Economic Theory (2023) 75:1071–1098
https://doi.org/10.1007/s00199-022-01438-0

RESEARCH ARTICLE

Choquet expected discounted utility

Lorenzo Bastianello[1] · **José Heleno Faro**[2]

Received: 16 March 2021 / Accepted: 2 June 2022 / Published online: 30 June 2022

Abstract

We study intertemporal decision making under uncertainty in a purely subjective framework. The concept of stationarity, introduced by Koopmans for deterministic discounted utility, is naturally extended to a framework with uncertainty and plays a central role for both attitudes towards time and uncertainty. We show that a strong stationarity axiom characterizes discounted expected utility. When considerations about correlations across time between uncertain outcomes are taken into account, a weaker stationarity axiom generalizes discounted expected utility to Choquet expected discounted utility, allowing for non-neutral attitudes towards subjective uncertainty.

Keywords Ambiguity · Intertemporal choice · (Choquet) Discounted expected utility · Stationarity · Serial correlation · Uncertainty aversion/seeking

JEL Classification D81 · D83 · D84

1 Introduction

When making economic decisions, agents usually need to take into account two fundamental dimensions that affect the outcome of their choice: *time* and *uncertainty*. Consider for instance a firm that wants to implement a project that will deliver a stochastic cash flow in the future or a government that needs to decide how to allocate its budget taking into account GDP growth in the following years. In both cases, decision makers (DMs henceforth) are required to make choices that involve uncertain outcomes occurring at future dates.

✉ Lorenzo Bastianello
lorenzo.bastianello@u-paris2.fr

José Heleno Faro
josehf@insper.edu.br

1 LEMMA, Université Paris 2 Panthéon-Assas, Paris, France

2 Insper, São Paulo, Brazil

One of the most popular models used in economics to evaluate uncertain streams of income or consumption is the *(exponential) discounted expected utility model*. The choice problem faced by the firm or the government can be formalized in the following way. Suppose that, at time $t = 0$, Nature chooses a state of the world $\omega \in \Omega$ which is hidden to the agent. At the same time, the agent should choose an alternative, called act, denoted by $h := (h_0, h_1, \dots)$. Each state of the world determines a path describing the amounts of income that the DM will receive across future periods, denoted by $h(\omega) := (h_0(\omega), h_1(\omega), \dots)$. Thus, given h and $\omega \in \Omega$, she gets $h_t(\omega)$ at time $t \geq 0$. The discounted expected utility model asserts that the stochastic stream $h := (h_0, h_1, \dots)$ is evaluated through the intertemporal utility function

$$V(h) = \sum_{t=0}^{\infty} \beta^t \mathbb{E}_P[u(h_t)], \tag{1}$$

where $u(\cdot)$ is an instantaneous utility index, the DM's attitude towards time is described through the discount factor β^t with $\beta \in (0, 1)$ and the expectation is taken with respect to a subjective probability P over Ω. Mathematically, the expression in (1) prescribes to compute the expected utility of each random variable h_t, actualize its value through the discount factor β^t and sum up the actualized expected utilities.

Exponential discounted utility was first axiomatized by Koopmans (1960, 1972) in the deterministic framework. We propose a generalization of his *stationarity* condition under uncertainty in order to characterize the formula in (1).[1] Roughly speaking, our axiom requires that a DM who is indifferent between two uncertain streams f and g should remain indifferent if, given an arbitrary period t, the random incomes of both streams are shifted one period ahead starting from t, and *the same* stochastic amount of income h_t is inserted in period t.

As it turns out, the implications of the stationarity axiom in presence of uncertainty are stronger than in the deterministic framework. Consider a firm that needs to decide today whether to invest in solar (s) technology or to build a carbon (c) plan. In year 2022 there will be elections (this is the source of uncertainty). If democrats win (event D), they will subsidize firms that invested in green technologies. If republicans win (event R), they will subsidize traditional ways of producing energy. The firm has therefore to choose between the two streams

$$\begin{array}{lccccc} & 2021 & 2022 & 2023 & 2024 & \dots \\ s_0, s_1, \dots \equiv & 0 & s_1 \begin{cases} 10 & \text{if } D \\ 0 & \text{if } R \end{cases} & 0 & \dots & \dots \\ c_0, c_1, \dots \equiv & 0 & c_1 \begin{cases} 0 & \text{if } D \\ 10 & \text{if } R \end{cases} & 0 & \dots & \dots \end{array}$$

If the manager of the firm thinks that democrats and republicans are equally likely to win the elections, then it seems reasonable to be indifferent between s and c. Suppose

[1] For an axiomatization of the discounted expected utility model with an infinite horizon *under risk* one can see Epstein (1983) and Peitler (2019). Under uncertainty, a charcterization of (1) can be derived from Kochov (2015), see discussions after Theorems 1 and 5.

now that the republicans are the incumbent. For budget reasons they are obliged to postpone any subsidy in favor of the technological sector to 2023, but they promise to reduces taxes on firms in 2022 in case they win the elections. Taking into account the shift of subsidies and tax reductions, the streams now look like

$$\begin{array}{cccccc} & 2021 & 2022 & 2023 & 2024 & \dots \\ s_0, h_1, s_1, \dots \equiv & 0 & h_1 \begin{cases} 0 & \text{if } D \\ 7 & \text{if } R \end{cases} & s_1 \begin{cases} 10 & \text{if } D \\ 0 & \text{if } R \end{cases} & 0 & \dots \\ c_0, h_1, c_1, \dots \equiv & 0 & h_1 \begin{cases} 0 & \text{if } D \\ 7 & \text{if } R \end{cases} & c_1 \begin{cases} 0 & \text{if } D \\ 10 & \text{if } R \end{cases} & 0 & \dots \end{array}$$

If stationarity holds, than the firm should remain indifferent between (the modified version of) s and c. The rationale being that in year 2022 the firm gets h_1 whatever the investment decision, and from 2023 on, the streams are the same as in the previous decision context.

However note that choosing to invest in carbon looks "more uncertain". In fact, if democrats win, the firm will not get anything for two consecutive years. On the other hand, investing on solar would allow the firm to hedge against any possible winner. If democrats win, the firm gets 10 in 2023 and if republicans win it gets 7 in 2022: whatever the winner, the firm has a positive result in at least one period. Choosing solar offers a possible hedge against the uncertainty of the electoral result and therefore one may expect that a cautious (or pessimistic) manager strictly prefers s to c. Note that, a confident (or optimistic) manager may very well prefer to invest in carbon rather than in solar.

The central idea is that h_1 is positively correlated with c_1 and negatively correlated with s_1. An uncertainty averse DM would prefer negatively correlated variables following one another in order to protect herself against uncertainty. At the same time, an uncertainty loving DM would appreciate positively correlated variables one after the other. Finally, it can also be the case that a DM would sometimes prefer negatively correlated variables and sometimes positively correlated variables following one another. Imposing stationarity when such considerations can be done may be demanding, and this is why, under uncertainty, we name this property *Strong Stationarity*.

The example above shows a particular case of a more general concept. Loosely speaking, two random variables ϕ, ψ are comonotonic if they "vary in the same direction": when ϕ takes relative high (low) values, ψ also takes relative high (low) values. For instance, in the example above, h_1 is comonotonic with c_1 but not with s_1. Comonotonicity is strictly linked to correlation, and DMs may have positive or negative attitude towards correlated variables. This pattern of correlation across time is called autocorrelation, serial correlation or temporal correlation. Several authors already remarked that attitude towards correlation is an important behavioral feature. For an early treatment of autocorrelation in the context of lotteries, see Epstein and Tanny (1980).[2] Recent papers studying this issue are Kochov (2015) and Bommier et al. (2017).

[2] In finance, the fact that the return on a stock presents serial correlation is a violation of the weak form of market-efficiency. See, for instance, Rosenberg and Rudd (1982) and Jegadeesh (1990).

In order to cope with possible non-neutral attitudes towards this kind of temporal correlations, we introduce a generalization of Strong Stationarity, called *Comonotonic Stationarity*, that restricts the shifts prescribed by Strong Stationarity only to cases in which considerations about correlations do not play a role. Technically, we require the stochastic income plugged at period t to be comonotonic with the variables that follow in both sequences. If variables involved in the shifts are comonotonic, there is no modification of the perceived uncertainty inherent to a sequence. Theorem 2, which is our main contribution, shows formally that substituting Strong Stationarity with Comonotonic Stationarity makes it possible to generalize discounted expected utility to the following representation

$$V(h) = \int \sum_{t=0}^{\infty} \beta^t u(h_t) dv. \tag{2}$$

The integral is a *Choquet integral*, taken with respect to a *capacity* v. While a precise mathematical definition will be given later, we just recall here that a capacity is a non-necessarily additive set function and the Choquet integral is a mathematical tool that allows to compute expectations with respect to it.

The Choquet expected utility model was introduced in economics in the atemporal setting proposed by Anscombe and Aumann (1963) in the seminal paper of Schmeidler (1989). This model generalizes expected utility and solves the famous Ellsberg (1961) paradox, which shows that agents cannot quantify uncertainty in terms of a (additive) probability measure. In this sense, our model of Choquet expected discounted utility in (2) parallels Schmeidler's work in the atemporal setting. Therefore, this paper provides novel foundations for two fundamental models of decision making: Choquet expected utility and discounted expected utility. Discounted expected utility is characterized by Strong Stationarity (together with other basic axioms) which implies a form of uncertainty neutrality. Our more general axiom, Comonotonic Stationarity, reflects the idea that agents' behaviour is affected by intertemporal correlations between random variables. The Choquet expected discounted utility model is a flexible tool that can be applied in order to study how serial correlation aversion or serial correlation seeking (or both at the same time) impact agents' decisions.

There is a rich literature concerning decision making under uncertainty, and several models have been proposed to address the Ellsberg paradox. A popular way to take into account agents' behavior towards uncertainty is through the MaxMin expected utility model of Gilboa and Schmeidler (1989). This model, developed in the atemporal setting, says that a DM has a set of probabilities in her mind and takes the minimum expected utility calculated with respect to the probabilities in this set. Our paper is closely related to a recent article of Kochov (2015), who provides a generalization of discounted expected utility (1) by considering an intertemporal version of the MaxMin expected utility model of Gilboa and Schmeidler (1989). In our paper we use the same framework of Kochov (2015) and we show that changing his axioms of Intertemporal Hedging and Path Stationarity into Comonotonic Stationarity allows us to obtain the Choquet expected discounted utility model. The importance of Comonotonic Stationarity is revealed by the weaknesses of Strong Stationarity (as explained in the beginning of this section and in Example 1). We will discuss further this relevant article in the main body of our paper.

Finally we recall that there are several interesting ways to discount the future, besides exponential discounting, that are outside the scope of this paper. While exponential discounting may be appropriate in the case of a firm, who can set $\beta = \frac{1}{1+r}$ where r is the interest rate, it may not be so compelling for other types of agents. For instance, interesting alternatives are the hyperbolic, quasi-hyperbolic, quasi-geometric and constant-sensitivity discounting, see Loewenstein and Prelec (1992), Laibson (1997), Montiel Olea and Strzalecki (2014), Phelps and Pollak (1968), Hayashi (2003) and Ebert and Prelec (2007). Another interesting approach appears in Chambers and Echenique (2018) where the authors study models with multiple discount factors. The analysis of this paper is bounded to exponential discounting in order to give a generalization of under uncertainty that is as close at possible to Koopmans' original model.

The rest of the paper is organized as follows. Section 2 introduces the framework and notation. In Sect. 3 we characterize the discounted expected utility model. Section 4 generalizes the previous section and contains our main result, the Choquet expected discounted utility model. Sections 5 treats the case of uncertainty aversion. We provide some additional discussions in Sect. 6. Section 7 concludes. All proofs are gathered in the Appendix.

2 Framework and mathematical preliminaries

Time is discrete and identified with $\mathbb{N} = \{0, 1, 2, \dots\}$. Let X be a connected, separable and first-countable topological space. It will be interpreted as the *space of outcomes*. For instance, if X is a convex subset of $\mathbb{R}^n$, then $x \in X$ may represent a bundle of goods. Let Ω be a non-empty set of *states of nature*. A *filtration* $(\mathcal{F}_t)_t$ over Ω is a sequence of algebras such that $\mathcal{F}_0 = \{\emptyset, \Omega\}$[3] and $\mathcal{F}_t \subseteq \mathcal{F}_{t+1}$ for all $t \in \mathbb{N}$. We denote by $\mathcal{F}$ the union of these algebras $\mathcal{F} := \cup_t \mathcal{F}_t$. A *stochastic process* $h := (h_t)_{t\in\mathbb{N}}$ is a sequence such that h_t is a $\mathcal{F}_t$-measurable function from Ω to X for all t. We sometimes call measurable functions *random variables*. The following technical assumption restricts the set of stochastic processes that we consider.

Assumption. Stochastic processes are *bounded* and *finite*. Boundedness means that for each act h there exists a compact set $K_h \subset X$ such that $\cup_t h_t(\Omega) \subset K_h$. Finiteness means that for each act h there is a finitely generated algebra $\mathcal{A}_h \subset \mathcal{F}$ such that h_t is $\mathcal{A}_h$-measurable for all $t \in \mathbb{N}$.

We denote by $\mathcal{H}$ the set of bounded and finite stochastic processes

$$\mathcal{H} := \{h = (h_t)_{t\in\mathbb{N}} | h_t : \Omega \to X, h_t \text{ is } \mathcal{F}_t\text{-measurable } \forall t \text{ and } h \text{ is finite and bounded}\}.$$

A sequence $h \in \mathcal{H}$ will be called *act*. The set $\mathcal{D} \subset \mathcal{H}$ denotes the set of *deterministic acts*: we have $d \in \mathcal{D}$ if and only if d_t is a constant random variable for all $t \in \mathbb{N}$ and

[3] $\mathcal{F}_0 = \{\emptyset, \Omega\}$ is a standard requirement in the literature. However, one can consider a larger algebra if the DM already has some information at $t = 0$.

there exists a compact set $K_d \subseteq X$ such that $d_t \in K_d$ for all $t \geq 0$. As usual we identify $\mathcal{D}$ with a subset of X^∞. For example, if $X = \mathbb{R}$ then $\mathcal{D} = l^\infty$, the set of real-valued bounded sequences. When $x \in X$ and $d \in \mathcal{D}$, (x, d) denotes the act $(x, d_0, d_1, \dots)$. Obviously the procedure can be repeated as in $(x, y, d) = (x, y, d_0, d_1, \dots)$ and so on. Note that we do not invoke the convex structure of a mixture space as in the standard Anscombe and Aumann (1963) setting. Instead, we use a purely subjective setup with some measurability and topological assumptions. This framework with the restrictions that defines $\mathcal{H}$ was proposed by Kochov (2015), which is a relatively standard setting that appears for instance (up to small differences) in Epstein and Wang (1995).

A *(normalized) capacity* v on the measurable space $(\Omega, \mathcal{F})$ is a set function $v : \mathcal{F} \mapsto \mathbb{R}$ such that $v(\emptyset) = 0$, $v(\Omega) = 1$ and for all $A, B \in \mathcal{F}$, $A \subset B \Rightarrow v(A) \leq v(B)$. Given a capacity v on $(\Omega, \mathcal{F})$, the *Choquet integral* of a real-valued, bounded, $\mathcal{F}$-measurable function $f : \Omega \to \mathbb{R}$ with respect to v is defined as

$$\int_\Omega f\, dv := \int_{-\infty}^{0} (v(\{f \geq t\}) - 1)\, dt + \int_0^{+\infty} v(\{f \geq t\})\, dt,$$

where $\{f \geq t\} = \{\omega \in \Omega | f(\omega) \geq t\}$. A capacity $v : \mathcal{F} \mapsto \mathbb{R}$ is *convex* if, for all $A, B \in \mathcal{F}$, $v(A \cup B) + v(A \cap B) \geq v(A) + v(B)$. A *(finitely additive) probability* $P : \mathcal{F} \mapsto \mathbb{R}$ is a capacity such that $A \cap B = \emptyset$ implies $P(A \cup B) = P(A) + P(B)$. The *core* of a capacity v is defined by $core(v) = \{P | P \text{ is a probability s.t. } P(A) \geq v(A)\, \forall A \in \mathcal{F}\}$. Finally, if P is a probability, then we denote the integral with respect to P of a real-valued, bounded, $\mathcal{F}$-measurable function f with the usual notation for expectation $\int_\Omega f\, dP = \mathbb{E}_P[f]$.

3 Uncertainty and discounted expected utility

Section 3 characterizes the discounted expected utility model. The main novelty that we introduce is an independence condition, called Strong Stationarity, that generalizes in the context of uncertainty Koopmans' original stationarity axiom. This section motivates the results of Sect. 4 in which we argue that Strong Stationarity is actually too restrictive and we propose a generalization.

Let $\succsim$ be a non-trivial (i.e. with a non-empty strict part $\succ$), complete and transitive binary relation over $\mathcal{H}$. Recall that an act $h \in \mathcal{H}$ is a bounded and finite stochastic process: at time $t = 0$, Nature chooses a state $\omega \in \Omega$ (hidden to the DM at any finite time t) and the DM receives $h_t(\omega)$ in period t. The relation $\succsim$ represents the preferences of the DM over acts.[4] While repeated choices are outside the scope of this paper, as we want to stick to Koopmans (1972) original framework, we discuss them briefly in Sect. 6.1. Preferences over outcomes in X are defined in the usual way, i.e. $x \succsim y$ if and only if $(x, x, \dots) \succsim (y, y, \dots)$.

The following three axioms are well known in the literature and are used for instance in Koopmans (1972), Bleichrodt et al. (2008) and Kochov (2015). We refer to them as *basic axioms*.

[4] Note that the DM is choosing only at time $t = 0$. As Koopmans' (1972) put it: "no question of consistency or inconsistency of orderings adopted at different points in time is raised".

CONTINUITY (C) For all compact sets $K \subset X$ and for all acts $h \in \mathcal{H}$ the sets $\{d \in K^{\infty} | d \succsim h\}$ and $\{d \in K^{\infty} | h \succsim d\}$ are closed in the product topology over K^{∞}.

When X is a subset of $\mathbb{R}^n$, continuity with respect to the product topology is a stronger requirement than continuity with respect to the sup-norm topology, as postulated by Koopmans (1972). However, as noted in Kochov (2015), axiom (C) (together with the other axioms and non-triviality of the preference relation) makes it possible to drop Koopmans' metrizability assumption and his postulates P.2 and P.5. We give a proof of this result in Proposition 4 in Appendix A.1 for sake of completeness.

The next axiom, Time Separability, is exactly the same as Koopmans' postulate P.3, see Appendix A.1. Several authors drop this axiom in order to get endogenous discounting, see for instance Epstein (1983) and Bommier et al. (2019). We keep Time Separability in order to obtain a constant exponential discount factor.

TIME SEPARABILITY (TS) For all $x, y, x', y' \in X$ and $d, d' \in \mathcal{D}$, $(x, y, d) \succsim (x', y', d)$ if and only if $(x, y, d') \succsim (x', y', d')$.

Before introducing Monotonicity we need a piece of notation. Let $h \in \mathcal{H}$ and $\omega \in \Omega$, then $h(\omega)$ denotes the (deterministic) sequence $(h_0(\omega), h_1(\omega), \dots) \in \mathcal{D}$.

MONOTONICITY (M) For all $h, g \in \mathcal{H}$, if $h(\omega) \succsim g(\omega)$ $\forall \omega \in \Omega$ then $h \succsim g$.

Suppose that the DM prefers the deterministic act $h(\omega)$ rather than the deterministic act $g(\omega)$ for every possible choice ω of Nature. Then (M) says that she should prefer h to g.

The last axiom, Strong Stationarity, represents the main novelty of this section. Let us first introduce the original stationarity axiom, proposed by Koopmans in the deterministic setting, in order to make a comparison.

P.4 (*K-Stationarity*)[5] For all $x \in X$ and $d, d' \in \mathcal{D}$, $d \succsim d'$ if and only if $(x, d) \succsim (x, d')$.

P.4 asserts the following. Suppose that a DM prefers a deterministic stream d to d'. Postpone all elements of the two sequences one period ahead (d_0 and d'_0 will be consumed in period 1, d_1 and d'_1 will be consumed in period 2 and so on) and introduce a common period-zero consumption bundle x. Then she will prefer (x, d) to (x, d'). The same reasoning can be done the other way around: if two streams have a common period zero consumption bundle, then it can be dropped, all the bundles can be shifted one period back, and preferences will not change.

Remark 3.1 Note that K-Stationarity is equivalent to the following (apparently) stronger stationarity property: for all $t \in \mathbb{N}$ and $d, d', c, c', w \in \mathcal{D}$, $(d_0, \dots, d_{t-1}, c_t, d) \succsim (d_0, \dots, d_{t-1}, c'_t, d')$ if and only if $(d_0, \dots, d_{t-1}, w_t, c_t, d) \succsim (d_0, \dots, d_t - 1, w_t, c'_t, d')$.

Consider two acts that may differ from period t on. Suppose that the DM prefers $(d_0, \dots, d_{t-1}, c_t, d)$ to $(d_0, \dots, d_{t-1}, c'_t, d')$. Then we can shift all bundles one period ahead starting from period t, introduce a common bundle w_t in period t and the DM's preferences will not change. The reasoning can be done the other way around: the common bundle w_t can be deleted and consumption can be anticipated.

Under uncertainty, consumption can be stochastic at some points in time. Consider the following generalization of P.4, inspired by Remark 3.1.

[5] K-Stationarity stands for Koopmans' stationarity.

Strong Stationarity (SS) For all $t \in \mathbb{N}$, $d, d' \in \mathcal{D}$ and $f, g, h \in \mathcal{H}$,

$$(d_0, \ldots, d_{t-1}, f_t, d_{t+1}, \ldots) \succsim (d_0, \ldots, d_{t-1}, g_t, d'_{t+1}, \ldots) \Leftrightarrow$$
$$(d_0, \ldots, d_{t-1}, h_t, f_t, d_{t+1}, \ldots) \succsim (d_0, \ldots, d_{t-1}, h_t, g_t, d'_{t+1}, \ldots)$$

The axiom of (SS) naturally generalizes K-Stationarity in light of Remark 3.1. Preferences ought to be the same after shifts since we are only postponing (anticipating) of one period the timing of consumption and inserting (deleting) a *common* variable h_t in both sequences, while keeping the information structure fixed. Note however that now acts can be stochastic in period t as well as the common random variable h_t inserted (or deleted) in both sequences in period t.

Clearly (SS) implies the property in Remark 3.1 and hence K-Stationarity (just consider (SS) over $\mathcal{D}$). A caveat is in order, since acts are adapted to the filtration $(\mathcal{F}_t)_t$, one needs to be careful not to shift back a $\mathcal{F}_t$-measurable variable to period $t-1$, as this variable may not be $\mathcal{F}_{t-1}$-measurable. Measurability is also the reason why we need to state (SS) for all periods t, as in Remark 3.1, and not only for period 0. If one states the axiom only for shifts starting at period 0, it would be possible to insert only random variables that are measurable with respect to $\mathcal{F}_0$, i.e. constant random variables (for more about this, see Sect. 5).

It is important to note that (SS) plays the role that the Independence axiom has in the Anscombe–Aumann (1963) model. Shifting variables one period ahead (or one period back) and inserting (or removing) a variable h_t can be done *independently* of the choice of the sequences and the variable h_t. As we argued in the Introduction and in Example 1, (SS) is too restrictive in presence of uncertainty. This motivates the generalization that is given in Sect. 4.

Finally, note that there are two possible ways of generalizing K-Stationarity. Our axiom, (SS), keeps the timing of resolution of uncertainty fixed while it changes the date in which consumption takes place. Another way would be to change both dates of resolution of uncertainty and consumption. The former generalization is considered in Kochov (2015) and Bommier et al. (2017), the latter in Bommier et al. (2019).

We are ready to state the main result of this section. The basic axioms (C), (M) and (TS) together with (SS) deliver the discounted expected utility representation in (1).

Theorem 1 *A preference relation $\succsim$ over $\mathcal{H}$ satisfies (C), (M), (TS) and (SS) if and only if there exists a probability $P : \mathcal{F} \to [0, 1]$, a continuous utility index $u : X \to \mathbb{R}$ and a discount factor $\beta \in (0, 1)$ such that $\succsim$ is represented by*

$$V(h) = \mathbb{E}_P\left[\sum_{t=0}^{\infty} \beta^t u(h_t)\right]$$

Moreover P and β are unique and u is unique up to a positive affine transformation.

Theorem 1 can be derived from Theorem 5 (Section 5), as remarked by Kochov (2015) (p. 244), by requiring indifference in the Intertemporal Hedging axiom included in his elegant characterization of an intertemporal version of the MaxMin expected utility model of Gilboa and Schmeidler (1989). We think that the main merit of using

Strong Stationarity in a representation result for the Discounted Expected Utility model is that it is a direct generalization of Koopmans' original axiom and that it can be interpreted as an independence condition. Moreover, this leads to the interpretation of such behavior as indicative of ambiguity neutrality in the sense of Ghirardato and Marinacci (2002). We will elaborate more on this after Theorem 5.

A last remark is in order. In the literature, the utility function of Theorem 1 is usually written

$$\sum_{t=0}^{\infty} \beta^t \mathbb{E}_P\left[u(h_t)\right]$$

where the sum and the expected value operator are exchanged. Clearly this cannot be done for the Choquet functional in (2) and for the MaxMin functional (which is defined later in equation (4)). However we show in Proposition 1 that for the discounted expected utility model of Theorem 1 both forms are possible.

Proposition 1 *For all* $h \in \mathcal{H}$, $\mathbb{E}_P\left[\sum_t \beta^t u(h_t)\right] = \sum_t \beta^t \mathbb{E}_P\left[u(h_t)\right]$.

4 Main result: Choquet expected discounted utility

While discounted expected utility, axiomatized in Theorem 1, is widely used in economic applications, we argue that its behavioral foundations are not exempt from criticisms. Example 1 below, already mentioned in the Introduction, shows that (SS) may be too restrictive (this is why we call this axiom *Strong* Stationarity) as it neglects possible hedging considerations made by the DM.

Example 1 Let X be an interval of $\mathbb{R}$. Acts are interpreted as stochastic flows of income. Consider an event $A \in \mathcal{F}_t$ (this was the event "Democrats win the elections" in the Introduction) and two acts $f \sim g$, with $f = (d_0, \ldots, d_{t-1}, f_t, d_{t+1}, \ldots)$ and $g = (d_0, \ldots, d_{t-1}, g_t, d_{t+1}, \ldots)$. These acts are constant in every period except in period t in which they are defined by

$$f_t(\omega) = \begin{cases} 10\$ & \text{if } \omega \in A \\ 0\$ & \text{if } \omega \in A^c \end{cases} \quad \text{and} \quad g_t(\omega) = \begin{cases} 0\$ & \text{if } \omega \in A \\ 10\$ & \text{if } \omega \in A^c. \end{cases}$$

If $\succsim$ has the discounted expected utility representation of Theorem 1 (with $u(0) = 0$) then $f \sim g$ implies $\beta^t \mathbb{E}_P[u(f_t)] = \beta^t \mathbb{E}_P[u(g_t)]$ and hence $P(A) = P(A^c) = \frac{1}{2}$, i.e. being indifferent reveals that the DM thinks that the two events A and A^c are equally likely.

Consider now an act $h \in \mathcal{H}$ such that

$$h_t(\omega) = \begin{cases} 0\$ & \text{if } \omega \in A \\ 7\$ & \text{if } \omega \in A^c. \end{cases}$$

If (SS) holds then one obtains

$$(d_0, \ldots, d_{t-1}, h_t, f_t, d_{t+1}, \ldots) \sim (d_0, \ldots, d_{t-1}, h_t, g_t, d_{t+1}, \ldots).$$

Of course, indifference is obvious if preference are of the discounted expected utility type.

However, as we have argued in the Introduction, deducing indifference in this latter situation from $f \sim g$ is not straightforward. We argue that we may actually observe

$$(d_0, \ldots, d_{t-1}, h_t, f_t, d_{t+1}, \ldots) \succ (d_0, \ldots, d_{t-1}, h_t, g_t, d_{t+1}, \ldots).$$

This happens because in case of bad luck, formally for $\omega \in A$, the act on the right-hand side of the preference relation makes the DM "poor" for 2 consecutive dates (she gets 0$ in t and 0$ in $t+1$). Whereas, by choosing the act on the left-hand side, she can be sure that she will be"rich" in at least one period. Introducing h_t in front of f_t gives a *temporal hedge* to the DM against any choice of nature. The reverse preference may be observed as well, namely $(d_0, \ldots, d_{t-1}, h_t, f_t, d_{t+1}, \ldots) \prec (d_0, \ldots, d_{t-1}, h_t, g_t, d_{t+1}, \ldots)$. This would be the choice of a DM who would like to leverage uncertainty in order to get a high pay-off in case of "luck". Finally, it could happen that a DM may want to hedge against uncertainty for some events $A \in \mathcal{F}_t$ and leverage uncertainty for other events $B \in \mathcal{F}_t$. This last attitude may be explained by the competence hypothesis: Heath and Tversky (1991) show in an experiment that DMs are uncertainty seeking whenever they are betting over events for which they feel competent, but not otherwise. All these attitudes are precluded by axiom (SS).

Recall that two random variables f, g from Ω to X are *comonotonic* if there is no ω and ω' in Ω such that $f(\omega) \succ f(\omega')$ and $g(\omega') \succ g(\omega)$. Note that a constant random variable is comonotonic with any other random variable. If X is a convex interval of $\mathbb{R}$, two random variables f, g are comonotonic if for all $\omega, \omega' \in \Omega$, $\left[f(\omega) - f(\omega')\right] \cdot \left[g(\omega) - g(\omega')\right] \geq 0$. Comonotonic random variables are "positively correlated".[6] It is not difficult to check that in Example 1 the variables h_t and g_t are comonotonic, whereas h_t and f_t are not. Therefore, putting h_t and f_t one after the other can lead to some considerations about (positive) correlation aversion/loving, which may reverse preferences if we perform the shifts prescribed by stationarity.

If two consecutive random variables are not comonotonic, the DM can temporally hedge or leverage uncertainty. On the other hand, two consecutive comonotonic random variables remove any such consideration.

The main novelty of our paper is represented by the axiom Comonotonic Stationarity. This axiom weakens (SS) by taking into account temporal correlations. The idea is simply to restrict the set of acts on which (SS) has a bite. The formal statement follows.

[6] We cannot properly speak about correlation since there is no probability defined on $(\Omega, \mathcal{A})$. However, if Ω is finite it is possible to show that two real-valued random variables f, g are comonotonic if and only if $cov(f, g) \geq 0$ for all probabilities over $(\Omega, \mathcal{A})$.

COMONOTONIC STATIONARITY (CS) For all $t \in \mathbb{N}$, $d, d' \in \mathcal{D}$ and $f, g, h \in \mathcal{H}$ such that h_t is comonotonic with f_t and g_t,

$$(d_0, \dots, d_{t-1}, f_t, d_{t+1}, \dots) \succsim (d_0, \dots, d_{t-1}, g_t, d'_{t+1}, \dots) \Leftrightarrow$$
$$(d_0, \dots, d_{t-1}, h_t, f_t, d_{t+1}, \dots) \succsim (d_0, \dots, d_{t-1}, h_t, g_t, d'_{t+1}, \dots)$$

Axiom (CS) allows the same type of shifts as (SS) only when the random variable h_t, inserted in period t, is comonotonic with the two random variables f_t and g_t. Therefore (CS) generalizes (SS) (in the sense that (SS) implies (CS)) limiting its range of action. Shifts can be performed only when no considerations about temporal correlations can be done. For instance all preferences described in Example 1 are admissible under (CS).

It is interesting to stress the conceptual similarity of axiom (CS) with the axiom Comonotonic Independence of Schmeidler (1989). Comonotonic Independence restricts the classical Independence axiom of expected utility to comonotonic acts. In the Anscombe and Aumann (1963) (atemporal) framework, acts are functions from states of the world to lotteries over X. The Independence axiom says that for any three acts f, g and h and any mixing weight $\alpha \in [0, 1]$, $f \succsim g$ if and only if $\alpha f + (1 - \alpha)h \succsim \alpha g + (1 - \alpha)h$. Comonotonic Independence retains this property only when the act h is comonotonic with both f and g, see Schmeidler (1989). In this latter case, no hedging can occur when h is mixed with f or g. In our framework, hedging is not achieved by probability mixing but through "time mixing". Hence (SS) plays the role of Independence while (CS) the one of Comonotonic Independence.

The following is the main result of this section.

Theorem 2 *A preference relation $\succsim$ over $\mathcal{H}$ satisfies (C), (M), (TS) and (CS) if and only if there exists a capacity $v : \mathcal{F} \to [0, 1]$, a continuous utility index $u : X \to \mathbb{R}$ and a discount factor $\beta \in (0, 1)$ such that $\succsim$ is represented by*

$$V(h) = \int \sum_{t=0}^{\infty} \beta^t u(h_t) dv.$$

Moreover v and β are unique and u is unique up to a positive affine transformation.

Theorem 2 generalizes Theorem 1 allowing non-neutral attitudes towards correlations and, therefore, towards uncertainty. Note however that Choquet expected discounted utility does not axiomatically impose any particular behavior. Section 5 focuses on the particular case of uncertainty aversion.[7]

[7] It is also interesting to note that the Choquet model has been fruitfully applied to an intertemporal setting at least since Gilboa (1989). Gilboa's paper however differs from ours since uncertainty is absent and the Choquet integral is used to model aversion (or love) for variability of payments across time. Other papers taking this approach are for instance Araujo et al. (2011) and Bastianello and Chateauneuf (2016).

5 Uncertainty aversion

We consider here the particular (but relevant) case of uncertainty aversion. As it will be clearer later, uncertainty aversion is akin to some form of pessimism. This justifies the name of the following axiom.

PESSIMISTIC STATIONARITY (PS) The preference relation $\succsim$ satisfies (CS). Moreover for all $t \in \mathbb{N}$, for all $d \in \mathcal{D}$ and for all $f, g, h \in \mathcal{H}$ such that h_t is comonotonic with g_t,

$$(d_0, \dots, d_{t-1}, f_t, d_{t+1}, \dots) \succsim (d_0, \dots, d_{t-1}, g_t, d'_{t+1}, \dots) \Rightarrow$$
$$(d_0, \dots, d_{t-1}, h_t, f_t, d_{t+1}, \dots) \succsim (d_0, \dots, d_{t-1}, h_t, g_t, d'_{t+1}, \dots)$$

It has the same interpretation of Wakker's (1990) axiom Pessimism Independence. See also Chateauneuf (1994). For a pessimistic DM shifting all variables one period ahead starting from period t and inserting a comonotonic variable h_t in front of g_t will decrease the appreciation of the act $(d_0, \dots, d_{t-1}, g_t, d'_{t+1}, \dots)$. On the other end, the possibly non-comonotonic variable h_t in front of f_t will make the stream $(d_0, \dots, d_{t-1}, f_t, d_{t+1}, \dots)$ more appealing since it may offer a temporal hedge. Consider again Example 1.

Example 1 - cont. Consider the acts of Example 1. If DM's preferences satisfy (PS) one will actually observe the path of choices described in Example 1, namely

$$f \sim g \Rightarrow (d_0, \dots, d_{t-1}, h_t, f_t, d_{t+1}, \dots) \succ (d_0, \dots, d_{t-1}, h_t, g_t, d_{t+1}, \dots).$$

Let $\succsim$ be represented by the Choquet expected discounted utility functional V of Theorem 2 (with $u(0) = 0$). First note that $f \sim g$ implies $V(f) = \beta^t u(10)v(A) + k = \beta^t u(10)v(A^c) + k = V(g)$ where $k = \sum_{i \neq t} \beta^i u(d_i)$; and hence $v(A) = v(A^c)$. Since v it is not required to be additive, we may have $v(A) \neq \frac{1}{2}$.

Call $\hat{f} = (d_0, \dots, d_{t-1}, h_t, f_t, d_{t+1}, \dots)$ and $\hat{g} = (d_0, \dots, d_{t-1}, h_t, g_t, d_{t+1}, \dots)$ and assume for the sake of calculations $\beta u(10) > u(7)$ (the opposite case is treated similarly and yields the same conclusions). Note that $V(\hat{f}) = \beta^t u(7) + [\beta^{t+1} u(10) - \beta^t u(7)]v(A) + k'$ and $V(\hat{g}) = [\beta^t u(7) + \beta^{t+1} u(10)]v(A^c) + k'$ where $k' = \sum_{i \neq t, t+1} \beta^i u(d_i)$. Therefore $\hat{f} \succ \hat{g}$ if and only if

$$u(7) + [\beta u(10) - u(7)]v(A) > [u(7) + \beta u(10)]v(A^c)$$
$$u(7)(1 - v(A)) > u(7)v(A^c) \quad [\text{since } v(A) = v(A^c)]$$
$$v(A) + v(A^c) < 1$$

and hence $v(A) < \frac{1}{2}$. If v is actually convex, one gets $v(A) + v(A^c) < 1$, and in this case one can observe $f \sim g$ and $\hat{f} \succ \hat{g}$. It is not difficult to find examples in which v is convex and actually $g \succ f$ and $\hat{f} \succ \hat{g}$. As Theorem 3 shows, (PS) forces the capacity v to be convex.

Note that (PS) implies (CS) and hence Theorem 2 remains valid if (CS) is replaced by (PS). Since this latter axiom is stronger, we can prove in Theorem 3 that the capacity appearing in the Choquet integral is convex.

Theorem 3 *A preference relation $\succsim$ over $\mathcal{H}$ satisfies (C), (M), (TS) and (PS) if and only if there exists a convex capacity $v : \mathcal{F} \to [0, 1]$, a continuous utility index $u : X \to \mathbb{R}$ and a discount factor $\beta \in (0, 1)$ such that $\succsim$ is represented by*

$$V(h) = \int \sum_{t=0}^{\infty} \beta^t u(h_t) dv.$$

Moreover v and β are unique and u is unique up to a positive affine transformation.

It is known that (see Schmeidler 1986) when v is a convex capacity the following equality holds

$$\int \sum_{t=0}^{\infty} \beta^t u(h_t) dv = \min_{P \in core(v)} \mathbb{E}_P \left[\sum_{t=0}^{\infty} \beta^t u(h_t) \right]. \tag{3}$$

This equality suggests a sharp interpretation of a Choquet integral with respect a convex capacity v and justifies why we call pessimist a DM with preferences as the ones in Theorem 3. When v is convex, an agent reasons as if she computes the discounted expected utility for all probabilities in $core(v)$ and then selects the minimal one.

If in (3) one replaces $core(v)$ with a convex and closed set $\mathcal{P}$ of probabilities, then one obtains the intertemporal version of the MaxMin expected utility model recently axiomatized by Kochov (2015). MaxMin expected utility is a popular[8] decision theoretic model axiomatized by Gilboa and Schmeidler (1989) in an atemporal framework in order to address Ellsberg (1961) paradox. Consider the following axiom proposed by Kochov (2015).

Intertemporal Hedging (IH) For all $t \in \mathbb{N}$, for all $d \in \mathcal{D}$ and for all $g, h \in \mathcal{H}$,

$$(d_0, \ldots, d_{t-1}, h_t, h_t, d_{t+2} \ldots) \sim (d_0, \ldots, d_{t-1}, g_t, g_t, d_{t+2} \ldots)$$
$$\Rightarrow (d_0, \ldots, d_{t-1}, g_t, h_t, d_{t+2} \ldots) \succsim (d_0, \ldots, d_{t-1}, h_t, h_t, d_{t+2} \ldots)$$

The interpretation of (IH) is that a DM prefers to smooth consumption through states rather than through time. This in turns implies that she is pessimistic *vis-à-vis* Nature's choice of the state of the world. We can note in fact that the act $(d_0, \ldots, d_{t-1}, g_t, h_t, d_{t+2} \ldots)$ allows for a temporal mix that may provide some hedging against uncertainty. As explained by Kochov (2015) this axiom is the intertemporal counterpart to the Uncertainty Aversion axiom of Gilboa and Schmeidler (1989).

We show now that the representation of Theorem 3 can be obtained also using (IH) and weakening (PS) to (CS).

8 There are several other famous models that deal with choice under uncertainty and may exhibit non neutral attitudes towards uncertainty. For instance the Choquet expected utility model of Schmeidler (1989), the smooth ambiguity model of Klibanoff et al. (2005), the variational model of Maccheroni et al. (2006), the confidence model of Chateauneuf and Faro (2009), prospect theory of Tversky and Kahneman (1992) etc.

Theorem 4 *A preference relation* $\succsim$ *over* $\mathcal{H}$ *satisfies (C), (M), (TS), (CS) and (IH) if and only if there exists a convex capacity* $v : \mathcal{F} \to [0, 1]$*, a continuous utility index* $u : X \to \mathbb{R}$ *and a discount factor* $\beta \in (0, 1)$ *such that* $\succsim$ *is represented by*

$$V(h) = \int \sum_{t=0}^{\infty} \beta^t u(h_t) dv.$$

Moreover v *and* β *are unique and* u *is unique up to a positive affine transformation.*

The utility function in Theorem 3 and Theorem 4 is a particular case of the intertemporal MaxMin model studied by Kochov (2015)

$$V(h) = \min_{P \in \mathcal{P}} \mathbb{E}_P \left[\sum_{t=0}^{\infty} \beta^t u(h_t) \right] \tag{4}$$

where $\mathcal{P}$ is a convex and closed (with respect to the weak* topology) set of probabilities. Actually, if one sets $\mathcal{P} = core(v)$, the functionals in (3) and (4) are exactly the same.

If one is willing to obtain Kochov's (2015) Theorem 1, in which he characterizes the utility function in (4), then (CS) should be weakened to the following axiom, called Path Stationarity in Bommier et al. (2017).[9]
Path Stationarity (PathS) For all $x \in X$, $d, d' \in \mathcal{D}$ and $f \in \mathcal{H}$,

$$(d_0, \ldots, d_{t-1}, f_t, d_{t+1}, \ldots) \succsim (d_0, \ldots, d_{t-1}, g_t, d'_{t+1}, \ldots) \Leftrightarrow$$
$$(x, d_0, \ldots, d_{t-1}, f_t, d_{t+1}, \ldots) \succsim (x, d_0, \ldots, d_{t-1}, g_t, d'_{t+1}, \ldots)$$

This axiom is weaker than (CS) since the outcome $x \in X$ can be identified with a constant random variable, which is comonotonic with all other variables. Note also that it is not needed to state the axiom for all periods of time $t \in \mathbb{N}$. A constant random variable is $\mathcal{F}_t$-measurable for all $t \in \mathbb{N}$ and therefore there are no measurability concerns. The interpretation of this axiom is the same as for K-Stationarity. Kochov (2015) underlines an interesting parallel between (PathS) and the Certainty Independence axiom of Gilboa and Schmeidler (1989).

Proposition 2 shows how the different stationarity axioms presented in the paper are linked one to another. We recall that (K-Stationarity) is due to Koopmans (1960, 1972) and characterizes discounted utility, see Proposition 5; (PathS) is due to Kochov (2015) and characterizes MaxMin discounted expected utility, see Theorem 5 below; finally axioms (SS), (CS) and (PS) are new and characterize discounted expected utility and Choquet expected discounted utility, see Theorems 1, 2 and 3.

Proposition 2 *The following implications hold:*

$$(SS) \Rightarrow (PS) \Rightarrow (CS) \Rightarrow (PathS) \Rightarrow (K\text{-}Stationarity).$$

[9] This axiom is weaker than the one used by Kochov (2015), but one can show that it is sufficient to derive his main result.

Weakening (CS) to (PathS) and adding (IH) to the other basic axioms, allow us to recover Kochov's main result (see Kochov 2015 Theorem 1, p. 245).

Theorem 5 (Kochov 2015) *A preference relation $\succsim$ over $\mathcal{H}$ satisfies (C), (M), (TS), (PathS) and (IH) if and only if there exists nonempty weak*-closed convex set $\mathcal{P}$ of probabilities, a continuous strictly increasing utility index $u : X \to \mathbb{R}$ and a discount factor $\beta \in (0, 1)$ such that $\succsim$ is represented by*

$$V(h) = \min_{P \in \mathcal{P}} \mathbb{E}_P \left[\sum_{t=0}^{\infty} \beta^t u(h_t) \right].$$

Moreover $\mathcal{P}$ and β are unique and u is unique up to a positive affine transformation.

Some remarks are in order. First, it is clear that the discounted expected utility model, obtained in Theorem 1, could be derived requiring indifference in (IH) and applying Theorem 5, as remarked by Kochov (2015). However we believe that (SS) is conceptually closer to the independence axiom of a-temporal expected utility and closer in spirit to Koopmans' original axiomatization. In some sense, deriving discounted expected utility from Theorem 5 would be as deriving expected utility from Gilboa and Schmeidler (1989), without knowing the Independence axiom of Anscombe and Aumann (1963). Second, the multiple prior model postulates uncertainty aversion (or seeking) as an axiom. In the intertemporal framework this is implied by (PS) or (IH). We believe that (CS) and therefore the Choquet expected discounted utility model with a general (e.g. not convex) capacity have a different flavor. This model obtains when (SS) is relaxed in order to perform the shifts of stationarity only when no considerations about temporal correlations arise. Hence, Choquet expected discounted utility does not take any stance about the attitude towards uncertainty of a DM. Note that this model can handle the competence hypothesis of Heath and Tversky (1991): this is done considering a capacity that is neither convex nor concave. For instance one can think about the neo-additive capacities defined by Chateauneuf et al. (2007).

6 Discussion

6.1 Dynamic choices and updating

While dynamic choices are outside the scope of this paper, as our main purpose is to provide a sound axiomatiozation of Koopmans' (1972) intertemporal model under uncertainty, we will briefly discuss in this section some issues about dynamic preferences and some open research questions.

In the previous sections, the DM was only deciding at time $t = 0$. Suppose now that the agent can express a preference at every time t. Clearly at time t she knows that the true state of nature ω belongs to $\mathcal{F}_t(\omega) = \cap\{A \in \mathcal{F}_t | \omega \in A\}$, i.e. $\mathcal{F}_t(\omega)$ denotes the intersection of all sets in $\mathcal{F}_t$ that contain ω.[10] Therefore for each ω and t she will have a preference relation $\succsim_{t,\omega}$ over $\mathcal{H}$.

Two classical axioms imposed on the class of preferences $\{\succsim_{t,\omega}\}_{t\in\mathbb{N},\omega\in\Omega}$ in the context of repeated choice are Consequentialism and Dynamic Consistency. Consequentialism says that preferences at a certain date t do not depend neither on previous dates nor on states of nature that do not belong to $\mathcal{F}_t(\omega)$. Formally,

Consequentialism (Co) For all $t \in \mathbb{N}$, $\omega \in \Omega$ and $f, g \in \mathcal{H}$ such that $f_k(\omega') = g_k(\omega')$ for all $k \geq t$ and for all $\omega' \in \mathcal{F}_t(\omega)$, $f \sim_{t,\omega} g$.

Dynamic Consistency postulates that, if two acts f and g are equal up to time t, and in period $t+1$ f is preferred to g in all states of the world, then f ought to be preferred to g at time t.

Dynamic Consistency (DC) For all $t \in \mathbb{N}$, $\omega \in \Omega$ and $f, g \in \mathcal{H}$ such that $f_k = g_k$ for all $k \leq t$ and such that $f \succsim_{t+1,\omega'} g$ for all $\omega' \in \mathcal{F}_t(\omega)$, $f \succsim_{t,\omega} g$.

When preferences are of the discounted expected utility type, as the ones axiomatized in Theorem 1, it is well known that (Co) and (DC) imply that $\succsim_{t,\omega}$ is characterized by the recursive formula

$$V_{t,\omega}(h) = u(h_t(\omega)) + \beta \mathbb{E}_{P_{t,\omega}}\left[V_{t+1,\cdot}(h)\right] \tag{5}$$

and one has

$$V_{t,\omega}(h) = \mathbb{E}_P\left[\sum_{k\geq t} \beta^{k-t} u(h_t) \mid \mathcal{F}_t(\omega)\right].$$

Hence $P_{t,\omega}$ is the Bayesian update of P given information $\mathcal{F}_t(\omega)$. See Kreps and Porteus (1978) and Skiadas (1998) for recursivity and discounted expected utility. For general results on (monotone) recursive preferences see a recent paper of Bommier et al. (2017).

When preferences are of the Choquet expected discounted utility type, as the ones axiomatized in Theorem 2, things get more complicated since there is not a unique way to update a capacity. The three most well known updating rules for capacities are the *Dempster–Shafer updating rule* (Dempster 1968; Shafer 1976), *naive Bayes' updating rule* (Gilboa and Schmeidler 1993), and *generalized Bayesian updating rule* (Dempster 1967; Fagin and Halpern 1991; Jaffray 1992). Let $v : \mathcal{F} \to [0, 1]$ be a capacity and consider a set $A \in \mathcal{F}$ such that $A \neq \emptyset, \Omega$, then the three updating rules are respectively

$$v_A^{DS}(E) = \frac{v((E \cap A) \cup A^c) - v(A^c)}{1 - v(A^c)}$$
$$v_A^{NB}(E) = \frac{v(E \cap A)}{v(A)}$$
$$v_A^{GB}(E) = \frac{v(E \cap A)}{v(E \cap A) + 1 - v(E \cup A^c)}$$

[10] In the study of repeated choice $\mathcal{F}_t$ is assumed to be finitely generated for all t.

were the denominators are assumed to be strictly positive. Note that if v is additive, the three updating rules boil down to Bayesian updating.

It is known that, if a capacity is updated with one of these rules, (DC) is violated in general. Eichberger et al. (2007) and Horie (2013) show that a weakening of (DC), adapted from the one proposed by Pires (2002) for the multiple prior model, leads to generalized Bayesian updating. Asano and Kojima (2019) propose two different relaxations of (DC) that yield Dempster–Shafer and naive Baye's updating rules. Another path is followed by Dominiak and Lefort (2011) who restrict (DC) to suitably defined non-ambiguous events and show that in this case the correct updating rule is the naive Baye's one. A different approach is the one of Siniscalchi (2011), who studies consistent-planning for preferences over decision trees. Finally, in a recent paper Gul and Pesendorfer (2018) propose a consequentialist and recursive model of updating for totally monotone capacities and derive a new updating rule called the *proxy rule*.

This variety of results indicates how delicate is the issue of updating a capacity and its consequences for (DC). Moreover, the axiom of (CS) do not ensure recursivity. For this reasons, we do not address the issue of the recursive formulation of the Choquet expected discounted utility model and its appropriate updating rules. This is an open area for future research.

6.2 About the stationarity conditions

Note that for acts involved in our stationarity axioms, consumption is stochastic in only two periods, t and $t+1$. These axioms are necessary and sufficient to derive our results. One natural question is whether our representation results imply some form of stationarity for acts in which consumption is stochastic in several (maybe infinitely many) periods. We propose in this section two stronger versions of (SS) and of (CS) in order to take into account this case. A strengthening of (PathS) was already given in Kochov (2015).

Consider the following two axioms.

STRONG STATIONARITY' (SS') For all $t \in \mathbb{N}$, $f, g, h \in \mathcal{H}$,

$$(h_0, \ldots, h_{t-1}, f_t, f_{t+1}, \ldots) \succsim (h_0, \ldots, h_{t-1}, g_t, g_{t+1}, \ldots) \Leftrightarrow$$
$$(h_0, \ldots, h_{t-1}, h_t, f_t, f_{t+1}, \ldots) \succsim (h_0, \ldots, h_{t-1}, h_t, g_t, g_{t+1}, \ldots)$$

COMONOTONIC STATIONARITY' (CS') For all $t \in \mathbb{N}$, $d \in \mathcal{D}$ and $f, g, h \in \mathcal{H}$ such that h_t is comonotonic with f_i and g_i for all $i \geq t$,

$$(d_0, \ldots, d_{t-1}, f_t, f_{t+1}, \ldots) \succsim (d_0, \ldots, d_{t-1}, g_t, g_{t+1}, \ldots) \Leftrightarrow$$
$$(d_0, \ldots, d_{t-1}, h_t, f_t, f_{t+1}, \ldots) \succsim (d_0, \ldots, d_{t-1}, h_t, g_t, g_{t+1}, \ldots)$$

Axiom (SS') is a strengthening of (SS) as consumption could be stochastic in every period. Note however that the two consumption streams must be equal in the first t periods. Axiom (CS') allows the same type of shifts as (SS') only when the random variable h_t, inserted in period t, is comonotonic with *all* random variables after period

t for both acts f and g. Moreover the first t periods of consumption must be equal and deterministic.

It is easy to see that (SS') and (CS') are strengthening of (SS) and (CS) respectively. Therefore these axioms are sufficient to derive our results. The next proposition shows that they are also necessary.

Proposition 3 *Formula (1) implies (SS') and formula (2) implies (CS').*

7 Conclusions

In this paper we make two contributions that can be of high interest for economists working with problems that involve decisions through time and under uncertainty.

Our first contribution is to show how a Strong Stationarity axiom is the key behavioral condition behind discounted expected utility in a purely subjective framework. Strong Stationarity plays the same role as the Independece axiom for decisions under uncertainty in an atemporal setting.

Second, we argue that Strong Stationarity neglects agents' hedging behavior. Strong Stationarity is subject to the same critiques as the Independece axiom. We solve this problem introducing a new axiom, Comonotonic Stationarity. This latter condition allows us to generalize the discounted expected utility model to the Choquet expected discounted utility model. Our axioms have a simple interpretation, akin to the original stationarity condition of Koopmans (1960, 1972). Testing these new axioms in the lab and studying dynamic choices will be the focus of subsequent explorations.

Acknowledgements We would like to thank Antoine Bommier, Alain Chateauneuf, Mehmet Ismail, Asen Kochov, Marcus Pivato, Jean-Marc Tallon, Stéphane Zuber and audiences at numerous seminars and conferences. José Heleno Faro is also grateful for financial support from the CNPq-Brazil (Grant No. 308183/2019-3).

A Appendix

Proposition 5 in Appendix A.1 derives a discounted utility representation over the set of deterministic streams $\mathcal{D}$ using the axioms of Sect. 3. This result is known since Koopmans (1960, 1972). There are however few differences (e.g. Koopmans assumes continuity w.r.t. the sup-norm topology while we use the product topology) and this is why we include it in this appendix. Appendix A.2 uses Proposition 5 in order to derive our main theorems.

A.1 The deterministic setting

We derive in Proposition 5 a discounted utility representation $V(d) = \sum_{t=0}^{\infty} \beta^t u(d_t)$ for $\succsim$ over the set $\mathcal{D}$ of deterministic acts. This is done in several steps. First, we show in Proposition 4 that the axioms (C), (TS), (M) and (SS) of Sect. 3 imply Koopmans' (1972) original postulates P.1–5. Second, we show in Lemma A.3 that P.1–5 imply the axioms of Bleichrodt et al. (2008). Finally we use Theorem 6 (proved in Bleichrodt et al. (2008)) to derive the representation. We detail the proof to make our paper as self-contained and explicit as possible. The reader who is familiar with Koopmans (1960, 1972) and Bleichrodt et al. (2008) can skip Appendix A.1.

The following postulates, P.1–5, where considered first by Koopmans (1972).

P.1 (*Continuity*) For all compact sets $K \subset X$ and for all deterministic acts $d' \in \mathcal{D}$, the sets $\{d \in K^\infty | d \succsim d'\}$ and $\{d \in K^\infty | d' \succsim d\}$ are closed in the product topology over K^∞.

P.2 (*Sensitivity*) There exist $x, y \in X$, $d \in \mathcal{D}$ such that $(x, d) \succ (y, d)$.

P.3 (*Time Separability*) For all $x, y, x', y' \in X$ and $d, d' \in \mathcal{D}$, $(x, y, d) \succsim (x', y', d)$ if and only if $(x, y, d') \succsim (x', y', d')$.

P.4 (*K-Stationarity*) For all $x \in X$ and $d, d' \in \mathcal{D}$, $d \succsim d'$ if and only if $(x, d) \succsim (x, d')$.

P.5 (*K-Monotonicity*) Let $d, d' \in \mathcal{D}$. If $d_t \succsim d'_t$ for all t, then $d \succsim d'$; if moreover $d_t \succ d'_t$ for some t then $d \succ d'$.

We list now the alternative axioms studied by Bleichrodt et al. (2008).

Notation: given $T \in \mathbb{N}$, $X_T = \{(x_0, x_1, \ldots, x_T, \alpha, \alpha) | x_0, \ldots, x_T, \alpha \in X\}$. Note that, for any $T \in \mathbb{N}$, there is a one-to-one function from X_T and the product X^{T+1}.

Ultimate Continuity (UC). $\succsim$ is continuous (with respect to the product topology) on each set X_T, i.e. the sets $\{x \in X_T | x \succsim y\}$ and $\{x \in X_T | y \succsim x\}$ are closed for all $y \in X_T$.

Constant equivalent (CE). $\succsim$ satisfies constant equivalence if for all $d \in \mathcal{D}$ there exists a constant sequence $x_d \in \mathcal{D}$ such that $d \sim x_d$.

Tail Robustness (TR). $\succsim$ satisfies tail robustness if for all constant sequence $x \in \mathcal{D}$, if $d \succ (\prec) x$ then there exists $t \in \mathbb{N}$ such that $(d_0, \ldots, d_T, x, x,) \succ (\prec) x$ for all $T \geq t$.

Theorem 6 (Bleichrodt et al. 2008) *Let $\succsim$ be defined over $\mathcal{D}' \supset \mathcal{D}$, a domain that contains all ultimately constant programs, then TFAE:*

(i) *Discounted Utility holds over $\mathcal{D}'$ with u continuous and not constant.*
(ii) *$\succsim$ satisfies P.2, P.3, P.5, UC, CE, TR.*

In Observation 3, Bleichrodt et al. (2008) noticed that “Tail robustness can also be replaced by monotonicity if [$\mathcal{D}'$] contains only bounded programs.” Boundededness means that for every $d \in \mathcal{D}'$ there exist $x, y \in X$ such that $x \succsim d \succsim y$.

Remark The definition of $\mathcal{D}$ (compactness) and P.5 imply boundedness. Therefore Observation 3 of Bleichrodt et al. (2008) applies.

We show now that the axioms of Sect. 3 imply P1–5.

Proposition 4 *(C), (TS), (M) and (SS) imply P.1–5.*

Proof It easy to see that $(C) \Rightarrow P.1$, $(TS) \Leftrightarrow P.3$, $(SS) \Rightarrow P.4$.

Lemma A.1 *(C), (M) and (SS) imply $P.2$.*

Proof See Lemma 5 in Kochov (2015). □

Lemma A.2 *(C), (TS) and (SS) imply $P.5$.*[11]

Proof Recall that if $x, y \in X$, we define $x \succsim y \Leftrightarrow (x, x, \dots) \succsim (y, y, \dots)$. We introduce the notation $({}_n x, d) = (\underbrace{x, \dots, x}_{n-\text{times}}, d)$ and $\bar{x} = (x, x, \dots)$.

Claim Let $x, y \in X$ and $d \in \mathcal{D}$. Then $(x, d) \succsim (\succ)(y, d) \Leftrightarrow x \succsim (\succ) y$.

$(\Rightarrow)$ We have $(x, d) \succsim (\succ)(y, d) \overset{(TS)}{\Longleftrightarrow} \bar{x} = (x, \bar{x}) \succsim (\succ)(y, \bar{x}) \overset{(SS)}{\Longleftrightarrow} (y, \bar{x}) \succsim (\succ) (y, y, \bar{x})$. Hence reasoning by induction and using transitivity of $\succsim$ we get $\bar{x} \succsim (\succ) (y, \bar{x}) \succsim (\succ)({}_n y, \bar{x})$. Since $({}_n y, \bar{x})$ converges to $\bar{y}$ as $n \to \infty$, by (C) we obtain $x \succsim (\succ) y$.

$(\Leftarrow)$ Suppose $x \succsim (\succ) y$. If there exists $d \in \mathcal{D}$ such that $(y, d) \succ (\succsim)(x, d)$, then by the first part of this proof $y \succ (\succsim) x$, which is impossible. Since $\succsim$ is complete we conclude that $(x, d) \succsim (\succ)(y, d)$.

This conclude the prof of the Claim.

Let $d, d' \in \mathcal{D}$ such that $d_t \succsim d'_t$ for all $t \in \mathbb{N}$. Fix $n \in \mathbb{N}$, since $d_n \succsim d'_n$, (SS) then implies $(d_{n-1}, d_n, d_n, \dots) \succsim (d_{n-1}, d'_n, d'_n, \dots)$. Moreover the Claim yields $(d_{n-1}, d'_n, d'_n, \dots) \succsim (d'_{n-1}, d'_n, d'_n, \dots)$ and by transitivity $(d_{n-1}, d_n, d_n, \dots) \succsim (d'_{n-1}, d'_n, d'_n, \dots)$. Applying n times this reasoning we obtain

$$(d_0, d_1, \dots, d_n, d_n, \dots) \succsim (d'_0, d'_1, \dots, d'_n, d'_n, \dots)$$

Since n was arbitrary, this is true for all $n \in \mathbb{N}$. When n tends to infinity, the sequence on the left hand-side of the preference converges to d, and the one to the right hand-side converges to d' in the product topology. Then (C) implies $d \succsim d'$.

Suppose now that in addition $d_t \succ d'_t$ for some t. By the Claim we have $(d_t, d_{t+1}, \dots) \succ (d'_t, d_{t+1}, \dots)$. Applying t times (SS) we obtain

$$(d_0, \dots, d_{t-1}, d_t, d_{t+1}, \dots) \succ (d_0, \dots, d_{t-1}, d'_t, d_{t+1}, \dots).$$

By the Claim, for all $n \geq t$ $(d_0, \dots, d_{t-1}, d'_t, d_{t+1}, \dots) \succsim (d'_0, \dots, d'_n, d_{n+1} \dots)$. Since the latter sequence converge to d', by (C) we have $(d_0, \dots, d_{t-1}, d'_t, d_{t+1}, \dots) \succsim d'$ and hence $d \succ d'$. □

Therefore we proved that (C), (TS), (M) and (SS) imply $P.1$–5. □

We show that the axioms P1–5 entail the representation with exponential discounted utility.

Proposition 5 *A preference relation $\succsim$ over $\mathcal{D}$ satisfies P.1–5 if and only if there exists a continuous function $u : X \to \mathbb{R}$ and a discount factor $\beta \in (0, 1)$ such that $\succsim$ is represented by*

$$V(d) = \sum_{t=0}^{\infty} \beta^t u(d_t).$$

[11] A similar statement appears without a formal proof in the proof of Lemma 7 in Kochov (2015). We provide a proof for sake of completeness.

Moreover β is unique and u is unique up to a positive affine transformation.

Proof of Proposition 5 In order to prove Proposition 5, we will show that postulates P.1–5 imply the axioms of Bleichrodt et al. (2008). By the Remark after Theorem 6, we only need to show UC and CE.

Lemma A.3 *If a preference relation $\succsim$ over $\mathcal{D}$ satisfies P.1–5 then it satisfies UC and CE.*

Proof We show that $\succsim$ satisfies UC. Fix $T \in \mathbb{N}$ and $y \in X_T$. We will show that the set $U_y = \{x \in X_T | x \succsim y\}$ is closed.

Let $(\hat{x}^n)_n$ be a sequence in U_y such that $\hat{x}^n \to \hat{x}$. Note that $\hat{x}^n = (\hat{x}_0^n, \hat{x}_1^n, \ldots, \hat{x}_T^n, \hat{\alpha}^n, \hat{\alpha}^n, \ldots)$. By definition $\hat{x}^n \to \hat{x}$ if and only if $\hat{x}_i^n \to \hat{x}_i$ for all $i \in \mathbb{N}$, with $\hat{x}_i^n = \hat{\alpha}^n$ for all $i \geq T+1$. Since $\hat{x}^n \in X_T$, $\hat{\alpha}^n \to \hat{\alpha}$ and this implies that $\hat{x} \in X_T$. The sets $C_i = \{\hat{x}_i^n | n \in \mathbb{N}\}$, $i = 0, \ldots, T$, and $C_\alpha = \{\hat{\alpha}^n | n \in \mathbb{N}\}$ are compact and hence $C = \left(\cup_{i=0}^T C_i\right) \cup C_\alpha$ is compact. Since $\hat{x} \in \mathcal{D}$, there is a compact set $\hat{K}$ such that $\hat{x}_t \in \hat{K}$ for all $t \in \mathbb{N}$. Therefore $K = C \cup \hat{K}$ is compact and $\hat{x}^n, \hat{x} \in K^\infty$ for all $n \in \mathbb{N}$. Consider now $U = \{d \in K^\infty | d \succsim y\}$. We have therefore that for all $n \in \mathbb{N}$, $\hat{x}^n \succsim y$ and $\hat{x}^n \in U$, moreover $\hat{x}^n \to \hat{x}$ and since U is closed by P.1 we obtain $\hat{x} \succsim y$. Hence U_y is closed.

We show that $\succsim$ satisfies CE. Fix $d \in \mathcal{D}$, and let K_d be a compact set such that $d_t \in K_d$ for all $t \in \mathbb{N}$. By compactness, there are $x_0, x_1 \in K_d$ such that $x_0 \succsim d_t \succsim x_1$ for all $t \in \mathbb{N}$. By P.5 we have $x_0 \succsim d \succsim x_1$. Consider $A = \{y \in co(K_d) | \bar{y} \succsim d\}$ and $B = \{y \in co(K_d) | d \succsim \bar{y}\}$, where $\bar{y}$ denotes the constant sequence $\bar{y} = (y, y, \ldots)$ and $co(K_d)$ is the convex hull of K_d. By P.1, A, and B are closed and since $x_0 \in A$ and $x_1 \in B$ they are both non empty. By connectedness of X we have that $co(K_d)$ is connected and therefore there exists $x_d \in co(K_d)$ such that $x_d \sim d$. □

Lemma A.3 and Theorem 6 of Bleichrodt et al. (2008) imply Proposition 5. □

A.2 Proof of the main results

Proof of Theorem 2 We first prove necessity of the axioms. Monotonicity (M) and Time Separability (TS) follow from the properties of monotonicity and comonotonic additivity of the Choquet integral. Continuity (C) is proved as in Kochov (2015). Given a compact set $K \subseteq X$, one has that $|\beta^t u(d_t)| \leq \beta^t M$ for some upper bound M of the function u (this bound exists since u is continuous over a compact set). By Rudin (1976), Theorem 7.10, the function $d \to \sum_{t=0}^n \beta^t u(d_t)$ converges uniformly on K^∞. Hence (C) follows from Rudin (1976), Theorem 7.11. Necessity of Comonotonic Stationarity (CS) is shown in Proposition 3.

We turn now to sufficiency. A simple modification of Proposition 4 shows that (C), (TS), (M) and (CS) imply P.1–5. Hence, by Proposition 5 there exists a continuous function $u : X \to \mathbb{R}$ and a discount factor $\beta \in (0, 1)$ such that the restriction of $\succsim$ over $\mathcal{D}$ is represented by the functional

$$U(d) = \sum_{t=0}^{\infty} \beta^t u(d_t).$$

Moreover $\beta \in (0, 1)$ is unique and u unique up to a positive affine transformation.

We can note that connectedness of X and continuity of u imply that $u(X)$ is an interval. By Lemma A.1 this interval has non-empty interior. Re-normalize w.l.g. u so that $[-1, 1] \subseteq u(X)$ and let $x^* \in X$ be such that $u(x^*) = 0$.

Lemma A.4 *For every $h \in \mathcal{H}$ there exists $d_h \in \mathcal{D}$ s.t. $h \sim d_h$.*

Proof See Lemma 8 of Kochov (2015). □

Define the function $V : \mathcal{H} \to \mathbb{R}$ by $V(h) := U(d_h)$. Since U represents preferences over $\mathcal{D}$, the function V is well defined and represents preferences over $\mathcal{H}$. Consider now the set

$$\mathcal{U} := \left\{ U \circ h := \sum_t \beta^t u(h_t) | h := (h_t)_t \in \mathcal{H} \right\}.$$

For every $h \in \mathcal{H}$, $U \circ h \in \mathcal{U}$ is a function $U \circ h : \Omega \to \mathbb{R}$ and will be denoted by capital letters.

Define now the function $I : \mathcal{U} \to \mathbb{R}$ as $I(H) := V(h)$ where $h \in \mathcal{H}$ is such that $U \circ h = H$. Note that I is well defined by monotonicity: if $H = U \circ h = U \circ h'$ then $h(\omega) \sim h'(\omega)$ for all $\omega \in \Omega$, by (M) $h \sim h'$ and therefore $V(h) = V(h')$.

Recall that $\mathcal{F} = \cup_t \mathcal{F}_t$. We denote $B^o := B^o(\Omega, \mathcal{F}, \mathbb{R})$, i.e. the set of simple, real-valued $\mathcal{F}$-measurable functions over Ω. Given a set $A \in \mathcal{F}$, $1_A \in B^o$ denotes the indicator function of the set A.

Lemma A.5 *For all $a \in B^o$, there exists $\delta > 0$ such that $a \in \delta\mathcal{U}$, i.e. $\mathcal{U}$ is an absorbing subset of B^o.*

Proof We need a slight modification of the proof of Lemma 9 in Kochov (2015) in order to take into account our definition of (SS). Consider his proof and let $h \in \mathcal{H}$ such that $h_k = x^*$ for all $k \neq t$ and $h_t = f$.[12] □

The next lemma extends $I : \mathcal{U} \to \mathbb{R}$ to $\tilde{I} : B^o \to \mathbb{R}$ and shows that $\tilde{I}$ is translation invariant and β-homogeneous.

Lemma A.6 ($\tilde{I}$ is translation invariant) *There exists a unique functional $\tilde{I} : B^o \to \mathbb{R}$ such that the restriction $\tilde{I}|_{\mathcal{U}}$ of $\tilde{I}$ on $\mathcal{U}$ is such that $\tilde{I}|_{\mathcal{U}} = I$. Moreover for every $a \in B^o$, for every $\alpha \in \mathbb{R}$, $\tilde{I}(\beta a) = \beta \tilde{I}(a)$ and $\tilde{I}(a + \alpha) = \tilde{I}(a) + \alpha$.*

Proof See Lemma 10, Lemma 11 and Lemma 12 of Kochov (2015). □

Lemma A.7 *Let $a, b, c \in B^o$ be such that c is comonotonic with a and b. Then $\tilde{I}(a) = \tilde{I}(b) \Leftrightarrow \tilde{I}(a + c) = \tilde{I}(b + c)$.*

Proof Fix $a, b, c \in B^o$ such that c is comonotonic with a and b. Since a, b, c are in B^o there exists $t_1 \in \mathbb{N}$ such that a, b, c are $\mathcal{F}_{t_1}$-measurable. Moreover there is $t_2 \in \mathbb{N}$

[12] Note that Kochov defines U as $U(d) = (1 - \beta) \sum_{t=0}^{\infty} \beta^t u(d_t)$.

such that the range of $\beta^{t_2}a$, $\beta^{t_2}b$ and $\beta^{t_2}c$ is included in $[-1, 1]$. Pick $n \geq \max\{t_1, t_2\}$ and define for all $t \in \mathbb{N}$ and for all $\omega \in \Omega$

$$f_t(\omega) := \begin{cases} x^* & \text{if } t \neq n \\ u^{-1}(\beta^n a(\omega)) & \text{if } t = n. \end{cases} \tag{6}$$

Note that $f \in \mathcal{H}$ since it is finite (because a is finite), and $u^{-1}(\beta^n a(\omega))$ is $\mathcal{F}_n$-measurable since $\beta^n a(\omega)$ is $\mathcal{F}_{t_1}$-measurable and $\mathcal{F}_{t_1} \subseteq \mathcal{F}_n$. In the same way define g and h using b and c respectively in the place of a. We have that $U \circ f = \beta^{2n}a$, $U \circ g = \beta^{2n}c$ and $U \circ h = \beta^{2n}c$. Hence

$$\begin{aligned}\tilde{I}(a) = \tilde{I}(b) &\Leftrightarrow \tilde{I}(\beta^{2n}a) = \tilde{I}(\beta^{2n}b) \Leftrightarrow I(U \circ f) = I(U \circ g) \Leftrightarrow V(f) \\ &= V(g) \Leftrightarrow f \sim g.\end{aligned}$$

Note now that h_n is comonotonic with f_n and g_n (and with x^*) and therefore by (CS) $f \sim g$ iff

$$f^h := (x^*, \ldots, x^*, \underbrace{h_n}_{n}, \underbrace{f_n}_{n+1}, x^*, \ldots) \sim (x^*, \ldots, x^*, \underbrace{h_n}_{n}, \underbrace{g_n}_{n+1}, x^*, \ldots) =: g^h$$

and therefore $V(f^h) = V(g^h)$. Since $U \circ f^h = \beta^{2n}c + \beta^{2n+1}a$ and $U \circ g^h = \beta^{2n}c + \beta^{2n+1}b$ then $\tilde{I}(\beta^{2n}c + \beta^{2n+1}a) = \tilde{I}(\beta^{2n}c + \beta^{2n+1}b)$. Therefore (using Lemma A.6) we proved that $\tilde{I}(a) = \tilde{I}(b) \Leftrightarrow \tilde{I}(c + \beta a) = \tilde{I}(c + \beta b)$. However by Lemma A.6 we have

$$\tilde{I}(a) = \tilde{I}\left(\frac{\beta}{\beta}a\right) = \beta\tilde{I}\left(\frac{1}{\beta}a\right) \Leftrightarrow \tilde{I}\left(\frac{1}{\beta}a\right) = \frac{1}{\beta}\tilde{I}(a)$$

and therefore

$$\begin{aligned}\tilde{I}(a) = \tilde{I}(b) &\Leftrightarrow \tilde{I}\left(\frac{1}{\beta}a\right) = \tilde{I}\left(\frac{1}{\beta}b\right) \Leftrightarrow \tilde{I}\left(c + \beta\frac{1}{\beta}a\right) \\ &= \tilde{I}\left(c + \beta\frac{1}{\beta}b\right) \Leftrightarrow \tilde{I}(a + c) = \tilde{I}(b + c)\end{aligned}$$

□

We will prove now that $\tilde{I}$ satisfies comonotonic additivity on B^o.

Lemma A.8 ($\tilde{I}$ satisfies comonotonic additivity) *Let $a, b \in B^o$ be comonotonic, then $\tilde{I}(a + b) = \tilde{I}(a) + \tilde{I}(b)$.*

Proof Take $a, b \in B^o$ s.t. a is comonotonic with b. By Lemma A.6 (translation invariance) $\tilde{I}(a) = \tilde{I}(0 + \tilde{I}(a))$. Since b is comonotonic with a and with the constant function $\tilde{I}(a)$, $\tilde{I}(a + b) = \tilde{I}(\tilde{I}(a) + b) = \tilde{I}(a) + \tilde{I}(b)$ the first equality coming from Lemma A.7 and the second one from Lemma A.6. □

Let $a, b \in B^o$, then $a \geq b$ means $a(\omega) \geq b(\omega)$ for all $\omega \in \Omega$. We will prove now that $\tilde{I}$ is monotone.

Lemma A.9 ($\tilde{I}$ is monotone) *Let $a, b \in B^o$ be such that $a \geq b$, then $\tilde{I}(a) \geq \tilde{I}(b)$.*

Proof By Lemma A.5 there is $n \in \mathbb{N}$ such that $\beta^n a, \beta^n b \in \mathcal{U}$. Let $f, g \in \mathcal{H}$ be such that $U \circ f = \beta^n a$ and $U \circ g = \beta^n b$. Then $U \circ f(\omega) \geq U \circ g(\omega)$ for all $\omega \in \Omega$ and by monotonicity $f \succsim g$. Hence $V(f) \geq V(g) \Leftrightarrow \tilde{I}(\beta^n a) \geq \tilde{I}(\beta^n a) \Leftrightarrow \tilde{I}(a) \geq \tilde{I}(a)$ (where the last equivalence comes from Lemma A.6). □

We will prove now that $\tilde{I}$ is positively homogeneous.

Lemma A.10 ($\tilde{I}$ is positively homogeneous) *For all $\alpha \geq 0$, for all $a \in B^o$ $\tilde{I}(\alpha a) = \alpha \tilde{I}(a)$.*

Proof This comes from Lemma A.8 and Lemma A.9 as noticed by Schmeidler (1986) in Remark 1 p. 256. □

We proved therefore that $\tilde{I} : B^o \to \mathbb{R}$ satisfies comonotonic additivity (Lemma A.8) and positive homogeneity (Lemma A.10) and thus defining defining $v(A) = \tilde{I}(1_A)$ for $A \in \mathcal{F}$, we can use Proposition 1 of Schmeidler (1986) and we can conclude that for all $a \in B^o$

$$\tilde{I}(a) = \int a dv.$$

Hence for all $f, g \in \mathcal{H}$,

$$f \succsim g \text{ iff } I(U \circ f) \geq I(U \circ g) \text{ iff } \int \sum_t \beta^t u(f_t) dv \geq \int \sum_t \beta^t u(g_t) dv.$$

We turn to uniqueness. The fact that that β is unique and u is unique up to positive affine transformation comes from Proposition 5.
Suppose now that the preference relation is represented by $J(U \circ f) := (1 - \beta) \int \sum_t \beta^t u(f_t) dv'$. Fix $A \in \mathcal{F}$, let $x^1 \in X$ be such that $u(x^1) = 1$ and consider the stream $f \in \mathcal{H}$ defined by

$$f_t(\omega) = \begin{cases} x^1 & \text{if } \omega \in A \\ x^0 & \text{otherwise} \end{cases}$$

for every $t \in \mathbb{N}$. Note that $U \circ f = \frac{1_A}{1-\beta}$ and therefore $J(U \circ f) = v'(A)$ and $I(U \circ f) = v(A)$. Take $x \in X$ such that $u(x) = v(A)$ and define $g \in \mathcal{H}$ by $g := (x, x, \ldots)$. Since $U \circ g = v(A)$ we obtain that $I(U \circ g) = v(A) = I(U \circ f)$ and therefore $f \sim g$. Note that $J(U \circ g) = v(A)$, and since we supposed that J represents the preference relation over $\mathcal{H}$, then $v(A) = J(U \circ g) = J(U \circ f) = v'(A)$. □

Proof of Theorem 1 Repeating the same steps as in Theorem 2, we can note that the functional $\tilde{I}$ is additive on B^o, i.e. for all $a, b \in B^o$

$$\tilde{I}(a+b) = \tilde{I}(a) + \tilde{I}(b).$$

This comes from the fact that (SS) does not restrict additivity to comonotonic acts. Consider now two sets $A, B \in \mathcal{F}$ s.t. $A \cap B = \emptyset$. We have that $1_{A\cup B} = 1_A + 1_B$. Therefore

$$v(A \cup B) = \tilde{I}(1_{A\cup B}) = \tilde{I}(1_A + 1_B) = \tilde{I}(1_A) + \tilde{I}(1_B) = v(A) + v(B).$$

Which implies that v is a probability. □

Proof of Theorem 3 Necessity is shown as in the proof of Theorem 2, using the fact that when v is convex, $\int (a+b)dv \geq \int adv + \int bdv$ for all $a, b \in B^o$ (see Proposition 3 of Schmeidler (1986)).

We prove sufficiency. Since (PS) implies (CS) the proof of Theorem 2 is valid. The only thing that is needed to show is that v is convex, i.e. $v(A \cup B) + v(A \cap B) \geq v(A) + v(B)$.

Let $\tilde{I} : B^o \to \mathbb{R}$ be the functional defined in the proof of Theorem 2.

Lemma A.11 *Let $a, b, c \in B^o$ be such that c is comonotonic with b. Then $\tilde{I}(a) = \tilde{I}(b) \Rightarrow \tilde{I}(a+c) \geq \tilde{I}(b+c)$.*

Proof It follows using (PS) and doing the proof as in Lemma A.7. □

Let $A, B \in \mathcal{F}$. Note that $\tilde{I}(1_A) = v(A) = \tilde{I}(v(A)1_\Omega)$ and $\tilde{I}(1_B) = v(B) = \tilde{I}(v(B)1_\Omega)$. Since 1_B is comonotonic with $v(A)1_\Omega$, by Lemma A.11 it follows $\tilde{I}(1_A + 1_B) \geq \tilde{I}(v(A)1_\Omega + 1_B) = v(A) + \tilde{I}(1_B) = v(A) + v(B)$. Note that $1_A + 1_B = 1_{A\cup B} + 1_{A\cap B}$, and moreover $1_{A\cup B}$ and $1_{A\cap B}$ are comonotone. Hence by comonotonic additivity of the Choquet integral

$$\begin{aligned} v(A \cup B) + v(A \cap B) &= \tilde{I}(1_{A\cup B}) + \tilde{I}(1_{A\cap B}) = \tilde{I}(1_{A\cup B} + 1_{A\cap B}) \\ &= \tilde{I}(1_A + 1_B) \geq v(A) + v(B). \end{aligned}$$

□

Proof of Theorem 4 The proof follows from Theorem 1 of Kochov (2015) and the Proposition in Schmeidler (1989) p. 582. □

Proof of Proposition 1 Fix $h \in \mathcal{H}$. Since h is bounded there exists a compact set $K_h \subseteq X$ such that $\cup_t h_t(\Omega) \subset K_h$. Since $u : X \to \mathbb{R}$ is continuous, we can find $M \in \mathbb{R}_+$ such that $|u(x)| \leq M$ for all $x \in K_h$. Hence $\forall t \in \mathbb{N}$ and $\forall \omega \in \Omega$, $\beta^t |u(h_t(\omega))| \leq \beta^t M$. Since $\sum_t \beta^t M$ converges to $\frac{M}{1-\beta}$, by Theorem 7.10 of Rudin (1976), the series of functions $\sum_t \beta^t u(h_t)$ converges uniformly on Ω. Define $H_n = \sum_{t=0}^{n} \beta^t u(h_t)$ and $H = \sum_{t=0}^{\infty} \beta^t u(h_t)$. Since H_n converges uniformly to H on Ω, for every $\epsilon > 0$, there exists $N \in \mathbb{N}$ s.t. $n \geq N$ implies $\sup_\omega |H_n(\omega) - H(\omega)| < \epsilon$. Hence, for $n \geq N$, using the fact that P is a probability

$$\left| \int H_n dP - \int H dP \right| \leq \int |H_n - H| \, dP < \int \epsilon dP = \epsilon$$

where the first inequality follows from Theorem 4.4.4 (ii) and (iii) of Rao and Rao (1983) (note that H_n and H are simple functions by the finiteness of acts). This implies that the series converges and $\lim_n \int H_n dP = \int H dP$. Rewriting explicitly we obtain

$$\begin{aligned} \sum_t \beta^t \mathbb{E}_P \left[u(h_t) \right] &= \lim_n \sum_{t=0}^n \beta^t \int u(h_t) dP \\ &= \lim_n \int \sum_{t=0}^n \beta^t u(h_t) dP = \int \sum_t \beta^t u(h_t) dP = \mathbb{E}_P \left[\sum_t \beta^t u(h_t) \right]. \end{aligned}$$

□

Proof of Proposition 3 We only prove that formula (2) implies (CS') as the other implication can be readily shown.
Fix $t \in \mathbb{N}$, $d \in \mathcal{D}$ and $f, g, h \in \mathcal{H}$ such that h_t is comonotonic with f_i and g_i for all $i \geq t$. Note that h_t is comonotonic with f_i if and only if for all $\omega, \omega' \in \Omega$

$$\left[u(h_t(\omega)) - u(h_t(\omega')) \right] \left[u(f_i(\omega)) - u(f_i(\omega')) \right] \geq 0,$$

and the same is true for g_i. Let ω and ω' be in Ω, we have

$$\begin{aligned} &\left[u(h_t(\omega)) - u(h_t(\omega')) \right] \left[\sum_{i \geq t} \beta^{i+1} u(f_i(\omega)) - \sum_{i \geq t} \beta^{i+1} u(f_i(\omega')) \right] \\ &= \lim_{n \to \infty} \sum_{i=t}^n \beta^{i+1} \left[u(h_t(\omega)) - u(h_t(\omega')) \right] \left[u(f_i(\omega)) - u(f_i(\omega')) \right] \geq 0. \end{aligned}$$

Therefore h_t is comonotonic with $\sum_{i \geq t} \beta^{i+1} u(f_i)$ and $\sum_{i \geq t} \beta^{i+1} u(g_i)$. Denoting $\delta = \sum_{i=1}^{t-1} \beta^i u(d_i)$, and using the fact that the Choquet integral satisfies comonotonic additivity and positive homogeneity we get

$$\begin{aligned} &(d_0, \ldots, d_{t-1}, h_t, f_t, f_{t+1}, \ldots) \succsim (d_0, \ldots, d_{t-1}, h_t, g_t, g_{t+1}, \ldots) \Leftrightarrow \\ &\int \delta + \beta^t u(h_t) + \sum_{i \geq t} \beta^{i+1} u(f_i) dv \geq \int \delta + \beta^t u(h_t) + \sum_{i \geq t} \beta^{i+1} u(g_i) dv \Leftrightarrow \\ &\delta + \int \beta^t u(h_t) dv + \beta \int \sum_{i \geq t} \beta^i u(f_i) dv \geq \delta + \int \beta^t u(h_t) dv + \beta \int \sum_{i \geq t} \beta^i u(g_i) dv \Leftrightarrow \end{aligned}$$

$$\int \sum_{i \geq t} \beta^i u(f_i) dv \geq \int \sum_{i \geq t} \beta^i u(g_i) dv \Leftrightarrow$$
$$\int \delta + \sum_{i \geq t} \beta^i u(f_i) dv \geq \int \delta + \sum_{i \geq t} \beta^i u(g_i) dv \Leftrightarrow$$
$$(d_0, \ldots, d_{t-1}, f_t, f_{t+1}, \ldots) \succsim (d_0, \ldots, d_{t-1}, g_t, g_{t+1}, \ldots).$$

□

References

Anscombe, F., Aumann, R.J.: A definition of subjective probability. Ann. Math. Stat. **34**(1), 199–205 (1963)
Araujo, A., Novinski, R., Páscoa, M.: General equilibrium, wariness and efficient bubbles. J. Econ. Theory **146**, 785–811 (2011)
Asano, T., Kojima, H.: Consequentialism and dynamic consistency in updating ambiguous beliefs. Econ. Theory **68**(1), 223–250 (2019)
Bastianello, L., Chateauneuf, A.: About delay aversion. J. Math. Econ. **63**, 62–77 (2016)
Bleichrodt, H., Rohde, K.I., Wakker, P.: Koopmans' constant discounting for intertemporal choice: a simplification and a generalization. J. Math. Psychol. **52**(6), 341–347 (2008)
Bommier, A., Kochov, A., Le Grand, F.: On monotone recursive preferences. Econometrica **85**(5), 1433–1466 (2017)
Bommier, A., Kochov, A., Le Grand, F.: Ambiguity and endogenous discounting. J. Math. Econ. **83**, 48–62 (2019)
Chambers, C.P., Echenique, F.: On multiple discount rates. Econometrica **86**(4), 1325–1346 (2018)
Chateauneuf, A.: Modeling attitudes towards uncertainty and risk through the use of Choquet integral. Ann. Oper. Res. **52**(1), 1–20 (1994)
Chateauneuf, A., Eichberger, J., Grant, S.: Choice under uncertainty with the best and worst in mind: neo-additive capacities. J. Econ. Theory **137**(1), 538–567 (2007)
Chateauneuf, A., Faro, J.H.: Ambiguity through confidence functions. J. Math. Econ. **45**(9–10), 535–558 (2009)
Dempster, A.P.: Upper and lower probabilities induced by a multi-valued mapping. Ann. Math. Stat. **38**, 325–339 (1967)
Dempster, A.P.: A generalization of Bayesian inference. J. Roy. Stat. Soc.: Ser. B (Methodol.) **30**(2), 205–232 (1968)
Dominiak, A., Lefort, J.P.: Unambiguous events and dynamic Choquet preferences. Econ. Theory **46**, 401–425 (2011)
Ebert, J.E., Prelec, D.: The fragility of time: time-insensitivity and valuation of the near and far future. Manag. Sci. **53**(9), 1423–1438 (2007)
Eichberger, J., Grant, S., Kelsey, D.: Updating Choquet beliefs. J. Math. Econ. **43**, 888–899 (2007)
Ellsberg, D.: Risk, ambiguity, and the savage axioms. Q. J. Econ. **75**, 643–669 (1961)
Epstein, L.G.: Stationary cardinal utility and optimal growth under uncertainty. J. Econ. Theory **31**(1), 133–152 (1983)
Epstein, L.G., Schneider, M.: Recursive multiple-priors. J. Econ. Theory **113**(1), 1–31 (2003)
Epstein, L.G., Tanny, S.M.: Increasing generalized correlation: a definition and some economic consequences. Can. J. Econ. **13**, 16–34 (1980)
Epstein, L.G., Wang, T.: Uncertainty, risk-neutral measures and security price booms and crashes. J. Econ. Theory **67**(1), 40–82 (1995)
Fagin, R., Halpern, J.Y.: A new approach to updating beliefs. In: Bonissone, P.P., Henrion, M., Kanal, L.N., Lemmer, J.F. (eds.) Uncertainty in Artificial Intelligence, vol. 6, pp. 347–374. Springer, Berlin (1991)
Gilboa, I.: Expectation and variation in multi-period decisions. Econometrica **57**, 1153–1169 (1989)
Ghirardato, P., Marinacci, M.: Ambiguity made precise: a comparative foundation. J. Econ. Theory **102**, 251–289 (2002)

Gilboa, I., Schmeidler, D.: Maxmin expected utility with non-unique prior. J. Math. Econ. **18**(2), 141–153 (1989)

Gilboa, I., Schmeidler, D.: Updating ambiguous beliefs. J. Econ. Theory **59**(1), 33–49 (1993)

Gul, F., Pesendorfer, W.: Evaluating ambiguous random variables from Choquet to maxmin expected utility. J. Econ. Theory **192**, 105129 (2018)

Hayashi, T.: Quasi-stationary cardinal utility and present bias. J. Econ. Theory **112**, 343–352 (2003)

Heath, C., Tversky, A.: Preference and belief: ambiguity and competence in choice under uncertainty. J. Risk Uncertain. **4**(1), 5–28 (1991)

Horie, M.: Reexamination on updating Choquet beliefs. J. Math. Econ. **49**, 467–470 (2013)

Jaffray, J.Y.: Bayesian updating and belief functions. IEEE Trans. Syst. Man Cybern. **22**(5), 1144–1152 (1992)

Jegadeesh, N.: Evidence of predictable behavior of security returns. J. Finance **45**, 881–898 (1990)

Kochov, A.: Time and no lotteries: an axiomatization of maxmin expected utility. Econometrica **83**, 239–262 (2015)

Klibanoff, P., Marinacci, M., Mukerji, S.: A smooth model of decision making under ambiguity. Econometrica **73**, 1849–1892 (2005)

Koopmans, T.C.: Stationary ordinal utility and impatience. Econometrica **28**, 287–309 (1960)

Koopmans, T.C.: Representation of preference orderings over time. Decis. Organ. **57**, 100 (1972)

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

Laibson, D.: Golden eggs and hyperbolic discounting. Q. J. Econ. **112**, 443–477 (1997)

Loewenstein, G., Prelec, D.: Anomalies in intertemporal choice: evidence and an interpretation. Q. J. Econ. **107**(2), 573–597 (1992)

Maccheroni, F., Marinacci, M., Rustichini, A.: Ambiguity aversion, robustness, and the variational representation of preferences. Econometrica **74**, 1447–1498 (2006)

Montiel Olea, J.L., Strzalecki, T.: Axiomatization and measurement of quasi-hyperbolic discounting. Q. J. Econ. **129**(3), 1449–1499 (2014)

Peitler, P.: Foundations of Intertemporal Choice under Uncertainty. Working paper (2019)

Phelps, E.S., Pollak, P.A.: On second-best national saving and game-equilibrium growth. Rev. Econ. Stud. **35**(2), 185–199 (1968)

Pires, C.P.: A rule for updating ambiguous beliefs. Theory Decis. **53**, 137–152 (2002)

Rao, K.B., Rao, M.B.: Theory of Charges: A Study of Finitely Additive Measures, vol. 109. Academic Press, Cambridge (1983)

Rosenberg, B., Rudd, A.: Factor-related and specific returns of common stocks: serial correlation and market inefficiency. J. Finance **37**, 543–554 (1982)

Rudin, W.: Principle of Mathematical Analysis, 3rd edn. McGraw-Hill, New York (1976)

Samuelson, P.: A note on measurement of utility. Rev. Econ. Stud. **4**, 155–161 (1937)

Sarin, R., Wakker, P.: Dynamic choice and nonexpected utility. J. Risk Uncertain. **17**, 87–119 (1998)

Savage, L.J.: The Foundations of Statistics. Courier Corporation, Chelmsford (1972)

Schmeidler, D.: Integral representation without additivity. Proc. Am. Math. Soc. **97**, 255–261 (1986)

Schmeidler, D.: Subjective probability and expected utility without additivity. Econometrica **57**, 571–587 (1989)

Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, Princeton (1976)

Siniscalchi, M.: Dynamic choice under ambiguity. Theor. Econ. **6**(3), 379–421 (2011)

Skiadas, C.: Recursive utility and preferences for information. Econ. Theory **12**(2), 293–312 (1998)

Tversky, A., Kahneman, D.: Advances in prospect theory: cumulative representation of uncertainty. J. Risk Uncertain. **5**(4), 297–323 (1992)

Wakker, P.: Characterizing optimism and pessimism directly through comonotonicity. J. Econ. Theory **52**, 453–463 (1990)