 to outcome x are denoted by $x \in P$. A lottery yielding outcome x with probability $p(x) = \lambda$ and outcome x' with probability $p(x') = 1 - \lambda$ is written as $\lambda x + (1-\lambda)x' \in P$. Note that the 'plus' sign between elements of X always characterizes a lottery.[3] Again more generally, the set of Borel probability measures on any compact metric space Y is denoted by $\Delta(Y)$. Finally, I denote with $\mathrm{I\!R}_+ = \{z \in \mathrm{I\!R} : z \geq 0\}$ and $\mathrm{I\!R}_{++} = \{z \in \mathrm{I\!R} : z > 0\}$ the set of all positive, respectively strictly positive, real numbers.

An *uncertainty aggregation rule* is defined as a functional $\mathcal{M} : P \times \mathcal{C}^0(X) \rightarrow \mathrm{I\!R}$. It takes as input the decision maker's perception of uncertainty, expressed by the probability measure p, and his evaluation of certain outcomes, expressed by his Bernoulli utility function u. For certain outcomes uncertainty aggregation rules are imposed to return the value of Bernoulli utility, i.e. $\mathcal{M}(x, u) = u(x)$. The uncertainty aggregation rule generated by the axiomatization in Traeger (2007a) is the following. For a strictly monotonic and continuous function $f : \mathrm{I\!R} \rightarrow \mathrm{I\!R}$ define $\mathcal{M}^f : P \times \mathcal{C}^0(X) \rightarrow \mathrm{I\!R}$ by

$$\mathcal{M}^f(p, u) = f^{-1}\left[\int_X f \circ u \, dp\right] , \tag{1}$$

where $f \circ u$ denotes the usual composition of two functions.[4] The composition sign will often be omitted. This shall not create confusion, as usual multiplication of two functions does not appear within this model. If the probability measure would be defined directly on the range of u, the expression in equation (1) would be known as the generalized or f-mean. It aggregates the utility values weighted by some function f and applies the inverse of f to normalize the resulting expression. The only difference between the mean and the uncertainty aggregation rule is that the latter takes the Bernoulli utility function as an explicit argument. If such a correspondence between a mean and an uncertainty aggregation rule holds, I say that the uncertainty aggregation rule (here $\mathcal{M}^f$) is induced by the mean (here generalized or f-mean).[5]

To illustrate the uncertainty aggregation rule $\mathcal{M}^f$ with some examples, let me consider

---

[3]As X is only assumed to be a compact metric space there is no immediate addition defined for its elements. In case it is additionally equipped with some vector space or field structure, the vector addition will not coincide with the "+" used here. The "+" sign used here alludes to the additivity of probabilities.

[4]Note that by continuity of $f \circ u$ and compactness of X Lesbeque's dominated convergence theorem (e.g. Billingsley 1995, 209) ensures integrability.

[5]Precisely this relation can be defined as follows. Let $p^u \in \Delta(U)$ denote the probability measure induced by p defined on X through the Bernoulli utility function $u \in \mathcal{C}^0(X)$ on its (compact) range U. Then an uncertainty aggregation rule $\mathcal{M}$ is said to be induced by a mean $\overline{\mathcal{M}} : \Delta(U) \rightarrow \mathrm{I\!R}$ whenever $\mathcal{M}(p, u) = \overline{\mathcal{M}}(p^u)\ \forall p \in P$. Mean inducedness implies that only the probability of x is used to weigh $u(x)$.

the subset of lotteries having finite support, i.e. the set of all simple probability measures $P^s \subset P$ on X. Then, equation (1) can be written as

$$\mathcal{M}^f(p,u) = f^{-1}\left[\sum_x p(x) f \circ u(x)\right].$$

The simplest uncertainty aggregation rule corresponds to the expected value operator, and is obtained for $f = \text{id}$:

$$\text{E}(p,u) \equiv \text{E}_p\, u(x) = \sum_x p(x)u(x)\,.$$

It is induced by the arithmetic mean. For Bernoulli utility functions with a range restricted to $U \subseteq \mathbb{R}_+$ another example of an uncertainty aggregation rule is induced by the geometric mean and corresponds to $f = \ln$:

$$\text{G}(p,u) = \prod_x u(x)^{p(x)}\,.$$

Both of the above uncertainty aggregation rules are, again assuming $U \subseteq \mathbb{R}_+$, contained as special cases in the following uncertainty aggregation rule achieved by $f(z) = z^\alpha$:

$$\mathcal{M}^\alpha(p,u) \equiv \mathcal{M}^{\text{id}^\alpha}(p,u) = \left[\sum_x p(x)u(x)^\alpha\right]^{\frac{1}{\alpha}}$$

defined for $\alpha \in \mathbb{R}$ with $\mathcal{M}^0(p,u) \equiv \lim_{\alpha\to 0} \mathcal{M}^\alpha(p,u) = \text{G}(p,u)$ and $\mathcal{M}^1(p,u) = \text{E}(p,u)$.[6] The corresponding mean is known as power mean. In the limit, where α goes to infinity respectively minus infinity, the uncertainty aggregation rule $\mathcal{M}^\alpha$ only considers the extreme outcomes (abandoning continuity in the probabilities): $\mathcal{M}^\infty(p,u) \equiv \lim_{\alpha\to\infty} \mathcal{M}^\alpha(p,u) = \max_x\; u(x)$ and $\mathcal{M}^{-\infty}(p,u) \equiv \lim_{\alpha\to-\infty} \mathcal{M}^\alpha(p,u) = \min_x\; u(x)$.

Moving the representation of preferences, the remainder of this section discusses the von Neumann-Morgenstern axioms. I represent preferences over lotteries in the usual way by a binary relation on P denoted $\succeq$. For two lotteries $p, p' \in P$ the interpretation of $p \succeq p'$ is that lottery p is weakly preferred with respect to lottery p'. The relation $\succeq$ $(\subset P \times P)$ is required to be reflexive.[7] The asymmetric part of the relation $\succeq$ is denoted by $\succ$ and interpreted as a strict preference. The symmetric part of the relation $\succeq$ is denoted by $\sim$ and interpreted as indifference. Nonindifference is denoted by $\not\sim$ and defined as $\not\sim \equiv P \times P \setminus \sim$. An uncertainty aggregation rule is said to represent the preference relation $\succeq$ over lotteries if

$$p \succeq p' \Leftrightarrow \mathcal{M}(p,u) \geq \mathcal{M}(p',u) \qquad \text{for all } p, p' \in P \tag{2}$$

and some $u \in \mathcal{C}^0(X)$. It is said to represent $\succeq$ for $u^* \in \mathcal{C}^0(X)$ if equation (2) holds with

[6] The easiest way to recognize the limit for $\alpha \to 0$ is to note that for any $\alpha > 0$ the function $f_\alpha(z) = \frac{z^\alpha - 1}{\alpha}$ is an affine transformation of $f(z) = z^\alpha$. However, affine transformations leave the uncertainty aggregation rule unchanged (see Traeger 2007*a*). Therefore the fact that $\lim_{\alpha\to 0} \frac{z^\alpha-1}{\alpha} = \ln(z)$ gives the result.

[7] Note that reflexivity is implied by completeness in axiom A1.

$u = u^*$.

The following axioms are close relatives to the ones suggested by von Neumann & Morgenstern (1944) for (atemporal) choice under uncertainty.

A1 (weak order) $\succeq$ is transitive and complete, i.e.:

- transitive: $\forall\, p, p', p'' \in P :\; p \succeq p'$ and $p' \succeq p'' \Rightarrow p \succeq p''$
- complete: $\forall\, p, p' \in P : p \succeq p'$ or $p' \succeq p$

Axiom A1 assumes that the decision maker can rank all lotteries (completeness). Moreover, if one is preferred to a second and the second is preferred to a third, then the first should also be preferred to the third (transitivity). Note that, within a normative context of deriving a principled approach to choice under uncertainty, A1 should be interpreted as "if a decision maker had the capacities to rank all possible outcomes, then his ranking should satisfy transitivity" rather than as an assumption that the decision maker has actually worked out a ranking of all possible outcomes.

A2 (independence) $\forall\, p, p', p'' \in P$:
$$p \sim p' \quad \Rightarrow \quad \lambda\, p + (1-\lambda)\, p'' \sim \lambda\, p' + (1-\lambda)\, p'' \quad \forall\, \lambda \in [0,1]$$

The independence axiom states the following. Let a decision maker be indifferent between a lottery p and another lottery p'. Now offer him two compound lotteries, which both start out with a coin toss. In both lotteries the decision maker enters the same third lottery p'' if head comes up. However, if tail comes up, the decision maker faces lottery p in the first compound lottery and the 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. More generally, the coin toss is replaced by an arbitrary binary lottery with outcome probabilities λ and $1-\lambda$, deciding which p-lottery in the compound lottery is to be 'played'.

A3 (continuity) $\forall p \in P : \{p' \in P : p' \succeq p\}$ and $\{p' \in P : p \succeq p'\}$ are closed in P

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).

2.2 Time and Temporal Lottery

Following Kreps & Porteus (1978) I extend von Neumann & Morgenstern's (1944) atemporal setup to multiple periods. Time is discrete with planning horizon $T \in \mathbb{N}$. Individual periods are usually denoted with time indices $t, \tau \in \{1, ..., T\}$. The set of all *certain* consumption paths from period t to period T is denoted by $\mathsf{X}^t = X^{T-t+1}$, where X^{T-t+1} denotes the T–t+1–fold Cartesian product of X with itself.[8] A consumption path is generally written with a calligraphic x and its period τ entry is denoted by x_τ. Whenever such a notation is unambiguous, I also label the entry x_t by x_t, yielding the notation $\mathsf{x} = (x_1, x_2, ..., x_T)$. Moreover, given a consumption paths $\mathsf{x} = (\mathsf{x}_t, \mathsf{x}_{t+1}, ..., \mathsf{x}_T) \in \mathsf{X}^t$ and an outcome $x \in X$, I define the reassembled consumption path $(\mathsf{x}_{-i}, x) = (\mathsf{x}_t, ..., \mathsf{x}_{i-1}, x, \mathsf{x}_{i+1}, ..., x_T) \in \mathsf{X}^t$, as the consumption path that coincides with x in all but the i^{th} period, when it renders outcome x.

Introducing *uncertainty* to the multiperiod setup, I employ Kreps & Porteus' (1978) framework of temporal lotteries. Instead of the more widespread framework of atemporal lotteries, corresponding to probability measures over consumption paths, this richer framework involves a recursive description of uncertainty. At the beginning of every period the decision maker faces uncertainty over the future, as well as over the outcome in the respective period. In the last period the decision maker has preferences over all lotteries on the space of outcomes $\tilde{X}_T \equiv X$, which are modeled as elements of $P_T \equiv \Delta(X)$. 'Half a time step' before, after the uncertainty in period $T-1$ has resolved, the decision maker faces pairs of certain outcomes from $x_T \in X$ in period $T-1$ and lotteries over the future P_T. The corresponding choice space is depicted by $\tilde{X}_{T-1} = X \times P_T$. In general however, before uncertainty in period $T-1$ is resolved, the decision makers choice will correspond to lotteries over the elements of $\tilde{X}_{T-1}$. These are modeled as elements of the set $P_{T-1} \equiv \Delta(\tilde{X}_{T-1}) = \Delta(X \times P_T)$. Period $T-1$ preferences on this space are represented by the binary relation $\succeq_{T-1}$ ($\subset P_{T-1} \times P_{T-1}$). Note the recursive structure of the definition. The uncertainty at the beginning of period $T-1$ is not modeled as a probability distribution over the Cartesian product of outcomes in $T-1$ and T. Rather, it is defined as uncertainty over the outcome in $T-1$ and the lottery faced in period T. In *general*, define $\tilde{X}_T = X$ and recursively $\tilde{X}_{t-1} = X \times \Delta(\tilde{X}_t)$ for all $t \in \{2, ..., T\}$. Equip the set of Borel probability measures on $\tilde{X}_t$, denoted by $P_t \equiv \Delta(\tilde{X}_t)$, with the Prohorov metric and the space $\tilde{X}_{t-1}$ with the product metric (making it compact). The elements p_t of P_t are called (period t) lotteries. Preferences in period t are defined on the set P_t and denoted by $\succeq_t$ ($\subset P_t \times P_t$). For a detailed introduction to recursive lotteries see also Kreps & Porteus (1978). For the setting of this paper, Traeger (2007*c*) elaborates the relation between these recursive lotteries and the special case where probability measures are defined directly on consumption paths.

An uncertainty aggregation rule in period t is defined as a functional $\mathcal{M}^{f_t} : P_t \times \mathcal{C}^0(\tilde{X}_t) \rightarrow$ $\mathbb{R}$ with $\mathcal{M}^{f_t}(p_t, \tilde{u}_t) = f_t^{-1} \int_{\tilde{X}_t} dp_t \, f_t \circ \tilde{u}_t(\tilde{x}_t)$. In order to match the widespread model of additively separable utility over time on certain consumption paths, I introduce the following assumption.

[8] There are $T-t+1$ periods from t to T for which consumption has to be specified. I do not distinguish different sets of outcomes for different periods. X can be thought of as the union of all possible outcomes perceivable in any period.

A4 (certainty additivity/coordinate independence)
For all $\mathsf{x}, \mathsf{x}' \in \mathsf{X}^1$, $x, x' \in X$ and $i \in \{1, ..., T\}$ it holds that

$$(\mathsf{x}_{-i}, x) \succeq_1 (\mathsf{x}'_{-i}, x) \Leftrightarrow (\mathsf{x}_{-i}, x') \succeq_1 (\mathsf{x}'_{-i}, x')$$

Axiom A4 requires that, whenever the outcome in some period i coincides for two consumption paths, then the preference over the two paths does not depend on period i. Axiom 4 is taken from Wakker (1988) and represents the main ingredient to allow for a certainty additive representation of the form $\sum_{t=1}^{T} u_t^{ca}(\mathsf{x}_t)$ for certain consumption paths. Note that so far, the resulting functions u_t^{ca} would be allowed to vary arbitrarily over time. In particular, tastes are even allowed to reverse between two periods.

Preferences in different periods are connected by the following consistency axiom.

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

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

It is a requirement for choosing between two consumption plans in period t that yield a degenerate lottery with a coinciding entry in the respective period. For these choice situations, axiom 5 demands that in period t the decision maker shall prefer the plan that gives rise to the lottery that is preferred in period $t+1$. The axiom stems from Kreps & Porteus (1978).

Finally, I assume that the decision maker faces at least three periods ($T \geq 3$) and that in every period there are at least two outcomes such that the decision maker is nonindifferent between the two. Precisely, I assume

A0 (nonindifference) For all $t \in \{1, ..., T \geq 3\}$ there exist $\mathsf{x} \in \mathsf{X}^1$ and $x \in X$ such that

$$(\mathsf{x}_{-t}, x) \not\sim_1 \mathsf{x}\,.$$

The assumption $T \geq 3$ is convenient to simplify the axiomatization of additive seperability over certain consumption paths. For the case $T = 2$, an additional assumption known as the Thomson condition is required (see Wakker 1988, definition 4.2 & theorem 4.3).[9]

Finally, I introduce the following notation regarding the codomains of the functions u and g in the multiperiod setting. Define $\underline{U}_t = \min_{x\in X} u_t(x)$, $\overline{U}_t = \max_{x\in X} u_t(x)$ and $U_t = [\underline{U}_t, \overline{U}_t]$, as well as $\underline{G}_t = g_t(\underline{U}_t)$, $\overline{G}_t = g_t(\overline{U}_t)$, $G_t = [\underline{G}_t, \overline{G}_t]$ and $\Delta G_t = \overline{G}_t - \underline{G}_t$ for all $t \in \{1, ..., T\}$. Moreover let $\Gamma_t = (\underline{G}_t, \overline{G}_t)$.

[9] Also, it would be sufficient to require condition A0 only for three different periods, instead of requiring nonindifference for all $t \in \{1, ..., T\}$.

2.3 The Representation

This section states the general representation theorem for preferences satisfying the axioms of the preceding sections. Special attention is payed to different possibilities of fixing (gauging) the evaluation function over the certain outcomes within a period. To characterize these functions precisely, I define an induced preference relation on these certain one period outcomes. For this purpose pick an arbitrary element $x^0 \in X$ and define for every $t \in \{1, ..., T\}$ and given preference relation $\succeq_t$ the binary relations $\succeq_t^*$ on X by:

$$x \succeq_t^* x' \quad \Leftrightarrow \quad (x, x^0, ..., x^0) \succeq_t (x', x^0, ..., x^0) \qquad \forall\, x, x' \in X$$

Axiom A4 makes this expression of preference independent of the particular choice for x^0. Then, I define the set of Bernoulli utility functions corresponding to the preference relation $\succeq_t$ by

$$B_{\succeq_t} = \{u_t \in \mathcal{C}^0(X) : x \succeq_t^* x' \Leftrightarrow u_t(x) \geq u_t(x')\, \forall\, x, x' \in X\}$$

For a given preference relation $\succeq_t$ every Bernoulli utility function will express the decision makers preference over period t outcomes in the sense that a higher value of u_t corresponds to a preferred choice. Obviously, with any $u_t \in B_{\succeq_t}$, also any strictly increasing continuous transformation is in $B_{\succeq_t}$ (and vice versa). The reason for the special attention paid to the choice of the Bernoulli utility function lies in the fact that different choices will give rise to different forms of the representation theorem. In particular, a convenient choice of Bernoulli utility can either render intertemporal aggregation *or* or uncertainty evaluation additive. However, rendering both aggregations, over time and over uncertainty, linear is generally not feasible. The following representation theorem holds.

Theorem 1: Let there be given a sequence of preference relations $(\succeq_t)_{t\in\{1,...,T\}}$ on $(P_t)_{t\in\{1,...,T\}}$ satisfying A0, and a sequence of Bernoulli utility functions $(u_t)_{t\in\{1,...,T\}}$ with $u_t \in B_{\succeq_t}$. The sequence of preference relations $(\succeq_t)_{t\in\{1,...,T\}}$ satisfies

i) A1-A3 for all $\succeq_t, t \in \{1, ..., T\}$ (vNM setting)

ii) A4 for $\succeq_1$ (certainty additivity)

iii) A5 (time consistency)

if and only if, for all $t \in \{1, ..., T\}$ there exist strictly increasing[10] and continuous functions $f_t : U_t \to \mathbb{R}$ and $g_t : U_t \to \mathbb{R}$ such that with defining

v) the normalization constants $\theta_T = 1$, $\vartheta_T = 0$ and for $t < T$

$$\theta_t = \frac{\Delta G_t}{\sum_{\tau=t}^T \Delta G_\tau} \quad \text{and} \quad \vartheta_t = \frac{\overline{G}_{t+1} \underline{G}_t - \underline{G}_{t+1} \overline{G}_t}{\Delta G_t} \quad \text{and}$$

[10]Alternatively the theorem can be stated replacing increasing by monotonic for $(f_t)_{t\in\{1,...,T\}}$ and demanding that either all $(g_t)_{t\in\{1,...,T\}}$ are strictly increasing or that all are strictly decreasing.

vi) recursively the functions $\tilde{u}_t : \tilde{X}_t \to \mathbb{R}$ by $\tilde{u}_T(x_T) = u(x_T)$ and

$$\tilde{u}_{t-1}(x_{t-1}, p_t) = g_{t-1}^{-1} \left[\theta_{t-1}\, g_{t-1} \circ u_{t-1}(x_{t-1}) + \frac{\theta_{t-1}}{\theta_t}\, g_t \circ \mathcal{M}^{f_t}(p_t, \tilde{u}_t) + \frac{\theta_{t-1}}{\theta_t} \vartheta_{t-1}\right]$$

it holds for all $t \in \{1, ..., T\}$ that

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

Moreover, $(u_t, f_t, g_t)_{t\in\{1,...,T\}}$ and $(u_t, f'_t, g'_t)_{t\in\{1,...,T\}}$ both represent $(\succeq_t)_{t\in\{1,...,T\}}$ in the above sense, if and only if, for all $t \in \{1, ..., T\}$ there exist constants $a_t^f \in \mathbb{R}_{++}$ and $b_t^f \in \mathbb{R}$ such that $f_t = a_t^f f'_t + b_t^f$, as well as constants $a^g \in \mathbb{R}_{++}$ and $b_t^g \in \mathbb{R}$ such that $g_t = a^g g'_t + b_t^g$.

A sequence of triples $(u_t, f_t, g_t)_{t\in\{1,...,T\}}$ as above is called a representation in the sense of theorem 1 for the set of preference relations $\succeq = (\succeq_t)_{t\in\{1,...,T\}}$. The representation theorem recursively constructs an aggregate utility $\tilde{u}_t$ that depends on the utility in the respective period $u_t(x_t)$, as well as the aggregate utility derived from a particular lottery p_{t+1} over the future. The representation is discussed in detail in Traeger (2007*a*). Note that at some cost in the freedom to normalize the function $(g_t)_{t\in\{1,...,T\}}$ one can eliminate the normalization constants ϑ_t (see Traeger 2007*d*). However in the present paper they will vanish 'for free' thanks to the stationarity assumption. For the purpose of this paper, just note that the functions g_t characterize intertemporal aggregation, while the functions f_t characterize uncertainty aggregation. Both are allowed to vary arbitrarily between different periods. As shown in Traeger (2007*a*), a handy choice of the Bernoulli utility functions $(u_t)_{t\in\{1,...,T\}}$ can render intertemporal aggregation linear, corresponding to $g_t = \text{id} \forall t \in \{1, ..., T\}$, and yield the representation characterized by the recursion relation

$$vi') \quad \tilde{u}_{t-1}(x_{t-1}, p_t) = \theta_{t-1} u_{t-1}(x_{t-1}) + \frac{\theta_{t-1}}{\theta_t} \mathcal{M}^{f_t}(p_t, \tilde{u}_t) + \frac{\theta_{t-1}}{\theta_t} \vartheta_{t-1}$$

Alternatively, one can choose Bernoulli utility in a way to render uncertainty aggregation linear, corresponding to $f_t = \text{id} \forall t \in \{1, ..., T\}$, and yield the representation characterized by the recursion relation

$$vi'') \quad \tilde{u}_{t-1}(x_{t-1}, p_t) = g_{t-1}^{-1} \left[\theta_{t-1}\, g_{t-1} \circ u_{t-1}(x_{t-1}) + \frac{\theta_{t-1}}{\theta_t}\, g_t \circ \mathrm{E}_{p_t} \tilde{u}_t + \frac{\theta_{t-1}}{\theta_t} \vartheta_{t-1}\right]$$

However, the linearization of uncertainty aggregation comes at the cost of making intertemporal aggregation nonlinear. Implicitly such a representational form is chosen in the related representation of Kreps & Porteus (1978). However, as pointed out in Traeger (2007*d*), a certainty additive representation is convenient for the economic interpretation. In particular, only with a normalization as in *vi'*) Bernoulli utility can be interpreted as welfare in the following sense. One unit of welfare more in one period and another unit of welfare less in another period leaves aggregate welfare unchanged. This interpretation in terms of welfare will prove particularly convenient when the concept of intertemporal risk aversion will be introduce in section 4. Note finally that a simultaneous linearization of uncertainty and intertemporal aggregation corresponds to the intertemporally additive expected utility model. However, the latter only accounts for the special case of the above representation corresponding to an intertemporally risk averse decision maker, a restriction that is by no means implied by the above axioms.[11]

3 Stationarity

3.1 Certainty Stationarity

The representation reviewed in the preceding section allows time and uncertainty aggregation to vary arbitrarily from period to period. In the following, I introduce different stationarity assumptions and elaborates the corresponding implications for the representation. First, the current section develops an axiom restricting choice under certainty that renders intertemporal aggregation and Bernoulli utility stationary. On certain consumption paths, it gives rise to the common discount utility representation. For reasons set out below I stick to a finite planning horizon. Therefore, the corresponding stationarity axiom have to be adapted from the forms usually put forth in the literature. Precisely, I split the axiom into two. The first part can be interpreted as the assumption that the mere passage of time does not change preference. The second part is the assumption that the ranking of two lotteries does not depend on a common certain outcome in the last period. The latter assumption comes for free in an infinite time horizon with a strictly positive discount rate.

Subsequently, in section 3.2, I introduce an axiom that renders uncertainty aggregation invariant over time. Strictly spoken, it is not a stationarity condition, but a condition that characterizes indifference with respect to the length of risk taking. The resulting representation contains the model of generalized isoelastic preferences, usually employed to disentangle (atemporal) risk aversion from intertemporal substitutability. Section 3.3 works out an alternative stationarity assumption for the evaluation of uncertainty. Analogously to the axiom for certain consumption paths derived in the current section, it builds on the assumptions that the ranking of two lotteries does not depend on a common certain outcome in the last period and that the mere passage of time does not change preference (this time over risky outcomes).

Stationarity, in the sense of the standard discount utility model is a ubiquitous assumption in economic modeling, and in particular in environmental economics. However, to the best of my knowledge, the assumption is expressed in terms of the underlying preference relations only for models featuring an infinite time horizon. In these models, the axiomatic

[11] In particular note that it is therefore wrong to call the intertemporally additive expected utility model, i. e. an evaluation of the form $\mathrm{E}_p \sum_{t=1}^{T} u_t$, the 'additively (time) separable expected utility model'.

characterization of stationarity requires that a decision maker prefers 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).[12] Such an axiomatization makes use of the fact that for an infinite time horizon, adding an additional outcome does not change the length of a consumption path. Precisely, both paths x and (x^0, x), are elements of X^∞ and can be compared by the same preference relation. On the contrary, for a finite time 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 reason for keeping the model in the finite time horizon is threefold. First, I want to avoid the assumption of a strictly positive rate of pure time preference at the outset of the model. However, the latter is required to apply the common techniques for analyzing infinite time horizon settings (contraction and fix point theorems). From a normative point of view, such a positive discount rate is not without controversy and several famous economists argued against such a rate. For example Ramsey states that such a positive rate of pure time preference is "ethically indefensible" Ramsey (1928, 543). Second, the reasoning on stationarity carried out in this chapter together with the reasoning on attitude with respect to the timing of uncertainty resolution carried out in section 5, make a strong point for choosing a zero rate of time preference for a time consistent approach to choice under uncertainty. Third, for most planning processes and scenario evaluations there exist reasonable upper bounds for the planning horizon.[13] Finally, from a descriptive perspective, finiteness of the planning horizon is more than likely to be satisfied. I will provide several comments with respect to the limit of an infinite time horizon. The following axiom is applicable in a finite time horizon setting and, there, yields the standard discount utility model for the evaluation of certain consumption paths.

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

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

On the right hand side of the equivalence, the decision maker faces a comparison between x and x' as consumption paths starting in period 2. On the left hand side of the equivalence, the decision maker faces a comparison between x and x' as consumption paths starting in period 1. The additional outcome x, which is commonly added to the paths x and x', makes (x, x) and (x', x) choice objects of the appropriate length, so that they can be compared in period 1 by the preference relation $\succeq_1$. The important property of the axiom is that the decision maker's preference over the (certain) consumption paths is independent of their

[12]See page 13 for details.

[13]While such an upper bound can be in the magnitude of several decades, note that taking as upper bound a point of time by which our sun has burned out or turned into a red giant still provides a finite upper bound (Sackmann, Boothroyd & Kraemer 1993).

starting point.[14]

I give an interpretation of axiom 6 by separating the underlying idea into two steps. 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) (*not* with their continuation). Furthermore, let him plan ahead the same amount of periods in period 2 as he does in period one, implying a time horizon $T+1$. Formally, I denote these preferences of the decision maker in period 2 with time horizon $T+1$ by $\succeq_{2|T+1}$. Then, given that nothing else changes between period 1 and period 2, I *assume* that the decision maker ranks (or plans to rank)[15] the consumption paths in both choice situations the same way. Requiring the latter for all consumption paths yields the condition

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

for all $\mathsf{x}, \mathsf{x}' \in \mathsf{X}^2$ and $x \in X$. Condition (3) most clearly captures the intuition of stationarity, in the sense that the mere passage of time should not change the evaluation. However, up to now the preference relations $\succeq_{\cdot|T}$ and $\succeq_{\cdot|T+1}$ are unrelated. In consequence, equation (3) on its own does not restrict the decision maker's preference relations $(\succeq_t)_{t\in\{1,\ldots,T\}} = (\succeq_{t|T})_{t\in\{1,\ldots,T\}}$ in any way. Thus, the second step in the reasoning has to relate the preference relation $\succeq_{2|T+1}$ to the relation $\succeq_2 = \succeq_{2|T}$. Both preference relations specify how the decision maker anticipates to evaluate choice objects from period 2 into the future. The relation $\succeq_{2|T}$ specifies his ranking when planning $T-2$ periods ahead (until period T), and the relation $\succeq_{2|T+1}$ states his ranking when he plans $T-1$ periods ahead (until period $T+1$). Accepting stationarity in the sense of equation (3), axiom A6 requires the following relation between $\succeq_{\cdot|T}$ and $\succeq_{\cdot|T+1}$:

$$\mathsf{x} \succeq_{2|T} \mathsf{x}' \quad \Leftrightarrow \quad (\mathsf{x}, x) \succeq_{2|T+1} (\mathsf{x}', x) \tag{4}$$

for all $\mathsf{x}, \mathsf{x}' \in \mathsf{X}^2$ and $x \in X$. In words, if two scenarios or consumption paths 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.

Let me point out the analogous reasoning to yield stationarity from the assumption expressed in equation (3) for the case of an infinite planning horizon. Denote the consumption paths corresponding to (x, x) and (x', x) simply by $\mathsf{x}^\infty, \mathsf{x}'^\infty \in X^\infty$. Then, by time consistency the right hand side of equation (3) is equivalent to $(x, \mathsf{x}^\infty) \succeq_{1|T+1} (x, \mathsf{x}'^\infty)$ for all $\mathsf{x}^\infty, \mathsf{x}'^\infty \in X^\infty$ and $x \in X$. Moreover, in the infinite horizon setting, it holds $\succeq_{1|T+1} = \succeq_{1|\infty} = \succeq_{1|T}$, a relation which makes equation (4) dispensable. That way, I arrive at

[14] Note the difference to time consistency. The latter is a condition on consumption paths starting in the same period that 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.

[15] Plans to rank refers to a normative interpretation of the axioms.

the standard axiom of stationarity for the infinite planning horizon: $\mathrm{x}^\infty \succeq_{1|\infty} \mathrm{x}'^\infty \Leftrightarrow (x, \mathrm{x}^\infty) \succeq_{1|\infty} (x, \mathrm{x}'^\infty)$ for all $x \in X$ and all $\mathrm{x}^\infty, \mathrm{x}'^\infty \in X^\infty$, dating back to Koopmans (1960, 294)[16]. Hence, at first sight, the second assumption, corresponding to equation (3), seems to come for free with an infinite time horizon. However, this is not the case. It is a necessary assumption in the standard framework with an infinite planning horizon that the decision maker applies a positive rate of pure time preference. Therefore, the weight given to future consumption converges to zero. Thus, the assumption that coinciding outcomes in the 'last' period of the planning horizon do not matter for the ranking of consumption paths is implicit in the infinite horizon setting. It is the combined result of the decision maker's intrinsic devaluation of the future and his infinite planning horizon.

Stationarity implies that the sets of Bernoulli utility functions coincide for different periods. Therefore, define $u \in B_\succeq \equiv B_{\succeq_1}$. Preference stationarity on certain consumption paths as formulated in axiom A6, together with the assumptions reviewed in section 2, yields the following representation.

Theorem 2: Let there be given a sequence of binary relations $\succeq = (\succeq_t)_{t\in\{1,...,T\}}$ on $(P_t)_{t\in\{1,...,T\}}$ satisfying A0, and a Bernoulli utility function $u \in B_\succeq$ with range U. The sequence $\succeq = (\succeq_t)_{t\in\{1,...,T\}}$ satisfies

i) A1-A3 for all $\succeq_t, t \in \{1, ..., T\}$ (vNM setting)

ii) A4 for $\succeq_1$ (certainty additivity)

iii) A5 (time consistency)

iv) A6 (certainty stationarity)

if and only if, there exist strictly increasing and continuous functions $f_t : U \to \mathbb{R}$ for all $t \in \{1, ..., T\}$ and $g : U \to \mathbb{R}$ as well as a discount factor $\beta \in \mathbb{R}_{++}$, such that with defining

v) the normalized discount weights

$\beta_t = \beta \, \frac{1-\beta^{T-t}}{1-\beta^{T-t+1}}$ for $\beta \neq 1$ and

$\beta_t = \frac{T-t}{T-t+1}$ for $\beta = 1$ and

vi) 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

$$\tilde{u}_{t-1} = g^{-1}\left[(1-\beta_{t-1})\, g \circ u(x_{t-1}) + \beta_{t-1}\, g \circ \mathcal{M}^{f_t}(p_t, \tilde{u}_t)\right] \tag{5}$$

[16]Koopmans (1960) actually formulates his postulates in terms of utility functionals. However the translation of his postulate 4 into the preference setup is immediate. His general axiomatic setting is translated into preferences in Koopmans (1972), again with stationarity corresponding to postulate 4.

it holds for all $t \in \{1, ..., T\}$ that

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

Moreover the functions g and f_t are unique up to nondegenerate positive affine transformations.

Certainty stationarity implies that the same Bernoulli utility function $u \in B_{\succeq}$ can be used in the representation for all periods. Moreover, it makes the functions g_t, characterizing intertemporal aggregation, coincide for adjacent periods up to a (common) multiplicative constant. This constant corresponds to the discount factor β. As shown in the proof, a representation $(u, f_t, g)_{t \in \{1,...,T\}}$ for $(\succeq_t)_{t \in \{1,...,T\}}$ in the sense of theorem 2 corresponds to a representation $(u_t, f_t, g_t)_{t \in \{1,...,T\}} = (u, f_t, \beta^t g)_{t \in \{1,...,T\}}$ for $(\succeq_t)_{t \in \{1,...,T\}}$ in the sense of the general non-stationary representation of theorem 1. Expressing this relation in words, the information characterizing intertemporal aggregation, which in the general setting is contained in the functions g_t for $t \in \{1, ..., T\}$, can be captured in the stationary setting by two quantities. The first piece of information is taken up by the function g, which now is common to all periods and describes the nonlinearity involved in intertemporal aggregation. The second piece of information characterizes the change of the functions g_t between different periods. This change is captured in a single parameter, the discount factor β, which describes the reduction in weight given to future outcomes.

For the limit of an infinite time horizon under the assumption $\beta < 1$, the normalized discount weights β_t used in the representation converge to the discount factor itself: $\lim_{T \to \infty} \beta_t = \beta$ for all t. Then, the weight given to the present as opposed to the future is constant. However, for a decision maker who plans with a finite time horizon, the weights β_t have to accommodate not only discounting, but also the weight that an individual period receives as opposed to the remaining future. The shorter the time horizon, or the closer the end of the time horizon, the higher must be the weight that the present period obtains as opposed to the remaining future.[17] Moreover, there exists a particular choice of Bernoulli utility such that the function $g = \text{id}$ and the recursive construction of aggregate utility in equation (5) becomes additive

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

For certain consumption paths this representation is ordinally equivalent to the widely applied evaluation

$$\mathrm{x}^t \succeq_t \mathrm{x}'^t \quad \Leftrightarrow \quad \sum_{\tau=t}^{T} \beta^t u(\mathrm{x}^t_\tau) \geq \sum_{\tau=t}^{T} \beta^t u(\mathrm{x}'^t_\tau). \tag{6}$$

[17] In particular, at the end of the time horizon, the weight given to the future has to be zero. Note, that this reasoning is necessary because the weights given to the present and to the future have to add up to one in the time aggregator $g^{-1}\left[(1 - \beta_{t-1})\, g(\cdot) + \beta_{t-1}(\cdot)\right]$. Otherwise the symmetric characterization of intertemporal aggregation and uncertainty aggregation by functions g_t and f_t would fail.

In difference to the intertemporal aggregation rules, the uncertainty aggregation rules are allowed to vary arbitrarily over time. The next two sections elaborate two different assumptions rendering the uncertainty aggregation rules stationary as well.

3.2 Constant Uncertainty Evaluation

In the preference framework of the preceding section, I assume stationarity in the evaluation of certain consumption paths. The assumption implies a close relation between intertemporal aggregation rules in different periods. In contrast, in the representation of theorem 2 uncertainty evaluation is allowed to vary arbitrarily over time.[18] It stands to reason that a decision maker who relates his evaluation of certain consumption paths between different periods, is also willing to relate his evaluation of uncertain consumption plans for different periods. An example of a preference representation which relates uncertainty evaluation between different periods is the generalized isoelastic model. It was developed independently by Epstein & Zin (1989) and Weil (1990) to disentangle (standard) risk aversion from intertemporal substitutability. The model has been used in many applications ranging from asset pricing (Attanasio & Weber 1989, Svensson 1989, Epstein & Zin 1991, Normandin & St-Amour 1998, Epaulard & Pommeret 2001) over measuring the welfare cost of volatility (Obstfeld 1994, Epaulard & Pommeret 2003*b*) to resource management[19] (Knapp & Olson 1996, Epaulard & Pommeret 2003*a*, Howitt et al. 2005) and evaluation of global warming scenarios (Ha-Duong & Treich 2004). Here, uncertainty aggragation is commonly characterized for all peirods by the function $f_t = z^\alpha \,, \forall\, t \in \{1, ..., T\}$. I will discuss the model further below.

In order to state an axiom that implies time constant uncertainty aggregation rules, it proves useful to introduce a special notation for constant consumption paths. Let $\bar{\mathrm{x}}^t = (\bar{x}, \bar{x}, ..., \bar{x})$ denote the certain constant consumption path that gives consumption $\bar{x}$ from t until T. Then $\frac{1}{2}\bar{\mathrm{x}}^t + \frac{1}{2}\bar{\mathrm{x}}'^t \in P_t$ is the lottery in period t that randomizes with probability $\frac{1}{2}$ between the constant consumption streams giving $\bar{x}$ and $\bar{x}'$. The following axiom demands that these randomized consumption streams relate to certain consumption streams the same way in different periods.

[18] Precisely, uncertainty evaluation is allowed to vary arbitrarily between different periods t and t'. By the requirement of time consistency, uncertainty aggregation has to be fixed for a given period t and, thus, independent of whether period t is τ or τ' periods into the future.

[19] While Knapp & Olson (1996) and Epaulard & Pommeret (2003*a*) solve theoretical models in order to obtain optimal rules for resource use, Howitt, Msangi, Reynaud & Knapp (2005) try to rationalize observed reservoir management in California, which cannot be explained by means of intertemporally additive expected utility.

A7 (constant risk evaluation) For all $t \in \{1, ..., T-1\}$ holds

$$\tfrac{1}{2}\bar{\mathrm{x}}^{\,t} + \tfrac{1}{2}\bar{\mathrm{x}}'^{\,t} \succeq_t \bar{\mathrm{x}}''^{\,t} \;\;\Leftrightarrow\;\; \tfrac{1}{2}\bar{\mathrm{x}}^{\,t+1} + \tfrac{1}{2}\bar{\mathrm{x}}'^{\,t+1} \succeq_{t+1} \bar{\mathrm{x}}''^{\,t+1} \quad \forall\, \bar{x}, \bar{x}', \bar{x}'' \in X\,.$$

The axiom can be conceived as an indifference requirement to the start and, thus, the duration of a taken risk. In particular, for a decision maker who is indifferent between the lottery $\frac{1}{2}\bar{x} + \frac{1}{2}\bar{x}'$ and a certain outcome $\bar{x}''$ in period T, axiom A7 requires that he is indifferent between the lottery $\frac{1}{2}(\bar{x}, \bar{x}) + \frac{1}{2}(\bar{x}', \bar{x}')$ and the certain consumption path $(\bar{x}'', \bar{x}'')$ in period $T-1$ as well. Be aware that in the lotteries of axiom A7 the outcomes in the different periods are perfectly correlated. In particular, in the above example, lottery $\frac{1}{2}(\bar{x}, \bar{x}) + \frac{1}{2}(\bar{x}', \bar{x}')$ is not the same as the lottery $\frac{1}{2}(\bar{x}, \frac{1}{2}\bar{x} + \frac{1}{2}\bar{x}') + \frac{1}{2}(\bar{x}', \frac{1}{2}\bar{x} + \frac{1}{2}\bar{x}')$, which would correspond to independent coin tosses in both periods. Adding axiom A7 to the assumptions of theorem 2 yields the following representation.

Theorem 3: Let there be given a sequence of binary relations $\succeq = (\succeq_t)_{t\in\{1,...,T\}}$ on $(P_t)_{t\in\{1,...,T\}}$ satisfying A0, and a Bernoulli utility function $u \in B_{\succeq}$ with range U. The sequence $\succeq = (\succeq_t)_{t\in\{1,...,T\}}$ satisfies

i) A1-A3 for all $\succeq_t, t \in \{1, ..., T\}$ (vNM setting)

ii) A4 for $\succeq_1$ (certainty additivity)

iii) A5 (time consistency)

iv) A6 & A7 (certainty stationarity & constancy of risk evaluation)

if and only if, there exist strictly increasing and continuous functions $f : U \to \mathbb{R}$ and $g : U \to \mathbb{R}$ as well as a discount factor $\beta \in \mathbb{R}_{++}$, such that with defining

v) the normalized discount weights

$$\beta_t = \beta\, \frac{1-\beta^{T-t}}{1-\beta^{T-t+1}} \quad \text{for } \beta \neq 1 \text{ and}$$

$$\beta_t = \frac{T-t}{T-t+1} \qquad \text{for } \beta = 1 \text{ and}$$

vi) 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

$$\tilde{u}_{t-1}(x_{t-1}, p_t) = g^{-1}\left[(1-\beta_{t-1})\, g \circ u(x_{t-1}) + \beta_{t-1}\, g \circ \mathcal{M}^f(p_t, \tilde{u}_t)\right]$$

it holds for all $t \in \{1, ..., T\}$ that

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

Moreover g and f are unique up to nondegenerate positive affine transformations.

In this representation a common function f characterizes uncertainty aggregation in all periods. Relating the representation to the general non-stationary setting, a representation

(u, f, g) in the sense of theorem 3 corresponds to the representation $(u, f, \beta^t g)$ in the sense of theorem 1. The following lemma is an immediate consequence of Traeger (2007*a*, lemma 4).

Lemma 1: Choose any strictly increasing and continuous functions $f^* : \mathbb{R} \to \mathbb{R}$ and $g^* : \mathbb{R} \to \mathbb{R}$. Let the set of preference relations $\succeq$ satisfy the assumptions of theorem 3. Then there exists a Bernoulli utility function $u \in B_\succeq$, such that uncertainty aggregation in the representation of theorem 3 is characterized by $f = f^*|_U$. Moreover, it exists a Bernoulli utility functions $u' \in B_\succeq$, such that intertemporal aggregation in the representation of theorem 3 is characterized by the function $g = g^*|_U$.

In particular, either f^* or g^* can be chosen as the identity. In general, however, the Bernoulli utility function that makes intertemporal aggregation linear ($g = \text{id}$) and the one that renders expected value ($f = \text{id}$) will not coincide. In the certainty additive representation the recursive construction of the representation in theorem 3 becomes

$$\tilde{u}_{t-1}(x_{t-1}, p_t) = (1 - \beta_{t-1})\, u(x_{t-1}) + \beta_{t-1}\, \mathcal{M}^f(p_t, \tilde{u}_t)\,.$$

In the Kreps Porteus form which is linear in uncertainty aggregation one obtains

$$\tilde{u}_{t-1}(x_{t-1}, p_t) = g^{-1}\left[(1 - \beta_{t-1})\, g \circ u(x_{t-1}) + \beta_{t-1}\, g \circ E_{p_t} \tilde{u}_t)\right]\,.$$

In the one commodity setting, there is another interesting choice for Bernoulli utility. Assuming $X \subset \mathbb{R}_+$ and nonsatiation in the interior, one can always pick u as the identity. Then, one obtains the framework that has been developed by Epstein & Zin (1989) and Weil (1990) in order to disentangle risk aversion from intertemporal substitutability:

$$\tilde{u}_{t-1}(x_{t-1}, p_t) = g^{-1}\left[(1 - \beta_{t-1})\, g(x_{t-1}) + \beta_{t-1}\, g \circ \mathcal{M}^f(p_t, \tilde{u}_t)\right]\,,$$

where $\tilde{u}_T = u = \text{id}$. In such a setting f can be interpreted as the characterization of standard or atemporal[20] risk attitude (see Epstein & Zin 1989, Traeger 2007*a*). In general the concavity of f characterizes risk aversion. For a twice differentiable function f, the Arrow-Pratt-measure of relative risk aversion correspond to $\text{RRA}(x) = -\frac{f''(x)}{f'(x)}\, x$. The advantage of the Arrow-Pratt-measure of relative risk aversion as opposed to f itself, is that it eliminates the affine indeterminacy of f that prevails by the moreover part of theorem 3. On the other hand, the function g can be interpreted as a parametrization of intertemporal substitutability. Both, Epstein & Zin (1989) and Weil (1990) assume that g is of the form $g(z) = z^\rho$, rendering a constant elasticity of intertemporal substitution $\sigma = \frac{1}{1-\rho}$. Weil (1990) moreover assumes that also risk aversion is characterized through a power function assuming $f(z) = z^\alpha$, rendering constant relative risk aversion $\text{RRA} = -\frac{f''(x)}{f'(x)} x = 1 - \alpha$.

[20] The wording 'atemporal' stems from Normandin & St-Amour (1998, 268), who point out the difference between the 'intertemporal' information contained in the parametrization of intertemporal substitutability, and the 'atemporal' nature of the risk attitude captured by f. In contrast, the concept of intertemporal risk aversion that will be developed in section 4 also comprises intertemporal information on evaluation.

For this specification the recursive construction of aggregate utility in the representation of theorem 3 becomes

$$\tilde{u}_{t-1}(x_{t-1}, p_t) = \left\{ (1-\beta_{t-1})\, x^\rho + \beta_{t-1} \left[\int_{\tilde{X}_t} \tilde{u}_t(\tilde{x}_t)^\alpha \, dp_t \right]^{\frac{\rho}{\alpha}} \right\}^{\frac{1}{\rho}} \tag{7}$$

Weil (1990) terms this framework the generalized isoelastic model. By now, it has become the standard workhorse for the disentanglement of (atemporal) risk aversion from intertemporal substitutability.[21] Assuming $\beta < 1$, the standard form of the aggregator is obtained for the limit of an infinite time horizon, where $\lim_{T\to\infty} \beta_t = \beta$ for all t.

In Traeger (2007*a*) I take the above model back into the multi-commodity world. While the individual characterizations of atemporal risk aversion and intertemporal substitutability are shown to depend on the particular good under observation, as well as their measure scale, I show that the function $f \circ g^{-1}$ stays invariant.[22] Section 4.2 will show that this invariant characterizes a form of risk aversion that I will introduce axiomatically under the name intertemporal risk aversion in section 4.1.

3.3 Risk Stationarity

The assumption on risk evaluation in the preceding section was motivated by the objective to obtain constant uncertainty aggregation rules. In this section, I derive a representation for risk stationary preferences. An analogous reasoning to the one carried out in section 3.1 yields a preference representation distinct from that given in theorem 3.

In section 3.1 I have motivated the axiom of certainty stationarity by splitting it up into two assumptions. The first requirement, corresponding to equation (3), expresses that the mere passage of time shall not change preferences. The second assumption, corresponding to equation (4), compares two scenarios or consumption paths yielding the same outcome in period $T+1$. For such paths, it requires that the adoption of a time horizon of $T+1$ and of T yield the same ranking of the two scenarios. In the following, I give an analogous reasoning for risky scenarios. As already in axioms A9 and A7, it proves sufficient to require risk stationarity only for 'coin toss' compositions of certain consumption paths, i.e. probability a half mixtures of type $\frac{1}{2}\mathrm{x} + \frac{1}{2}\mathrm{x}'$. Moreover, it is enough to have the decision

[21] Note that these formulations usually are stated in terms of a probability distribution over future welfare, rather then over outcomes. Define the probability measure p_t^* as the measure that is induced by p_t on U through the function $\tilde{u}_t$ and denote the elements of $U \subset \mathrm{I\!R}$ by $\mathfrak{u}$. Then equation (7) writes as

$$\tilde{u}_{t-1}(x_{t-1}, p_t^*) = \left\{ (1-\beta_{t-1})\, x^\rho + \beta_{t-1} \left[\int_U \mathfrak{u}^\alpha \, dp_t^* \right]^{\frac{\rho}{\alpha}} \right\}^{\frac{1}{\rho}} .$$

Moreover, among the papers cited above, Svensson (1989) translates the isoelastic model to continuous time, which is also used in Epaulard & Pommeret (2003*b*) and Epaulard & Pommeret (2003*a*).

[22] One has to take care, however, of the affine transformations allowed for f and g.

maker rank these lotteries with respect to certain alternatives. Then, the requirement analogous to equation (3) becomes

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

for all $\mathsf{x},\mathsf{x}',\mathsf{x}''\in\mathsf{X}^{t+1}$, $x\in X$ and $t\in\{1,...,T\}$. In words, the mere passage of time does not change the ranking between the different scenarios. Demanding that equation 8 holds for all periods is equivalent to the requirement that the condition holds for lotteries where uncertainty resolves at any point in the future.[23]

The second step to arrive at the axiom of risk stationarity, is to connect the relations $\succeq_{\cdot|T}$ and $\succeq_{\cdot|T+1}$. As in section 3.1, this is achieved by requiring that scenarios whose outcomes coincide in the last period of a finite planning horizon $T+1$ are ranked the same way when applying a planning horizon of T. This statement formally translates into the 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 $t\in\{1,...,T\}$. As the right hand side of the requirements in equations (8) and (9) coincides, together, the equations 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 $x\in X$:
$$\tfrac{1}{2}(\mathsf{x},x)+\tfrac{1}{2}(\mathsf{x}',x)\succeq_{t}(\mathsf{x}'',x)\;\;\Leftrightarrow\;\;\tfrac{1}{2}\mathsf{x}+\tfrac{1}{2}\mathsf{x}'\succeq_{t+1}\mathsf{x}''\quad\forall\,\mathsf{x},\mathsf{x}',\mathsf{x}''\in\mathsf{X}^{t+1}.$$

In short, the decision maker ranks lotteries of the form $\frac{1}{2}\mathsf{x}+\frac{1}{2}\mathsf{x}'$ the same way 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.

Before I come to the representation, let me briefly point out the analogous reasoning to yield risk stationarity from the assumption expressed in equation (8) in the case of an infinite planning horizon. Denote the consumption paths corresponding to (x,x) and (x',x) simply by $\mathsf{x}^\infty,\mathsf{x}'^\infty\in X^\infty$, yielding the notation $\frac{1}{2}\mathsf{x}^\infty+\frac{1}{2}\mathsf{x}'^\infty$ for the lotteries considered in the infinite horizon version of equation (8). Moreover, in the infinite horizon setting, it is $\succeq_{1|T+1}=\succeq_{1|\infty}=\succeq_{1|T}$. 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 trans-

[23]Thus, one can formulate the requirement as well by only considering preference in periods 1 and 2, as done in equation (3) in section 3.1. Such a reformulation is straight forward, once it is recognized that time consistency A4 allows to carry over all the requirements in equation (8) into the first two periods, by adding common outcomes to the beginning of all consumption plans which start in later periods.

formed using the corresponding statement for $t = 1$ into the requirement:

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

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

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

for all $\mathrm{x}^\infty, \mathrm{x}'^\infty, \mathrm{x}''^\infty \in X^\infty$, $t \in \mathbb{N}$ and $\mathrm{x}^t \in X^t$. A corresponding[24] axiom for stationarity of risk attitude is found in Chew & Epstein (1991, 356).

Preference stationarity for the evaluation of lotteries together with the assumptions of section 2 yields the following representation.

Theorem 4: Let there be given a sequence of binary relations $\succeq = (\succeq_t)_{t\in\{1,...,T\}}$ on $(P_t)_{t\in\{1,...,T\}}$ satisfying A0, and a Bernoulli utility function $u \in B_\succeq$ with range U. The sequence $\succeq = (\succeq_t)_{t\in\{1,...,T\}}$ satisfies

i) A1-A3 for all $\succeq_t, t \in \{1, ..., T\}$ (vNM setting)

ii) A4 for $\succeq_1$ (certainty additivity)

iii) A5 (time consistency)

iv) A8 (risk stationarity)

if and only if, there exists a strictly increasing and continuous function $g : U \to \mathbb{R}$ and a discount factor $\beta \in \mathbb{R}_{++}$ as well as a function $h \in \left\{\exp, \mathrm{id}, \frac{1}{\exp}\right\}$, such that with defining the functions $\tilde{w}_t : \tilde{X}_t \to \mathbb{R}$ for $t \in \{1, ..., T\}$ by

v) $\tilde{w}_T(x_T) = g \circ u(x_T)$ and recursively

$$\tilde{w}_{t-1}(x_{t-1}, p_t) = g \circ u(x_{t-1}) + \beta\, \mathcal{M}^h(p_t, \tilde{w}_t) \text{ or by} \tag{11}$$

it holds for all $t \in \{1, ..., T\}$ that

$$p_t \succeq_t p'_t \quad \Leftrightarrow \quad \mathcal{M}^h(p_t, \tilde{w}_t) \geq \mathcal{M}^h(p'_t, \tilde{w}_t) \quad \forall\, p_t, p'_t \in P_t. \tag{12}$$

Moreover, if the representation employs $h \in \left\{\exp, \frac{1}{\exp}\right\}$, then two functions g and g' both represent $(\succeq_t)_{t\in\{1,...,T\}}$ in the above sense, if and only if, there exists $b \in \mathbb{R}$ such that $g = g' + b$. In a representation employing $h = \mathrm{id}$, two functions g and g' both represent $(\succeq_t)_{t\in\{1,...,T\}}$ in the above sense, if and only if, there exist $a \in \mathbb{R}_{++}$ and $b \in \mathbb{R}$ such that $g = ag' + b$.

Note that axiom A8 not only relates the functions f_t characterizing uncertainty aggregation in theorem 2 for different periods, but also implies a close relation between the functions

[24] In difference to 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.

f_t and the function g characterizing intertemporal aggregation. In order to exploit this relation, the functions $\tilde{w}_t$ and h are introduced and replace $\tilde{u}_t$ and f_t.[25] Note that for $h = \frac{1}{\exp}$ the characterization of the uncertainty aggregation rule corresponds to the function $h(z) = \frac{1}{\exp(z)} = \exp(-z)$.[26] In difference to the requirements of intertemporally additive expected utility, the relation between 'risk aversion' and 'intertemporal substitutability' implied by axiom A8 leaves one degree of freedom.[27] This freedom breaks the representation up into the three classes, corresponding to $h \in \left\{ \exp, \mathrm{id}, \frac{1}{\exp} \right\}$. The next section elaborates that these classes correspond to intertemporally risk seeking, intertemporally risk neutral and intertemporally risk averse behavior.

4 Intertemporal Risk Aversion

4.1 Axiomatic Characterization of Intertemporal Risk Aversion

In Traeger (2007*a*) I have introduced the concept of intertemporal risk aversion in a general framework with non-stationary preferences. This section gives a slightly simplified axiomatic definition for decision makers that exhibit (at least certainty) stationary behavior. At the end of section 3.2 I have pointed out that the assignment of atemporal risk aversion and intertemporal substitutability to the functions g and f generally fails or at least becomes good-dependent in the multi-commodity setting. This problem can also be inferred from the fact that both functions depend on the choice of Bernoulli utility (see lemma 1). In fact, Traeger (2007*a*) elaborates that the good- and scale-dependence of these characterizations of intertemporal evaluation can be translated into a dependence

[25] I introduce a new name for the functions employed in the recursive construction of the representation in equation (11), as they are not complete analogues to the ones used in earlier representations. Precisely, the functions $\tilde{w}_t$ used in equation (11) relate to the functions $\tilde{u}_t$ used in the earlier theorems as $\tilde{w}_t = \frac{1}{1-\beta_t}\, g \circ \tilde{u}_t$. This rephrasing of the recursion relation greatly simplifies the representation (see part IV of the proof). Moreover, the function h characterizes the relation between f and g imposed by axiom A8. In section 4 the function h will be seen to be closely related to the notion of intertemporal risk aversion. The functions f_t characterizing uncertainty aggregation in the sense of the earlier representation theorems are affine transformations of $h \circ g$ (the affine transformation is negative for the case where $h = \frac{1}{\exp}$). However, the mentioned relation between f_t and g, which is used to simplify the representation, only holds for particular choices of g within its affine indeterminedness. In consequence, in order to exploit the relation, I have to give up part of the affine freedom for the choice of g. For this reason, in the cases where $h \in \left\{ \exp, \frac{1}{\exp} \right\}$, the function g is no longer free up to affine transformations, but only up to a translational constant.

[26] I avoid the notation $h = \exp^{-1}$ because h^{-1} is used to denote the inverse.

[27] Recall that in the intertemporally additive expected utility model, the coefficient of relative risk aversion is always fixed to the inverse of the elasticity of intertemporal substitution. I put quotation marks, as this interpretation of f and g is meaningful only in a one commodity Epstein Zin form of the representation.

on the choice of Bernoulli utility in the representation. Seeking for a measure of risk that is invariant under the choice of Bernoulli utility, and thus good-independent, I derive a concept termed intertemporal risk aversion. The following is an axiomatic characterization of the concept for the stationary setup.

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

A9$^{\text{w}}$ (weak intertemporal risk aversion) For all $\bar{\mathrm{x}}, \mathrm{x} \in \mathrm{X}^t$ holds

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

A decision maker is said to exhibit *strict intertemporal risk aversion* in period t, if and only if, the following axiom is satisfied:

A9$^{\text{s}}$ (strict intertemporal risk aversion) For all $\bar{\mathrm{x}}, \mathrm{x} \in \mathrm{X}^t$ holds

$$\bar{\mathrm{x}} \sim_t \mathrm{x} \quad \wedge \quad \exists\, \tau \in \{t, ..., T\} \text{ s.th. } \mathrm{x}_\tau \ \not\sim^*_\tau \ \bar{x}$$

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

I start with the interpretation of the strict axiom. The first part of the premise in axiom A9$^{\text{s}}$ states that a decision maker is indifferent between a constant consumption path delivering outcome $\bar{x}$ in every period and the consumption paths x. The second part of the premise requires that the path x exhibits some variation, making at least one of the outcomes better or worse than $\bar{x}$. Without loss of generality assume that outcome x_τ is strictly preferred to outcome $\bar{x}$ (i.e. $\mathrm{x}_\tau \succ^*_\tau \bar{x}$). Then[28], by the first part of the premise, there also has to exist a period τ', in which the outcome $\mathrm{x}_{\tau'}$ is judged inferior to the outcome $\bar{x}$. Moving to the implication of axiom A9$^{\text{s}}$, recall that 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 with the outcome x_i. Then, the lottery $\sum_{i=t}^{T} \frac{1}{T-t+1} (\bar{\mathrm{x}}_{-i}, \mathrm{x}_i)$ delives a consumption path with outcomes $\bar{x}$ in all but one period. In that period, i.e. in period i, the outcome $\bar{x}$ is replaced by outcome x_i. The lottery draws with equal probability the period in which the outcome is replaced.[29] As seen above, some of these outcomes x_i are better than $\bar{x}$, while others are worse. However, altogether the outcomes x_i that are considered superior and those that are considered inferior with respect to $\bar{x}$ balance each other, in the sense that receiving all of these outcomes with certainty leaves the decision maker indifferent with respect to receiving $\bar{x}$ in every period. The second line of axiom A9$^{\text{s}}$ demands that for consumption satisfying the above conditions, an intertemporally risk averse decision maker should prefer the constant consumption path $\bar{\mathrm{x}}$ with certainty over the lottery that yields with equal probability any of the consumption

[28] Assuming that the axioms of choice given in the section 2 prevail.

[29] Note that for $i \in \{t, ..., T\}$ there are $T - t + 1$ different consumption paths $(\mathrm{x}_{-i}, \mathrm{x}'_i)$, each of length $T - t + 1$.

paths $(\bar{\mathrm{x}}_{-i}, \mathrm{x}_i)$, some of which make him better off and some of which make him worse off. The intuition for such a choice is that the decision maker might be better off in the lottery than with the certain consumption path $\bar{x}$, but he might as well be worse off. This differs from the decision problem in the premise, where the decision maker can be certain to get the higher outcome x_τ in period τ whenever he receives the lower outcome $\mathrm{x}_{\tau'}$ in period τ'. In other words, the fear of receiving an outcome which is judged inferior with respect to $\bar{x}$ makes the intertemporally risk averse decision maker prefer the certain and constant consumption path to the lottery.

The interpretation of the *weaker* axiom A9$^{\text{w}}$ is analogous, only that the consumption path x is allowed to coincide with $\bar{\mathrm{x}}$, and the implication only requires that the lottery is not strictly preferred to the certain and constant consumption path. If axiom A9$^{\text{s}}$ (A9$^{\text{w}}$) is satisfied with $\succ_t$ ($\succeq_t$) replaced by $\prec_t$ ($\preceq_t$), the decision maker is called a strong (weak) *intertemporal risk seeker*. If a decision maker's preferences satisfy weak intertemporal risk aversion as well as weak intertemporal risk seeking, the decision maker is called *intertemporally risk neutral*.

In the following I give an alternative interpretation of axiom A9$^{\text{s}}$, relating it to the representation of theorem 2. For this purpose, I choose the representation where Bernoulli utility is picked in a way to render intertemporal aggregation linear ($g = \text{id}$). Then, 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 (equation 6). The second part of the premise requires that at least in one period the utility gained from consumption path $\bar{x}$ differs from the utility gained from the outcome $\bar{x}$. Then, the lottery in axiom A9$^{\text{s}}$ renders in expectation the same utility as the certain consumption path $\bar{\mathrm{x}}$. A decision maker is defined to be strictly intertemporally risk averse when preferring 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. Note that this interpretation only works for the certainty additive choice of Bernoulli utility corresponding to $g = \text{id}$. As (only) in such a certainty additive representation a utility gain of one unit (after discounting) in some period and a utility loss of one unit (after discounting) in another period leaves the aggregate utility unchanged, I suggest in Traeger (2007*d*) to identify certainty additive Bernoulli utility with the decision maker's welfare. Then, intertemporal risk aversion can be interpreted as risk aversion with respect to welfare gains and losses.

4.2 Functional Characterization of Intertemporal Risk Aversion

This section characterizes intertemporal risk aversion in terms of the functional representations derived in sections 3.1 through 3.3. I start with a characterization for the certainty stationary case in terms of *representation theorem 2*.

Theorem 5: Let the sequence of triples $(u, f_t, g)_{t\in\{1,...,T\}}$ represent the set of preferences $\succeq = (\succeq_t)_{t\in\{1,...,T\}}$ in the sense of theorem 2. Furthermore let $t \in \{1, ..., T-1\}$. Then the following assertions hold:
a) A decision maker is strictly intertemporally risk averse [seeking] in period t in the sense of axiom A9$^{\text{s}}$, if and only if, $f_t \circ g^{-1}(z)$ is strictly concave [convex] in $z \in \Gamma_t$.
b) A decision maker is weakly intertemporally risk averse [seeking] in period t in the sense of axiom A9$^{\text{w}}$, if and only if, $f_t \circ g^{-1}(z)$ is concave [convex] in $z \in \Gamma_t$.
c) A decision maker is intertemporally risk neutral in period t, if and only if, $f_t \circ g^{-1}(z)$ is linear in $z \in \Gamma_t$.

Intertemporal risk attitude is described by the second order characteristics of the function $f_t \circ g^{-1}$. I refer to the latter as the stationary characterization of intertemporal risk attitude. Note that the function $f_t \circ g^{-1}$ is concave if and only if the function $f_t \circ g_t^{-1}(z) = f_t \circ g^{-1}(\beta^{-t} z)$ is concave. As elaborated in Traeger (2007*a*), the latter characterizes intertemporal risk aversion in the general non-stationary representation of theorem 1. In the certainty stationary setting of representation theorem 2, the functions f_t are allowed to vary arbitrarily over time. Therefore, the decision maker's intertemporal risk attitude may also differ arbitrarily between different periods. This feature changes for the representations worked out in sections 3.2 and 3.3. Concerning the *representation of theorem 3*, observe that it corresponds to the special case of theorem 2, where uncertainty aggregation is constant over time, i.e. $f_t = f \; \forall t \in \{1, ..., T\}$. In consequence, theorem 5 applies with $f_t g^{-1} = f g^{-1}$ independent of the period. Thus, the decision maker is either intertemporally risk averse, risk neutral or risk seeking in *all* periods. The same is true for a decision maker whose preferences satisfy risk stationarity and can be *represented by theorem 4*. Infer from the proof of theorem 4 that h in the representation of theorem 4 corresponds to an affine transformation of $f_t \circ g^{-1}$. Therefore, the decision maker is intertemporally risk averse, risk neutral or risk seeking depending on whether h is respectively $\frac{1}{\exp}$, id or exp.[30]

In order to derive quantitative characterizations of risk attitude, I follow Traeger (2007*a*) and define for a twice differentiable function $f_t \circ g_t^{-1} : \Gamma_t \rightarrow \mathbb{R}$ the measures of *relative* and *absolute intertemporal risk aversion* as the functions :

$$\text{RIRA}_t : \Gamma_t \rightarrow \mathbb{R}$$

$$\text{RIRA}_t(z) = -\frac{\left(f_t \circ g_t^{-1}\right)''(z)}{\left(f_t \circ g_t^{-1}\right)'(z)} \, z \,, \tag{13}$$

[30]Note again that in the case where $h = \frac{1}{\exp}$, h is a negative affine transformation of $f_t \circ g^{-1}$, making the latter concave (see footnote 25). Also note that intertemporal risk attitude going along with preferences that satisfy axiom A8 can be observed better in corollary 1 following below.

and

$$\text{AIRA}_t : \Gamma_t \to \mathbb{R}$$

$$\text{AIRA}_t(z) = -\frac{\left(f_t \circ g_t^{-1}\right)''(z)}{\left(f_t \circ g_t^{-1}\right)'(z)} \,. \tag{14}$$

However, due to the affine freedom in the representations of theorems 2 to 4, these quantities are not well defined, yet. As I elaborate in detail in Traeger (2007*a*), the standard (Arrow-Pratt) measures of absolute and relative atemporal risk aversion are not well defined, unless a cardinal scale for the measurement of the consumption commodity has been fixed. Similarly, the measures of absolute and relative intertemporal risk aversion are not well defined, unless a cardinal scale for the measurement of the intertemporal consumption consumption trade of has been fixed. This is best understood in the perspective given at the end of the preceding section. Here certainty additive Bernoulli utility, which specifies the intertemporal trade off for consumption, has been given the interpretation of the decision maker's welfare. Then, intertemporal risk aversion can be interpreted as risk aversion with respect to welfare gains and losses. However, the (certainty additive) Bernoulli utility function is only defined up to affine transformations.[31] In consequence, to render the measure of absolute intertemporal risk aversion unique, the unit of welfare has to be fixed. Similarly, to render the measure of relative intertemporal risk aversion unique, the zero level of welfare has to be fixed. This intuitive reasoning which is elaborated in more detail in Traeger (2007*a*) is verified by the upcoming lemma. Before stating the latter, I have to decide for which period I fix the scale for the intertemporal trade off measure which I refer to as welfare. I adopt the convention to fix the measure scale always for the first period. In consequence, fixing $\text{range}(u_1^{ca}) = W^*$ implies that the range for later periods is fixed to $\text{range}(u_t^{ca}) = \beta^{t-1}\text{range}(u_1^{ca}) = \beta^{t-1}W^*$. Note that for different representation, which are not linear in intertemporal aggregation, the relevant intertemporal trade off measure that has to be fixed in scale is characterized by the function $g \circ u$.

Lemma 2: Let there be given a sequence of preference relations $\succeq = (\succeq_t)_{t\in\{1,\ldots,T\}}$ satisfying axioms A1-A3, A4, A5 and A6 or A8. In addition, choose
i) a number $w^* \in \mathbb{R}_{++}$,
ii) an outcome $x^{\text{zero}} \in X$ or
iii) a nondegenerate closed interval $W^* \subset \mathbb{R}$.
Then, for representations in the sense of theorems 2 or 3 with twice differentiable functions $f_t \circ g^{-1}$ and for representations in the sense of theorem 4, which

[31] Recall that the functions g characterizing intertemporal aggregation were only unique up to affine transformations in any of the stated representations. Choosing Bernoulli utility in a way to make intertemporal aggregation linear, thus, still leaves an affine freedom to the choice of the certainty additive Bernoulli utility function. This reasoning is carried out rigorously in Traeger (2007*a*).

a) satisfy $\Delta G = w^*$, the risk measures AIRA_t
b) satisfy $g \circ u(x^{\text{zero}}) = 0$, the risk measures RIRA_t
c) satisfy $\Delta G = w^*$ and $g \circ u(x^{\text{zero}}) = 0$, the risk measures AIRA_t and RIRA_t
d) satisfy $G = W^*$, the risk measures AIRA_t and RIRA_t
are determined uniquely and independent of the choice of the Bernoulli utility function.

Independence of Bernoulli utility implies that, once the corresponding welfare information has been fixed, the measures RIRA_t and AIRA_t are determined uniquely, independent of the choice of Bernoulli utility in the corresponding representation theorem and, thus, independent of any other information on goods or measure scales. In assertion a) the required welfare information is obtained from fixing the unit of measurement, by prescribing a numerical value to the difference in welfare between the best and the worst outcome, i.e. $u_1^{ca}(x^{\max}) - u_1^{ca}(x^{\min}) = g \circ u(x^{\max}) - g \circ u(x^{\min}) = \overline{G} - \underline{G} = w^*$. Such a partial specification of the measure scale for welfare makes the measures of *absolute* intertemporal risk aversion unique. Assertion b) fixes the 'zero welfare level', by choosing an outcome that shall correspond to zero welfare. The information is enough to render the measures of *relative* intertemporal risk aversion unique. Assertion c) fixes the welfare unit and the zero welfare level together. This step completely eliminates the freedom in the choice of measure scale for welfare. In consequence, both measures of intertemporal risk aversion are determined uniquely. Assertion d) offers an alternative way to eliminate the indeterminacy of the measure scale for welfare, by specifying the range of the function g and, thus, the welfare levels corresponding to the best and the worst outcomes. The latter approach is taken in the subsequent corollaries.

For preferences satisfying risk stationarity as formulated in axiom A8, it is worthwhile to take a closer look at the representations that fix the degree of freedom in the measure scale for welfare. For this purpose define the uncertainty aggregation rule $\mathcal{M}^{\exp^\xi}$ for the case $\xi = 0$ by limit, yielding[32]

$$\mathcal{M}^{\exp^0}(p_t, \tilde{w}_t) \equiv \lim_{\xi \to 0} \mathcal{M}^{\exp^\xi}(p_t, \tilde{w}_t) = \lim_{\xi \to 0} \frac{1}{\xi} \ln \left[\int dp_t \exp(\xi \tilde{w}_t) \right] = E_{p_t} \tilde{w}_t \,.$$

The limit is a simple application of l'Hospital's rule, as shown in the proof of corollary 1. Choosing Bernoulli utility in the representation of theorem 4 in a way to yield $g = \text{id}$ and fixing the range of the intertemporal trade off measure, which for $g = \text{id}$ corresponds to the range of u, one obtains the following representation.

Corollary 1 ($g = \mathbf{id}^+$**−gauge**) : Choose a nondegenerate closed interval $W^* \subset \mathbb{R}$. A sequence of binary relations $\succeq = (\succeq_t)_{t \in \{1,\dots,T\}}$ on $(P_t)_{t \in \{1,\dots,T\}}$ satisfying A0 satisfies

[32] Note that the characterization of the uncertainty aggregation rule by $f = \exp^\xi$ is equivalent to $f(z) = \exp(\xi\, z)$.

i) A1-A3 for all $\succeq_t, t \in \{1, ..., T\}$ (vNM setting)

ii) A4 for $\succeq_1$ (certainty additivity)

iii) A5 (time consistency)

iv) A8 (risk stationarity)

if and only if, there exists a continuous and surjective function $u : X \to W^*$, a discount factor $\beta \in \mathbb{R}_{++}$ and $\xi \in \mathbb{R}$, such that with defining the functions

v) $\tilde{w}_t : \tilde{X}_t \to \mathbb{R}$ for $t \in \{1, ..., T\}$ by $\tilde{w}_T(x_T) = u^{ca}(x_T)$ and recursively

$$\tilde{w}_{t-1}(x_{t-1}, p_t) = u^{ca}(x_{t-1}) + \beta\, \mathcal{M}^{\exp^\xi}(p_t, \tilde{w}_t) \tag{15}$$

it holds for all $t \in \{1, ..., T\}$ that

$$p_t \succeq_t p'_t \quad \Leftrightarrow \quad \mathcal{M}^{\exp^\xi}(p'_t, \tilde{w}_t) \geq \mathcal{M}^{\exp^\xi}(p'_t, \tilde{w}_t) \quad \forall\, p_t, p'_t \in P_t. \tag{16}$$

Moreover, the function u is determined uniquely. With the convention that $g_1 = g$, the uniquely defined measures of intertemporal risk aversion are calculated to $\text{AIRA}_t = -\frac{\xi}{\beta^{t-1}(1-\beta_t)}$ and $\text{RIRA}_t = -\frac{\xi}{\beta^{t-1}(1-\beta_t)}\,\text{id}$.

Fixing the welfare range eliminates the affine freedom of g, here corresponding to the freedom of u ($g = \text{id}-$gauge). In the representation of theorem 4, part of this freedom was employed to carry information over the relation between the functions g and f_t. Fixing g and its range exogenously, this information gives rise to the new parameter ξ. It parametrizes intertemporal risk aversion and corresponds to a degree of freedom between the function g, characterizing intertemporal aggregation, and the functions f_t, characterizing uncertainty aggregation. In the particular case where $\xi = 0$, the coefficient of relative atemporal risk aversion (defined in a one commodity setting) is confined to the inverse of the intertemporal elasticity of substitution and intertemporal risk neutrality prevails. In this case, equations (15) and (16) recursively define the intertemporally additive expected utility framework.

The representation in theorem 4 features three different representations corresponding to $h \in \left\{\frac{1}{\exp}, \text{id}, \exp\right\}$ that appear to be disconnected. Corollary 1 shows how the coefficient of absolute intertemporal risk aversion, which is proportional to ξ, connects these classes continuously, allowing for a wide range of intertemporal risk attitude. Moreover, axiom A8 implies that the coefficient of absolute intertemporal risk aversion is constant in the utlity levels respectively outcomes. However, the coefficient of absolute intertemporal risk aversion is not constant over time. In the discussion of theorem 5, it had already been observed that for risk stationary representations in the sense of axiom A7 only the expression $f_t \circ g^{-1}$, which I referred to as the stationary characterization of intertemporal risk aversion, stays constant over time. The general characterization $f_t \circ g_t^{-1}$ picks up the discount rate from $g_t = \beta^{t-1} g$. The same happens for risk stationarity in the sense of axiom A8. The interpretation is as follows. The function $f_t \circ g^{-1}$ characterizes intertemporal risk aversion in period t with respect to a welfare scale that is fixed in period t to range(g) = range(u^{ca}) = W^*. One could formulate this characterization as a measurement in terms of a 'current value measure scale for welfare'. With respect to such a constant 'current value measure scale', the characterizing functions of intertemporal risk aversion are constant over time.[33] In contrast, the measures AIRA$_t$ and RIRA$_t$ are defined with respect to the characterizing functions $f_t \circ g_t^{-1}$. Fixing range(g_1) = range(g) = range(u^{ca}) = W^*, implies that range(g_t) = β^{t-1}range(g) = $\beta^{t-1}W^*$. Thus, in these measures intertemporal risk aversion is measured with respect to a 'present value measure scale for welfare' and discounting shrinks the range of welfare that serves as basis for the measurement of intertemporal risk aversion in period t. Then, as the range of the welfare measure scale (in present value) becomes smaller and smaller over time due to discounting, the coefficient of intertemporal risk aversion has to increase (in absolute terms) in order to keep up a stationary aversion to risk. Therefore, the coefficients of intertemporal risk aversion AIRA$_t$ and RIRA$_t$ include the factor β^t in the denominator.

Moreover, risk stationarity implies another dependence of absolute intertemporal risk aversion on time. In the denominator of AIRA$_t$ appears as well the time-dependent normalized discount factor β_t. Recall that the latter takes account of the relative weight given to a single period as opposed to the remaining future, a weight changing over time when a finite planning horizon is approached. For a representation satisfying risk stationarity in the sense of axiom A8, this change of weight enters into the characterization of intertemporal risk aversion. It implies that the stationary part of intertemporal risk aversion, characterized by $f_t \circ g^{-1}$, slowly decreases over time as the term $1 - \beta_t$ increases to unity for the last period. Leaving this term unconsidered, yields a representation in the sense of theorem 3, satisfying axiom A7. In other words, *disregarding the adjustment of intertemporal risk aversion by the change of weight that the remaining future obtains as opposed to the present period*, in a setting with a finite planning horizon, *makes the corresponding decision maker indifferent to the length of risk taking (axiom A7)*. For an infinite time horizon this weight is obviously constant, precisely it holds $(1 - \beta_t) = (1 - \beta)$, and a representation in the (limiting) sense of theorem 4 also is a representation in the (limiting) sense of theorem 3.[34]

Alternatively to corollary 1, the representation can also be stated in the Kreps and Porteus form, where uncertainty aggregation is rendered linear. A corresponding corollary is stated in appendix A. Similarly, in order to relate to the generalized isoelastic model,

[33] Except for the normalization factor $\frac{1}{1-\beta_t}$ in the case of risk stationarity in the sense of axiom A8. This term will be discussed further below.

[34] Note that a constant term $(1 - \beta)$ can be absorbed into the parameter ξ and makes no difference for the comparison between different classes of representations.

one can render uncertainty aggregation in equation (16) isoelastic. If this is done for risk stationary preferences, the corresponding intertemporal aggregation turns out to be multiplicative (see appendix A). In consequence, for a one-commodity setting, the only overlap between the generalized isoelastic model and the risk stationary model is obtained for an intertemporal elasticity of zero. Interpreting once more certainty additive Bernoulli utlity as welfare, this setting corresponds to a logarithmic welfare function. The latter is an assumption frequently put forth in certainty additive macroeconomic models.

5 Timing Indifference and the Rate of Pure Time Preference

5.1 Indifference to the Timing of Risk Resolution

A particular feature of the recursive utility models employed in the preceding sections, is that they allow for an intrinsic preference for early or late resolution of uncertainty. This preference is intrinsic in the sense that a decision maker can strictly prefer an early resolution of uncertainty, even if the information obtained from the early resolution is known not to affect his plans and, thus, his future outcomes. In Traeger (2007*b*) I analyze the relation between intertemporal risk aversion and such a preference for the timing of uncertainty resolution. Assuming indifference to the timing of uncertainty resolution implies that the evaluation of uncertainty can be reduced to the evaluation of atemporal lotteries expressing uncertainty as probability measures on consumption paths. In particular, from a normative point of view, such an indifference can be motivated by the fact that a decision maker who has a strict preference for early of late resolution of uncertainty is willing to give up welfare for pre-drawing or postponing information of which he knows that it has no effect on his future behavior or payoffs. In this section I analyze the effects of adding such an indifference assumption to the setting laid out so far. The according axiom is stated as follows.

A10 (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 $\lambda \in [0, 1]$ it holds

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

In words, a decision maker who is indifferent to the timing of risk resolution in the sense of axiom A10, does not distinguish between a lottery where the uncertainty about the future faced in period $t+1$ - described by lottery p_{t+1} versus lottery p'_{t+1} - resolves in period t (lottery on the left) or in period $t+1$ (lottery on the right). Here, the uncertainty about the future faced in period $t+1$ is described by the probability mixture λ and $1-\lambda$.

Under indifference to the timing of risk resolution the preference representation only makes use of the atemporal lotteries $p_t^{\mathsf{X}}, p_t'^{\mathsf{X}} \in \Delta(\mathsf{X}^t)$ that are defined non-recursively over consumption paths. In Traeger (2007*b*) I show how these probability measures can be derived from their recursive counterparts $p_t, p_t' \in P_t$. This relation, however, is only needed to axiomatize the representation within the more general setting. For an application of the representation theorem below, it is sufficient to describe the uncertain future directly by the measures $p_t^{\mathsf{X}} \in \Delta(\mathsf{X}^t)$. The following representation analyzes the consequences of certainty stationarity in such a non-recursive setting. I state the representation in the certainty additive form, where it coincides with the standard discount utility evaluation on certain consumption paths. Moreover I immediately fix the measure scale of the intertemporal trade-off (to the exogenously given interval W^*) in order to render the measures of intertemporal risk aversion unique.

Theorem 6 ($g = \textbf{id}^+$−gauge, certainty stationary) : Choose a nondegenerate closed interval $W^* \subset \mathrm{I\!R}_{++}$. A sequence of binary relations $\succeq = (\succeq_t)_{t\in\{1,...,T\}}$ on $(P_t)_{t\in\{1,...,T\}}$ satisfying A0, satisfies

i) A1-A3 for all $\succeq_t, t \in \{1, ..., T\}$ (vNM setting)

ii) A4 for $\succeq_1$ (certainty additivity)

iii) A5 (time consistency)

iv) A6 & A10 (certainty stationarity & timing indifference)

if and only if, there exists a continuous and surjective function $u^{ca} : X \to W^*$, a discount factor $\beta \in \mathrm{I\!R}_{++}$ and $\xi \in \mathrm{I\!R}$, such that with defining

v) the functions $\tilde{u}_t : \tilde{X}_t \to \mathrm{I\!R}$ for $t \in \{1, ..., T\}$ by

$$\tilde{u}_t(\mathsf{x}^t) = \sum_{\tau=t}^{T} \beta^{\tau-1} u^{ca}(\mathsf{x}_\tau^t)$$

it holds for all $t \in \{1, ..., T\}$ that

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

Moreover, the function u is determined uniquely, as are the measures of intertemporal risk aversion $\mathrm{AIRA}_t = -\frac{\xi}{1-\beta_t}$ and $\mathrm{RIRA}_t = -\frac{\xi}{1-\beta_t}\,\mathrm{id}$.

In view of the axioms note that, in the above setting, a result by Chew & Epstein (1989, 110) would allow to replace the independence axiom A3 by a collection of weaker assumptions. The representation evaluates outcomes in all periods by means of a common, certainty additive Bernoulli utility function u which, once more, I will identify with the decision maker's welfare. Overall evaluation of a particular consumption path is performed by taking the discounted sum of per period welfare. 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 $\mathcal{M}^{\exp^\xi}$, which is parametrized (up to a normalization factor) by the coefficient of absolute intertemporal risk seeking, i.e. the negative of absolute intertemporal risk aversion. For the limit of an infinite time horizon, the normalization constant $1-\beta_t$ (depicting the relative weight of an individual period as opposed to the remaining future) becomes constant over time. In consequence, also the coefficient of intertemporal risk aversion $\lim_{T\to\infty} \text{AIRA}_t = -\frac{\xi}{1-\beta}$ becomes constant. For a finite time horizon, however, as the end of the planning horizon is approached, the decreasing length of the welfare paths under consideration goes along with a coefficient of absolute intertemporal risk aversion AIRA_t that decreases over time to $-\xi$ for the last period. Note that, in accordance with the convention underlying lemma 2, the measure scale for welfare has been fixed to W^* in period 1, implying ranges $\beta^{t-1}W^*$ for welfare measurement in later periods. In particular, theorem 6 shows that it is possible to disentangle atemporal risk aversion from intertemporal substitutability without assuming an intrinsic preference for early or late resolution of uncertainty. For details see Traeger (2007*b*). In addition, the model is compatible with the widespread discount utility model for the evaluation of individual consumption paths.

5.2 Implications for Discounting

In the preceding section I have shown that the requirement of indifference to the timing of uncertainty resolution is compatible with strict intertemporal risk aversion and a discount utility evaluation of certain consumption paths. This section analyzes the consequences of merging the assumption of indifference to the timing of uncertainty resolution with that of *risk stationarity* formulated in axiom A8.

In theorem 6, I have described how certainty stationarity determines the time development of intertemporal risk aversion for a decision maker who has no intrinsic preference for early or late resolution of uncertainty. The coefficient of absolute intertemporal risk aversion was seen to be constant in welfare and to adapt to the length of the planning horizon lying ahead of the decision maker. It was calculated to $\text{AIRA}_t = -\frac{\xi}{1-\beta_t}$. Similarly, the assumption of risk stationarity formulated in axiom A8 gives rise to a coefficient of absolute intertemporal risk aversion that is constant in welfare. Moreover, the respective representation stated in corollary 1 exhibits the same adaption of the coefficients of intertemporal risk aversion to the length of the remaining planning horizon through the factor $\frac{1}{1-\beta_t}$. However, in contrast to the representation of the preceding section, for a decision maker who complies with risk stationarity, the coefficient of absolute intertemporal risk aversion also depends on the discount factor β^t. As I elaborated have elaborated in section 3.3, under the assumption of axiom A8 only the functions $f_t \circ g^{-1}$ stay constant over time (up to the normalization by $\frac{1}{1-\beta_t}$). These functions, to which I have referred as the stationary characterization of intertemporal risk attitude, measure intertemporal risk aversion with respect to a 'current value measure scale for welfare'. In contrast, the coefficient AIRA_t expresses intertemporal risk aversion with respect to the 'present value measure scale for welfare'. That is, if the measure scale for period 1 is fixed to range$(u_1^{ca}) = W^*$, then the measure scale of welfare in period t shrinks down to the range$(u_t^{ca}) = \beta^{t-1}W^*$. But then, as the range of welfare measurement (in present value) becomes smaller and smaller over time due to discounting, the coefficient of intertemporal risk aversion has to increase in order to keep up a stationary aversion to risk. However, this is not allowed by axiom A10. If indifference to the timing of uncertainty resolution should prevail, the latter requires intertemporal risk aversion to be constant over time (up to the normalization by $\frac{1}{1-\beta_t}$). Otherwise, a decision maker would be willing to give up welfare in order to have uncertainty resolved in the period with the lowest intertemporal risk aversion, even if the information obtained is known to be of no use.

In consequence, risk stationary devaluation of the future, which implies by axiom A8 a decreasing coefficient of absolute intertemporal risk aversion, is not compatible with the demand of axiom A10, i.e. the lack of an intrinsic preference of uncertainty resolution. Precisely, there is only one situation where such a devaluation of the future is compatible with both axioms. For a decision maker who is intertemporally risk neutral, the assumption of risk stationarity has no more bite than the assumption of certainty stationarity. Here, the coefficient $\text{AIRA}_t = 0$ for all $t \in \{1, ..., T\}$ is constant over time and, thus, the intertemporally additive expected utility model trivially satisfies the requirements implied by both axioms. However, for a nontrivial model of intertemporally risk averse decision making, the following result obtains.

Theorem 7: A sequence of binary relations $\succeq = (\succeq_t)_{t\in\{1,\ldots,T\}}$ on $(P_t)_{t\in\{1,\ldots,T\}}$ satisfying axiom A0, satisfies

i) A1-A3 for all $\succeq_t, t \in \{1, ..., T\}$ (vNM setting)

ii) A4 for $\succeq_1$ (certainty additivity)

iii) A5 (time consistency)

iv) A9$^{\text{s}}$ (strict intertemporal risk aversion)

v) A8 (risk stationarity II)

vi) A10 (timing indifference)

if and only if, there exists a representation in the sense of theorem 6 with $\xi < 0$ and $\beta = 1$.

In words, a decision maker who accepts the above axioms does not discount the future due to an intrinsic timing preference. However, he does devaluate uncertain welfare in the sense of the certainty additive Bernoulli utility u^{ca}. In consequence, if uncertainty increases over time, future welfare gains less weight than current welfare. The remainder of this paragraph renders the latter intuition precise. For this purpose let $p_1^{\mathrm{X}} \in \Delta(\mathrm{X}^1)$ be a product measure $p_1^{\mathrm{X}} = \mathbb{P}_1^{x_1} \otimes ... \otimes \mathbb{P}_T^{x_T}$ with $\mathbb{P}_\tau^{x_\tau} \in \Delta(X)$ and $\tau \in \{1, ..., T\}$, so that the outcomes in different periods are independently distributed. Assume that expected welfare is the same in all periods, i.e. $E_{\mathbb{P}_\tau^{x_\tau}} \; u^{ca}(x_\tau) = \mathfrak{u}^* \in U = W^* \, \forall \, \tau \in \{1, ..., T\}$. To define what it means that uncertainty over welfare is increasing over time, I employ Rothschild & Stiglitz's (1970) definition of increasing risk. The authors define a random variable to be riskier than another, if the corresponding probability distribution has more weight on the tails.[35] In particular, this condition is satisfied for a mean preserving spread. Now, consider the probability distributions $\mathbb{P}_\tau^u$ over U that are induced by the measures $\mathbb{P}_\tau^{x_\tau}$ through the certainty additive Bernoulli utility function u^{ca}. Then, uncertainty of welfare increases over time, if $\mathbb{P}_t^u$ has more weight in the tails than $\mathbb{P}_{t'}^u$ for all $t, t' \in \{1, ..., T\}$ satisfying $t > t'$.[36] For such an uncertainty specification it follows from theorem 2 in Rothschild & Stiglitz (1970, 237) that the certainty equivalent of welfare in period t is lower than the certainty equivalent of welfare in period t'. As the expected welfare is the same in both periods, the difference in weights exhibits some resemblance to discounting. Note, in particular, that the intertemporally additive expected utility model does not allow for intertemporal risk aversion and, thus, not for risk aversion on welfare and the reasoning I carried out above. Therefore, the only possibility it permits to capture a difference in the weighting of expected welfare is by introducing a positive rate of pure time preference.

To my knowledge, the only consideration in the literature which is concerned with a relation between discounting and stationarity that is somewhat comparable to the one derived in theorem 7, is due to Epstein (1992, 16). Motivating models of recursive utility, he points out a contradiction between a disentanglement of risk aversion and intertemporal substitutability in a non-recursive model on the one hand, and the positiveness of the discount rate on the other. He concludes that a disentanglement is not possible, at least in a stationary setting. The preceding section has elaborated how such a disentanglement is possible in a non-stationary and in a certainty stationary setting. Theorem 7 confirms Epstein's (1992, 16) assertion, but with a very different interpretation. Having analyzed the reasons and consequences of an intrinsic timing preference in Traeger (2007*b*), I suggest that a non-recursive evaluation can be a desirable normative characterisitic (axiom A10).

[35] An equivalent characterization is that the riskier random variable can be obtained from the less risky random variable by adding some noise. For a formal definition compare footnote 36.

[36] Formally let P_t denote the cumulative distribution function characterizing the measure $\mathbb{P}_t^u$ for $t \in \{1, ..., T\}$. P_t is said to have more weight in the tales than $\mathrm{P}_{t'}$, if $\int_{\underline{U}}^{\mathfrak{u}} \mathrm{P}_t(y) - \mathrm{P}_{t'}(y) \, dy \geq 0 \; \forall \, \mathfrak{u} \in [\underline{U}, \overline{U}]$.

In that case, a risk stationary decision maker in the sense of axiom A8 has to accept that he does not have the freedom to devaluate the future for sheer impatience, without violating any of the other axioms. Furthermore, theorem 7 together with theorem 6 show that, for a decision maker with a finite planning horizon, it is well possible to disentangle atemporal risk aversion from intertemporal substitutability, without violating any of the axioms. Moreover, also in the limit of an infinite planning horizon, a factor $\beta = 1$ does not necessarily imply that aggregate welfare diverges. Due to intertemporal risk aversion, an increase in uncertainty over time can still yield a finite evaluation of scenarios.[37] Naturally, instead of accepting the consequences of theorem 7, the underlying axioms can be dropped. Since Chew & Epstein (1989, 110) have shown that under the assumption of axiom A10 the independence axiom can be replaced by a collection of much weaker axioms, it is not a promising candidate to give up in order to avoid the implication of a zero rate of pure time preference. However, one could abandon risk stationarity and allow for an (anticipated) change of preference over time.

I close by pointing out an important application of the modeling framework derived in this section. In relation to climate change, Nordhaus' (1993,1994) integrated assessment model for climate change and its critical discussions and alternative assessments have shown the importance of carefully quantifying the social discount rate for the derivation of an optimal greenhouse gas abatement path (see e.g. Toth 1995, Plambeck, Hope & Anderson 1997). In particular, Plambeck et al. (1997, 85) point out that a reduction of the pure rate of time preference from 3%, as assumed by Nordhaus (1993), to 0% (corresponding to $\beta = 1$), would result in an optimal abatement path that cuts emissions by 50% from the baseline to the year 2100, as opposed to 10% in the assessment of Nordhaus (1993). To the best of my knowledge, so far a zero rate of pure time preference has only been argued for in terms of moral consideration. Theorem 7 states formal axioms dealing with consistency aspects of evaluation under uncertainty, and shows that their acceptance alone suffices to call for a zero rate of pure time preference. In difference to the evaluations used in current climate models, however, the representation implied by theorem 7 goes along with an intertemporally risk averse decision maker. Therefore, uncertainty has a higher cost than in the above climate models, which apply the intertemporally risk neutral standard model when they consider uncertainty at all.[38] In consequence, an evaluation of global climate change under the assumptions of theorem 7, implies an additional preference for scenarios that give rise to a less uncertain future. Since uncertainty is likely to increase

[37]Note however, that increasing uncertainty can also make the evaluation functional converge to zero. Preliminary analysis shows that convergence to finite non-zero evaluation are knife-edge in the assumptions on the probability distributions and their evolvement over time.

[38]With the exception of the stylized simulation by Ha-Duong & Treich (2004) that features two possible damage states in a generalized isoelastic framework.

in the perturbation of the climate system, which increases with the amount of greenhouse gas emissions, a first conjecture is that the additional effect caused by intertemporal risk aversion in an evaluation in the sense of theorem 7, yields an even higher abatement recommendation than the one pointed out by Plambeck et al. (1997, 85). A closer analysis of this aspect constitutes an interesting area of future research.

6 Conclusions

The paper has imposed different *stationarity* assumptions on a general recursive evaluation framework featuring nontrivial intertemporal risk aversion and a finite planning horizon. The axioms offer an alternative to the standard stationarity axioms that rely on an infinite time horizon and a positive rate of pure time preference. Certainty stationarity has been characterized and shown to imply the standard discount utility model on certain consumption paths. To this purpose, the standard stationarity assumption has been decomposed into two more basic assumptions. The first of which sololey states that the mere passage of time does not affect preferences. The second assumes that the ranking of two different consumption paths does not depend on a common outcome in the last period. I put forth an axiom implying constancy of the functions characterizing (atemporal) uncertainty aggregation. The resulting representation includes the generalized isoelastic model, However, the corresponding axiom does not express the idea that the mere passage of time should not affect preference orderings. A careful translation of the latter assumption to the evaluation of risky outcomes, implies that constancy of atemporal risk attitude is only supported for an infinite time horizon. For a finite planning horizon, the corresponding axiom no longer admits the whole class of generalized isoelastic evaluation rules. Moreover, for risk stationary preferences, the measure of absolute intertemporal risk aversion has been shown to be characterized by a single parameter, to be constant in welfare, and to increases over time.

I have refined the model by assuming *indifference to the timing of uncertainty resolution*. The corresponding axiom is only required to hold for situations, where the information resulting from an early resolution of uncertainty cannot be used to alter outcomes. As a result, evaluation can be expressed non-recursively over consumption paths and uncertainty can be modeled by means of probability distributions over consumption paths instead of the more general temporal lotteries. I have shown that in such a framework, a stationary evaluation over risky outcomes implies a zero *rate of pure time preference*. Precisely, a devaluation of the future for reasons of pure time preference is only possible for an intertemporally risk neutral decision maker, where the axiom of *risk* stationarity has no additional bite. When uncertainty is increasing over time, also an intertemporally risk averse decision maker values (expected) future welfare less than current welfare.

From a *normative perspective*, the modeling framework suggests that a decision maker, who accepts the assumption that were shown to result in a zero rate of pure time preference, has to attach a high weight to the long-run in model-based scenario evaluations. While the certain future is treated equal to the present, the uncertain future gains the more importance, the more the decision maker can know about it. In this connection, note that the idea of precaution and, thus, intertemporal risk averse evaluation of the future is often put forth in the sustainability debate. Moreover, in the same context, the current generation is expected to sustain living conditions for future generations that somehow resemble their own. Usually, the latter condition is directly translated into a requirement concerning the discount rate. However, as the current paper points out, translating the condition simply into the requirement of stationarity, i.e. a similarity between todays needs and future needs, already can be sufficient to yield a strong implication for the choice of the discount rate when combined with the concept of precaution. As a corresponding application of the derived representational framework I suggested a reevaluation of climate change, deducing a first conjecture with regard to the qualitative effects. Finally, the paper offers an alternative framework to disentangle atemporal risk aversion from intertemporal substitutability. Rather than imposing isoelastic evaluation on intertemporal trade-offs and risk, it motivates a constant coefficient of intertemporal risk aversion. From a *descriptive perspective*, a comparison of the models ability to explain observed phenomena has to be contrasted with that of the generalized isoelastic model.

Notation

Symbol	Explanation	Page
$\succeq$	weak preference relation	5
$\succ$	strong preference relation	5
$\sim$	indifference relation	5
$\succeq_t$	preference relation in period t on P_t	7
$\succeq_t^*$	derived preference relation over period t outcomes in X	9
$\succeq\|_{\mathsf{X}}$	restriction of $\succeq$ to the set of certain consumption paths	
$\equiv$	defining equality	
A	$\mathsf{A} = \{\mathfrak{a} : \mathbb{R} \to \mathbb{R} : \mathfrak{a}(z) = a\,z + b\,,\ a, b \in \mathbb{R}, a \neq 0\}$, group of affine transformations	
A^+	$\mathsf{A}^+ = \{\mathfrak{a}^+ : \mathbb{R} \to \mathbb{R} : \mathfrak{a}^+(z) = a\,z + b\,,\ a, b \in \mathbb{R}, a > 0\}$	
AIRA_t	measure of absolute intertemporal risk aversion in period t	26
$B_{\succeq}$	$B_{\succeq} \equiv B_{\succeq_1}$, stationary setting	14
$B_{\succeq_t}$	set of Bernoulli utility functions, non-stationary setting $B_{\succeq_t} = \{u_t \in \mathcal{C}^0(X) : x \succeq_t^* x' \Leftrightarrow u_t(x) \geq u_t(x')\,\forall\, x, x' \in X\}$	9
$\mathcal{C}^0(X)$	space of all continuous functions from X to $\mathbb{R}$	3
$\Delta(Y)$	space of Borel probability measures on Y	3
ΔG_t	$\Delta G_t = \overline{G}_t - \underline{G}_t$	8
exp	exponential function	
$f g^{-1}$	$f \circ g^{-1}$, composition	
$\underline{G}_t$	$\underline{G}_t = g_t(\underline{U}_t)$	8
$\overline{G}_t$	$\overline{G}_t = g_t(\overline{U}_t)$	8
G_t	$G_t = [\underline{G}_t, \overline{G}_t]$	8
Γ_t	$\Gamma_t = (\underline{G}_t, \overline{G}_t)$	8
id	identity	
$\lambda x + (1-\lambda)x'$	lottery over outcomes x and x' with respective probabilities λ and $1-\lambda$, $\lambda x + (1-\lambda)x' \in P$	4
ln	natural logarithm	
$\mathcal{M}^f$	uncertainty aggregation rule, $\mathcal{M}^f : \Delta(Y) \times \mathcal{C}^0(Y) \to \mathbb{R}$ with $\mathcal{M}^f(p,u) = f^{-1}\left[\int_Y f \circ u\, dp\right]$, $f : \mathbb{R} \to \mathbb{R}$ strictly monotonic and continuous	4
$\mathcal{M}^\alpha$	shorthand for $\mathcal{M}^{\text{id}^\alpha}(p,u) = \left[\int_Y u^\alpha dp\right]^{\frac{1}{\alpha}}$	5
$\mathcal{M}^0$	shorthand for $\lim_{\alpha\to 0} \mathcal{M}^\alpha(p,u) = \exp\left[\int_Y \ln(u)\, dp\right]$	5
$\mathcal{M}^{f_t}$	uncertainty aggregation rule in period t	7

Symbol	Explanation	Page
P	space of Borel probability measures on X	3
p	uncertain outcomes or lotteries, $p \in P$	3
P_t	general choice space in period t	7
p_t	period t lottery, $p_t \in P_t$	7
$\mathbb{R}_+$	$\mathbb{R}_+ = \{z \in \mathbb{R} : z \geq 0\}$	4
$\mathbb{R}_{++}$	$\mathbb{R}_{++} = \{z \in \mathbb{R} : z > 0\}$	4
range	range of a function	
RIRA_t	measure of relative intertemporal risk aversion in period t	25
RRA	Arrow-Pratt-measure of relative risk aversion	18
θ_t	normalization constant, non-stationary representation	9
ϑ_t	normalization constant, non-stationary representation	9
u	Bernoulli utility function, $u \in \mathcal{C}^0(X)$	3, 9
$\underline{U}$	$\min_{x \in X} u(x)$	3
$\overline{U}$	$\max_{x \in X} u(x)$	3
U	$\mathrm{range}(u) = [\underline{U}, \overline{U}]$	3
$\underline{U}_t$	$\min_{x \in X} u_t(x)$	8
$\overline{U}_t$	$\max_{x \in X} u_t(x)$	8
U_t	$[\underline{U}_t, \overline{U}_t] = \mathrm{range}(u_t)$	8
u^{ca}	certainty additive Bernoulli utility function, also interpreted as the decision maker's welfare	
X	connected compact metric space of outcomes	3
x	consumption levels or (certain) outcomes, $x \in X$	3
X^t	space of consumption paths starting in period t	6
x	consumption path	7
x^t	consumption path from period t to period T, $\mathsf{x}^t \in \mathsf{X}^t$	7
x^t_τ	period τ entry of consumption path x^t	7
$(\mathsf{x}_{-i}, \mathsf{x}'_i)$	$(\mathsf{x}_t, ..., \mathsf{x}_{i-1}, \mathsf{x}'_i, \mathsf{x}_{i+1}, ..., x_T) \in \mathsf{X}^\tau$	7
$\tilde{X}_t$	degenerate choice space in period t	7
$\tilde{x}_t$	degenerate period t lottery	7
Y	connected compact metric space, used for general definitions	

A Additional Specifications of the Risk Stationary Represenation

An alternative representation of risk stationary preferences that fixes the measures of intertemporal risk aversion uniquely is the following.

Corollary 2 **($f=$ id$^+$-gauge)** : Choose a nondegenerate closed interval $U^* \subset \mathbb{R}_{++}$. A sequence of binary relations $\succeq=(\succeq_t)_{t\in\{1,...,T\}}$ on $(P_t)_{t\in\{1,...,T\}}$ satisfying A0 satisfies

i) A1-A3 for all $\succeq_t, t\in\{1,...,T\}$ (vNM setting)

ii) A4 for $\succeq_1$ (certainty additivity)

iii) A5 (time consistency)

iv) A8 (risk stationarity)

if and only if, there exists a continuous and surjective function $u : X \rightarrow U^*$, a discount factor $\beta \in \mathbb{R}_{++}$ and $\xi \in \mathbb{R}$, such that defining the functions

v) $\tilde{u}_t : \tilde{X}_t \rightarrow \mathbb{R}$ for $t\in\{1,...,T\}$ by $\tilde{u}_T(x_T)=u(x_T)$ and recursively

- for $\xi>0$: $\tilde{u}_{t-1}(x_{t-1},p_t)=u(x_{t-1})^\xi \big(\mathrm{E}_{p_t}\tilde{u}_t\big)^\beta$ and
- for $\xi=0$: $\tilde{u}_{t-1}(x_{t-1},p_t)=\ln u(x_{t-1})+\beta\,\mathrm{E}_{p_t}\tilde{u}_t$ and
- for $\xi<0$: $\tilde{u}_{t-1}(x_{t-1},p_t)=-u(x_{t-1})^\xi \big(-\mathrm{E}_{p_t}\tilde{u}_t\big)^\beta$

it holds for all $t\in\{1,...,T\}$ that

$$p_t \succeq_t p'_t \quad \Leftrightarrow \quad \mathrm{E}_{p_t}\tilde{u}_t \geq \mathrm{E}_{p'_t}\tilde{u}_t \quad \forall\, p_t, p'_t \in P_t\,.$$

Moreover, the function u is determined uniquely. With the convention that $g_1=g$,[39] the uniquely defined measures of intertemporal risk aversion are calculated to $\mathrm{AIRA}_t = -\frac{\xi}{\beta^{t-1}(1-\beta_t)}$ and $\mathrm{RIRA}_t = -\frac{\xi}{\beta^{t-1}(1-\beta_t)}\,\mathrm{id}$.

Here, uncertainty is evaluated by taking the expected value. However, this comes at the cost of a nonlinear aggregation of Bernoulli utlity over time. In the above representation, the functions $\tilde{u}_t$ are the same as those used in representation theorem 1.

Observe the particular nonlinear form for intertemporal aggregation that arises when uncertainty aggregation is required to be linear. For decision makers that are not intertemporally risk neutral, it is 'almost multiplicative'. But it depends on the exponent ξ. Translating the exponent ξ back into the uncertainty aggregation rule and establishing a purely multiplicative intertemporal aggregation yields the following representation.

[39] This notation relates to the underlying representing triples in the sense of theorem 1. In corollary 2 the assumption implies that the measure scale of welfare is fixed for the first period to range(u_1^{welf}) = range$(g\circ u)=\ln U^*$. See also the discussion below and the first remark in the proof of corollary 2.

Corollary 3 (isoelastic uncertainty evaluation): Choose a nondegenerate closed interval $U^* \subset \mathbb{R}_{++}$. A sequence of binary relations $\succeq = (\succeq_t)_{t\in\{1,...,T\}}$ on $(P_t)_{t\in\{1,...,T\}}$ satisfying A0 satisfies

i) A1-A3 for all $\succeq_t, t \in \{1, ..., T\}$ (vNM setting)

ii) A4 for $\succeq_1$ (certainty additivity)

iii) A5 (time consistency)

iv) A8 (risk stationarity)

if and only if, there exists a continuous and surjective function $u : X \to U^*$, a discount factor $\beta \in \mathbb{R}_{++}$ and $\xi \in \mathbb{R}$, such that defining the functions

v) $\tilde{v}_t : \tilde{X}_t \to \mathbb{R}$ for $t \in \{1, ..., T\}$ by $\tilde{v}_T(x_T) = u(x_T)$ and recursively

$$\tilde{v}_{t-1}(x_{t-1}, p_t) = u(x_{t-1}) \left(\mathcal{M}^{\alpha=\xi}(p_t, \tilde{v}_t)\right)^\beta \tag{17}$$

it holds for all $t \in \{1, ..., T\}$ that

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

Moreover, the function u is determined uniquely. With the convention that $g_1 = g$,[40] the uniquely defined measures of intertemporal risk aversion are calculated to $\text{AIRA}_t = -\frac{\xi}{\beta^{t-1}(1-\beta_t)}$ and $\text{RIRA}_t = -\frac{\xi}{\beta^{t-1}(1-\beta_t)}$ id.

Note, that the recursive construction (17) of the representation for $\xi = 0$ is equivalent to the intertemporally additive expected utility setting. The above representation is particularly interesting, because it points out a special case that closely relates to the generalized isoelastic framework analyzed in section 3.2. In the one commodity setting and for $u = \text{id}$ the following recursive characterization of the decision maker's evaluation is obtained:

$$\tilde{v}_{t-1}(x_{t-1}, p_t) = x_{t-1} \left(\mathcal{M}^{\alpha=\xi}(p_t, \tilde{v}_t)\right)^\beta . \tag{18}$$

It corresponds to an intertemporal elasticity of substitution of unity ($\rho = 0$) and uses the isoelastic uncertainty aggregation rule $\mathcal{M}^\alpha$. Adopting once more the interpretation of certainty additive Bernoulli utility as welfare, the case $\rho = 0$, i.e. $g = \ln$, corresponds to logarithmic welfare, which is a widespread assumption in macroeconomics and popular also in environmental economic modeling. It is the only specification for which risk stationarity in the sense of axiom A8 allows an isoelastic uncertainty aggregation.[41] Observe

[40] See footnote 39.

[41] In order to render f_t a power function, intertemporal risk aversion has to be of the form $f_t \circ g^{-1} = \left(g^{-1}\right)^\alpha$. This expression can only be proportional to $\exp^\xi$, characterizing up to proportionality intertemporal risk aversion in the representation of theorem 4, if g^{-1} is proportional to exp. But then g has to be proportional to ln which corresponds to the case $\rho = 0$.

that the setting (18) only coincides with the (corresponding special case of the) isoelastic representation of section 3.2 for an infinite planning horizon. As pointed out on page 29, the representations in the sense of theorems 3 and 4 differ in the way they take account of the approaching end of the planning horizon. More precisely, only the representation based on axiom A8 incorporates the change in weight that the present receives as opposed to the remaining future 'which shortens over time'. In contrast, axiom A7 characterizes what it needs to make atemporal uncertainty aggregation constant in a setting with a finite time horizon. This condition can be expressed as a form of indifference to the length of risk taking.

B Proofs for Section 3

Remark: I apologize for that fact that some of the references to theorems and equations in the accompanying papers Traeger (2007*a*) and Traeger (2007*b*) produce **??** and have not yet been revised.

Proof of theorem 1: See Traeger (2007*a*).

Proof of theorem 2: The proof is divided into four parts. Axioms A1-A3, A4 and A5 assure the existence of a representation in the sense of theorem 1. In the first part I show that axiom A6 allows to pick the same Bernoulli utility for all periods. In the second part I work out a relation between the functions g_t in different periods that has to hold in such a representation by axiomA6. Part three calculates the corresponding normalization constants and brings about the representation stated in the theorem. Finally, part four proves the necessity of the axioms.

Part I ("⇒"): I show that axiom A6 implies that there exists a strictly monotonic and continuous transformation s_t such $u_{t-1} = s_t \circ u_t$ for any $t \in \{1, .., T\}$. To this end, translate axiom A6 into the representation of theorem 1 using equation (**??**).

$$
\begin{aligned}
(\mathrm{x}^2, x^0) \quad &\succeq_1 \quad (\mathrm{x}'^2, x^0) \qquad \Leftrightarrow \qquad \mathrm{x}^2 \quad \succeq_2 \mathrm{x}'^2 \\
\rightsquigarrow \tilde{u}_1\big((\mathrm{x}^2, x^0)\big) \;&\geq\; \tilde{u}_1\big((\mathrm{x}'^2, x^0)\big) \quad \Leftrightarrow \quad \tilde{u}_1(\mathrm{x}^2) \;\geq\; \tilde{u}_1(\mathrm{x}'^2) \\
\rightsquigarrow g_1^{-1} &\left(\theta_1 \sum_{\tau=2}^{T} g_{\tau-1} \circ u_{\tau-1}(\mathrm{x}^2_\tau) + g_T \circ u_T(x^0) + \xi_1 \right) \\
&\geq g_1^{-1} \left(\theta_1 \sum_{\tau=2}^{T} g_{\tau-1} \circ u_{\tau-1}(\mathrm{x}'^2_{\tau-1}) + g_T \circ u_T(x^0) + \xi_1 \right) \\
\Leftrightarrow g_2^{-1} &\left(\theta_2 \sum_{\tau=2}^{T} g_\tau \circ u_\tau(\mathrm{x}^2_\tau) + \xi_2 \right) \geq g_2^{-1} \left(\theta_2 \sum_{\tau=2}^{T} g_\tau \circ u_\tau(\mathrm{x}'^2_\tau) + \xi_2 \right)
\end{aligned}
$$

Considering in particular the consumption paths x^2, x'^2 satisfying $x^2_\tau = x'^2_\tau \, \forall \, \tau \neq t$ yields

$$\begin{aligned} &\rightsquigarrow g_{t-1} \circ u_{t-1}(x^2_t) \geq g_{t-1} \circ u_{t-1}(x'^2_t) \Leftrightarrow g_t \circ u_t(x^2_t) \geq g_t \circ u_t(x'^2_t) \\ &\rightsquigarrow \quad u_{t-1}(x^2_t) \quad \geq \quad u_{t-1}(x'^2_t) \quad \Leftrightarrow \quad u_t(x^2_t) \quad \geq u_t(x'^2_t) \end{aligned}$$

for all $x^2_t = x'^2_t \in X$. Therefore, as in the proof of proposition **??**, it has to exist a strictly monotonic and continuous transformation s_t such that $u_{t-1} = s_t \circ u_t$. But then, by induction it is $B_{\succeq_1} = B_{\succeq_2} = ... = B_{\succeq_T} \equiv B_{\succeq}$ and I can pick a common Bernoulli utility function $u \in B_{\succeq}$ for all periods.

Part II ("⇒"): In this part, I derive an affine relation between the functions g_t in different periods. To this end, I translate axiom A6 into the particular representation in the sense of theorem 1, which applies the same Bernoulli utility function u for all periods. Using again equation (**??**) I obtain the condition

$$\begin{aligned} \sum_{\tau=2}^{T} g_{\tau-1} \circ u_{\tau-1}(x^2_\tau) + \cancel{g_T \circ u_T(x^0)} &\geq \sum_{\tau=2}^{T} g_{\tau-1} \circ u_{\tau-1}(x'^2_{\tau-1}) + \cancel{g_T \circ u_T(x^0)} \\ \Leftrightarrow \sum_{\tau=2}^{T} g_\tau \circ u_\tau(x^2_\tau) &\geq \sum_{\tau=2}^{T} g_\tau \circ u_\tau(x'^2_\tau) \end{aligned}$$

for all $x^2, x'^2 \in X^2$. The above equivalence implies that both, $\sum_{\tau=2}^{T} g_\tau \circ u_\tau(x^2_\tau)$ and $\sum_{\tau=2}^{T} g_{\tau-1} \circ u_{\tau-1}(x^2_\tau)$, are representations for $\succeq_2 |_{X^2}$. In consequence, by the moreover part of theorem 1 there exist $a \in \mathbb{R}_{++}$ and $b_t \in \mathbb{R}, t \in \{1, ..., T-1\}$, such that $g_t = a g_{t+1} + b_t$ for all $t \in \{1, ..., T-1\}$.[42] Use the freedom in the uniqueness of $(g_t)_{t \in \{1,...,T\}}$ to define $\tilde{g}_t = g_t - \sum_{\tau=t}^{T-1} a^{\tau-t} b_\tau$ for $t \in \{1, ...T-1\}$ without loosing the representative character of the sequence $(u, f_t, \tilde{g}_t)_{t \in \{1,...,T\}}$ for $(\succeq_t)_{t \in \{1,...,T\}}$. Observe that $\tilde{g}_t = g_t - \sum_{\tau=t}^{T-1} a^{\tau-t} b_\tau = a g_{t+1} + b_t - b_t - a \sum_{\tau=t+1}^{T-1} a^{\tau-t} b_\tau = a \tilde{g}_{t+1}$. Set $g = a^{T-1} \tilde{g}_T$. Moreover let $\beta = a^{-1}$. Then the sequence of triples $(u, f_t, a^{T-t} \tilde{g}_T) = (u, f_t, \beta^{t-T} \beta^{T-1} g) = (u, f_t, \beta^{t-1} g)$ for $t \in \{1, ..., T\}$ represents $(\succeq_t)_{t \in \{1,...,T\}}$ in the sense of theorem 1.

Note 1: Expressing the triples with respect to $\tilde{g}_\tau$ instead of g yields the equivalent representation triples $(u, f_t, \beta^{t-\tau} \tilde{g}_\tau)_{t \in \{1,...,T\}}$ and in particular for $\tau = T$ the representation $(u, f_t, \beta^{t-T} \tilde{g}_T)_{t \in \{1,...,T\}}$.

Part III ("⇒"): Calculating the corresponding normalization constants for the representing tuples derived in the previous step, yields the representation stated in the theorem. In the usual convention denote $\Delta G_t = \overline{G}_t - \underline{G}_t$ and $G = [\underline{G}, \overline{G}] = [g(\min_{x \in X} u(x)),$

[42] Here it is $g'_t = g_{t+1}$. Note that it is immediate from the proof of the moreover part in theorem 1 that coincidence of the representations (only) on the certain outcome paths is enough to assure the uniqueness result for $(g_t)_t \in \{1, ..., T\}$.

$g(\max_{x\in X} u(x))]$ and find

$$\theta_t = \frac{\Delta G_t}{\sum_{\tau=t}^T \Delta G_\tau} = \frac{\beta^t \Delta G}{\sum_{\tau=t}^T \beta^\tau \Delta G} = \frac{1}{1+\beta+\beta^2+...+\beta^{T-t}} = \frac{1-\beta}{1-\beta^{T-t+1}} \quad \text{for } \beta \neq 1,$$

$$\theta_t = \frac{\Delta G_t}{\sum_{\tau=t}^T \Delta G_\tau} = \frac{\Delta G}{\sum_{\tau=t}^T \Delta G} = \frac{\Delta G}{(T-t+1)\Delta G} = \frac{1}{T-t+1} \quad \text{for } \beta = 1 \text{ and}$$

$$\vartheta_t = \frac{\overline{G}_{t+1}\underline{G}_t - \underline{G}_{t+1}\overline{G}_t}{\Delta G_t} = \frac{\beta^{t+1}\overline{G}\,\beta^t\underline{G} - \beta^{t+1}\underline{G}\,\beta^t\overline{G}}{\beta^t \Delta G} = 0.$$

Using equation (??) it is straight forward to calculate the aggregate intertemporal utility functions. In the case $\beta \neq 1$ they are

$$\begin{aligned}\tilde{u}_t(\cdot,\cdot) &= \tilde{g}_t^{-1}\Big[\theta_t \tilde{g}_t \circ u(\cdot) + (1-\theta_t)\frac{\Delta\tilde{G}_t}{\Delta\tilde{G}_{t+1}}\Big(\tilde{g}_{t+1}\circ \mathcal{M}^{f_{t+1}}(\cdot,\tilde{u}_{t+1}) + 0\Big)\Big\}\Big] \\ &= g^{-1}\Big[\beta^{-t+1}\Big\{\theta_t\beta^{t-1} g\circ u(\cdot) + (1-\theta_t)\beta^{-1}\Big(\beta^t g\circ\mathcal{M}^{f_{t+1}}(\cdot,\tilde{u}_{t+1})\Big)\Big\}\Big] \\ &= g^{-1}\Big[\theta_t g\circ u(\cdot) + (1-\theta_t) g\circ \mathcal{M}^{f_{t+1}}(\cdot,\tilde{u}_{t+1})\Big].\end{aligned}$$

Defining $\beta_t = 1-\theta_t = 1-\frac{1-\beta}{1-\beta^{T-t+1}} = \frac{1-\beta^{T-t+1}-1+\beta}{1-\beta^{T-t+1}} = \beta\frac{1-\beta^{T-t}}{1-\beta^{T-t+1}}$ gives the representation stated in the theorem. For $\beta = 1$ find

$$\tilde{u}_t(\cdot,\cdot) = g^{-1}\Big[\frac{1}{T-t+1} g\circ u(\cdot) + (1-\frac{1}{T-t+1}) g\circ\mathcal{M}^{f_{t+1}}(\cdot,\tilde{u}_{t+1})\Big]$$

and define $\beta_t = 1-\theta_t = 1-\frac{1}{T-t+1} = \frac{T-t+1-1}{T-t+1} = \frac{T-t}{T-t+1}$ to get the stated representation.

Note 2: For the evaluation of certain consumption paths equation (??) together with $\vartheta_t = 0$ and hence $\xi_t = 0$ yields:

$$\tilde{u}_t(\mathrm{x}^t) = \tilde{g}_t^{-1}\Big[\theta_t \sum_{\tau=t}^T \tilde{g}_\tau \circ u_\tau(\mathrm{x}_\tau^t)\Big] = g^{-1}\Big[(1-\beta_t)\sum_{\tau=t}^T \beta^{\tau-t} g\circ u(\mathrm{x}_\tau^t)\Big]. \tag{19}$$

Part IV ("⇐"): Axioms A1-A5 follow immediately from "⇐" of theorem 1. To see that axiom A6 holds, take a look at equation (19) and note that g^{-1} and the x^0 term cancel in the representation of A6 (for any x^0).

Moreover part: The moreover part is an immediate consequence of the moreover part of theorem 1. □

Proof of theorem 3: "⇒": Axioms A1-A6 assure the existence of a representation in the sense of theorem 2. Axiom A7 implies furthermore that the uncertainty aggregation rules in different periods can be characterized by the same function. Using equation (19) to translate axiom A7 into the representation of theorem 2 yields for the first expression

$$
\begin{aligned}
& \frac{1}{2}\bar{\mathrm{x}}^{\,t} + \frac{1}{2}\bar{\mathrm{x}}'^{\,t} && \succeq_t && \bar{\mathrm{x}}''^{\,t} \\
\Leftrightarrow\quad & f_t^{-1}\left[\frac{1}{2} f_t \circ \tilde{u}_t(\bar{\mathrm{x}}^{\,t}) + \frac{1}{2} f_t \circ \tilde{u}_t(\bar{\mathrm{x}}'^{\,t})\right] && \geq && \tilde{u}_t(\bar{\mathrm{x}}''^{\,t}) \\
\Leftrightarrow\quad & f_t^{-1}\Big[\frac{1}{2} f_t \circ g^{-1}\big[(1-\beta_t)\textstyle\sum_{\tau=t}^{T}\beta^{\tau-t} g \circ && u(\bar{x})\big] + \frac{1}{2} f_t \circ g^{-1}\big[(1-\beta_t)\textstyle\sum_{\tau=t}^{T}\beta^{\tau-t} g \circ u(\bar{x}')\big]\Big] \\
& && \geq && g^{-1}\big[\,(1-\beta_t)\textstyle\sum_{\tau=t}^{T}\beta^{\tau-t} g \circ u(\bar{x}'')\,\big] \\
\Leftrightarrow\quad & f_t^{-1}\Big[\frac{1}{2} f_t \circ g^{-1}\big[\frac{1-\beta}{1-\beta^{T-t+1}}\,\frac{1-\beta^{T-t+1}}{1-\beta}\, g \circ u(\bar{x})\big] + \frac{1}{2} f_t \circ g^{-1}\big[\frac{1-\beta}{1-\beta^{T-t+1}}\,\frac{1-\beta^{T-t+1}}{1-\beta}\, g \circ u(\bar{x}')\big]\Big] \\
& && \geq && g^{-1}\big[\,\frac{1-\beta}{1-\beta^{T-t+1}}\,\frac{1-\beta^{T-t+1}}{1-\beta}\, g \circ u(\bar{x}'')\,\big] \\
\Leftrightarrow\quad & f_t^{-1}\Big[\frac{1}{2} f_t \circ u(\bar{x}) + \frac{1}{2} f_t \circ u(\bar{x}')\Big] && \geq && u(\bar{x}'')\,,
\end{aligned}
$$

and analogously for the second expression

$$
\begin{aligned}
& \frac{1}{2}\bar{\mathrm{x}}^{\,t+1} + \frac{1}{2}\bar{\mathrm{x}}'^{\,t+1} && \succeq_{t+1} \quad \bar{\mathrm{x}}''^{\,t+1} \\
\Leftrightarrow\ & f_{t+1}^{-1}\left[\frac{1}{2} f_{t+1} \circ \tilde{u}_{t+1}(\bar{\mathrm{x}}^{\,t+1}) + \frac{1}{2} f_{t+1} \circ \tilde{u}_{t+1}(\bar{\mathrm{x}}'^{\,t+1})\right] && \geq \quad \tilde{u}_{t+1}(\bar{\mathrm{x}}''^{\,t+1}) \\
\Leftrightarrow\ & f_{t+1}^{-1}\Big[\frac{1}{2} f_{t+1} \circ u(\bar{x}) + \frac{1}{2} f_{t+1} \circ u(\bar{x}')\Big] && \geq \quad u(\bar{x}'')\,.
\end{aligned}
$$

For all $\bar{x}, \bar{x}' \in X$ there is an outcome $\bar{x}'' \in X$ such that the above relations hold with equality (compare proof of theorem **??**). This fact implies that the following equality has to hold for all $\bar{x}, \bar{x}' \in X$:

$$
\begin{aligned}
& f_t^{-1}\Big[\frac{1}{2} f_t \circ u(\bar{x}) + \frac{1}{2} f_t \circ u(\bar{x}')\Big] = f_{t+1}^{-1}\Big[\frac{1}{2} f_{t+1} \circ u(\bar{x}) + \frac{1}{2} f_{t+1} \circ u(\bar{x}')\Big] \\
\Leftrightarrow\ & f_{t+1} f_t^{-1}\Big[\frac{1}{2} f_t u(\bar{x}) + \frac{1}{2} f_t u(\bar{x}')\Big] = \frac{1}{2} f_{t+1}\, f_t^{-1} f_t\, u(\bar{x}) + \frac{1}{2} f_{t+1}\, f_t^{-1} f_t\, u(\bar{x}')\,.
\end{aligned}
$$

Defining $h_t = f_{t+1} \circ f_t^{-1}$ and the interval $\mathsf{F}_t = f_t(U)$, this condition translates into the equation

$$
h_t\left(\frac{1}{2}y + \frac{1}{2}y'\right) = \frac{1}{2}h_t\,(y) + \frac{1}{2}h_t\,(y') \qquad \forall y, y' \in \mathsf{F}_t\,.
$$

Therefore h_t has to be linear on F_t (Hardy, Littlewood & Polya 1964, refinement of theorem 83 on p.74). Hence the expression $f_{t+1} \circ f_t^{-1}$ is linear on $f_t(U)$ implying that there exists

$\mathfrak{a}_t \in \mathsf{A}$ such that with $z = f_t^{-1}(y) \in U$ it is

$$\begin{aligned} f_{t+1}f_t^{-1}(y) &= \mathfrak{a}_t^{-1}\,y \\ \Leftrightarrow \quad f_t^{-1}(y) &= f_{t+1}^{-1}\mathfrak{a}_t^{-1}\,y \\ \Leftrightarrow \quad f_t(z) &= \mathfrak{a}_t f_{t+1}(z)\,. \end{aligned}$$

By the fact that f_t and f_{t+1} are both increasing it follows that $\mathfrak{a}_t$ has to be *positive* affine, i.e. $\mathfrak{a}_t \in \mathsf{A}^+$. But as each f_t in the representation is determined only up to positive affine transformations, setting $f_t = f_{t+1} = f$ still yields a representation of $\succeq_{t\in\{1,\ldots,T\}}$.
"⇐": Axioms A1-A5 follow immediately from "⇐" of theorem 1. As seen in the first part of the proof, for constant consumption paths it is $\tilde{u}_t(\bar{\mathrm{x}}^t) = u(\bar{x}) \forall t \in \{1, ..., T\}$. Therefore axiom A7 is seen to hold by observing that for all $\bar{x}, \bar{x}', \bar{x}'' \in X$:

$$\begin{aligned} & \frac{1}{2}\bar{\mathrm{x}}^{\,t} + \frac{1}{2}\bar{\mathrm{x}}'^{\,t} && \succeq_t && \bar{\mathrm{x}}''^{\,t} \\ \Leftrightarrow \quad & f^{-1}\left[\frac{1}{2}f \circ \tilde{u}_t(\bar{\mathrm{x}}^{\,t}) + \frac{1}{2}f \circ \tilde{u}_t(\bar{\mathrm{x}}'^{\,t})\right] && \geq && \tilde{u}_t(\bar{\mathrm{x}}''^{\,t}) \\ \Leftrightarrow \quad & f^{-1}\left[\frac{1}{2}f \circ \tilde{u}_{t+1}(\bar{\mathrm{x}}^{\,t+1}) + \frac{1}{2}f \circ \tilde{u}_{t+1}(\bar{\mathrm{x}}'^{\,t+1})\right] && \geq && \tilde{u}_{t+1}(\bar{\mathrm{x}}''^{\,t+1}) \\ \Leftrightarrow \quad & \frac{1}{2}\bar{\mathrm{x}}^{\,t+1} + \frac{1}{2}\bar{\mathrm{x}}'^{\,t+1} && \succeq_{t+1} && \bar{\mathrm{x}}''^{\,t+1}\,. \end{aligned}$$

Moreover part: The moreover part is an immediate consequence of the moreover part of theorem 1. □

Proof of lemma 1: Except for admitting decreasing functions f and g, when changing "increasing" into "monotonic" in theorem 3, the statements are special cases of lemma **??**, corollary **??** and corollary **??**. The decreasing functions come in the same way as in the proofs for chapter **??**, by noting that $\mathcal{M}^{\mathfrak{a}f} = \mathcal{M}^f$ and analogously $\mathcal{N}^{\mathfrak{a}g} = \mathcal{N}^g$ for all $\mathfrak{a} \in \mathsf{A}$. Therefore, if the triple (u, f, g) represents $\succeq_{t\in\{1,\ldots,T\}}$ in the sense of theorem 3, then so do the triples $(u, -f, g)$ and $(u, f, -g)$, if f and g are admitted to be decreasing in the representation. □

Proof of theorem 4: The proof is divided into five parts. First, I translate axiom A8 into the representation of theorem 2. This step yields a requirement for the representing functions f_t and g that is solved in the second part under the assumption of differentiability of $f_t \circ g^{-1}$. The third part shows that the derived solution has to hold as well without assuming differentiability. Part four translates the solution into the representation stated in the theorem. Finally, part five proofs the necessity of the axioms for the representation.
Part I ("⇒"): First note that axiom A8 implies axiom A6 by choosing $\mathrm{x} = \mathrm{x}'$. Therefore

a representation in terms of theorem 2 has to exist. In order to translate 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 $(\mathsf{x}, x^0) \in \mathsf{X}^t$ corresponds to $(\mathsf{x}, x^0)_\tau = \mathsf{x}_{\tau+1}$ for $\tau \in \{t, ..., T-1\}$. Then, using equation (19), the left hand side of the equivalence in axiom A8 translates into

$$\begin{aligned}
&\frac{1}{2}(\mathsf{x}, x^0) + \frac{1}{2}(\mathsf{x}', x^0) \succeq_t (\mathsf{x}'', x^0) \\
\Leftrightarrow\ & f_t^{-1}\Big\{\frac{1}{2} f_t g^{-1}\Big[(1-\beta_t)\textstyle\sum_{\tau=t}^{T-1} \beta^{\tau-t} gu(\mathsf{x}_{\tau+1}) + (1-\beta_t)\beta^{T-t} gu(x^0)\Big] \\
&\qquad + \frac{1}{2} f_t g^{-1}\Big[(1-\beta_t)\textstyle\sum_{\tau=t}^{T-1} \beta^{\tau-t} gu(\mathsf{x}'_{\tau+1}) + (1-\beta_t)\beta^{T-t} gu(x^0)\Big]\Big\} \\
&\geq g^{-1}\Big[(1-\beta_t)\textstyle\sum_{\tau=t}^{T-1} \beta^{\tau-t} gu(\mathsf{x}''_{\tau+1}) + (1-\beta_t)\beta^{T-t} gu(x^0)\Big] \\
\Leftrightarrow\ & g f_t^{-1}\Big\{\frac{1}{2} f_t g^{-1}\Big[(1-\beta_t)\textstyle\sum_{\tau=t+1}^{T} \beta^{\tau-(t+1)} gu(\mathsf{x}_\tau) + (1-\beta_t)\beta^{T-t} gu(x^0)\Big] \\
&\qquad + \frac{1}{2} f_t g^{-1}\Big[(1-\beta_t)\textstyle\sum_{\tau=t+1}^{T} \beta^{\tau-(t+1)} gu(\mathsf{x}'_\tau) + (1-\beta_t)\beta^{T-t} gu(x^0)\Big]\Big\} \\
&\geq \Big[(1-\beta_t)\textstyle\sum_{\tau=t+1}^{T} \beta^{\tau-(t+1)} gu(\mathsf{x}''_\tau) + (1-\beta_t)\beta^{T-t} gu(x^0)\Big].
\end{aligned}$$

Define the sum $S = \sum_{\tau=t+1}^{T} \beta^{\tau-(t+1)} gu(\mathsf{x}_\tau)$ and similarly $S' = \sum_{\tau=t+1}^{T} \beta^{\tau-(t+1)} gu(\mathsf{x}'_\tau)$ and $S'' = \sum_{\tau=t+1}^{T} \beta^{\tau-(t+1)} gu(\mathsf{x}''_\tau)$ as well as $A = (1-\beta_t)\beta^{T-t} gu(x^0)$. Then, varying the consumption paths x, x' and x'' in X^{t+1}, goes along with varying S, S' and S'' in the interval $[\frac{1-\beta^{T-t}}{1-\beta}\,\underline{G}, \frac{1-\beta^{T-t}}{1-\beta}\,\overline{G}]$. Similarly, as x^0 is varied in X, the value A takes on any number in the interval $[(1-\beta_t)\beta^{T-t}\,\underline{G}, (1-\beta_t)\beta^{T-t}\,\overline{G}]$. In the introduced notation, the above inequality corresponding to the left hand side of the equivalence in axiom A8 writes as

$$g f_t^{-1}\Big\{\frac{1}{2} f_t g^{-1}\Big[(1-\beta_t)S + A\Big] + \frac{1}{2} f_t g^{-1}\Big[(1-\beta_t)S' + A\Big]\Big\} - A \geq (1-\beta_t)S''. \tag{20}$$

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

$$g f_{t+1}^{-1}\Big\{\frac{1}{2} f_{t+1} g^{-1}\Big[(1-\beta_{t+1})S\Big] + \frac{1}{2} f_{t+1} g^{-1}\Big[(1-\beta_{t+1})S'\Big]\Big\} \geq (1-\beta_{t+1})S''. \tag{21}$$

As derived in the proof of theorem 1 (induction hypothesis H??), for every lottery $p_{t+1} \in P_{t+1}$ there exists a certain consumption path as certainty equivalent. In consequence, for any $\mathsf{x}, \mathsf{x}' \in \mathsf{X}^{t+1}$, there exists a certainty equivalent $\mathsf{x}'' \in \mathsf{X}^{t+1}$ for the lottery $\frac{1}{2}\mathsf{x} + \frac{1}{2}\mathsf{x}' \in P_{t+1}$, such that equation (21) holds with equality. Then, by axiom A8 also equation (20) has to hold with equality. Equating the two equations by S'' yields the requirement

$$\begin{aligned}
&g f_t^{-1}\Big\{\frac{1}{2} f_t g^{-1}\Big[(1-\beta_t)S + A\Big] + \frac{1}{2} f_t g^{-1}\Big[(1-\beta_t)S' + A\Big]\Big\} - A \qquad (22)\\
&= \frac{(1-\beta_t)}{(1-\beta_{t+1})} g f_{t+1}^{-1}\Big\{\frac{1}{2} f_{t+1} g^{-1}\Big[(1-\beta_{t+1})S\Big] + \frac{1}{2} f_{t+1} g^{-1}\Big[(1-\beta_{t+1})S'\Big]\Big\}
\end{aligned}$$

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

Part II ("⇒"): In this part, I establish the general solution to equation (22), under the assumption that $h_t = f_t \circ g^{-1}$ is differentiable for all $t \in \{1, ..., T\}$. First, observe that the right hand side of equation (22) is independent of A.[43] Thus, the left hand side has to be constant in A. Taking the first derivative with respect to A, the latter requirement yields

$$\frac{\partial}{\partial A}\; h_t^{-1}\Big\{\frac{1}{2}h_t\,[\,(1-\beta_t)S + A] + \frac{1}{2}h_t\,[\,(1-\beta_t)S' + A]\,\Big\} - A = 0$$

$$\Leftrightarrow\; h_t^{-1\prime}\Big\{\frac{1}{2}h_t\,[\,(1-\beta_t)S + A] + \frac{1}{2}h_t\,[\,(1-\beta_t)S' + A]\,\Big\}\cdot$$

$$\Big\{\frac{1}{2}h_t'\,[\,(1-\beta_t)S + A] + \frac{1}{2}h_t'\,[\,(1-\beta_t)S' + A]\,\Big\} = 1$$

$$\Leftrightarrow\; \frac{1}{2}h_t'\,[\,(1-\beta_t)S + A] + \frac{1}{2}h_t'\,[\,(1-\beta_t)S' + A]$$

$$= h_t'\Big\{h^{-1}\Big[\frac{1}{2}h_t\,[\,(1-\beta_t)S + A] + \frac{1}{2}h_t\,[\,(1-\beta_t)S' + A]\,\Big]\Big\},$$

where the prime at the function h_t (and only the one at the function h_t) denotes a derivative. Defining $y = h_t\,[(1-\beta_t)S + A]$ and $y' = h_t\,[(1-\beta_t)S' + A]$, both in $\mathsf{F}_t = \big(f_t(\underline{U}), f_t(\overline{U})\big)$, the latter equation becomes

$$\frac{1}{2}h_t'h_t^{-1}(y) + \frac{1}{2}h_t'h_t^{-1}(y') = h_t'h_t^{-1}\left(\frac{1}{2}y + \frac{1}{2}y'\right).$$

By Hardy et al. (1964, refinement of theorem 83 on p.74) it follows that the composition $h_t'h^{-1}$ has to be linear. Therefore, I obtain the following differential equation for h_t, where $a_t, b_t \in \mathbb{R}$ and $z = h_t^{-1}(y) \in \Gamma_t = (\underline{G}, \overline{G})$:

$$h_t'h_t^{-1}(y) = a_t y + b_t \qquad \forall y \in \mathsf{F}_t$$

$$\Leftrightarrow \quad h_t'(z) = a_t h_t(z) + b_t \quad \forall z \in \Gamma_t. \tag{23}$$

For $a_t = 0$ the solution to $h_t'(z) = b_t$ is obviously $h_t(z) = b_t\,z + k_t$ with $k_t \in \mathbb{R}$. I will come back to this solution below (case 2). In the meanwhile (case 1), assume $a_t \neq 0\,\forall t \in \{1, ..., T\}$.

Case 1, $a_t \neq 0\,\forall t \in \{1, ..., T\}$**:**

First the differential equation (23) is solved using variation of the constant. Solving the

[43] Note that a functional equation that corresponds to the requirement that the left hand side of equation (22) is independent of A is solved in a different way by Aczél (1966, 153) by relating it to a Cauchy equation.

homogeneous differential equation for period t yields

$$\begin{aligned}
\int \tfrac{1}{h_t}\, dh_t &= \int a_t\, dz \\
\Leftrightarrow \quad \ln h_t &= a_t z + \tilde{c}_t \qquad \text{with } \tilde{c}_t \in \mathbb{R} \\
\Leftrightarrow \quad h_t(z) &= c_t \exp(a_t z) \qquad \text{with } c_t = \exp(\tilde{c}_t) \in \mathbb{R}_{++}.
\end{aligned}$$

Taking the integration constant c_t as a function of z renders the ansatz $h_t(z) = c_t(z)\exp(a_t z)$ for the inhomogeneous equation:

$$\begin{aligned}
h_t'(z) &= a_t h_t(z) + b_t \\
\Rightarrow c_t'(z)\exp(a_t z) + c_t(z)\, a_t \exp(a_t z) &= a_t\, c_t(z)\exp(a_t z) + b_t \\
\Rightarrow \quad c_t'(z)\exp(a_t z) &= b_t \\
\Rightarrow \quad \int dc_t &= \int b_t \exp(-a_t z) dz \\
\Rightarrow \quad c_t(z) &= -\frac{b_t}{a_t}\exp(-a_t z) + k_t \qquad \text{with } k_t \in \mathbb{R}.
\end{aligned}$$

Therefore $h_t(z) = [-\frac{b_t}{a_t}\exp(-a_t z) + k]\exp(a_t z) = -\frac{b_t}{a_t} + k_t \exp(a_t z)$ *with* $k_t \in \mathbb{R}$ *is the general solution to equation* (23) *with* $a_t, b_t \in \mathbb{R}, a_t \neq 0$. Note, however, that it is also known by theorem 2 that h_t has to be strictly increasing. Thus, whenever for $a_t > 0$ it has to hold as well $k_t > 0$ and for $a_t < 0$ it has to hold as well $k_t < 0$. Furthermore denote $d_t = -\frac{b_t}{a_t} \in \mathbb{R}$, and determine the inverse of h_t to be $h_t^{-1}(y) = \frac{1}{a_t}\ln\left[\frac{-d_t+y}{k_t}\right]$.[44]

Second, I substitute the solution for h_t and h_{t+1} back into equation (22) to find for the left hand side

$$\begin{aligned}
& h_t^{-1}\Big\{\frac{1}{2}h_t\Big[(1-\beta_t)S + A\Big] + \frac{1}{2}h_t\Big[(1-\beta_t)S' + A\Big]\Big\} - A \\
&= \frac{1}{a_t}\ln\left[\frac{1}{k_t}\Big\{ - d_t + \frac{1}{2}d_t + \frac{1}{2}k_t\exp\Big[a_t\{(1-\beta_t)S + A\}\Big] \right. \\
&\qquad \left. + \frac{1}{2}d_t + \frac{1}{2}k_t \exp\Big[a_t\{(1-\beta_t)S' + A\}\Big]\Big\}\right] - A \\
&= \frac{1}{a_t}\ln\left[\frac{1}{2}\exp\Big[a_t\{(1-\beta_t)S + A\}\Big] + \frac{1}{2}\exp\Big[a_t\{(1-\beta_t)S' + A\}\Big]\right] - A \\
&= \frac{1}{a_t}\ln\left[\frac{1}{2}\exp\Big[a_t\,(1-\beta_t)S\Big] + \frac{1}{2}\exp\Big[a_t\,(1-\beta_t)S'\Big]\right], \\
&= \ln\left[\frac{1}{2}\exp(S)^{a_t\,(1-\beta_t)} + \frac{1}{2}\exp(S')^{a_t\,(1-\beta_t)}\right]^{\frac{1}{a_t}},
\end{aligned}$$

[44] Note that the relation holds also holds $k_t < 0$. Then the nominator inside the logarithm $-d_t + y = k_t \exp(a_t z)$ is negative as well.

and analogously for the right hand side

$$\frac{(1-\beta_t)}{(1-\beta_{t+1})}\, h_{t+1}^{-1}\Big\{\frac{1}{2}h_{t+1}\Big[(1-\beta_{t+1})S\Big]+\frac{1}{2}h_{t+1}\Big[(1-\beta_{t+1})S'\Big]\Big\}$$
$$=\frac{(1-\beta_t)}{(1-\beta_{t+1})}\,\frac{1}{a_{t+1}}\ln\left[\frac{1}{2}\exp\Big[a_{t+1}\,(1-\beta_{t+1})S\Big]+\frac{1}{2}\exp\Big[a_{t+1}\,(1-\beta_{t+1})S'\Big]\right]$$
$$=\ln\left[\frac{1}{2}\exp(S)^{a_{t+1}\,(1-\beta_{t+1})}+\frac{1}{2}\exp(S')^{a_{t+1}\,(1-\beta_{t+1})}\right]^{\frac{(1-\beta_t)}{a_{t+1}(1-\beta_{t+1})}}. \tag{24}$$

Therefore, equation (22) requires that for a continuum of values S and S' it has to hold

$$\left[\frac{1}{2}\exp(S)^{a_t\,(1-\beta_t)}+\frac{1}{2}\exp(S')^{a_t\,(1-\beta_t)}\right]^{\frac{1}{a_t(1-\beta_t)}}$$
$$=\left[\frac{1}{2}\exp(S)^{a_{t+1}\,(1-\beta_{t+1})}+\frac{1}{2}\exp(S')^{a_{t+1}\,(1-\beta_{t+1})}\right]^{\frac{1}{a_{t+1}(1-\beta_{t+1})}}.$$

Necessary and sufficient for this equality is the condition $a_t(1-\beta_t)=a_{t+1}(1-\beta_{t+1})\equiv\xi$.[45] As $a_t\in\mathbb{R}\backslash\{0\}$ and $1-\beta_t\neq 0$ for all $t\in\{1,...,T\}$, there exists a solution if and only if $\xi\in\mathbb{R}\backslash\{0\}$. Summarizing, in the case that $a_t\neq 0$ for all $t\in\{1,...,T\}$, *equation* (22) *implies that there exists* $\xi\in\mathbb{R}\backslash\{0\}$ *such that for every* t *it is* $h_t(z)=f_t\circ g^{-1}(z)=d_t+k_t\exp(\frac{\xi}{1-\beta_t}z)$ *with* $d_t,k_t\in\mathbb{R},k_t\neq 0$. In addition for $\xi>0$ it has to hold $k_t>0$ and for $\xi<0$ it has to hold $k_t<0$.

Case 2, $\exists\, t\in\{1,...,T\}$ with $a_t=0$**:**

The solution to equation (23) for $a_t=0$ is $h_t(z)=b_t\,z+k_t$ with $k_t\in\mathbb{R}$. By theorem 2 it is known that h_t has to be strictly increasing. Thus, the constant b_t has to be strictly positive. But then, the constants b_t and k_t correspond to positive affine transformations of f_t, which are known not to affect the representation. Therefore, wlog I can set $b_t=1$ and $k_t=0$. Then h_t *is the identity* and the left hand side of equation (22) becomes

$$gf_t^{-1}\Big\{\frac{1}{2}f_tg^{-1}\Big[(1-\beta_t)S+A\Big]+\frac{1}{2}f_tg^{-1}\Big[(1-\beta_t)S'+A\Big]\Big\}-A$$
$$=\frac{1}{2}\Big[(1-\beta_t)S+A\Big]+\frac{1}{2}\Big[(1-\beta_t)S'+A\Big]-A=\frac{1}{2}(1-\beta_t)\Big[S+S'\Big]. \tag{25}$$

Let me first consider the case where $a_{t+1}\neq 0$. Then, equation (24) gives the right hand side of equation (22). Define $s=\exp(S)$ and $s'=\exp(S')$ and find that equation (22)

[45] For the necessity see for example Hardy et al. (1964, 26).

yields the following condition:

$$\begin{aligned}
\cancel{(1-\beta_t)}\Big[\frac{1}{2}S+\frac{1}{2}S'\Big] &= \ln\Big[\frac{1}{2}\exp(S)^{a_{t+1}(1-\beta_{t+1})}+\frac{1}{2}\exp(S')^{a_{t+1}(1-\beta_{t+1})}\Big]^{\frac{\cancel{(1-\beta_t)}}{a_{t+1}(1-\beta_{t+1})}}\\
\Leftrightarrow\quad \Big[\frac{1}{2}\ln s+\frac{1}{2}\ln s'\Big] &= \ln\Big[\frac{1}{2}s^{a_{t+1}(1-\beta_{t+1})}+\frac{1}{2}s^{a_{t+1}(1-\beta_{t+1})}\Big]^{\frac{1}{a_{t+1}(1-\beta_{t+1})}}\\
\Leftrightarrow\quad s^{\frac{1}{2}}s'^{\frac{1}{2}} &= \Big[\frac{1}{2}s^{a_{t+1}(1-\beta_{t+1})}+\frac{1}{2}s^{a_{t+1}(1-\beta_{t+1})}\Big]^{\frac{1}{a_{t+1}(1-\beta_{t+1})}}
\end{aligned}$$

for a continuum of s and s'. However, the above equality does not hold for $a_{t+1}(1-\beta_{t+1})\neq 0$ (Hardy et al. 1964, 15,26). As it is $1-\beta_t\neq 0$, this fact implies a contradiction to the assumption that $a_{t+1}\neq 0$. Evaluating equation (22) for period $t-1$ the same reasoning brings about a contradiction to the assumption $a_{t-1}\neq 0$. Therefore, if $a_t=0$ for some t it necessarily follows that $a_t=0$ for all $t\in\{1,...T\}$.

In the case $a_t=0$ for all $t\in\{1,...,T\}$ use (25) to see that equation (22) simplifies to the tautology

$$\begin{aligned}
\frac{1}{2}\Big[(1-\beta_t)S+A\Big]+\frac{1}{2}\Big[(1-\beta_t)S'+A\Big]-A &= \frac{(1-\beta_t)}{(1-\beta_{t+1})}\frac{1}{2}\Big[(1-\beta_{t+1})S+(1-\beta_{t+1})S'\Big]\\
\Leftrightarrow\quad \frac{1}{2}(1-\beta_t)\Big[S+S'\Big] &= \frac{1}{2}(1-\beta_t)\Big[S+S'\Big],
\end{aligned}$$

which implies no further restrictions on the functional form of h_t.

Part III (“$\Rightarrow$”): In this part I show that the solution to equation (22) derived in part two has to hold as well if only continuity of $h_t=f_t\circ g^{-1}$ is assumed.[46] Other than differentiability, the latter is assured by theorem 2. Assume that some continuous function h_t satisfies equation (22). Expecting that the general solution will be of the form derived in part two, I define for all $t\in\{t,...,T\}$ the continuous functions $r_t:\mathbb{R}\to\mathbb{R}$ by $r_t(y)=h_t\left[(1-\beta_t)\ln(y)\right]\;\Leftrightarrow\; h_t(z)=r_t\circ\exp(\frac{1}{1-\beta_t}z)$. Then the left hand side of equation (22) becomes

$$\begin{aligned}
&h_t^{-1}\Big\{\frac{1}{2}h_t\Big[(1-\beta_t)S+A\Big]+\frac{1}{2}h_t\Big[(1-\beta_t)S'+A\Big]\Big\}-A\\
&=(1-\beta_t)\ln\circ r_t^{-1}\Big\{\frac{1}{2}r_t\circ\exp\Big[\frac{1}{1-\beta_t}\{(1-\beta_t)S+A\}\Big]\\
&\qquad\qquad+\frac{1}{2}r_t\circ\exp\Big[\frac{1}{1-\beta_t}\{(1-\beta_t)S'+A\}\Big]\Big\}-A\\
&=(1-\beta_t)\ln\circ r_t^{-1}\Big\{\frac{1}{2}r_t\circ\exp\Big[S+\frac{A}{1-\beta_t}\Big]+\frac{1}{2}r_t\circ\exp\Big[S'+\frac{A}{1-\beta_t}\Big]\Big\}-A
\end{aligned}$$

[46] I.e. there are no further continuous solutions to equation (23).

and with defining $s = \exp[S]$, $s' = \exp[S']$ and $a = \exp\left[\frac{A}{1-\beta_t}\right]$ the relation writes as

$$= (1-\beta_t)\ln\circ r_t^{-1}\Big\{\frac{1}{2}\,r_t\,(s\,a) + \frac{1}{2}\,r_t\,(s'\,a)\Big\} - (1-\beta_t)\ln a$$

$$= (1-\beta_t)\ln\left[\frac{1}{a}\,r_t^{-1}\Big\{\frac{1}{2}\,r_t\,(s\,a) + \frac{1}{2}\,r_t\,(s'\,a)\Big\}\right].$$

Analogously, the right hand side of equation (22) becomes

$$\frac{(1-\beta_t)}{(1-\beta_{t+1})}\cdot(1-\beta_{t+1})\ln\left[r_{t+1}^{-1}\Big\{\frac{1}{2}\,r_{t+1}\,(s) + \frac{1}{2}\,r_{t+1}\,(s')\Big\}\right].$$

Using these expressions equation (22) translates into the requirement

$$\frac{1}{a}\,r_t^{-1}\Big\{\frac{1}{2}\,r_t\,(s\,a) + \frac{1}{2}\,r_t\,(s'\,a)\Big\} = r_{t+1}^{-1}\Big\{\frac{1}{2}\,r_{t+1}\,(s) + \frac{1}{2}\,r_{t+1}\,(s')\Big\}$$

for a continuum of values s, s' and a. First of all, this relation implies that the left hand side has to be constant in a for all values of s and s'. By Hardy et al. (1964, 66,68) it follows that r_t *has to be either an affine transformation of* $r_t(z) = z^{\xi_t}$ *for some* $\xi_t \in \mathbb{R}\backslash\{0\}$ *or an affine transformation of* ln. I will associate the latter case with $\xi = 0$. In the first case equation (22) becomes

$$\frac{1}{\cancel{a}}\Big\{\frac{1}{2}\,(s\cancel{a})^{\xi_t} + \frac{1}{2}\,(s'\cancel{a})^{\xi_t}\Big\}^{\frac{1}{\xi_t}} = \Big\{\frac{1}{2}\,(s)^{\xi_{t+1}} + \frac{1}{2}\,(s')^{\xi_{t+1}}\Big\}^{\frac{1}{\xi_{t+1}}} \tag{26}$$

which implies $\xi_t = \xi_{t+1} \equiv \xi$ for all $t \in \{1, ..., T-1\}$ (Hardy et al. 1964, 26). The case where $r_t = \ln$ corresponds to taking the limit $\xi_t \to 0$ in (26), and the same reasoning on ξ_t and ξ_{t+1} holds true, i.e if some r_t is an affine transformation of ln then all have to be an affine transformation of ln.

In consequence, the following solutions of equation (22) for h_t are possible. In the case $\xi \in \mathbb{R}\backslash\{0\}$ I find for all $t \in \{1, ..., T-1\}$

$$h_t^*(z) = k_t\left(\exp(\frac{1}{1-\beta_t}z)\right)^{\xi} + d_t = k_t\exp(\frac{\xi}{1-\beta_t}z) + d_t\,, \tag{27}$$

with $d_t, k_t \in \mathbb{R}$ and, in order to assure strict increasingness of $h_t^*(z)$, $k_t\,\xi > 0$. In the case $\xi = 0$ I find for all $t \in \{1, ..., T-1\}$

$$h_t^*(z) = \tilde{b}_t\ln\left(\exp(\frac{1}{1-\beta_t}z)\right) + d_t = \tilde{b}_t\frac{1}{1-\beta_t}z + d_t\,,$$

with $\tilde{b}_t, d_t \in \mathbb{R}$ and, in order to assure strict increasingness of $h_t^*(z)$, $\tilde{b}_t\,\xi > 0$. With $\tilde{b}_t = \frac{b_t}{1-\beta_t}$ this solution is seen to correspond to case two in part two. Thus, giving up the differentiability assumption for $f_t g^{-1}$ yields no further solutions to equation (22), than those already found in part two.

Part IV ("⇒"): In part four I substitute the relations found in parts two and three for $h_t = f_t \circ g^{-1}$ back into the representation of $\succeq = (\succeq_t)_{t\in\{1,...,T\}}$ in the sense of theorem 2. I start with the case $f_t \circ g^{-1}(y) = d_t + k_t \exp(\frac{\xi}{1-\beta_t} y)$ with $d_t, k_t \in \mathbb{R}$ and $k_t \, \xi > 0$. Taking g as given, the function f_t follow as

$$f_t \circ g^{-1}(y) = d_t + k_t \exp(\frac{\xi}{1-\beta_t} y)$$

$$\Leftrightarrow f_t(\cdot) = d_t + k_t \exp(\frac{\xi}{1-\beta_t} g(\cdot)) \, .$$

Then the functions $\tilde{u}_t$ in the representation of theorem 2 become

$$\begin{aligned} \tilde{u}_t(x_t, p_{t+1}) &= g^{-1} \Big\{ (1-\beta_t) \, g \circ u(x_t) + \beta_t \, g \circ f_{t+1}^{-1} \Big[\textstyle\int dp_{t+1}^{(x_{t+1}, p_{t+2})} f_{t+1} \circ \tilde{u}_{t+1} \Big] \Big\} \\ &= g^{-1} \Big\{ (1-\beta_t) \, g \circ u(x_t) + \beta_t \, \tfrac{1-\beta_{t+1}}{\xi} \ln \Big[\tfrac{1}{k_{t+1}} \Big\{ \\ &\qquad\qquad -d_{t+1} + \textstyle\int dp_{t+1}^{(x_{t+1}, p_{t+2})} f_{t+1} \circ \tilde{u}_{t+1} \Big\} \Big] \Big\} \, . \end{aligned}$$

Define the functions $\tilde{w}_t = \frac{1}{1-\beta_t} \, g \circ \tilde{u}_t$. Due to the relation between g and f_t, imposed by axiom A8, a recursive formulation employing these strictly monotonic transformation of the functions $\tilde{u}_t$, largely simplifies the representation.

$$\begin{aligned} \tilde{w}_T(x_T) &= gu(x_T) \quad \text{and} \\ \tilde{w}_{t-1}(x_{t-1}, p_t) &= \frac{1}{1-\beta_{t-1}} \, g \circ \tilde{u}_{t-1}(x_{t-1}, p_t) \\ &= gu(x_{t-1}) + \tfrac{\beta_{t-1}}{\xi} \, \tfrac{1-\beta_t}{(1-\beta_{t-1})} \ln \Big[\tfrac{1}{k_t} \Big\{ -d_t + \textstyle\int dp_t^{(x_t, p_{t+1})} f_t \tilde{u}_t \Big\} \Big] \, . \end{aligned}$$

Using the relation $\frac{1-\beta_{t+1}}{1-\beta_t} = \frac{1-\beta^{T-t+1}}{1-\beta^{T-(t+1)+1}} = \beta \beta_t^{-1}$ further yields

$$\begin{aligned} \tilde{w}_{t-1}(x_{t-1}, p_t) &= gu(x_{t-1}) + \tfrac{\beta}{\xi} \ln \Big[\tfrac{1}{k_t} \Big\{ -d_t + \textstyle\int dp_t^{(x_t, p_{t+1})} f_t g^{-1} \, g\tilde{u}_t \Big\} \Big] \\ &= gu(x_{t-1}) + \tfrac{\beta}{\xi} \ln \Big[\tfrac{1}{k_t} \Big\{ -d_t + \textstyle\int dp_t^{(x_t, p_{t+1})} d_t + k_t \exp \Big(\tfrac{\xi}{1-\beta_t} \cdot \\ &\qquad (1-\beta_t) \tilde{w}_t \Big) \Big\} \Big] \\ &= gu(x_{t-1}) + \tfrac{\beta}{\xi} \ln \Big[\textstyle\int dp_t^{(x_t, p_{t+1})} \exp \Big(\xi \tilde{w}_t \Big) \Big] && (28) \\ &= gu(x_{t-1}) + \beta \mathcal{M}^{\exp^\xi}(p_t, \tilde{w}_t) \, , && (29) \end{aligned}$$

where the uncertainty aggregation rule is characterized by the function $r(z) = \exp(\xi z) = \exp(z)^\xi$. Expression (29) will be used for the $g^+ - gauge$ of the representation in corollary 1, where the range of g has been fixed. Here however, the parameter ξ can be absorbed into the function g. To this end, define $\tilde{w}_t^* = |\xi| \, \tilde{w}_t$, $g^* = |\xi| \, g$ and $\text{sgn}(\xi)$ as the sign of ξ.

Then line (28) yields

$$\begin{aligned}\tilde{w}^*_{t-1}(x_{t-1},p_t) &= |\xi|\, gu(x_{t-1}) + \tfrac{|\xi|}{\xi}\beta \ln\left[\int dp_t^{(x_t,p_{t+1})} \exp\left(\xi \tilde{w}_t\right)\right]\\ &= g^* u(x_{t-1}) + \operatorname{sgn}(\xi)\ \beta \ln \left[\int dp_t^{(x_t,p_{t+1})} \exp\left(\ \operatorname{sgn}(\xi)\ \tilde{w}^*_t\right)\right]\\ &= g^* u(x_{t-1}) + \beta \mathcal{M}^{\exp^{\operatorname{sgn}(\xi)}}(p_t, \tilde{w}^*_t)\,. \end{aligned} \tag{30}$$

Expression (30) yields equation (11) for the cases $f \in \{\exp, \frac{1}{\exp}\}$. To obtain the representing equation (12) first observe that

$$\begin{aligned}\mathcal{M}^{f_t}(p_t,\tilde{u}_t) &= f_t^{-1}\Big[\int dp_t^{(x_t,p_{t+1})} f_t \circ \tilde{u}_t\Big]\\ &= f_t^{-1}\Big[\int dp_t^{(x_t,p_{t+1})} f_t g^{-1}\Big((1-\beta_t)\,\tilde{w}_t\Big)\Big]\\ &= f_t^{-1}\Big[\int dp_t^{(x_t,p_{t+1})} d_t + k_t \exp\left(\tfrac{\xi}{1-\beta_t}(1-\beta_t)\,\tilde{w}_t\right)\Big]\\ &= f_t^{-1}\Big[\int dp_t^{(x_t,p_{t+1})} d_t + k_t \exp\left(\ \operatorname{sgn}(\xi)\ \tilde{w}^*_t\right)\Big]\,.\end{aligned}$$

Then, recalling that $\operatorname{sgn}(k_t) = \operatorname{sgn}(\xi)$, find that the strictly increasing transformation

$$\begin{aligned}\mathcal{M}^{f}(p_t,\tilde{w}^*_t) &= \ln\left[\ \operatorname{sgn}(\xi) \int dp_t^{(x_t,p_{t+1})} \exp\left(\ \operatorname{sgn}(\xi)\ \tilde{w}^*_t\right)\right]\\ &= \mathcal{M}^{\exp^{\operatorname{sgn}(\xi)}}\left(p_t,\tilde{w}^*_t\right)\,.\end{aligned}$$

yields the expression representing the preferences in equation (12).

In the remaining case it is $f_t \circ g^{-1}(y) = b_t\, z + k_t$ with $b_t, k_t \in \mathbb{R}$ and $b_t > 0$. Taking g as and, thus, $f_t = b_t\, z + k_t$ an analogous reasoning to the one carried out above yields

$$\begin{aligned}\tilde{u}_t(x_t,p_{t+1}) &= g^{-1}\left\{(1-\beta_t)\, g\circ u(x_t) + \beta_t\, g \circ f_{t+1}^{-1}\left[\int dp_{t+1}^{(x_{t+1},p_{t+2})} f_{t+1}\circ \tilde{u}_{t+1}\right]\right\}\\ &= g^{-1}\left\{(1-\beta_t) g\circ u(x_t) + \beta_t \left[\tfrac{1}{b_{t+1}}\left\{-k_{t+1} + \int dp_{t+1}^{(x_{t+1},p_{t+2})} f_{t+1}\circ\tilde{u}_{t+1}\right\}\right]\right\}.\end{aligned}$$

And defining the functions

$$\tilde{w}_T(x_T) = gu(x_T) \quad \text{and}$$

$$\begin{aligned}\tilde{w}_{t-1}(x_{t-1},p_t) &= \frac{1}{1-\beta_{t-1}}\, g\circ \tilde{u}_{t-1}(x_{t-1},p_t)\\ &= gu(x_{t-1}) + \tfrac{\beta_{t-1}}{(1-\beta_{t-1})}\left[\tfrac{1}{b_t}\left\{-k_t + \int dp_t^{(x_t,p_{t+1})} f_t \tilde{u}_t\right\}\right]\\ &= gu(x_{t-1}) + \tfrac{\beta_{t-1}}{(1-\beta_{t-1})}\left[\tfrac{1}{b_t}\left\{-k_t + \int dp_t^{(x_t,p_{t+1})} b_t(1-\beta_t)\tilde{w}_t + k_t\right\}\right]\\ &= gu(x_{t-1}) + \beta\left[\int dp_t^{(x_t,p_{t+1})} \tilde{w}_t\right],\end{aligned}$$

where the latter expression corresponds to the recursion (11) stated in the theorem for the cases $f = \mathrm{id}$. The representing equation (12) follows from

$$\begin{aligned}\mathcal{M}^{f_t}(p_t, \tilde{u}_t) &= f_t^{-1}\Big[\int dp_t^{(x_t,p_{t+1})} f_t \circ \tilde{u}_t\Big]\\ &= f_t^{-1}\Big[\int dp_t^{(x_t,p_{t+1})} f_t g^{-1}\Big((1-\beta_t)\,\tilde{w}_t\Big)\Big]\\ &= f_t^{-1}\Big[\int dp_t^{(x_t,p_{t+1})} k_t + b_t(1-\beta_t)\,\tilde{w}_t\Big]\end{aligned}$$

which is a strictly increasing transformation of

$$\mathcal{M}^{f}(p_t, \tilde{w}_t) = \mathrm{E}_{p_t}\tilde{w}_t = \mathcal{M}^{\mathrm{id}}(p_t, \tilde{w}_t)\,.$$

Part V ("⇐"): As shown above, the representation is a special case of theorem 2. Therefore axioms A1-A5 follow immediately from "⇐" of theorem 2. The following calculation shows that axiom A8 is satisfied as well. Hereto note that for certain consumption paths $\mathsf{x} \in \mathsf{X}^t$ it is $\tilde{w}_t(\mathsf{x}) = \sum_{\tau=t}^{T}\beta^{\tau-t} g\circ u(\mathsf{x}_\tau)$. For the case $h = \exp$ define $k = 1$ and for the case $h = \frac{1}{\exp}$ define $k = -1$. Then, for $h \in \{\exp, \frac{1}{\exp}\}$ and for all $t \in \{1, ..., T-1\}$, $x^0 \in X$ and $\mathsf{x}, \mathsf{x}', \mathsf{x}'' \in \mathsf{X}^{t+1}$ it holds

$$\begin{aligned}&\frac{1}{2}(\mathsf{x}, x^0) + \frac{1}{2}(\mathsf{x}', x^0) \;\succeq_t\; (\mathsf{x}'', x^0)\\ \Leftrightarrow\; & k\,\ln\Bigg(\tfrac{1}{2}\exp\Big[k\textstyle\sum_{\tau=t}^{T-1}\beta^{\tau-t} g\circ u(\mathsf{x}_{\tau+1})\Big]\exp\Big[k\,\beta^T g\circ u(x^0)\Big]\\ &\quad +\tfrac{1}{2}\exp\Big[k\textstyle\sum_{\tau=t}^{T-1}\beta^{\tau-t} g\circ u(\mathsf{x}'_{\tau+1})\Big]\exp\Big[k\,\beta^T g\circ u(x^0)\Big]\Bigg)\\ &\qquad\qquad \geq \textstyle\sum_{\tau=t}^{T-1}\beta^{\tau-t} g\circ u(\mathsf{x}''_{\tau+1}) + \beta^T g\circ u(x^0)\\ \Leftrightarrow\; & k\,\ln\Bigg(\tfrac{1}{2}\exp\Big[k\textstyle\sum_{\tau=t+1}^{T}\beta^{\tau-(t+1)} g\circ u(\mathsf{x}_{\tau})\Big] + \tfrac{1}{2}\exp\Big[k\textstyle\sum_{\tau=t+1}^{T}\beta^{\tau-(t+1)} g\circ u(\mathsf{x}'_{\tau})\Big]\Bigg)\\ &\qquad\qquad \geq \textstyle\sum_{\tau=t+1}^{T}\beta^{\tau-(t+1)} g\circ u(\mathsf{x}''_{\tau})\\ \Leftrightarrow\; & \frac{1}{2}\tilde{v}_{t+1}(\mathsf{x}) + \frac{1}{2}\tilde{v}_{t+1}(\mathsf{x}') \geq \tilde{v}_{t+1}(\mathsf{x}'')\\ \Leftrightarrow\; & \frac{1}{2}\mathsf{x} + \frac{1}{2}\mathsf{x}' \;\succeq_{t+1}\; \mathsf{x}''\,.\end{aligned}$$

The case $h = \mathrm{id}$ makes both sides of the above inequalities linear in the term $\beta^T g\circ u(x^0)$, so that it cancels as well and A8 is satisfied.

Moreover part: "⇒": Assume that g and g' both represent the sequence of preference relations $\succeq = (\succeq_t)_{t\in\{1,\ldots,T\}}$ (the prime in g' does *not* indicate a derivative!). By the representation of $\succeq = (\succeq_t)_{t\in\{1,\ldots,T\}}$ on certain paths, the freedom of g is limited to positive affine transformations as in theorem 2, i.e. it have to exist $a, b \in \mathbb{R}, a > 0$ such that $g = ag' + b$. However, the dependence of f_t on g destroys part of this freedom when considering choice over lotteries.[47] Precisely, find that the function $\tilde{w}'_T$ corresponding to the choice g' is

$$\tilde{w}'_T(x_T) = g' \circ u(x_T) = a\, g \circ u(x_T) + b$$

Define again $k = 1$ for the case $h = \exp$ and $k = -1$ for the case $h = \frac{1}{\exp}$. Then, for the case $h \in \{\exp, \frac{1}{\exp}\}$, the fact that $\tilde{w}_T(x_T)$ as well as $\tilde{w}'_T(x_T)$ are to represent the same preferences over period T lotteries implies

$$\begin{aligned}
& k \int dp_T \exp(k\, \tilde{w}_T) && \geq\ k \int dp_T \exp(k\, \tilde{w}_T) \\
\Leftrightarrow\ & k \ln\left[\int dp_T \exp(k\, \tilde{w}_T)\right] && \geq\ k \ln\left[\int dp_T \exp(k\, \tilde{w}_T)\right] \\
\Leftrightarrow\ & \mathcal{M}^h(p_T, \tilde{w}_T) && \geq\ \mathcal{M}^h(p'_T, \tilde{w}_T) \\
\Leftrightarrow\ & p_T && \succeq_t\ p'_T \\
\Leftrightarrow\ & \mathcal{M}^h(p_T, \tilde{w}'_T) && \geq\ \mathcal{M}^h(p'_T, \tilde{w}'_T) \\
\Leftrightarrow\ & k \ln\left[\int dp_T \exp(k\, \tilde{w}'_T)\right] && \geq\ k \ln\left[\int dp_T \exp(k\, \tilde{w}'_T)\right] \\
\Leftrightarrow\ & k \int dp_T \exp(k\, \tilde{w}'_T) && \geq\ k \int dp_T \exp(k\, \tilde{w}'_T)
\end{aligned}$$

for all $p_t, p'_t \in P_t$. In consequence there have to exist constants $c, d \in \mathrm{I\!R}, c > 0$ such that

$$\begin{aligned}
\exp(k\, \tilde{w}_T) &= c \exp(k\, \tilde{w}'_T) + d \\
&= c \exp(k\, a\, \tilde{w}'_T + kb) + d \\
&= c \exp(kb) \exp(k\, \tilde{w}'_T)^a + d\,.
\end{aligned}$$

Thus, defining the constant $\tilde{c} = c \exp(kb)$ and the variable $z = \exp(k\, \tilde{w}'_T(x_t))$ the relation

$$z = \tilde{c} z^a + d$$

has to hold for all $z \in [\exp(\underline{G}), \exp(\overline{G})]$. The relation can only be satisfied if the right hand side is linear and, thus, $a = 1$. In consequence, if g and g' both represent the preferences $\succeq = (\succeq_t)_{t\in\{1,...,T\}}$, it has to exist $b \in \mathrm{I\!R}$ such that $g = g' + b$.

For the case $h = \text{id}$, corresponding to a maximizer of intertemporally additive expected utility, the above reasoning yields no further restrictions on the constants a or b. In that case, if g and g' both represent the preferences $\succeq = (\succeq_t)_{t\in\{1,...,T\}}$ all that can be claimed is the existence $a, b \in \mathrm{I\!R}, a > 0$, such that $g = ag' + b$.

"$\Leftarrow$": For the case $h \in \{\exp, \frac{1}{\exp}\}$, let $g = g' + b$ and g be part of a representation of $\succeq = (\succeq_t)_{t\in\{1,...,T\}}$. Define as before $k = 1$ for the case $h = \exp$ and $k = -1$ for the case $h = \frac{1}{\exp}$. I claim that for every $t \in \{1, ..., T\}$ it exists $\gamma_t \in \mathrm{I\!R}$ such that $\tilde{w}'_t = \tilde{w}_t + \gamma_t$. The

[47]Without the dependence of f on g an affine transformation $\mathfrak{a}$ of g cancels out. However, when f depends on g as in the representation of theorem 4, at the same time $f^{-1} \to \mathfrak{a} f^{-1}$, corresponding to an affine transformation of the inverse of f. Such a transformation is, in general, not compatible with the freedom in the choice of the representing functions.

proof is by backwards induction. For $t = T$ it holds

$$\tilde{w}'_T(x_T) = g' \circ u(x_T) = g \circ u(x_T) + b = \tilde{w}_T(x_T) + \gamma_T$$

with $\gamma_T = b$. The induction step from t to $t-1$ works as follows:

$$\begin{aligned}
\tilde{w}'_{t-1}(x_{t-1}, p_t) &= g' u(x_{t-1}) + \beta \mathcal{M}^{\exp^k}(p_t, \tilde{w}'_t) \\
&= gu(x_{t-1}) + b + k\, \beta \ln \left[\int dp_t^{(x_t, p_{t+1})} \exp\left(k\tilde{w}_t + k\gamma_t \right) \right] \\
&= gu(x_{t-1}) + b + k\, \beta \ln \left[\exp(k\gamma_t) \int dp_t^{(x_t, p_{t+1})} \exp\left(k\gamma_t \tilde{w}_t \right) \right] \\
&= gu(x_{t-1}) + k\, \beta \ln \left[\int dp_t^{(x_t, p_{t+1})} \exp\left(k\gamma_t \tilde{w}_t \right) \right] + b + \beta\gamma_t \\
&= \tilde{w}_{t-1}(x_{t-1}, p_t) + \gamma_{t-1}, + \beta\gamma_t
\end{aligned}$$

with $\gamma_{t-1} = b + \beta\gamma_t$. Next I show, that such an additive constant in $\tilde{w}_t$ cancels out in the representing equation (12):

$$\begin{aligned}
& \mathcal{M}^h(p_t, \tilde{w}'_t) && \geq \mathcal{M}^h(p'_t, \tilde{w}'_t) \\
\Leftrightarrow \quad & \mathcal{M}^h(p_t, \tilde{w}_t + \gamma_t) && \geq \mathcal{M}^h(p'_t, \tilde{w}_t + \gamma_t) \\
\Leftrightarrow \quad & k \ln \left[\int dp_t \exp(k\, \tilde{w}_t + \gamma_t) \right] && \geq k \ln \left[\int dp_t \exp(k\, \tilde{w}_t + \gamma_t) \right] \\
\Leftrightarrow \quad & k \ln \left[\int dp_t \exp(k\, \tilde{w}_t) \right] && \geq k \ln \left[\int dp_t \exp(k\, \tilde{w}_t) \right] \\
\Leftrightarrow \quad & \mathcal{M}^h(p_t, \tilde{w}_t) && \geq \mathcal{M}^h(p'_t, \tilde{w}_t) \,.
\end{aligned}$$

Thus, if g represents preferences $\succeq = (\succeq_t)_{t \in \{1,\dots,T\}}$ in the sense of theorem 4 with $h \in \{\exp, \frac{1}{\exp}\}$, then so does $g' = g + b$.

In the case $h = \text{id}$, let $g = ag' + b$ and g be part of a representation of $\succeq = (\succeq_t)_{t \in \{1,\dots,T\}}$ in the sense of theorem 4. I claim that for every $t \in \{1, ..., T\}$ it exists $\gamma_t \in \mathbb{R}$ such that $\tilde{w}'_t = a\tilde{w}_t + \gamma_t$. Proof is by backwards induction. For $t = T$ it holds

$$\tilde{w}'_T(x_T) = g' \circ u(x_T) = ag \circ u(x_T) + b = a\tilde{v}_T(x_T) + \gamma_T \,,$$

with $\gamma_T = b$. The induction step from t to $t-1$ is as follows:

$$\begin{aligned}
\tilde{w}'_{t-1}(x_{t-1}, p_t) &= g' \circ u(x_{t-1}) + \beta E_{p_t} \tilde{w}'_t(x_t, p_{t+1}) \\
&= a\, g \circ u(x_{t-1}) + b + \beta E_{p_t}\, a\, \tilde{w}_t(x_t, p_{t+1}) + \beta\, \gamma_t \\
&= a\, \tilde{w}_{t-1}(x_{t-1}, p_t) + b + \beta\, \gamma_t \,.
\end{aligned}$$

Setting $\gamma_{t-1} = b + \gamma_t$ closes the induction step. But then, the representation in equation (12) stays unchanged:

$$E_{p_t}\, \tilde{w}'_t \;\geq\; E_{p'_t}\, \tilde{w}'_t \quad \Leftrightarrow \quad E_{p_t}\, a\, \tilde{w}_t + \gamma_t \;\geq\; E_{p'_t}\, a\, \tilde{w}_t + \gamma_t \quad \Leftrightarrow \quad E_{p_t}\, \tilde{w}_t \geq E_{p'_t}\, \tilde{w}_t \,.$$

□

C Proofs for Section 4

Proof of theorem 5: The proof resembles that of theorem **??**. Part one translates axiom A9$^{\text{s}}$ into the representation of theorem 2. Then I show in the second part that the equation derived in the first locally implies concavity of $f_t \circ g^{-1}$. Part three extends this result to concavity on the entire set Γ_t. The necessity of axiom A9$^{\text{s}}$ is implied by theorem **??**. The difference to the proof of theorem **??**, i.e. the stronger prerequisite in axiom A9$^{\text{s}}$, mainly affects the first step in part two. Subsequently the proof follows that of theorem **??** and the reader is referred to the latter.

Part I ("⇒"): In this part I translate axiom A9$^{\text{s}}$ into the representation of theorem 2. I start with the first line, i.e the premise, and use equation (19) to find

$$\begin{aligned} \bar{\mathrm{x}}^t \quad & \sim_t \quad \mathrm{x}^t \\ \Rightarrow g^{-1}\Big[(1-\beta_t)\sum_{\tau=t}^{T} \beta^{\tau-t} u(\bar{x})\Big] \; &= \; g^{-1}\Big[(1-\beta_t)\sum_{\tau=t}^{T} \beta^{\tau-t} u(\mathrm{x}^t_\tau)\Big]\,. \end{aligned} \tag{31}$$

The existence of $\tau \in \{t, ..., T\}$ such that $[\mathrm{x}^t_\tau] \not\sim_\tau [\bar{x}]$ translates into

$$u(\mathrm{x}^t_\tau) \neq u(\bar{x}) \text{ for some } \tau \in \{t, ..., T\}. \tag{32}$$

The second line of axiom A9$^{\text{s}}$ becomes

$$\begin{aligned} \bar{\mathrm{x}}^t \quad & \succ_T \quad \textstyle\sum_{i=t}^{T} \frac{1}{T-t+1} \, (\bar{\mathrm{x}}^t_{-i}, \mathrm{x}^t_i). \\ \Rightarrow \; & g^{-1}\Big[(1-\beta_t)\textstyle\sum_{\tau=t}^{T} \beta^{\tau-t} u(\bar{x})\Big] \\ & > f_t^{-1}\Big[\textstyle\sum_{i=t}^{T} \frac{1}{T-t+1} f_t g^{-1}\Big[(1-\beta_t)\sum_{\tau=t}^{T} \beta^{\tau-t} u\left((\bar{\mathrm{x}}^t_{-i}, \mathrm{x}^t_i)_\tau\right)\Big]\Big] \\ \Rightarrow \; & f_t g^{-1}\Big[(1-\beta_t)\textstyle\sum_{\tau=t}^{T} \beta^{\tau-t} u(\bar{x})\Big] \\ & > \quad \textstyle\sum_{i=t}^{T} \frac{1}{T-t+1} f_t g^{-1}\Big[(1-\beta_t)\sum_{\tau=t}^{T} \beta^{\tau-t} u\left((\bar{\mathrm{x}}^t_{-i}, \mathrm{x}^t_i)_\tau\right)\Big]. \end{aligned}$$

Using equation (31) the left hand side can be transformed as follows:

$$f_t g^{-1}\left[\tfrac{T-t}{T-t+1}\left[(1-\beta_t)\sum_{\tau=t}^{T}\beta^{\tau-t}u(\bar{x})\right]+\tfrac{1}{T-t+1}\left[(1-\beta_t)\sum_{\tau=t}^{T}\beta^{\tau-t}u(\mathsf{x}^t_\tau)\right]\right]$$

$$> \quad \sum_{i=t}^{T}\tfrac{1}{T-t+1}f_t g^{-1}\Big[(1-\beta_t)\sum_{\tau=t}^{T}\beta^{\tau-t}u\left((\bar{\mathsf{x}}^t_{-i},\mathsf{x}^t_i)_\tau\right)\Big]$$

$$\Rightarrow f_t g^{-1}\left[\tfrac{1}{T-t+1}\left[(1-\beta_t)\sum_{i=t}^{T}\sum_{\tau=t}^{T}\beta^{\tau-t}u\left((\bar{\mathsf{x}}^t_{-i},\mathsf{x}^t_i)_\tau\right)\right]\right]$$

$$> \quad \sum_{i=t}^{T}\tfrac{1}{T-t+1}f_t g^{-1}\Big[(1-\beta_t)\sum_{\tau=t}^{T}\beta^{\tau-t}u\left((\bar{\mathsf{x}}^t_{-i},\mathsf{x}^t_i)_\tau\right)\Big]$$

$$\Rightarrow f_t g^{-1}\left[\sum_{i=t}^{T}\tfrac{1}{T-t+1}\left[(1-\beta_t)\sum_{\tau=t}^{T}\beta^{\tau-t}u\left((\bar{\mathsf{x}}^t_{-i},\mathsf{x}^t_i)_\tau\right)\right]\right] \tag{33}$$

$$> \quad \sum_{i=t}^{T}\tfrac{1}{T-t+1}f_t g^{-1}\Big[(1-\beta_t)\sum_{\tau=t}^{T}\beta^{\tau-t}u\left((\bar{\mathsf{x}}^t_{-i},\mathsf{x}^t_i)_\tau\right)\Big].$$

Define the function $\tilde{z}:\mathsf{X}^t\to\Gamma_t$ by $\tilde{z}(\mathsf{x}^t)=(1-\beta_t)\sum_{\tau=t}^{T}\beta^{\tau-t}u\left(\mathsf{x}^t_\tau\right)$. Restricting the domain to those consumption paths that satisfy condition (32) the function is onto $\Big((1-\beta_t)\sum_{\tau=t}^{T}\underline{G}\,,\,(1-\beta_t)\sum_{\tau=t}^{T}\overline{G}\Big)=\left(\underline{G}\,,\,\overline{G}\right)=\Gamma_t$. In particular define $z_i=\tilde{z}\big((\mathsf{x}^t_{-i}\mathsf{x}'^t_i)\big)$. In this notation equation (33) becomes

$$f_t g^{-1}\left(\sum_{i=t}^{T}\tfrac{1}{T-t+1}\,z_i\right) \quad > \quad \sum_{i=t}^{T}\tfrac{1}{T-t+1}f_t g^{-1}(z_i). \tag{34}$$

If equation (34) had to hold for all $z_i\in\Gamma_t$ it would be a straight forward condition for strict convexity of $f_t\circ g^{-1}$. However axiom A9$^{\mathrm{s}}$ does not immediately imply that the equation has to be met for every choice $(z_i)_{i\in\{t,\dots,T\}}$, $z_i\in\Gamma_t$. Only for combination $(z_i)_{i\in\{t,\dots,T\}}$ stemming from consumption paths $(\bar{\mathsf{x}}^t_{-i},\mathsf{x}^t_i)$ for which $\mathsf{x}^t\in\mathsf{X}^t$ and $\bar{x}\in X$ satisfy the premise of axiom A9$^{\mathrm{s}}$. In what follows I proceed to show that this restricted demand is enough to imply strict convexity of of $f_t\circ g^{-1}$ on Γ_t.

Part II ("⇒"): Let $z^o\in\Gamma_t$. In this part I show that for every such z^o there exists an open neighborhood $N_{z^o}\subset\Gamma_t$ such that equation (34) implies strict concavity of $f_t\circ g^{-1}$ on N_{z^o}.

In the first step I define a certain consumption path $\bar{\mathsf{x}}^t$ with $\bar{x}\in X$ such that $\tilde{z}(\bar{\mathsf{x}}^t)=z^o$. The fact $z^o\in\Gamma_t$ is equivalent to $\underline{G}<z^o<\overline{G}$. By connectedness of X and continuity of $g\circ u$ there exists an outcome $x^o\in u^{-1}\left[g^{-1}\left(z^o\right)\right]$ such that $z^o=u\circ g(x^o)$. Define $\mathsf{x}^{ot}=\bar{\mathsf{x}}^{ot}=(x^o,\dots,x^o)$ and find that $\tilde{z}(\mathsf{x}^{ot})=z^0$. Note that the difference between the stationary and the non-stationary setting is that only in the stationary setting it is guaranteed that any $z^o\in\Gamma_t$ can be attained by evaluating a constant consumption path.

From step two on the proof (including **Part III**) follows exactly the one laid out for theorem **??** on page **??** with $G^o_\tau=z^o$ for all $\tau\in\{t,\dots,T\}$ and $\epsilon=\min\{z^o-\underline{G},\overline{G}-z^o\}$.

Part IV ("⇐"): "⇐" is implied by theorem **??** for $\mathsf{x}^t=\bar{\mathsf{x}}^t$. □

Proof of lemma 2: The lemma is an immediate consequence of lemma **??** with the convention $g_1 \circ u_1 = g \circ u$. Then, the representing triples $(u, f_t, g)_{t\in\{1,\ldots,T\}}$ in the sense of theorems 2 correspond to the representing triples $(u_t = u, f_t, g_t = \beta^{t-1} g)_{t\in\{1,\ldots,T\}}$ in the sense of theorem 1. Therefore, imposing the unit, the zero level or the range of $g \circ u$ determines the according values for $g_t \circ u_t$ in the sense of theorem 1 for all periods. Thus, the statements in a), b), c) and d) in lemma **??** imply the assertions a), b), c) and d) in lemma 2. As theorems 3 and 4 are special cases of theorem 2, the reasoning holds true as well for representations in the sense of theorems 3 and 4. □

Proof of corollary 1: To the most part, the g^+-gauge of the representation in theorem 4 has already been derived in the proof of the latter theorem.

"⇒": Before absorbing the parameter ξ into the function g in the proof of theorem 4, the recursive construction of $\tilde{w}_t$ for the case corresponding to $h \in \left\{\exp, \frac{1}{\exp}\right\}$ was given by equation (29), which states

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

Simply defining the new utility function $u^* = g \circ u$ yields the $g = \text{id}$-gauge. Once the range of u^*, i.e. g, is fixed, a transformation absorbing the free parameter ξ into the function g, i.e. u^*, as carried out to arrive at the final representation stated in theorem 4, is no longer possible.

The representing equation (16) is obtained as follows. The representation that is known to hold by theorem 1 for the specifications of theorem 4 is

$$\begin{aligned}
\mathcal{M}^{f_t}(p_t, \tilde{u}_t) &= f_t^{-1}\Big[\int dp_t^{(x_t, p_{t+1})} f_t \circ \tilde{u}_t\Big] \\
&= f_t^{-1}\Big[\int dp_t^{(x_t, p_{t+1})} f_t g^{-1}\Big((1-\beta_t)\,\tilde{w}_t\Big)\Big] \\
&= f_t^{-1}\Big[\int dp_t^{(x_t, p_{t+1})} d_t + k_t \exp\Big(\xi\,\tilde{w}_t\Big)\Big]\,.
\end{aligned}$$

But, recalling that $k_t\,\xi > 0$, the latter expression is easily recognized as a strictly increasing transformation of

$$\mathcal{M}^{\exp^\xi}(p_t, \tilde{w}_t) = \frac{1}{\xi} \ln\Big[\int dp_t^{(x_t, p_{t+1})} \exp\Big(\xi\,\tilde{w}_t\Big)\Big]\,.$$

Therefore, also $\mathcal{M}^{\exp^\xi}(p_t, \tilde{w}_t)$ represents the preferences $\succeq = (\succeq_t)_{t\in\{1,\ldots,T\}}$.

For the case corresponding to $h = \text{id}$ in the representation of theorem 4, the proof of the latter theorem has derived the following representation

$$\tilde{w}_T(x_T) = gu(x_T) \quad \text{and}$$

$$\tilde{w}_{t-1}(x_{t-1}, p_t) = gu(x_{t-1}) + \beta \mathrm{E}_{p_t} \tilde{w}_t\,,$$

with

$$p_t \succeq_t p'_t \quad \Leftrightarrow \quad \mathrm{E}_{p_t}\tilde{w}_t \geq \mathrm{E}_{p'_t}\tilde{w}_t \, .$$

Thus, with the definition $\mathcal{M}^{\exp^0}(p_t, \tilde{w}_t) = E_{p_t}\tilde{w}_t$, the claimed representation also holds for $\xi = 0$.

Finally, observe that, as stated in the text, the above definition of $\mathcal{M}^{\exp^0}$ corresponds to the limit $\xi \to 0$. To see this, simply apply l'Hospital's rule:

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

"$\Leftarrow$": Implied by theorem 4.

Moreover part: By lemma 2 the function $g \circ u$ in theorem 4 is uniquely determined, once its range has been fixed. As seen above, the representing utility function in the corollary corresponds to the function $u^* = g \circ u$. Thus, fixing its range determines the function uniquely. Moreover lemma 2 implies that the measures of intertemporal risk aversion are determined uniquely.

Equation (14) defines the measure of absolute intertemporal risk aversion in period t as the function

$$\mathrm{AIRA}_t(z) = -\frac{\left(f_t \circ g_t^{-1}\right)''(z)}{\left(f_t \circ g_t^{-1}\right)'(z)} \, .$$

As derived in the proof of theorem 4, the case $\xi \neq 0$ corresponds to $f_t \circ g^{-1} = k_t \exp(\frac{\xi}{1-\beta_t}z) + d_t$, with $d_t, k_t \in \mathbb{R}$ and $k_t\,\xi > 0$ (compare 27). Then, with $g_1 = g$ and $g_t = \beta^{t-1}g$, the measure of absolute intertemporal risk aversion is calculated to

$$\text{AIRA}_t(z) = -\frac{\frac{d^2}{dz^2} f_t \circ g^{-1}(\beta^{-t+1} z)}{\frac{d}{dz} f_t \circ g^{-1}(\beta^{-t+1} z)} = -\frac{\frac{d^2}{dz^2} k_t \exp\left(\frac{\xi}{1-\beta_t}\beta^{-t+1} z\right) + d_t}{\frac{d}{dz} k_t \exp\left(\frac{\xi}{1-\beta_t}\beta^{-t+1} z\right) + d_t}$$

$$= -\frac{\left(\frac{\xi}{1-\beta_t}\beta^{-t+1}\right)^2 \exp\left(\frac{\xi}{1-\beta_t}\beta^{-t+1} z\right)}{\frac{\xi}{1-\beta_t}\beta^{-t+1} \exp\left(\frac{\xi}{1-\beta_t}\beta^{-t+1} z\right)} = -\frac{\xi}{\beta^{t-1}(1-\beta_t)} \,,$$

yielding the constant coefficient of absolute intertemporal risk aversion $\frac{-\xi}{\beta^{t-1}(1-\beta_t)}$. In the case $\xi = 0$ it as

$$\text{AIRA}_t(z) = -\frac{\frac{d^2}{dz^2} f_t \circ g^{-1}(\beta^{-t+1} z)}{\frac{d}{dz} f_t \circ g^{-1}(\beta^{-t+1} z)} = -\frac{\frac{d^2}{dz^2}\, b_t\, z + k_t}{\frac{d}{dz}\, b_t\, z + k_t} = 0 \,,$$

coinciding with the general expression $\text{AIRA}_t(z) = -\frac{\xi}{\beta^{t-1}(1-\beta_t)}$.

The measure of relative intertemporal risk aversion in period t is defined in equation (13) as the function

$$\text{RIRA}_t(z) = -\frac{\left(f_t \circ g_t^{-1}\right)''(z)}{\left(f_t \circ g_t^{-1}\right)'(z)}\, z \,.$$

In consequence it holds $\text{RIRA}_t(z) = \text{AIRA}_t(z) \cdot z$, yielding $\text{RIRA}_t = -\frac{\xi}{\beta^{t-1}(1-\beta_t)}\,\text{id}$. □

Proof of corollary 2: The proof is divided into two parts. The first part derives a representation triple in the sense of theorem 1, in which the functions f_t correspond to the identity, and which satisfies the requirements of corollary 2. The second part works out the corresponding representation as stated in the corollary. The necessity of the axioms is immediate by theorem 4.

Part I: First, observe that corollary 1 with Bernoulli utility u^* implies, with the definition $u = \exp(u^*) \Leftrightarrow u^* = \ln u$, the representation for the case $\xi = 0$ ($h = \text{id}$ in theorem 4). The logarithm is introduced because the representation for the case $\xi \neq 0$ fixes the measure scale for welfare to $\ln u^*$, as it will be observed in the remark at the end of this part of the proof. In the following, I work out the proof for the case where

$$h_t(z) = f_t \circ g^{-1}(z) = k_t \exp\left(\frac{\xi}{1-\beta_t} z\right) + d_t \,,$$

with $d_t, k_t \in \mathbb{R}$ and $k_t\, \xi > 0$, corresponding to equation (27) and case two of the proof of theorem 4. As I want to gauge the functions f_t to identity, I have to allow the functions g_t to vary over time. Therefore, I express the preferences $\succeq = (\succeq_t)_{t\in\{1,\ldots,T\}}$ in a representation in the sense of theorem 1. Recall, that a certainty stationary representation, as the one above, corresponds to a representation $(u, f_t, \beta^{t-1} g)$ in the sense of theorem 1. I take the functions f_t as given. Then, the requirement (27) for $f_t g^{-1}$ restated above implies

$$g_t = \beta^{t-1} g = \beta^{t-1} \frac{1-\beta_t}{\xi} \ln\left(\frac{1}{k_t}(f_t - d_t)\right) .$$

In consequence, the sequence of triples

$$\left(u \,,\, f_t \,,\, g_t = \beta^{t-1} \frac{1-\beta_t}{\xi} \ln\left(\frac{1}{k_t}(f_t - d_t)\right)\right)_{t\in\{1,\ldots,T\}}$$

represents the preferences described in theorem 4, in the sense of the non-stationary representation theorem 1. By gauge lemma **??** it is known that the same preferences are represented by the sequence of triples

$$\left(u'_t = f_t \circ u \,,\, f'_t = f_t \circ f_t^{-1} \,,\, g'_t = \beta^{t-1} \frac{1-\beta_t}{\xi} \ln\left(\frac{1}{k_t}(f_t \circ f_t^{-1} - d_t)\right)\right)_{t\in\{1,\ldots,T\}}$$

$$= \left(u'_t = f_t \circ u \,,\, f'_t = \mathrm{id} \,,\, g'_t = \beta^{t-1} \frac{1-\beta_t}{\xi} \ln\left(\frac{1}{k_t}(\mathrm{id} - d_t)\right)\right)_{t\in\{1,\ldots,T\}} . \tag{35}$$

As desired, uncertainty aggregation corresponding to the above representation is linear. However, observe that

$$u'_t = f_t \circ u = k_t \exp(\frac{\xi}{1-\beta_t} g \circ u) + d_t \,.$$

The relation implies that it is impossible to fix u'_t to a given range independent of ξ.[48] Therefore, define the functions

$$u^*_t = \left(\frac{1}{k_t}(u'_t - d_t)\right)^{\frac{1-\beta_t}{\xi}} = \exp(\frac{\xi}{1-\beta_t} g \circ u)^{\frac{1-\beta_t}{\xi}} = \exp(g \circ u) \tag{36}$$

Then $u^* = u^*_t$ is independent of ξ and moreover constant in time. Note also, that u^*_t is always positive. Using this definition, the representing triples (35) write as

$$\left(k_t {u^*}^{\frac{\xi}{1-\beta_t}} + d_t \,,\, \mathrm{id} \,,\, \beta^{t-1} \frac{1-\beta_t}{\xi} \ln\left(\frac{1}{k_t}(\mathrm{id} - d_t)\right)\right)_{t\in\{1,\ldots,T\}} .$$

Finally, the moreover part of corollary **??** allows to eliminate the constants k_t and d_t from the above triples, up to the sign of k_t (choose $\mathfrak{a}^+_t = \frac{1}{k_t}(\mathrm{id} - d_t)$ and note that $f_t = \mathrm{id}$). I obtain the representing sequence of triples

$$\left(u''_t = \mathrm{sgn}(\xi)\; {u^*}^{\frac{\xi}{1-\beta_t}} \,,\, f''_t = \mathrm{id} \,,\, g''_t = \beta^{t-1} \frac{1-\beta_t}{\xi} \ln\big(\mathrm{sgn}(\xi)\ \mathrm{id}\big)\right)_{t\in\{1,\ldots,T\}} . \tag{37}$$

The function u^* in expression (37) corresponds to the utility function u stated in corollary 2. For preference representations in theorem 4, Bernoulli utility lies in the class $u : X \to \mathrm{I\!R}$. By equation (36), the latter class for u corresponds to functions u^* lying in the class of

[48] Or from a different perspective, $g'_t \circ u'_t = \beta^{t-1} \frac{1-\beta_t}{\xi} \ln\left(\frac{1}{k_t}(u'_t - d_t)\right)$ depends on ξ. Thus fixing the range of u'_t as is, would not allow for a common measure scale for welfare.

continuous functions from X into the positive real numbers, i.e. $u^* \in \{u^* : X \to \mathbb{R}_{++}\}$.
Remark: The requirement of corollary 2 that u, i.e. u^* in the representing triples above, is onto the interval U^* corresponds to setting the range for the measurement of welfare in period t to the range of

$$g''_t \circ u''_t = \beta^{t-1} \frac{1-\beta_t}{\xi} \ln \left(u^{* \frac{\xi}{1-\beta_t}} \right) = \beta^{t-1} \ln (u^*) \tag{38}$$

and, thus, $g_1 \circ u_1 = \ln u^*$.
Part II: In the following I calculate the representation expressed by the sequence of triples in (37). Let $U^* = \left[\underline{U}^*, \overline{U}^*\right]$. Then, observe that $\Delta G''_t = \beta^t \ln \frac{\overline{U}^*}{\underline{U}^*}$ and $\vartheta'_t = 0$. Thus, with the same definition for $\beta_t = 1 - \theta''_t$ as in theorem 2 (compare page 44), it holds

$$\begin{aligned}
\tilde{u}_{t-1}(x_{t-1}, p_t) &= g''^{-1}_{t-1} \left[\theta''_{t-1}\, g''_{t-1} \circ u''_{t-1}(x_{t-1}) + (1-\theta''_{t-1}) \frac{\Delta G''_{t-1}}{\Delta G''_t}\, g''_t \circ \mathcal{M}^{f''_t}(p_t, \tilde{u}_t) \right] \\
&= g''^{-1}_{t-1} \left[(1-\beta_{t-1})\, g''_{t-1} \circ u''_{t-1}(x_{t-1}) + \beta_{t-1}\, \beta^{-1}\, g''_t \circ \mathcal{M}^{f''_t}(p_t, \tilde{u}_t) \right] \\
&= \text{sgn}(\xi)\ \exp \left(\frac{\xi}{\beta^{t-2}(1-\beta_{t-1})} \left[(1-\beta_{t-1})\, \beta^{t-2}\, \frac{1-\beta_{t-1}}{\xi} \ln \Big(\text{sgn}(\xi) \cdot \right.\right. \\
&\qquad \left.\left. \text{sgn}(\xi)\ u^*(x_{t-1})^{\frac{\xi}{1-\beta_{t-1}}} \Big) + \beta_{t-1}\, \beta^{t-2}\, \frac{1-\beta_t}{\xi} \ln \big(\text{sgn}(\xi)\ \text{E}_{p_t}\, \tilde{u}_t \big) \right] \right) \\
&= \text{sgn}(\xi)\ \exp \Big(\ln \big(u^*(x_{t-1})^{\xi} \big) \Big) \exp \left(\beta_{t-1} \frac{1-\beta_t}{1-\beta_{t-1}} \ln \big(\text{sgn}(\xi)\ \text{E}_{p_t}\, \tilde{u}_t \big) \right) \\
&= \text{sgn}(\xi)\ u^*(x_{t-1})^{\xi} \big(\text{sgn}(\xi)\ \text{E}_{p_t}\, \tilde{u}_t \big)^{\beta} .
\end{aligned} \tag{39}$$

Where I have used the relation $\frac{1-\beta_t}{1-\beta_{t-1}} = \beta \beta^{-1}_{t-1}$ to arrive at the last line. Distinguishing the two cases where $\text{sgn}(\xi) > 0$ and $\text{sgn}(\xi) < 0$, equation (39) corresponds to the representation stated in the theorem.
Moreover part: Equation (38) in the remark shows that the demand of u^*, corresponding to u in the corollary, being onto the given interval U^* fixes also the range for the measurement of welfare $g'' \circ u''_t$. Therefore, the moreover part follows as in corollary 1. □

Proof of corollary 3: The representation is a simple transformation of corollary 2.
"⇒": For $\xi \neq 0$ define $\tilde{v}_t : \tilde{X}_t \to \mathbb{R}$ for $t \in \{1, ..., T\}$ by $\tilde{v}_t = (\text{sgn}(\xi)\ \tilde{u}_t)^{\frac{1}{\xi}}$, where $\tilde{u}_t$ defines the recursive construction of the representation in corollary 2. Then it is

$$\begin{aligned}
\tilde{v}_{t-1}(x_{t-1}, p_t) &= u(x_{t-1}) \big(\text{E}_{p_t}\ \text{sgn}(\xi)\ \tilde{u}_t \big)^{\frac{\beta}{\xi}} \\
&= u(x_{t-1}) \big(\text{E}_{p_t}\, \tilde{v}_t^{\xi} \big)^{\frac{\beta}{\xi}} \\
&= u(x_{t-1})\ \left(\mathcal{M}^{\alpha=\xi}(p_t, \tilde{v}_t) \right)^{\beta} ,
\end{aligned}$$

yielding the stated construction of $\tilde{v}_t$. Then the representation of corollary 2 translates into

$$\begin{aligned}
& p_t && \succeq_t p'_t \\
\Leftrightarrow & \quad \mathrm{E}_{p_t} \tilde{u}_t && \geq \mathrm{E}_{p'_t} \tilde{u}_t \\
\Leftrightarrow & \operatorname{sgn}(\xi)\, \mathrm{E}_{p_t} \operatorname{sgn}(\xi)\, \tilde{u}_t && \geq \operatorname{sgn}(\xi)\, \mathrm{E}_{p'_t} \operatorname{sgn}(\xi)\, \tilde{u}_t \\
\Leftrightarrow & \quad \operatorname{sgn}(\xi)\, \mathrm{E}_{p_t} \tilde{v}_t^{\xi} && \geq \operatorname{sgn}(\xi)\, \mathrm{E}_{p'_t} \tilde{v}_t^{\xi}
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow & \quad \left(\mathrm{E}_{p_t} \tilde{v}_t^{\xi}\right)^{\xi} \geq \left(\mathrm{E}_{p'_t} \tilde{v}_t^{\xi}\right)^{\xi} \\
\Leftrightarrow & \mathcal{M}^{\alpha=\xi}(p'_t, \tilde{v}_t) \geq \mathcal{M}^{\alpha=\xi}(p'_t, \tilde{v}_t)
\end{aligned}$$

for all $p_t, p'_t \in P_t$.

For the case $\xi = 0$ the stated representation corresponds to

$$\begin{aligned}
\mathcal{M}^0(p_t, \tilde{v}_t) &= \exp\left(\int dp_t \ln \tilde{v}_t\right) \\
&= \exp\left(\int dp_t \ln \left[u(x_t)\left(\exp\left[\mathrm{E}_{p_t}\ \ln(\tilde{v}_{t+1})\right]\right)^{\beta}\right]\right) \\
&= \exp\left(\int dp_t \ln u(x_t) + \beta\left[\mathrm{E}_{p_t}\ \ln(\tilde{v}_{t+1})\right]\right)
\end{aligned}$$

Define $u^* = \ln u$ and $\tilde{v}_t^* = \ln \tilde{v}_t$. Then the representation is ordinally equivalent to

$$\int dp_t u^*(x_t) + \beta\left[\mathrm{E}_{p_t}\ \ln(\tilde{v}_{t+1})\right]$$

For the case $\xi = 0$ the stated representation corresponds to

$$\begin{aligned}
\tilde{v}_{t-1}(x_{t-1}, p_t) &= u(x_{t-1})\left(\exp\left[\mathrm{E}_{p_t}\ \ln(\tilde{v}_t)\right]\right)^{\beta} \\
&= u(x_{t-1}) \exp\left[\mathrm{E}_{p_t} \beta \ln(\tilde{v}_t)\right].
\end{aligned}$$

Defining $\tilde{v}_t^* = \ln \tilde{v}_t$ and $u^* = \ln u \Leftrightarrow u = \exp u^*$ yields the representation

$$\begin{aligned}
\tilde{v}_{t-1}^*(x_{t-1}, p_t) &= \ln\left(\exp[u^*(x_{t-1})] \exp\left[\mathrm{E}_{p_t} \beta \tilde{v}_t^*\right]\right) \\
&= u^*(x_{t-1}) + \mathrm{E}_{p_t} \beta\, \tilde{v}_t^*.
\end{aligned}$$

But the latter construction of aggregate welfare, corresponds to that of corollary 1 for preferences corresponding to $\xi = 0$ (intertemporally additive expected utility). Moreover, the uncertainty evaluation

$$\begin{aligned}\mathcal{M}^0(p_t, \tilde{v}_t) &= \exp\big(\textstyle\int dp_t \ln \tilde{v}_t\big) \\ &= \exp\big(\textstyle\int dp_t \ \tilde{v}^*_t\big)\end{aligned}$$

is a strictly increasing transformation of $\mathrm{E}_{p_t}\tilde{v}^*_t$. Therefore, the representation for the case $\xi = 0$ is equivalent to the formulation in corollary 1.
"$\Leftarrow$": Immediate consequence of corollary 2.
Moreover part: Is implied by the moreover part of corollary 2. Again, the measure scale for welfare is fixed for the first period to the range $\ln U^*$. □

D Proofs for Section 5

Proof of theorem 6: "$\Rightarrow$": Adding certainty stationarity to the assumptions of theorem **??** implies, as shown in the proof of theorem 2, that Bernoulli utility can be picked identical in all periods. Moreover, in that case it exist $\beta \in \mathrm{I\!R}_{++}$ and $g : \mathsf{X} \to \mathrm{I\!R}$ such that the functions g_t can be chosen as $g_t = \beta^{t-1}g$. Then, in the representation of theorem **??**, the construction of aggregate utility simplifies to the form

$$\tilde{u}_t(\mathsf{x}^t) = \sum_{\tau=t}^{T} g_\tau \circ u_\tau(\mathsf{x}^t_\tau) = \sum_{\tau=t}^{T} \beta^{\tau-1} g \circ u(\mathsf{x}^t_\tau) \equiv \sum_{\tau=t}^{T} \beta^{\tau-1} u^*(\mathsf{x}^t_\tau)\,,$$

where the simple redefinition of Bernoulli utility as $u^*(\mathsf{x}^t_\tau) = g \circ u(\mathsf{x}^t_\tau)$ yields the $g = \mathrm{id}-$gauge. Moreover, in the formulation of the theorem, the range of welfare $u^*(\mathsf{x}^t_\tau) = g \circ u(\mathsf{x}^t_\tau)$, i.e. u in the notation of the theorem, is fixed exogenously. Therefore, as in corollary **??**, the parameter ξ in equation (**??**) stemming from the relation

$$h_t(z) = g_t \circ f_t^{-1}(z) = a_t \exp(\frac{\xi}{1-\beta_t}\, z) + b_t$$

cannot be absorbed by the function u^*. In consequence, for the case $\xi \neq 0$, the representation (**??**) prevails, just as for the g^+-gauge in corollary **??**. Also as in the previous g^+- corollaries 1 and **??**, it is found that the case $\xi = 0$ is covered by the representation using the uncertainty aggregation rule $\mathcal{M}^{\exp^0}(p^{\mathsf{X}}_t, \tilde{u}_t) = \mathrm{E}_{p^{\mathsf{X}}_t}\, \tilde{u}_t$.
"$\Leftarrow$": Implied by theorems 2 and **??**.
Moreover part: By lemma 2, the choice of the range of $u = u_1 = g_1 \circ u_1$ as W^* fixes the measure scale of welfare for all periods. Therefore, corollary **??** covers the moreover part with $\theta_t = 1 - \beta_t$ (see proof of theorem 6). □

Proof of theorem 7: The assertion follows immediately from comparing the functions

characterizing intertemporal risk aversion in the representations of theorem 4 and theorem 6. These imply that the two representations can only coincide for the case where $\beta = 1$.
"$\Rightarrow$": Preferences satisfying the stated axioms have to be representable in the sense of theorems 4 and 6.[49] Choose a nondegenerate closed interval $W^* \subset \mathrm{I\!R}_{++}$ and require that $u = u^{\text{welf}}$ is onto W^*. Then, due to risk stationarity, by corollary 1 there have to exist ξ and β such that the functions $f_t \circ g_t$ characterizing intertemporal risk aversion are specified by the coefficients $\text{AIRA}_t = -\frac{\xi}{\beta^{t-1}(1-\beta_t)}$. Analogously, due to timing indifference, by theorem 6 there have to exist ξ' and β' such that the functions $f_t \circ g_t$ characterizing intertemporal risk aversion are specified by the coefficients $\text{AIRA}_t = -\frac{\xi'}{1-\beta'_t}$.

Both representations, that of corollary 1 and that of theorem 6, are special cases of the certainty stationary representation in theorem 2. For given preferences $\succeq = (\succeq_t)_{t\in\{1,...,T\}}$, coincidence of the representations on certain consumption paths implies that $\beta = \beta'$. In consequence, it also holds that $\beta_t = \beta'_t$. As the measure scale for welfare is fixed to W^* in the first period, lemma 2 states that the characterizations AIRA_t of intertemporal risk aversion are unique for all $t \in \{1, ..., T\}$. Therefore, comparison of the measures of intertemporal risk aversion for period one implies that $\xi = \xi'$. Then, the requirement that furthermore $\text{AIRA}_t \stackrel{!}{=} -\frac{\xi}{\beta^{t-1}(1-\beta_t)} \stackrel{!}{=} -\frac{\xi}{(1-\beta_t)}$ for all $t > 1$, cannot be satisfied unless $\beta = 1$ or $\xi = 0$. However, the requirement of strict intertemporal risk aversion as formulated in axiom A9^{s} implies implies $\xi < 0$. Therefore it has to hold that $\beta = 1$.
"$\Leftarrow$": Except for axiom A9^{s} all of the stated axioms are implied by theorems 4 and 6. Axiom A9^{s} is implied by theorem 5, case a). □

[49]Recall that axiom A8 implies certainty stationarity as described in axiom A6.

References

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

Attanasio, O. P. & Weber, G. (1989), 'Intertemporal substitution, risk aversion and the euler equation for consumption', *The Economic Journal* **99**(395), 59–73.

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

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.

Epaulard, A. & Pommeret, A. (2001), 'Agents' preferences, the equity premium, and the consumption-saving trade-off: An application to french data', *IMF Working Paper* (WP/01/117).

Epaulard, A. & Pommeret, A. (2003*a*), 'Optimally eating a stochastic cake: a recursive utility approach', *Resource and Energy Economics* **25**, 129–139.

Epaulard, A. & Pommeret, A. (2003*b*), 'Recursive utility, endogeneous growth, and the welfare cost of volatility', *Review of Economic Dynamics* **6**, 672–684.

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.

Ha-Duong, M. & Treich, N. (2004), 'Risk aversion, intergenerational equity and climate change', *Environmental and Resource Economics* **28**(2), 195–207.

Hardy, G., Littlewood, J. & Polya, G. (1964), *Inequalities*, 2 edn, Cambridge University Press. first puplished 1934.

Howitt, R. E., Msangi, S., Reynaud, A. & Knapp, K. C. (2005), 'Estimating intertemporal preferences for natural resource allocation', *American Journal of Agricultural Economics* **87**(4), 969–983.

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

Knapp, K. C. & Olson, L. J. (1996), 'Dynamic resource management: Intertemporal substitution and risk aversion', *American Journal of Agricultural Economics* **78**(4), 1004–1014.

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

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

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

Nordhaus, W. D. (1993), 'Rolling the 'dice': An optimal transition path for controlling greenhouse gases', *Resource and Energy Economics* **15**, 27–50.

Nordhaus, W. D. (1994), *Managing the Global Commons: The Economics of the Greenhouse Effect*, MIT Press, Cambridge.

Normandin, M. & St-Amour, P. (1998), 'Substitution, risk aversion, taste shocks and equity premia', *Journal of Applied Econometrics* **13**(3), 265–281.

Obstfeld, M. (1994), 'Evaluating risky consumption paths: The role of intertemporal substitutability', *European Economic Review* **38**, 1471–1486.

Plambeck, E. L., Hope, C. & Anderson, J. (1997), 'The Page95 model: Integrating the science and economics of global warming', *Energy Economics* **19**, 77–101.

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

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

Sackmann, I.-J., Boothroyd, A. I. & Kraemer, K. E. (1993), 'Our sun. III. Present and future', *The Astrophysical Journal* **418**, 457–468.

Svensson, L. E. (1989), 'Portfolio choice with non-expected utility in continuous time', *Economics Letters* **30**, 313–317.

Toth, F. L. (1995), 'Discounting in integrated assessments of climate change', *Energy Policy* **23**(4-5), 403–409.

Traeger, C. (2007*a*), Intertemporal risk aversion I - The concept. Unpublished Working Paper.

Traeger, C. (2007*b*), Intertemporal risk aversion II - Temporal resolution of uncertainty and timing indifference. Unpublished Working Paper.

Traeger, C. (2007*c*), Intertemporal risk aversion III - Stationarity and time preference.

Traeger, C. (2007*d*), 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.

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.

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