MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Foundations of Intertemporal Choice under Uncertainty"

verfasst von / submitted by

Philipp Peitler, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2018 / Vienna 2018

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 913
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Volkswirtschaftslehre
Betreut von / Supervisor:	Univ.-Prof. Dipl.-Math. Karl Schlag, PhD

Foundations of Intertemporal Choice under Uncertainty

Philipp Peitler

June 2018

Abstract

The purpose of this paper is to extend the von Neumann-Morgenstern evaluation of lotteries over single outcomes to lotteries over finite sequences of outcomes. We provide an axiomatization that leads to the typical approach of comparing sequences by the exponentially discounted sum of per period von Neumann-Morgenstern utilities. Most of our additional axioms, like Time consistency, Time invariance and Planning in stages, are motivated by dynamic decision making. Unlike most axiomatizations for exponential discounting, see Koopmans (1960) and Epstein (1983), we do neither require technical assumptions that assure the continuity of the utility function in outcomes nor do we require the use of the typical existences theorems for additive utility. Our proof is constructive and easy to follow.

Keywords: von Neumann-Morgenstern, exponential discounting, lotteries, finite, sequences, time consistency

1 Introduction

A driver faces a decision whether to buy car insurance or not. One could simply state that the driver prefers one action to the other, without modeling the obvious uncertainty involved. However, since the driver can also compare each of the possible consequences, "insured and accident", "not insured and accident", "insured and no accident" and "not insured and no accident", it seems natural to derive her preferences for the actions from her preferences for the possible consequences. Here, the most commonly used approach is to compare expected utility, axiomatized by von Neumann and Morgenstern (1944). Now consider a decision where consequences span over time. An individual has to live off 10000$ for the next year and has to allocate money to each month in advance. Again, one could simply state preferences over all allocations, but since the decision maker can also compare each individual month living off x\$, $x \in [0, 10000]$, it makes sense to model time explicitly and derive preferences for allocations from preferences for living of x\$ in a single month. The purpose of this paper is to find a reasonable set of axioms that allows us to derive preferences for consequences that span over time from preferences over partitions of those consequences. Since we model time discretely, we will refer to those consequences, if deterministic, as sequences of outcomes, where the individual receives an outcome in each time period. If consequences are uncertain, we refer to them as lotteries over sequences of outcomes.

The most commonly used evaluation method for sequences of outcomes is exponential discounting. The utility of a sequence $(x_1, x_2, ..., x_T)$ is determined by some discount factor $\delta \in (0,1)$ and some value function $V : A \mapsto \mathbb{R}$, where $x_1, x_2, ..., x_T \in A$, in the following way:

$$U(x_1, x_2, ..., x_T) = \sum_{t=1}^{T} \delta^{t-1} V(x_t). \tag{1}$$

This form of evaluation was first proposed by Samuelson (1937) and later axiomatized by Koopmans (1960) for deterministic sequences and infinite time horizon. Note that with Koopmans approach one could not simply evaluate lotteries over sequences of outcomes by evaluating the deterministic sequences and then taking expected utility, since neither δ nor V are unique. An axiomatization of exponential discounting for lotteries over infinite sequences was later done by Epstein (1983). Epstein (1983) and another related publication are discussed in Section 2.

In this paper we derive exponential discounting for lotteries over finite sequences. To our knowledge, this has never been done before. Furthermore we propose an alternative interpretation of outcomes and show how time consistent and time invariant von Neumann-Morgenstern preferences, in combination with a few weaker axioms, lead to exponential discounting. The axioms are presented in Section 3. In Section 4 we provide a simple and straightforward proof of our theorem that does not rely on outside references. Finally, in Section 5 we derive preferences for the infinite horizon case from preferences over finite sequences.

2 Related Literature

In the following we present two publications that have axiomatized exponential discounting under uncertainty and show how they differ from our approach.

Epstein (1983) derives exponential discounting for a discrete and infinite time environment, where the outcome space for each period is a subset of the Euclidean space. He assumes the existence of a continuous von Neumann-Morgenstern utility function, Stationarity and Risk independence. Note the similarities to our axioms in the Section 3. Assumptions 2 and 3 ensure the existence of a von Neumann-Morgenstern representation and Assumptions 7 and 9 are the equivalent of Risk independence. Unlike us, Epstein does not require an Impatience assumption, but Impatience is derived from his axioms. The obvious key difference to our paper is that preferences for a finite horizon case are not derived, whereas our approach provides results for both the finite and infinite horizon case. The other key difference is Epstein's restrictive assumption that the utility function is not merely continuous in probability, but continuous in outcomes. Our approach on the other hand provides results even for a discontinuous utility function, as long as there exists a best and a worst outcome in the specific decision problem.

Anchugina (2016) derives exponential discounting for a discrete and finite time environment, where the outcome space is a mixture set. She assumes Mixture independence, Mixture continuity, Monotonicity, Impatience and Stationarity. An interpretation for an element from the n-fold Cartesian product of a mixture set is a sequence of independent lotteries of outcomes. Therefore, Anchugina excludes the case of a sequence of correlated lotteries or a lottery over sequences of outcomes. Our approach on the other hand, dealing with lotteries over sequences, can easily provide results for sequences of lotteries. Furthermore, the axioms Mixture independence and Mixture continuity are not the standard von Neumann-Morgenstern axioms of Independence and Continuity assumed by us (Assumptions 2 and 3). Especially Mixture independence is a much stronger assumption and is the key factor in deriving additivity of utility by period.

3 Model

An agent faces a decision problem under uncertainty in an environment with discrete time and finite horizon $T \in \mathbb{N}$. Preferences in an infinite horizon environment are later derived from finite preferences. The agent receives an outcome from a finite set of outcomes A in each period, which we call a sequence of outcomes from the set A^T.[1] Although the purpose of this paper is to derive preferences for lotteries over sequences of outcomes from $\triangle\left(A^T\right)$, some of our axioms are motivated by dynamic decision

[1]Our results also apply if A is not finite, as long as there is a best and a worst outcome in the specific decision problem.

making where the agent can choose an action at multiple time periods $t \in \{1, 2, .., T\}$. Since modelling such a problem explicitly would be an unnecessary complication for our purpose, we assume that the agent does not care through which path of actions outcomes are realized, but only considers the final distribution of sequences of outcomes that results from her plan of actions, should she stick to it. At any time t, where outcomes have already been realized for the periods 1 to $t-1$, the agent has preferences for distributions over sequences of outcomes for the remaining periods t to T. Therefore $\succsim_t^T$ is a binary relation on $A^{(t-1)} \times \triangle\left(A^{T-t+1}\right)$ and represents the agents preferences in a dynamic decision problem at time t in an environment with time horizon T.

In the following, the notation $(x_1, ..., x_{t-1}, X_{T-t+1})$ can either refer to an element of $A^{(t-1)} \times \triangle\left(A^{T-t+1}\right)$, where $x_1, ..., x_{t-1} \in A$ and $X_{T-t+1} \in \triangle\left(A^{T-t+1}\right)$, or to an element of $\triangle\left(A^T\right)$, where there is only strictly positive probability on sequences that start with $x_1, ..., x_{t-1}$ and those sequences are distributed according to X_{T-t+1}.

Assumption 1 (Time consistency). $(x_1, ..., x_{t-1}, Y_{T-t+1}) \succsim_1^T (x_1, ..., x_{t-1}, Z_{T-t+1})$ *if and only if* $(x_1, ..., x_{t-1}, Y_{T-t+1}) \succsim_t^T (x_1, ..., x_{t-1}, Z_{T-t+1})$ *for all* $t \in \{1, 2, .., T\}$, $x_1, ..., x_{t-1} \in A$ *and* $Y_{T-t+1}, Z_{T-t+1} \in \triangle\left(A^{T-t+1}\right)$.

Time consistency usually refers to the concept that if a plan of actions is optimal in the first period, it is also optimal in any future period t (Strotz, 1955). We assume a slightly stronger notion of Time consistency where a strict preference in the initial period must imply a strict preference at t. Note that indifference in the initial period must imply indifference at t even under the weaker definition. We prove in Lemma 6 in the Appendix that Assumption 1 is equivalent to stating those two assumptions separately if preferences are complete. Furthermore, note that the concept, that if a plan is optimal at t then is also optimal in the initial period, is not Time consistency but better described as Backward reasoning. The agent anticipates her decision at t and chooses accordingly in the initial period. This distinction is oftentimes ignored in the literature, see for example Kreps and Porteus (1978). Now, since Time consistency assures that the agent will stick to her initial plan of actions, we can reduce the dynamic decision problem to the problem of finding the optimal plan in the first period. We define $\succsim^T := \succsim_1^T$.

In the following let $X_T, Y_T, Z_T \in \triangle\left(A^T\right)$ be distributions of sequences of outcomes. Note that if we formulate preference relations in terms of sequences of outcomes, we actually refer to the distribution where this sequence is realized with probability 1.

Assumption 2 (Rationality). $\succsim^T$ *is complete and transitive.*

Assumption 3 (Continuity)**.** *If* $X_T \succ^T Z_T$ *and* $X_T \succsim^T Y_T \succsim^T Z_T$ *then there exists a unique* $p \in [0,1]$ *such that* $p\,[X_T] + (1-p)\,[Z_T] \sim^T Y_T$.

Assumption 4 (Independence)**.** $X_T \succsim^T Y_T$ *if and only if* $p\,[X_T] + (1-p)\,[Z_T] \succsim^T p\,[Y_T] + (1-p)\,[Z_T]$ *for all* $p \in (0,1)$.

Assumption 5 (Non-triviality)**.** *There exists* $x, y \in A$ *such that* $x \succ^1 y$.

Assumption 6 (Impatience)**.** $x \succsim^1 y$ *if and only if* $(x,y) \succsim^2 (y,x)$ *where* $x, y \in A$.

Assumptions 2 to 4 are the standard von Neumann-Morgenstern axioms. Assumption 5 and 6 are not restrictive. If preferences do not satisfy Non-triviality, the agent is indifferent between any sequence, as long as all other assumptions hold. Impatience could be replaced by an assumption of over-patience, in the sense that $x \succsim^1 y$ if and only if $(y,x) \succsim^2 (x,y)$. Our proof concept would still be able to derive results. Only for the case of time indifference, $(y,x) \sim^2 (x,y)$, we would need to make additional assumptions.

In order to justify the next assumption, we suggest the following interpretation for outcomes in A. Let B be a finite set of events. Instead of an outcome only entailing the actual event at the given time, let it also entail the memory of the agent of the events of previous periods. We therefore define

$$A = \left\{ {}_{\beta}\alpha : \alpha \in B, \beta \in \bigcup_{\gamma=0}^{n} B^{\gamma} \right\}$$

for some $n \in \mathbb{N}$. Note that A is still finite. For any ${}_{\beta}\alpha \in A$ we call α the **actual event** and β the **memory**.[2]

Definition 1 (Memory true)**.** *A sequence of outcomes* $(x_1, x_2, ..., x_T)$ *is memory true if and only if* $m\,(x_t) = (e\,(x_1)\,, ..., e\,(x_{t-1}))$ *for all* $t \in \{1, 2, ..., T\}$, *where* $e : A \mapsto B$ *is the event function with* $e\left({}_{\beta}\alpha\right) = \alpha$ *and* $m : A \mapsto \bigcup_{\gamma=0}^{n} B^{\gamma}$ *is the memory function with* $m\left({}_{\beta}\alpha\right) = \beta$.

Note that we allow for preferences over sequences that are not memory true, since such sequences are conceivable.[3] The use of sequences that are not memory true in our proof must be viewed as a hypothetical scenario that allows us to derive preferences for memory true sequences.

Since the history of past outcomes is accounted for by inclusion of a memory component within the future outcomes, we can assume that a decision at some period t is independent of the past outcomes $x_1, ..., x_{t-1}$.

[2] One potential downside of this approach is the enormous number of outcomes that would result for example from a repeated 2x2 game. To deal with this issue one could make the strong assumption that outcomes are indifferent as long as they have the same actual event.

[3] Consider a scenario in which the agent looses her memory or has wrong memory of past events.

Assumption 7 (Independence of past)**.** $(x_1, ..., x_{t-1}, Y_{T-t+1}) \succsim_t^T (x_1, ..., x_{t-1}, Z_{T-t+1})$ *if and only if* $(a_1, ..., a_{t-1}, Y_{T-t+1}) \succsim_t^T (a_1, ..., a_{t-1}, Z_{T-t+1})$ *for all* $t \in \{1, 2, .., T\}$, $x_1, ..., x_{t-1} \in A$ *and* $Y_{T-t+1}, Z_{T-t+1} \in \triangle\left(A^{T-t+1}\right)$.

Note the similarities to the assumption from Epstein (1983) that the future is risk independent from the past. While Epstein uses the term "past" loosely to distinguish early periods from later ones, we actually refer to time periods that have already passed.

Assumption 8 (Time invariance)**.** *For all* $t \in \{1, 2, .., T\}$ *and* $Y_{T-t+1}, Z_{T-t+1} \in \triangle\left(A^{T-t+1}\right)$, *if* $(x_1, ..., x_{t-1}, Y_{T-t+1}) \succ_t^T (x_1, ..., x_{t-1}, Z_{T-t+1})$ *for all* $x_1, ..., x_{t-1} \in A$ *then* $Y_{T-t+1} \succ_1^{T-t+1} Z_{T-t+1}$ *and if* $(x_1, ..., x_{t-1}, Y_{T-t+1}) \sim_t^T (x_1, ..., x_{t-1}, Z_{T-t+1})$ *for all* $x_1, ..., x_{t-1} \in A$ *then* $Y_{T-t+1} \sim_1^{T-t+1} Z_{T-t+1}$.

Time invariance is usually an assumption made for preferences on $A \times \{1, 2, ..., T\}$, where the agent only receives one outcome at one future point in time, and it says that preferences do not change if both the timing of the decision and the timings of the outcomes are equally shifted (Halevy, 2015). In our model, a sequence of outcomes can not be shifted into the future and leave time periods empty, as the agent has to receive an outcome at every period starting from 1. In order to apply this axiom to our model, let us first consider the following hybrid environment. Let $\Theta = \{\theta\}$ be the set containing a neutral outcome, which can be interpreted as "nothing of interest is happening in this period". The agent has preferences $\succsim_t'^T$ on $\Theta^k \times A \times \Theta^{T-k-1}$ for all $k \in \{0, 1, ..., T-1\}$.

Definition 2 (Time invariance in $\Theta^k \times A \times \Theta^{T-k-1}$ environment)**.** *Preferences on* $\Theta^k \times A \times \Theta^{T-k-1}$ *are time invariant if it holds that* $(a_1, ..., a_{t+\zeta-1}, x, a_{t+\zeta+1}, ..., a_{t+\eta}) \succsim_t'^{t+\eta} (o_1, ..., o_{t+\eta-1}, y)$ *if and only if* $(a_1, ..., a_\zeta, x, a_{\zeta+2}, ..., a_{1+\eta}) \succsim_1'^{1+\eta} (o_1, ..., o_\eta, y)$ *for all* $t \in \mathbb{N}$, $\eta \in \mathbb{N}_0$, $\zeta \in \{0, 1, ..., \eta\}$, $x, y \in A$ *and* $o_1, ..., o_{t+\eta-1}, a_1, ..., a_{t+\eta} \in \Theta$.

Note that we relate a choice at period t to a choice in the initial period, instead of another general period t'. This is without loss of generality, as one could first relate t to 1 and then t' to 1. In order to find an equivalent formulation for our environment, consider that all outcomes come from A instead of Θ and that $a_i = o_i$ for all $i \in \{0, 1, ..., t-1\}$, since it makes sense to only compare sequences with a common past up to the decision. We now face the problem that the resulting statement would imply Assumption 7. This was not an issue in the hybrid environment, as there exists only one possible past, consisting of neutral outcomes. Note that by Lemma 6 in the Appendix there exists an equivalent formulation for Definition 2 in terms of "if - then" statements for strict preferences and indifference. We use this formulation and by requiring preferences at t to hold for all histories, we can remove the implied Independence of past assumption, resulting in Assumption 8.

In the following let $(Y_S, X_R) \in \triangle\left(A^S\right) \times \triangle\left(A^R\right)$ denote a distribution of sequences of outcomes where the outcomes for the first S periods are distributed according to Y_S and the outcomes of the R periods after are distributed independently of the first S periods and according to X_R.

Assumption 9 (Planning in stages)**.** $(Y_S, X_R) \succsim^T (Z_S, X_R)$ *if and only if* $Y_S \succsim^S Z_S$ *for all* $X_R \in \triangle\left(A^R\right)$*,* $Y_S, Z_S \in \triangle\left(A^S\right)$ *and* $S, R \in \mathbb{N}$ *such that* $S + R = T$*.*

In order to justify this assumption, consider that the agent faces a special decision problem with horizon $T = S + R$, where her action set in time period $S + 1$ and all action sets thereafter do not depend on her previous actions in the first S periods. At $S+1$ her choice does not depend on past realizations of outcomes due to Assumption 7 and she will always choose the same plan of actions for the following R periods, no matter her previous actions. Assumption 9 says that the agent anticipates this and simplifies her problem at time period 1 by only planning for the first S periods.

Theorem 1. *Let* $M \in \{3, 4, ...\}$ *be the maximal time horizon. For all* $T \in \{1, 2, ..., M\}$*, let* U_T *be a von Neumann-Morgenstern representation of* $\succsim^T$ *on* $\triangle\left(A^T\right)$*. If preferences* $\succsim^T$ *on* $\triangle\left(A^T\right)$ *satisfy Assumptions 1-9 for all* $T \in \{1, 2, ..., M\}$*, then* U_T *exists for all* $T \in \{1, 2, ..., M\}$ *and there exists a unique* $\delta \in (0, 1)$ *such that* $U_T(x_1, ..., x_T) = \sum_{t=1}^{T} \delta^{t-1} U_1(x_t)$*.*

In the following we derive the utility for any sequence of outcomes and prove Theorem 1. As mentioned before, although we denote sequences of outcomes, we actually refer to a distribution where this sequence is realized with probability 1. Because all von Neumann-Morgenstern axioms are assumed, the utility of any distribution of sequences of outcomes can be simply calculated by expected utility.

4 Proof of Theorem 1

Lemma 1. *Assumptions 1, 7 and 8 imply that* $(x_1, Y_{T-1}) \succsim^T (x_1, Z_{T-1})$ *if and only if* $Y_{T-1} \succsim^{T-1} Z_{T-1}$*.*

Proof. Assume Assumptions 1, 7 and 8 hold. First we prove that if $(x_1, Y_{T-1}) \succsim^T (x_1, Z_{T-1})$ then $Y_{T-1} \succsim^{T-1} Z_{T-1}$. Consider $(x_1, Y_{T-1}) \succsim^T (x_1, Z_{T-1})$. From Assumption 1 follows that $(x_1, Y_{T-1}) \succsim_2^T (x_1, Z_{T-1})$. Assumption 7 says that this relation must hold for all $x_1 \in A$. Since either $(x_1, Y_{T-1}) \succ_2^T (x_1, Z_{T-1})$ or $(x_1, Y_{T-1}) \sim_2^T (x_1, Z_{T-1})$ for all $x_1 \in A$, Assumption 8 rules out that $Z_{T-1} \succ^{T-1} Y_{T-1}$. Therefore

$Y_{T-1} \succsim^{T-1} Z_{T-1}$ which concludes the first part of the proof. Next we prove that if $Y_{T-1} \succsim^{T-1} Z_{T-1}$ then $(x_1, Y_{T-1}) \succsim^{T} (x_1, Z_{T-1})$. Consider $Y_{T-1} \succsim^{T-1} Z_{T-1}$. We assume that $(x_1, Y_{T-1}) \succsim^{T} (x_1, Z_{T-1})$ does not hold and find a contradiction. If $(x_1, Y_{T-1}) \succsim^{T} (x_1, Z_{T-1})$ does not hold $(x_1, Z_{T-1}) \succ^{T} (x_1, Y_{T-1})$ must be the case. From Assumption 1 follows that $(x_1, Z_{T-1}) \succ_2^{T} (x_1, Y_{T-1})$ and since this relation must hold for all $x_1 \in A$ due to Assumption 7, Assumption 8 gives $Z_{T-1} \succ^{T-1} Y_{T-1}$ which contradicts $Y_{T-1} \succsim^{T-1} Z_{T-1}$. □

Definition 3. *Let $b \in A$ be the best outcome in the sense that $b \succsim^1 x$ for all $x \in A$. Let $w \in A$ be the worst outcome in the sense that $x \succsim^1 w$ for all $x \in A$.*

Lemma 2. *For all $T \in \mathbb{N}$ there exists a von Neumann-Morgenstern representation U_T of $\succsim^T$ on $\triangle\left(A^T\right)$ such that $U_T(b, w, ..., w) = 1$ and $U_T(w, ..., w) = 0$.*

Proof. All von Neumann-Morgenstern axioms are assumed (Assumption 2 to 4) and A is finite. We now show that $(b, w, ..., w) \succ^T (w, ..., w)$. $b \succ^1 w$ due to Assumption 5. Adding w to the end of both sides of the relation $T-1$ times based on Assumption 9 gives $(b, w, ..., w) \succ^T (w, ..., w)$. □

Lemma 3. *There exists a unique $\delta \in (0, 1)$ such that*

$$(w, b) \sim^2 \delta\left[(b, w)\right] + (1 - \delta)\left[(w, w)\right]. \tag{2}$$

Proof. If $(b, w) \succ^1 (w, b) \succ^1 (w, w)$, then the lemma follows from Assumption 3. Since $b \succ^1 w$ due to Assumption 5, $(b, w) \succ^1 (w, b)$ follows directly from Assumption 6. Adding w to the beginning of $b \succ^1 w$ based on Lemma 1, gives $(w, b) \succ^1 (w, w)$. □

From now on δ refers to the one from Lemma 3.

Lemma 4. $U_2(x_1, x_2) = U_1(x_1) + \delta U_1(x_2)$.

Proof. First we derive the values for $U_2(w, w)$, $U_2(w, b)$, $U_2(b, w)$ and $U_2(b, b)$. $U_2(w, w) = 0$ and $U_2(b, w) = 1$ due to Lemma 2. $U_2(w, b) = \delta$ follows from Lemma 3. We define $U_2(b, b) := \alpha$ for simplicity and state

$$(w, b) \sim^2 \frac{\delta}{\alpha}\left[(b, b)\right] + \left(1 - \frac{\delta}{\alpha}\right)\left[(w, w)\right], \tag{3}$$

which is justified by the fact that expected utility of both sides is δ. We first add b to the beginning of (2) based on Lemma 1 to get

$$(b, w, b) \sim^3 \frac{\delta}{\alpha}\left[(b, b, b)\right] + \left(1 - \frac{\delta}{\alpha}\right)\left[(b, w, w)\right] \tag{4}$$

and then add w to the beginning of both sides of (2) to get

$$(w,w,b) \sim^3 \frac{\delta}{\alpha}\left[(w,b,b)\right] + \left(1-\frac{\delta}{\alpha}\right)\left[(w,w,w)\right]. \tag{5}$$

A distribution that gives the left side of (3) with probability δ^2 and the left side of (4) otherwise must be indifferent to a distribution that gives the right side of (3) with probability δ^2 and the right side of (4) otherwise, since expected utility of both distributions is the same. Therefore

$$\begin{aligned}\delta^2\left[(b,w,b)\right] + \left(1-\delta^2\right)\left[(w,w,b)\right] \sim^3& \\ &\delta^2\left(\frac{\delta}{\alpha}\left[(b,b,b)\right] + \left(1-\frac{\delta}{\alpha}\right)\left[(b,w,w)\right]\right) + \\ &\left(1-\delta^2\right)\left(\frac{\delta}{\alpha}\left[(w,b,b)\right] + \left(1-\frac{\delta}{\alpha}\right)\left[(w,w,w)\right]\right) = \\ &\frac{\delta}{\alpha}\left(\delta^2\left[(b,b,b)\right] + \left(1-\delta^2\right)\left[(w,b,b)\right]\right) + \\ &\left(1-\frac{\delta}{\alpha}\right)\left(\delta^2\left[(b,w,w)\right] + \left(1-\delta^2\right)\left[(w,w,w)\right]\right).\end{aligned} \tag{6}$$

Next we transform the right side of (5) such that every sequences ends with b.

$$\begin{aligned}\delta^2\left[(b,w,w)\right] + \left(1-\delta^2\right)\left[(w,w,w)\right] =& \\ &\delta\left(\delta\left[(b,w,w)\right] + (1-\delta)\left[(w,w,w)\right]\right) + (1-\delta)\left[(w,w,w)\right]\end{aligned} \tag{7}$$

and $\delta\left[(b,w,w)\right] + (1-\delta)\left[(w,w,w)\right] \sim^3 (w,b,w)$, which follows from Lemma 3 when w is added to the end based on Assumption 9. We substitute the term in (6) and derive

$$\delta^2\left[(b,w,w)\right] + \left(1-\delta^2\right)\left[(w,w,w)\right] \sim^3 \delta\left[(w,b,w)\right] + (1-\delta)\left[(w,w,w)\right]. \tag{8}$$

$\delta\left[(w,b,w)\right] + (1-\delta)\left[(w,w,w)\right] \sim^3 (w,w,b)$ follows from Lemma 3 when w is added to the beginning based on Lemma 1. We substitute the term in (7), further substitute this result in (5) and find

$$\begin{aligned}\delta^2\left[(b,w,b)\right] + \left(1-\delta^2\right)\left[(w,w,b)\right] \sim^3& \\ &\frac{\delta}{\alpha}\left(\delta^2\left[(b,b,b)\right] + \left(1-\delta^2\right)\left[(w,b,b)\right]\right) + \left(1-\frac{\delta}{\alpha}\right)\left[(w,w,b)\right].\end{aligned} \tag{9}$$

b can now be removed from the end of (8) based on Assumption 9 and we are left with

$$\begin{aligned}\delta^2\left[(b,w)\right] + \left(1-\delta^2\right)\left[(w,w)\right] \sim^3& \\ &\frac{\delta}{\alpha}\left(\delta^2\left[(b,b)\right] + \left(1-\delta^2\right)\left[(w,b)\right]\right) + \left(1-\frac{\delta}{\alpha}\right)\left[(w,w)\right].\end{aligned} \tag{10}$$

We equate the expected utility of both side of (9) and solve for α, which gives

$$\alpha = U_2(b,b) = 1 + \delta. \tag{11}$$

This concludes the first step of the proof. Next consider the following indifference relations,

$$x_1 \sim^1 U_1(x_1)[b] + (1 - U_1(x_1))[w] \tag{12}$$

and

$$x_2 \sim^1 U_1(x_2)[b] + (1 - U_1(x_2))[w]. \tag{13}$$

Adding x_2 to the end of (11) based on Assumption 9 gives

$$(x_1, x_2) \sim^1 U_1(x_1)[(b, x_2)] + (1 - U_1(x_1))[(w, x_2)]. \tag{14}$$

Indifference relations for (b, x_2) and (w, x_2) can be found by adding b to the beginning of (12) and w to the beginning of (12) based on Lemma 1. We substitute the results in (13), which gives

$$\begin{aligned}(x_1, x_2) \sim^2 {} & U_1(x_1)U_1(x_2)[(b,b)] + U_1(x_1)(1 - U_1(x_2))[(b,w)] + \\ & U_1(x_2)(1 - U_1(x_1))[(w,b)] + (1 - U_1(x_1))(1 - U_1(x_2))[(w,w)].\end{aligned} \tag{15}$$

Finally, we equate the expected utility of both sides of (14) and it follows that

$$U_2(x_1, x_2) = U_1(x_1) + \delta U_1(x_2). \tag{16}$$

□

Lemma 5. *$U_T(x_1, ..., x_T) = \sum_{t=1}^{T} \delta^{t-1} U_1(x_t)$ for all $T \in \mathbb{N}$.*

Proof. We prove this by induction. It is obvious that the proposition holds for $T = 1$. The proposition holds for $T = 2$ by Lemma 4. Therefore, we only have to show that if the proposition holds for some general T it also holds for $T + 1$. Note that Lemma 2 defines $U_{T+1}(b, w, ..., w) = 1$ and $U_{T+1}(w, ..., w) = 0$. We consider (11), add w to the end T times based on Assumption 9 and equate the expected utility of both sides to receive

$$U_{T+1}(x_1, w, ..., w) = U_1(x_1). \tag{17}$$

Next consider

$$(x_1, ..., x_T) \sim^T \frac{U_T(x_1, ..., x_T)}{U_T(b, ..., b)}[(b, ..., b)] + \left(1 - \frac{U_T(x_1, ..., x_T)}{U_T(b, ..., b)}\right)[(w, ..., w)], \tag{18}$$

which must hold since expected utility of both sides is equal. Adding w to the end of (17) based on Assumption 9 gives

$$(x_1, ..., x_T, w) \sim^{T+1} \frac{U_T(x_1, ..., x_T)}{U_T(b, ..., b)}[(b, ..., b, w)] + \left(1 - \frac{U_T(x_1, ..., x_T)}{U_T(b, ..., b)}\right)[(w, ..., w)]. \tag{19}$$

Note that $U_T(x_1,...,x_T) = \sum_{t=1}^{T} \delta^{t-1} U_1(x_t)$ by assumption and $U_1(b) = 1$, $U_1(w) = 0$ by Lemma 2. We equate the expected utility of both sides of (18), set $x_2,...,x_T = w$ and can solve for $U_{T+1}(b,...,b,w)$ due to (16). Equating expected utility of both sides of (18) once again, but leaving $x_2,...,x_T$ general, gives

$$U_{T+1}(x_1,...,x_T,w) = U_T(x_1,...,x_T). \tag{20}$$

Adding x_0 to the beginning of (17) based on Lemma 1 gives

$$(x_0,...,x_T) \sim^{T+1} \frac{U_T(x_1,...,x_T)}{U_T(b,...,b)} [(x_0,b,...,b)] + \left(1 - \frac{U_T(x_1,...,x_T)}{U_T(b,...,b)}\right) [(x_0,w,...,w)]. \tag{21}$$

We equate the expected utility of both sides of (20), set $x_T = w$ and can solve for $U_{T+1}(x_0,b,...,b)$ due to (16) and (19). Equating expected utility of both sides of (20) once again, but leaving x_T general, gives

$$U_{T+1}(x_0,...,x_T) = U_1(x_0) + \delta U_T(x_1,...,x_T) = \sum_{t=0}^{T} \delta^t U_1(x_t) \tag{22}$$

and therefore

$$U_{T+1}(x_1,...,x_{T+1}) = \sum_{t=1}^{T+1} \delta^{t-1} U_1(x_t). \tag{23}$$

$\square$

5 Infinite Horizon

We now consider a decision problem without horizon such that there is an infinite number of future time periods $t \in \mathbb{N}$. The set of distribution of sequences of outcomes is $\triangle(A^\infty)$ and $\succcurlyeq^\infty$ is a binary relation on $\triangle(A^\infty)$ representing the agents preferences. In the following let $(x_t)_{t=1}^{\infty} \in A^\infty$ be an infinite sequence of outcomes and $(x_t)_{t=1}^{T} \in A^T$ the truncated sequence of $(x_t)_{t=1}^{\infty}$ after period $T \in \mathbb{N}$, such that the first T entries of $(x_t)_{t=1}^{\infty}$ are identical to $(x_t)_{t=1}^{T}$. Note that although we denote sequences of outcomes, we actually refer to distributions where those sequences are realized with probability 1.

Definition 4. *Let $\lambda_x(T)$ be defined by $(x_t)_{t=1}^{T} \sim^T \lambda_x(T) \left[(b)_{t=1}^{T}\right] + (1-\lambda_x(T)) \left[(w)_{t=1}^{T}\right]$ for any T and $(x_t)_{t=1}^{\infty}$.*

We know that $\lambda_x(T)$ must exist and be unique for any T and $(x_t)_{t=1}^{\infty}$ due to Assumption 3 and the fact that $U_T(b,...,b) = \frac{1-\delta^T}{1-\delta}$, $U_T(w,...,w) = 0$ and $\frac{1-\delta^T}{1-\delta} \geq U_T(x_1,...,x_T) \geq 0$ by Theorem 1.

Assumption 10. $(x_t)_{t=1}^{\infty} \succsim^{\infty} (y_t)_{t=1}^{\infty}$ *if and only if there exists no* $\epsilon \in (0, \infty)$ *and* $M \in \mathbb{N}$ *such that* $\lambda_y(T) - \lambda_x(T) > \epsilon$ *for all* $T \geq M$.

Note that the λ's are proportional to the utility of their sequence. Therefore an infinite sequence $(x_t)_{t=1}^{\infty}$ is weakly preferred to another infinite sequence $(y_t)_{t=1}^{\infty}$ either if $(x_t)_{t=1}^{T}$ is strictly preferred to $(y_t)_{t=1}^{T}$ for all large enough T or if at least the difference between the utility of $(x_t)_{t=1}^{T}$ and $(y_t)_{t=1}^{T}$ goes to zero as T becomes large. Assumption 10 coincides with the commonly used limit of discounted sum of utilities $\sum_{t=1}^{\infty} \delta^{t-1} U_1(x_t)$, which we will prove in the following proposition. Note that U_T and δ still refer to the ones defined in Lemma 2 and 3 respectively.

Proposition 1. *If preferences* $\succsim$ *satisfy Assumptions 1-10, then if* $\sum_{t=1}^{\infty} \delta^{t-1} U_1(y_t) > \sum_{t=1}^{\infty} \delta^{t-1} U_1(x_t)$ *then* $(y_t)_{t=1}^{\infty} \succ^{\infty} (x_t)_{t=1}^{\infty}$ *and if* $\sum_{t=1}^{\infty} \delta^{t-1} U_1(x_t) = \sum_{t=1}^{\infty} \delta^{t-1} U_1(y_t)$ *then* $(x_t)_{t=1}^{\infty} \sim^{\infty} (y_t)_{t=1}^{\infty}$.

Proof. First we prove that if $\sum_{t=1}^{\infty} \delta^{t-1} U_1(y_t) > \sum_{t=1}^{\infty} \delta^{t-1} U_1(x_t)$ then $(y_t)_{t=1}^{\infty} \succ^{\infty} (x_t)_{t=1}^{\infty}$. We derive $U_T\left((x_t)_{t=1}^{T}\right) = \lambda_x(T)\, U_T\left((b)_{t=1}^{T}\right)$ by taking expected utility of both sides of the relation in Definition 3. Since $\sum_{t=1}^{\infty} \delta^{t-1} U_1(x_t) = \lim_{T\to\infty} U_T\left((x_t)_{t=1}^{T}\right)$ by Proposition 1 we can reformulate the inequality $\sum_{t=1}^{\infty} \delta^{t-1} U_1(y_t) > \sum_{t=1}^{\infty} \delta^{t-1} U_1(x_t)$ in terms of λ and receive $\lim_{T\to\infty} \lambda_y(T) > \lim_{T\to\infty} \lambda_x(T)$. Note that the limit of $U_T\left((b)_{t=1}^{T}\right)$ is finite and cancels out. We define $\lambda_x := \lim_{T\to\infty} \lambda_x(T)$. By the definition of the limit, there must exist an $\epsilon < \frac{\lambda_y - \lambda_x}{2}$ and some $M \in \mathbb{N}$ such that $|\lambda_y(T) - \lambda_y| < \epsilon$ for all $T \geq M$. The same holds for λ_x. Therefore there exists an ϵ and M such that $\lambda_y(T) - \lambda_x(T) > \epsilon$ for all $T \geq M$. By Assumption 10, $(x_t)_{t=1}^{\infty} \succsim^{\infty} (y_t)_{t=1}^{\infty}$ is ruled out and therefore $(y_t)_{t=1}^{\infty} \succ^{\infty} (x_t)_{t=1}^{\infty}$. This concludes the first part of the proof. Next we prove that if $\sum_{t=1}^{\infty} \delta^{t-1} U_1(x_t) = \sum_{t=1}^{\infty} \delta^{t-1} U_1(y_t)$ then $(x_t)_{t=1}^{\infty} \sim^{\infty} (y_t)_{t=1}^{\infty}$. We assume that an ϵ and M exists such that $\lambda_y(T) - \lambda_x(T) > \epsilon$ for all $T \geq M$ and find a contradiction. From the first part of the proof we know that $\sum_{t=1}^{\infty} \delta^{t-1} U_1(x_t) = \sum_{t=1}^{\infty} \delta^{t-1} U_1(y_t)$ is equivalent to $\lambda_x = \lambda_y$. By the definition of the limit, there must exist an $\epsilon' < \frac{\lambda_y(M) - \lambda_x(M)}{2}$ and some $N \in \mathbb{N}$ such that $|\lambda_y(T) - \lambda_y| < \epsilon'$ for all $T \geq N$. The same holds for λ_x. Therefore $\lambda_y(N) - \lambda_x(N) < \epsilon' < \epsilon$, which contradicts the assumption that $\lambda_y(T) - \lambda_x(T) > \epsilon$ for all $T \geq M$. It follows from Assumption 10 that $(x_t)_{t=1}^{\infty} \succsim^{\infty} (y_t)_{t=1}^{\infty}$. By the same argument it follows that there exists no ϵ and M such that $\lambda_x(T) - \lambda_y(T) > \epsilon$ for all $T \geq M$ and therefore it follows from Assumption 10 that $(y_t)_{t=1}^{\infty} \succsim^{\infty} (x_t)_{t=1}^{\infty}$. This rules out a strict preference and therefore $(y_t)_{t=1}^{\infty} \sim^{\infty} (x_t)_{t=1}^{\infty}$. $\square$

6 References

Anchugina, N. (2016). A simple framework for the axiomatization of exponential and quasi-hyperbolic discounting. *Theory and Decision*, 82(2), pp.185-210.

Epstein, L. (1983). Stationary cardinal utility and optimal growth under uncertainty. *Journal of Economic Theory*, 31(1), pp.133-152.

Halevy, Y. (2015). Time Consistency: Stationarity and Time Invariance. *Econometrica*, 83(1), pp.335-352.

Koopmans, T. (1960). *Stationary Ordinal Utility and Impatience.* Econometrica, 28(2), p.287.

Kreps, D. and Porteus, E. (1978). Temporal Resolution of Uncertainty and Dynamic Choice Theory. *Econometrica*, 46(1), p.185.

Samuelson, P. (1937). A Note on Measurement of Utility. *The Review of Economic Studies*, 4(2), p.155.

Strotz, R. (1955). Myopia and Inconsistency in Dynamic Utility Maximization. *The Review of Economic Studies*, 23(3), p.165.

Von Neumann, J. and Morgenstern, O. (1944). *Theory of games and economic behavior.* Princeton: Princeton University Press.

7 Appendix

Lemma 6. *Let $\succsim^x$ be a preference relation on the set X and $\succsim^y$ be a preference relation on the set Y. The following statements are equivalent if $\succsim^x$ and $\succsim^y$ are complete:*

(a) $x_1 \succsim^x x_2$ if and only if $y_1 \succsim^y y_2$, where $x_1, x_1 \in X$ and $y_1, y_1 \in Y$.

(b) If $x_1 \succ^x x_2$ then $y_1 \succ^y y_2$ and if $x_1 \sim^x x_2$ then $y_1 \sim^y y_2$, where $x_1, x_1 \in X$ and $y_1, y_1 \in Y$.

Proof. First we show that (a) implies (b). Assume $x_1 \succ^x x_2$. It follows from (a) that $y_1 \succsim^y y_2$. $y_1 \sim^y y_2$ is not possible because it would imply $x_2 \succsim^x x_1$ by (a). Therefore $y_1 \succ^y y_2$. Next assume $x_1 \sim^x x_2$. This implies both $y_1 \succsim^y y_2$ and $y_2 \succsim^y y_1$ by (a). Therefore $y_1 \sim^x y_2$. This concludes the first part of the proof. Next we show that (b) implies (a). Assume $y_1 \succ^y y_2$. $x_1 \sim^x x_2$ is not possible as it implies $y_1 \sim^y y_2$

by (b). $x_2 \succ^x x_1$ is not possible as it implies $y_2 \succ^y y_1$ by (b). Therefore, the only possibility left is $x_1 \succ^x x_2$. Next assume $y_1 \sim^y y_2$. Neither $x_1 \succ^x x_2$ nor $x_2 \succ^x x_1$ are possible, as they would imply a strict preference for y_1 and y_2 by (b). Therefore, the only possibility left is $x_1 \sim^x x_2$. $\square$

Zusammenfassung

Ziel dieser Arbeit ist es, die von-Neumann-Morgenstern Bewertung von Lotterien über einzelne Alternativen auf die von Lotterien über endliche Sequenzen von Alternativen zu erweitern. Wir liefern eine Axiomatisierung, die zu dem typischen Ansatz führt, Sequenzen durch die exponentiell abgezinste Summe von von-Neumann-Morgenstern Nutzen pro Periode zu vergleichen. Die meisten unserer zusätzlichen Axiome, wie Zeitkonsistenz, Zeitinvarianz und Planung in Etappen, sind durch dynamische Entscheidungen motiviert. Anders als die meisten Axiomatisierungen für die exponentielle Diskontierung, siehe Koopmans (1960) und Epstein (1983), benötigen wir weder technische Annahmen, die die Kontinuität der Nutzenfunktion in den Ergebnissen sicherstellen, noch die Verwendung der typischen Existenztheoreme für additiven Nutzen. Unser Beweis ist konstruktiv und einfach zu verstehen.

Schlagwörter: von-Neumann-Morgenstern, exponentiell, Diskontierung, Lotterien, endlich, Sequenzen, Zeitkonsistenz