

האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM

**ADDITIVE VALUATIONS OF STREAMS OF
PAYOFFS THAT OBEY THE TIME-VALUE
OF MONEY PRINCIPLE:
CHARACTERIZATION AND ROBUST
OPTIMIZATION**

By

ABRAHAM NEYMAN

מרכז פדרמן לחקר הרציונליות

**THE FEDERMANN CENTER FOR
THE STUDY OF RATIONALITY**

**Feldman Building, Edmond J. Safra Campus,
Jerusalem 91904, Israel**
PHONE: [972]-2-6584135 FAX: [972]-2-6513681
E-MAIL: ratio@math.huji.ac.il
URL: <http://www.ratio.huji.ac.il/>

**ADDITIVE VALUATIONS OF STREAMS OF PAYOFFS
THAT OBEY THE TIME-VALUE OF MONEY
PRINCIPLE: CHARACTERIZATION AND ROBUST
OPTIMIZATION**

ABRAHAM NEYMAN*
APRIL 29, 2018

ABSTRACT. This paper characterizes the preferences over bounded infinite utility streams that satisfy the time-value of money principle and an additivity property, and preferences that in addition are impatient. Based on this characterization, the paper introduces a concept of optimization that is robust to a small imprecision in the specification of the preference, and proves that the set of feasible streams of payoffs of a finite Markov decision process admits such a robust optimization.

Date: April 29, 2018.

* Institute of Mathematics, and Federmann Center for the Study of Rationality, The Hebrew University of Jerusalem, Givat Ram, Jerusalem 91904, Israel. *e-mail:* aneyman@math.huji.ac.il

This research was supported in part by Israel Science Foundation grant 1596/10. The author wishes to thank Elon Kohlberg for many valuable comments throughout the writing of this paper, and to thank Elchanan Ben-Porath, Sergiu Hart, and Eilon Solan for comments on an earlier draft.

1. INTRODUCTION

In many decision problems a decision maker must choose between different possible streams of payoffs. The decision maker can be an individual, a firm, or a community of individuals. The stream of payoffs may be a stream of equal payoffs (called a *perpetuity*) or of payoffs that vary over time.

The first objective of this paper is to characterize all preferences, over bounded streams of payoffs, that satisfy a few plausible assumptions.

The starting assumption is that any stream is equivalent to a perpetuity and the higher the perpetuity's (constant) payoff, the better. A preference on bounded streams of payoffs that satisfies this assumption is represented by a unique ordinal utility function that assigns to each stream of bounded payoffs its equivalent perpetuity's payoff.

The second assumption is that the preference obeys the *time-value of money* principle. The time-value of money principle reflects the preference of expediting the receipt of positive payoffs: the faster the accumulation of payoffs, the better. In other words, this principle states that a unit payoff in a given period is preferable to its being spread out over future periods. This principle is natural when saving is costless.

The third assumption is *additivity*. The *additivity property* states that if the streams A and B are equivalent to the perpetuities C and D, respectively, then the sum of the streams A and B is equivalent to the sum of the perpetuities C and D.

The unique ordinal utility function that represents a preference that satisfies these three assumption and assigns to each stream its equivalent perpetuity's payoff is called a *valuation*.

A valuation is *impatient* if the contribution of payoffs in the distant future is negligible. It is *patient* if it is neutral to the timing of payoffs.

We characterize the impatient valuations, the patient valuations, and the set of all valuations.

The characterization shows that (1) any impatient valuation is a weighted average of the periods' payoffs with averaging weights that are nonincreasing in time, (2) any patient valuation is a linear function that assigns to each stream a value that is between the limit inferior and the limit superior of the averages of the first n payoffs in the stream, and (3) any valuation is a weighted average of an impatient one and a patient one.

Two classical examples of impatient valuations are the n -th *Cesàro average valuation*, which is denoted by u_n , and the r -discounted valuation ($0 < r \leq 1$), which is denoted by u_r : For a stream $f = (f_1, f_2, \dots)$,

$$u_n(f) = \frac{f_1 + \dots + f_n}{n} \quad \text{and} \quad u_r(f) = \sum_{t=1}^{\infty} r(1-r)^{t-1} f_t.$$

The t -th period's averaging weight of u_n is 1 if $t \leq n$ and is 0 if $t > n$, and the t -th period's averaging weight of u_r is $r(1-r)^{t-1}$.

We now turn to our second topic: optimization that is robust to a small imprecision in the specification of the preference.

As there is a one-to-one correspondence between the preferences (that satisfy our assumptions) and the valuations, it suffices to study optimization that is robust to a small imprecision in the specification of the valuation.

Optimization that is robust to small changes in the valuation is common in a bank's selection of its portfolio. A few considerations in selecting the portfolio are discussed as an illustration of the importance of robust optimization in selecting a proper feasible stream of payoffs.

A bank's portfolio includes assets that are composed mainly of a collection of loans, each with a different maturity and a different payment schedule, and liabilities that are composed of customers' (including other banks') deposits, bonds issued by the bank, etc.

The economic value of the bank is the present value of the stream of its portfolio payoffs. It is a function of the yield curve, which specifies the interest rate as a function of time.

The bank's set of feasible portfolios depends on market and competitive conditions, as well as on regulatory constraints. One of the regulatory constraints, as well as an important consideration in the bank's selection of its portfolio, is the sensitivity of its economic value to changes in the yield curve.¹

The objective of maximizing the value of the bank's portfolio while ensuring that the losses due to given changes in the yield curve remain within prescribed limits is essentially an approximate optimization that is robust to given imprecision in the specification of the valuation.

The yield curve, and hence also the valuation, changes over time. Therefore, an additional desired property of the bank's portfolio is that it can be modified gradually as the yield curve changes.

¹Obviously, there are other important sensitivity issues. We mention the sensitivity to the yield curve as the yield curve specifies the valuation.

We now continue with the introduction of the formal concept of a robust optimizer in a set F of streams of payoffs.

For any valuation v , the maximum (or more precisely, the supremum) of $v(g)$ over all streams g in F , is called the *v-optimal value of F* and is denoted by $v(F)$.

An imprecise specification of a valuation is modeled as a set U of valuations. The maximum (or more precisely, the supremum) of $u(g)$ over all streams g in F and valuation u in U , is called the *U-optimal value of F* and is denoted by $U(F)$.

Fix a nonnegative number $\varepsilon \geq 0$, a valuation v , a set of valuations U , a set of streams of payoffs F , and a stream f in F .

The stream $f \in F$ is an *ε -optimizer for v with respect to F* if $v(f)$ (which is at most the v -optimal value of F) is within ε of the v -optimal value of F (i.e., $v(f) \geq v(g) - \varepsilon$ for any $g \in F$).

The stream $f \in F$ is an *ε -optimizer for U with respect to F* if for any valuation u in U (f is an ε -optimizer for u , and moreover) $u(f)$ (which is at most the U -optimal value of F) is within ε of the U -optimal value of F (i.e., $u(f) \geq w(g) - \varepsilon$ for any valuation w in U and any stream g in F).

It follows that if the set F of streams of payoffs has an ε -optimizer for a set of valuations U , then the oscillation of the u -optimal value of F , where u ranges over all valuations in U , is at most ε .

An imprecision in the specification of a valuation is often expressed by stating that a fixed valuation v is a good proxy of the “true” valuation. Such an imprecise specification of the valuation u is modeled as the set of all valuations that are sufficiently similar to the fixed valuation v . This leads to the following important concept of robust optimization.

The stream $f \in F$ is a *robust ε -optimizer at v with respect to F* if there is a neighborhood² W of v such that f is an ε -optimizer for W with respect to F .

It follows that if the set F of streams of payoffs has, for every $\varepsilon > 0$, a robust ε -optimizer at a valuation v , then the u -optimal value is continuous at v .

A neighborhood of a patient valuation contains, for all sufficiently large n and all sufficiently small r , the n -th Cesàro average valuation

²The formal definition of a neighborhood depends on the topology on the space of valuation, which is defined in Section 2.9. This topology is defined by the semi-metric d , where $d(u, v)$ is the max norm of the difference between the averaging weights of u and v .

u_n and the r -discounted valuation u_r . Therefore, if $f \in F$ is a robust ε -optimizer at a patient valuation v with respect to F , then, for all sufficiently large n and all sufficiently small r , $f \in F$ is an ε -optimizer for u_n and for u_r with respect to F and the oscillation of the u_r -optimal and the u_n -optimal value of F is at most ε .

Therefore, the notion of robustness at a patient valuation provides a unifying view of earlier studies of robust optimization of a patient decision maker. Here we study robust optimization at any valuation, namely, at any mixture of a patient valuation and a patient one.

The ability to select a robust ε -optimizer at v (with respect to F) that can be changed gradually as the valuation v changes corresponds to the existence of a robust ε -optimizer at v (with respect to F) that varies continuously as a function of the valuation v . Theorem 4 shows that if F is convex and has, for any valuation v , a robust ε -optimizer at v with respect to F , then there is a robust ε -optimizer at v with respect to F that depends continuously on v .

One may argue that impatience is a natural assumption of a preference over streams of payoffs and that it is therefore sufficient to confine the analysis to impatient valuations.

However, in order to model the imprecision in the specification of the impatient valuation, it may be advantageous to fix a non-impatient valuation, and then consider all the impatient valuations in its neighborhood.

For example, consider the imprecise specification of an impatient valuation that is obtained by specifying that its averaging weights are sufficiently small, e.g., less than one hundredth. This can be modeled as the set U of all impatient valuations that are one hundredth close to a patient valuation v .

We illustrate the importance of approximate optimizers and the advantage of patient valuations by considering the set F of feasible streams of payoffs that consists of the perpetuity $\mathbf{1}$, with a constant payoff 1, and the streams f^k , $k \geq 0$, where the payoff is 2 in the first period and in each of the first k even periods, and the payoff is 0 in all other periods.

If our objective is to select the best stream in F , given that the impatient valuation places a very small weight on each individual period, then it seems intuitive that we should select the perpetuity $\mathbf{1}$. But while for any patient valuation $\mathbf{1}$ is the unique optimizer and no other

stream in F is even a 0.99-optimizer for³ U with respect to F , $\mathbf{1}$ is not an optimizer for any specific impatient valuation but is a 0.02-optimizer for any⁴ $u \in U$ with respect to F .

Other examples arise in modeling a preference of an impatient decision maker who has a pretty good idea of the “interest rate” between successive points in time, as long as these are not too distant; however, as regards the very distant future, he cannot tell much beyond the fact that the interest rates remain nonnegative; furthermore, he wants to give the very distant future a non-zero weight, say 30%. Such preferences arise naturally in decision problems that involve pollution, global warming, etc.

The advantage of using valuations that are not impatient in describing a small imprecision in the specification of an impatient valuation is analogous to the advantage of using boundary points of an interval, or a circle, in describing a small imprecision in the specification of an interior point, e.g., an interior point that is sufficiently close to a fixed boundary point.

We now turn to our third topic: existence of robust optimization in a Markov decision process (MDP); see Section 2.12.

In many decision problems, e.g., in a MDP, the decision maker faces stochastic randomness. Thus a choice of a policy does not determine a single stream of payoffs but rather a distribution over streams of payoffs. The decision problem, then, is to choose from a set of distributions over a bounded set of streams of payoffs.

By assigning to each distribution the stream of expected periods’ payoffs, we can transform a decision problem where the choice is from a set of distributions to one where the choice is from a set of streams of payoffs.

This assignment, along the earlier introduced concepts of robust optimization, enables us to analyze robust optimization in the model of a MDP.

A policy π in a MDP determines a distribution P_π on the streams of payoffs. The distribution P_π determines the stream of expected (periods’) payoffs \hat{P}_π .

³For any patient valuation v , $v(\mathbf{1}) = 1$ and $v(f^k) = 0$ for all k (since the limit of the average of the first n payoffs is 1 or 0, respectively.)

⁴By the time value of money principle, a stream of alternating 0’s and 2’s is worth at most as much as a constant stream of 1’s. For a valuation $u \in U$, the extra 2 in the first period contributes at most $2/100$; hence, we have $u(f^k) \leq u(1) + 0.02$.

The set of feasible streams of payoffs in a MDP is the set of all streams \hat{P}_π , where π ranges over all policies of the decision maker.

A policy π in a MDP is an ε -optimal policy for v , respectively, a robust ε -optimal policy at v , if \hat{P}_π is an ε -optimizer for v with respect to $\{\hat{P}_\pi : \pi \text{ a policy}\}$, respectively, a robust ε -optimizer at v with respect to $\{\hat{P}_\pi : \pi \text{ a policy}\}$.

Theorem 5 shows that any finite MDP has, for every $\varepsilon > 0$ and every valuation v , a robust ε -optimal policy at v .

Theorem 6 states that a finite MDP has, for every $\varepsilon > 0$ and every valuation v , a robust v -optimal policy at ε that satisfies a stringent robustness property.

2. FORMAL DEFINITIONS AND STATEMENTS OF THE RESULTS

2.1. Streams of payoffs. A stream of payoffs is a sequence $g = (g_1, g_2, \dots)$ of real numbers. It is bounded if $\|g\| := \sup_t |g_t| < \infty$. The linear space of all bounded streams of payoffs is denoted by ℓ_∞ .

For $g, h \in \ell_\infty$ and $a \in \mathbb{R}$, $g + h$ is the element $(g_1 + h_1, g_2 + h_2, \dots)$ of ℓ_∞ , i.e., the t -th coordinate of $g + h$ is $g_t + h_t$, and ag is the element (ag_1, ag_2, \dots) of ℓ_∞ , i.e., the t -th coordinate of ag is ag_t .

2.2. Linearity. The t -th coordinate, g_t , of the stream g is often interpreted as the utility of consumption at stage t , and several classical sets of axioms (see [2, 3]) lead to a presentation of a utility over infinite streams of consumption that is a linear function of the stream g .

A real-valued function u that is defined on ℓ_∞ is *additive* if for every $g, h \in \ell_\infty$, $u(g + h) = u(g) + u(h)$. As $\mathbf{0} + \mathbf{0} = \mathbf{0}$, where $\mathbf{0} = (0, 0, \dots)$, an additive u satisfies $u(\mathbf{0}) = 0$.

A real-valued function u that is defined on ℓ_∞ is *linear* if it is additive and $u(ag) = au(g)$ for every $g \in \ell_\infty$ and $a \in \mathbb{R}$.

2.3. The time-value of money principle. This principle captures two desirable properties of a real-valued function $u : \ell_\infty \rightarrow \mathbb{R}$ that represents a preference over streams of payoffs.

The first is monotonicity: the higher the stage payoffs the better. For an additive u , the monotonicity is equivalent to the property that a stream of nonnegative payoffs is at least as desirable as the stream of zero payoffs.

The second desirable property of u expresses the fact that the earlier the payments the better. In particular, a unit payoff in a given period

is at least as desirable as its spread over later periods. In particular, $u(\mathbf{e}_t) \geq u(\mathbf{e}_{t+1})$ for all t , where \mathbf{e}_t is the t -th unit vector in ℓ_∞ .

An additive u satisfies $u(\mathbf{e}_t) \geq u(\mathbf{e}_{t+1})$ for all t iff for any two streams g and h with finitely many periods of nonzero payoffs such that $\sum_{t=1}^s g_t \geq \sum_{t=1}^s h_t \forall s$, $u(g) \geq u(h)$.

Definition 1. A real-valued function u that is defined on ℓ_∞ satisfies the time-value of money principle if:

For every two streams g and h such that $\sum_{t=1}^s g_t \geq \sum_{t=1}^s h_t \forall s$,
 $u(g) \geq u(h)$.

Remark 1. A function $u : \ell_\infty \rightarrow \mathbb{R}$ that satisfies the time-value of money principle is monotonic, i.e., $u(g) \geq u(h)$ whenever $g_t \geq h_t \forall t$, and satisfies $u(\mathbf{e}_t) \geq u(\mathbf{e}_{t+1})$ for all t .

Remark 2. An additive and monotonic function $u : \ell_\infty \rightarrow \mathbb{R}$ satisfies $u(\mathbf{e}_t) \geq 0$ and $\sum_{t=1}^\infty u(\mathbf{e}_t) < \infty$, and therefore $u(\mathbf{e}_t)$ goes to zero as t goes to infinity.

2.4. Valuations.

Definition 2. A real-valued function $u : \ell_\infty \rightarrow \mathbb{R}$ is normalized if: $u(\mathbf{1}) = 1$ where $\mathbf{1} = (1, 1, \dots)$.

Definition 3. A normalized additive real-valued function $u : \ell_\infty \rightarrow \mathbb{R}$ that satisfies the time-value of money principle is called a valuation.

Two classical examples of valuations are:
 The r -discounted valuation: $u_r(g) = \sum_{t=1}^\infty r(1-r)^{t-1}g_t$, $0 < r < 1$,
 and the k -th Cesàro average valuation: $u_k(g) = \bar{g}_k := \frac{1}{k} \sum_{t=1}^k g_t$, $k = 1, 2, \dots$

2.5. Impatient valuations. Let $\mathbf{1}_{>n}$ be the stream of payoffs $g = (g_1, g_2, \dots)$ with $g_t = 1 \forall t > n$ and $g_t = 0 \forall t \leq n$.

Definition 4. An impatient valuation is a valuation u such that

$$u(\mathbf{1}_{>n}) \rightarrow_{n \rightarrow \infty} 0.$$

Remark 3. If u is an impatient valuation then $u(g_1, g_2, \dots, g_n, 0, 0, \dots)$ converges to $u(g)$ as n goes to infinity, where $(g_1, g_2, \dots, g_n, 0, 0, \dots)$ stands for the stream whose t -coordinate equals g_t if $t \leq n$ and equals 0 if $t > n$.

Moreover, if u is an impatient valuation then $u(g_1, \dots, g_n, h_{n+1}, \dots)$, where $g, h \in \ell_\infty$ and $(g_1, \dots, g_n, h_{n+1}, \dots)$ stands for the stream whose t -coordinate equals g_t if $t \leq n$ and equals h_t if $t > n$, converges to $u(g)$ as n goes to infinity.

The two above-mentioned classical valuations are impatient valuation.

A normalized, impatient, and additive $u : \ell_\infty \rightarrow \mathbb{R}$ that obeys $u(\mathbf{e}_t) \geq u(\mathbf{e}_{t+1})$ for all t , satisfies the time-value of money property. See Lemma 1.

Therefore, a real-valued function that is defined on ℓ_∞ is an impatient valuation iff it is normalized, linear, and $u(\mathbf{e}_t) \geq u(\mathbf{e}_{t+1})$ for all t .

The first result characterizes all impatient valuations.

Theorem 1. *A real-valued function u that is defined on ℓ_∞ is an impatient valuation iff there are weights ω_t , where $t \geq 1$ ranges over the positive integers, with $\omega_t \geq \omega_{t+1} \geq 0$ and $\sum_{t=1}^{\infty} \omega_t = 1$, such that*

$$u(g) = \sum_{t=1}^{\infty} \omega_t g_t.$$

The weights representing the r -discounted valuation u_r are $\omega_t = r(1-r)^t$, and those representing the k -th Cesàro average valuation u_k are $\omega_t = 1/k$ if $t \leq k$ and $\omega_t = 0$ if $t > k$.

2.6. Convergence of impatient valuations. Next, we define convergence of a sequence of impatient valuations.

Definition 5. *A sequence u^k of impatient valuations converges if for every positive integer t the sequence $u^k(\mathbf{e}_t)$ converges as $k \rightarrow \infty$.*

The subspace of ℓ_∞ of all converging sequences $g \in \ell_\infty$, i.e., the limit of g_t exists as t goes to infinity, is denoted by c . An equivalent definition of convergence of a sequence of impatient valuations follows.

Remark 4. *A sequence v^k of impatient valuations converges iff $v^k(g)$ converges for every $g \in c$.*

It follows that the limit of a converging sequence of impatient valuations defines a real-valued function on c . On this restricted domain, the “limit” v satisfies the following properties of a valuation: linearity, $v(\mathbf{1}) = 1$, and the time-value of money property.

Examples of converging sequences of impatient valuations: the k -th Cesàro average valuations, u_k , converge as k goes to infinity and the r -discounted valuations, u_r , converge as $r > 0$ goes to zero.

The “limit” v of a sequence of impatient valuations need not coincide with the restriction of an impatient valuation to the domain c . For example, if v is the “limit” of u_k , then, for every fixed n , the sequence $u_k(\mathbf{1}_{>n})$ converges to 1 as k goes to infinity, and therefore $v(\mathbf{1}_{>n}) = 1$; hence, v is not impatient.

2.7. Patient valuations.

Definition 6. *A patient valuation is a valuation u such that*

$$u(\mathbf{1}_{>n}) = 1.$$

The second result characterizes the patient valuations.

Theorem 2. *A real-valued function u that is defined on ℓ_∞ is a patient valuation iff it is a linear function on the bounded streams of payoffs such that*

$$(4) \quad \liminf_{n \rightarrow \infty} \bar{g}_n \leq u(g) \leq \limsup_{n \rightarrow \infty} \bar{g}_n.$$

The lower and upper bounds in (4), $\liminf_{n \rightarrow \infty} \bar{g}_n$ and $\limsup_{n \rightarrow \infty} \bar{g}_n$, are tight. The tightness follows from Lemma 2 that shows that for any stream $g \in \ell_\infty$ there are patient valuations v and w such that $v(g) = \liminf_{n \rightarrow \infty} \bar{g}_n$ and $w(g) = \limsup_{n \rightarrow \infty} \bar{g}_n$.

In the characterization of patient valuations it is impossible to replace the time-value of money principle with the condition that $u(\mathbf{e}_t) \geq u(\mathbf{e}_{t+1})$ for all t : there are normalized, monotonic, and linear functions $u : \ell_\infty \rightarrow \mathbb{R}$ that satisfy $u(\mathbf{e}_t) \geq u(\mathbf{e}_{t+1})$ for all t , but do not satisfy the time-value of money principle. See Lemma 3.

A patient valuation can be viewed informally as a limit of the k -th Cesàro average valuation as k goes to infinity and of the r -discounted valuations as $0 < r < 1$ goes to zero. This informal view will be made formal at a later stage.

2.8. Characterization of valuations. There are other possible informal limits of impatient valuations. For example a weighted average $\beta v + (1 - \beta)w$, $0 \leq \beta < 1$, of an impatient valuation w and a patient one v is the informal limit, as k goes to infinity, of the impatient valuations $\beta u_k + (1 - \beta)w$.

The next result characterizes all valuations by showing that the weighted averages, of an impatient valuation and a patient one, are all the valuations.

Theorem 3. *A real-valued function u that is defined on ℓ_∞ is a valuation iff it is a convex combination of an impatient valuation and a patient one.*

2.9. The topology on the set of valuations. In order to define nearby valuations, as well as the proximity of one valuation to another one, we need to define a topology on the set V of valuations.

The coarser the topology the larger are the neighborhoods of a point.

The topology that we define is the minimal topology \mathcal{T} such that for every $g \in c$, the function $v \mapsto v(g)$ on the set V of valuations is continuous. This topology is the minimal topology such that the denumerably many functions $v \mapsto v(\mathbf{e}_t)$, $t \geq 1$, are continuous. Therefore, V is a pseudometric (semi-metric) space.

Namely, there is a function $d : V \times V \rightarrow \mathbb{R}_+$, e.g., $d(u, v) = \max_{t \geq 1} |v(\mathbf{e}_t) - u(\mathbf{e}_t)|$, such that (i) $d(u, v) + d(v, w) \geq d(u, w) \forall u, v, w \in V$, (ii) for every neighborhood U of a valuation u there is $\varepsilon > 0$ such that any valuation v with $d(v, u) < \varepsilon$ is in U , and (iii) for every valuation v and a positive $\varepsilon > 0$, $\{u : d(u, v) < \varepsilon\} \in \mathcal{T}$.

By defining the equivalence relation \sim on V by $u \sim v$ if and only if $v(\mathbf{e}_t) = u(\mathbf{e}_t) \forall t$, the space of equivalence classes V/\sim is a metrizable space.

Remark 5. *The topological space (V, \mathcal{T}) is compact.*

The impatient valuations are dense in V .

A sequence v^k of valuations converges iff the sequence $v^k(\mathbf{e}_t)$ converges $\forall t$.

For any two distinct impatient valuations $v, u \in V$, there is a converging sequence $g \in c$ s.t. $v(g) \neq u(g)$.

For any two patient valuations $v, u \in V$, and for any converging sequence $g \in c$, $v(g) = u(g)$. Therefore, any neighborhood of a patient valuation includes all patient valuations.

Note that for any neighborhood W of a patient valuation there is a positive integer k_0 and a positive $0 < r_0 < 1$ such that for all $k \geq k_0$ and $0 < r \leq r_0$ the impatient valuations u_r and u_k are in W .

2.10. Local robust optimization. Let F be a set of bounded streams of payoffs and v a valuation.

Recall that the v -optimal value of F , $v(F)$, is defined by

$$v(F) = \sup_{f \in F} v(f),$$

and that

Definition 7. *An element $f \in F$ is a robust ε -optimizer at v with respect to F , $\varepsilon \geq 0$, if there is $\delta > 0$ such that*

(5) $u(f) \geq v(F) - \varepsilon$ for all valuations u, w that are δ -close to v ;

equivalently, if there is a neighborhood U of v such that f is an ε -optimizer for U with respect to F , i.e.,

$$(6) \quad u(f) \geq w(F) - \varepsilon \quad \forall u, w \in U.$$

A related robustness property of an element $f \in F$ is the existence of a neighborhood U of v such that f is an ε -optimizer for any $u \in U$ with respect to F , i.e.,

$$(7) \quad u(f) \geq u(F) - \varepsilon \quad \forall u \in U.$$

If F is bounded and v is an impatient valuation, then for every $\varepsilon > 0$ there is $\delta > 0$ such that $u(F) > w(F) - \varepsilon$ for any two valuations u and w that are δ -close to v . Therefore, the related robustness property at v is closely related to our robustness property at v whenever v is an impatient valuation.

The next remark shows that if v is an impatient (and hence there are no other valuations that are 0-close to v) and $v(f) \geq v(F) - \varepsilon$, then f is a robust ε' -optimizer at v with respect to F for any $\varepsilon' > \varepsilon$. However, this is not the case if v is not an impatient valuation.

Remark 6. *If inequality (5) holds for any u that is 0-close to v , then, for every $\varepsilon' > \varepsilon$, f is a robust ε' -optimizer at v with respect to F .*

The following example demonstrates that the continuity of the optimal value of F is insufficient for the existence of a robust ε -optimizer at a non-impatient valuation v .

Example 1. *Let F_3 be the set of all streams $f = (f_1, f_2, \dots)$ with $f_t \in \{-1, 1\}$, $\liminf_{t \rightarrow \infty} \bar{f}_t = -1$, and $\limsup_{t \rightarrow \infty} \bar{f}_t = 1$.*

For any valuation v the v -optimal value of F_3 , $v(F_3)$, equals 1; see Section 4.2. Therefore, the function $v \mapsto v(F_3)$ is a constant function and thus continuous. However, if v is a non-impatient valuation, then F_3 does not have a robust ε -optimizer at v with respect to F for some $\varepsilon > 0$.

The next example demonstrates that the existence of a u -optimizer at any valuation u is insufficient for continuity of the optimal value at any non-impatient valuation v .

Example 2. *Let F_1 be a set that consists of a single stream of payoffs g such that $\liminf_{n \rightarrow \infty} \bar{g}_n + 2\varepsilon < \limsup_{n \rightarrow \infty} \bar{g}_n$, where $\varepsilon > 0$.*

The set F_1 consists of a single element. Therefore, it has, for every valuation u , a (unique) u -optimizer. However, it does not have a robust ε -optimizer at any patient valuation v . Moreover, if $v = (1 - \beta)w + \beta u$

where u is a patient valuation, $\beta > 0$, and w is a valuation, then F_1 does not have a robust $\beta\varepsilon$ -optimizer at v .

The next example demonstrates the (obvious) need to consider approximate optimization (rather than exact optimization) in the study of robust optimization.

Example 3. Let F_2 be the set that consists of the stream $f = (0, 2, 0, 2, \dots)$, i.e., a payoff of zero in the odd periods and a payoff of two in the even periods, and of the streams $(0, 2, \dots, 0, 2, 1, 0, 0, \dots)$, i.e., with the same pattern up to some even period, followed by a payoff of one in the following period, and thereafter the periods' payoffs are zero.

For any $1 > \varepsilon > 0$ and a patient valuation v , the stream f is the unique ε -optimizer for v with respect to F . Moreover, it is a robust ε -optimizer at v with respect to F . Note that if u_n is the n -th Cesàro average valuation, then the stream f is not a u_n -optimizer in F if n is odd. Since the n -th Cesàro average valuation converges, as n goes to infinity, to the patient valuation v , there is no robust 0 - v -optimizer in F .

The following proposition provides a “minmax=maxmin”-type condition on a set F that is equivalent to F having a robust ε -optimizer at v for every $\varepsilon > 0$.

Proposition 1. F has a robust ε -optimizer at v for every $\varepsilon > 0$, if and only if,

$$\sup_{f \in F, W \in \mathcal{N}(v)} \inf_{u \in W} u(f) = \inf_{W \in \mathcal{N}(v)} \sup_{h \in F, u \in W} u(h),$$

where $\mathcal{N}(v)$ denotes the set of all neighborhoods of a valuation v .

The next proposition is a simple corollary of the definition of a robust ε -optimizer at a valuation v .

Proposition 2. If the set F of feasible streams of bounded payoffs has a robust ε -optimizer at every fixed valuation v and every $\varepsilon > 0$, then the function $u \mapsto u(F)$ is continuous at v .

The following proposition shows that a bounded set of streams of payoffs F admits robust optimization at every impatient valuation v .

Proposition 3. Let F be a bounded set of streams of payoffs and let v be an impatient valuation. If f is an ε -optimizer for v with respect to F , then, for every $\varepsilon' > \varepsilon \geq 0$, f is a robust ε' -optimizer at v with respect to F . Therefore, F has, for every $\varepsilon > 0$, a robust ε -optimizer at v .

2.11. Global robust optimization. In this section we state the implications of a bounded set of streams of payoffs F having a robust ε -optimizer at every valuation v .

Theorem 4. *Assume that the set F of feasible streams of bounded payoffs has a robust ε -optimizer at every valuation v . Then, there is a finite list f^1, f^2, \dots, f^k in F such that*

- a) *for every valuation v there is an index $1 \leq i \leq k$ such that f^i is a robust ε -optimizer at v with respect to F , and*
- b) *there is a continuous function $v \mapsto f^v$ with values in the convex hull of $\{f^1, \dots, f^k\}$ such that every valuation v has a neighborhood U such that $u(f^v) \geq w(F) - \varepsilon \forall u, w \in U$; hence, if f^v is in F then f^v is a robust ε -optimizer at v with respect to F .*

The next proposition demonstrates that the condition that F has a robust ε -optimizer at every valuation v is essential for the conclusions of Theorem 4 and Proposition 2.

Proposition 4. *For every non-impatient valuation u and a neighborhood U of u , there is a bounded set of streams of payoffs $F \subset c$ such that:*

- (1) *The optimal value of F is not continuous at u . Moreover, there is a sequence of impatient valuations v_n that converges to u such that the sequence $v_n(F)$ does not converge.*
- (2) *The optimal value of F is continuous at any valuation $v \notin U$.*
- (3) *$\exists \eta > 0$ such that for every finite subset $G \subset F$ there is an impatient valuation w such that $w(F) - \eta > 1 + \eta > \max_{g \in G} w(g)$.*

2.12. Markov decision process (MDP). In a discrete-time finite Markov decision process (MDP), play proceeds in stages. At each stage, the process is in one of finitely many states, and the decision maker chooses an action from a finite set of possible actions. The action and the state jointly determine a payoff to the decision maker and transition probabilities to the succeeding state.

Before making the choice, the decision maker observes the current state.

A finite MDP is defined by the list $\Gamma = (S, A, r, p)$ where S is the finite set of states, A is the finite set of actions, $r : S \times A \rightarrow \mathbb{R}$ is the payoff function, and $p : S \times A \rightarrow \Delta(S)$ is the transition function. If action $a \in A$ is taken at stage t and the state in stage t is $s \in S$,

then the payoff at stage t is $r(s, a)$ and the (conditional) probability distribution of the state at stage $t + 1$ is $p(s, a)$.

A pure (respectively, behavioral) policy π of the decision maker specifies the action (respectively, the probability distribution over actions) at stage t as a function of the current state and past states and actions. Namely, $\pi : \cup_{t \geq 1} (S^t \times A^{t-1}) \rightarrow A$ (respectively, $\rightarrow \Delta(A)$).

Given an initial state $s_1 = s$, a policy π defines a probability distribution P_π^s over the sequences s_1, a_1, \dots of states and actions. The expectation w.r.t. P_π^s is denoted by E_π^s . For simplicity, we use the same symbol P_π^s to denote also the distribution over the streams of payoffs $g_t = r(s_t, a_t)$.

The set F^s of feasible distributions over streams of payoffs, as a function of the initial state s , is defined by $F^s = \{P_\pi^s : \pi \text{ a behavioral policy}\}$. It equals the convex hull of the sets $\{P_\pi^s : \pi \text{ a pure policy}\}$.

The expectation with respect to the probability distribution P_π^s is denoted by E_π^s . The set \hat{F}^s of feasible streams of payoffs, as a function of the initial state s , is the set of streams of payoffs $\hat{P}_\pi^s = g^{s, \pi}$, where $g_t^{s, \pi} = E_\pi^s r(s_t, a_t)$ and π ranges over all policies in the finite MDP. It equals the convex hull of the sets $\{\hat{P}_\pi^s : \pi \text{ a pure policy}\}$.

Theorem 5. *Let $\Gamma = (S, A, r, p)$ be a finite MDP. For every probability distribution $q \in \Delta(S)$, the set $\sum_{s \in S} q(s) \hat{F}^s$ has, for every $\varepsilon > 0$ and every valuation v , a robust ε -optimizer at v with respect to $\sum_{s \in S} q(s) \hat{F}^s$.*

In fact, we prove a stronger result. In order to state this stronger result we introduce the following notation. For a valuation u and a stream of payoffs g we denote by $\underline{u}(g)$, respectively, by $\bar{u}(g)$, the infimum, respectively, the supremum, of $u'(g)$ over all valuations u' that are 0-close to u .

Note that as $u(g)$ need not be measurable in g and therefore the expectation of $u(g)$ with respect to the probability P_π^s (where π is a policy) need not exist. However, $\underline{u}(g)$ and $\bar{u}(g)$ are measurable in g and therefore the expectation of $\underline{u}(g)$ and $\bar{u}(g)$ with respect to the probability P_π^s exists.

As $u(\hat{P}_\pi^s) \geq E_\pi^s \underline{u}(g)$ and $u(\hat{P}_\pi^s) \leq E_\pi^s \bar{u}(g)$, the next theorem implies Theorem 5.

Theorem 6. *For every finite MDP, a valuation v , and $\varepsilon > 0$, there is a policy π and $\delta > 0$, such that for all valuations u and w that are δ -close to v and any policy σ ,*

$$E_\pi^s \underline{u}(g) \geq E_\sigma^s \bar{w}(g) - \varepsilon.$$

The proof of Theorem 6 proves also the following stronger property of a finite MDP. The normed space of all sequence $\omega = (\omega_1, \omega_2, \dots)$ with $\|\omega\|_1 := \sum_{t=1}^{\infty} |\omega_t| < \infty$ is denoted by ℓ_1 .

Theorem 7. *For every finite MDP, $\omega \in \ell_1$, a patient valuation v , and $\varepsilon > 0$, there is a policy π and $\delta > 0$, such that for all impatient valuations u and w that are δ -close to v , any policy σ , and any $\omega' \in \ell_1$ with $\|\omega - \omega'\|_1 < \delta$,*

$$E_{\pi}^s\left(\sum_{t=1}^{\infty} \omega_t g_t + \underline{u}(g)\right) \geq E_{\sigma}^s\left(\sum_{t=1}^{\infty} \omega'_t g_t + \bar{w}(g)\right) - \varepsilon.$$

3. PROOFS OF THE THEOREMS

Note that an additive function $u : \ell_{\infty} \rightarrow \mathbb{R}$ that is monotonic is (by classical arguments) linear. Indeed, by additivity of u , $u(-g) = -u(g)$ and $u(\alpha g) = \alpha u(g)$ for every rational α . By additivity and monotonicity of u , for every $g, h \in \ell_{\infty}$, $|u(g) - u(h)| \leq \|g - h\|u(\mathbf{1})$ and therefore $u(\alpha g)$, $\alpha \in \mathbb{R}$, is continuous in α ; thus $u(\alpha g) = \alpha u(g) \forall \alpha \in \mathbb{R}$.

3.1. Proof of Theorem 1. Assume that u is an impatient valuation.

Define $\omega_t = u(\mathbf{e}_t)$.

By the additivity of u , $u(\mathbf{0}) = 0$. The time-value of money property of a valuation along the definition of ω_t implies that $u(\mathbf{0}) = 0 \leq \omega_t = u(\mathbf{e}_t) \geq u(\mathbf{e}_{t+1}) = \omega_{t+1}$.

Note that $-\|g\|\mathbf{1}_{>n} \leq g - \sum_{t=1}^n g_t \mathbf{e}_t \leq \|g\|\mathbf{1}_{>n}$ and, therefore, using the linearity of u , the definition of ω_t , monotonicity (which follows from the time-value of money property), and the impatience of u , we have

$$\left|u(g) - \sum_{t=1}^n \omega_t g_t\right| = \left|u(g) - u\left(\sum_{t=1}^n g_t \mathbf{e}_t\right)\right| \leq u(\|g\|\mathbf{1}_{>n}) \rightarrow_{n \rightarrow \infty} 0.$$

Therefore, $u(g) = \sum_{t=1}^{\infty} \omega_t g_t$. In particular, using the normalization assumption $u(\mathbf{1}) = 1$, we have $u(\mathbf{1}) = \sum_{t=1}^{\infty} \omega_t = 1$. This completes the proof of the “only if” part of the theorem.

Assume that $u(g) = \sum_{t=1}^{\infty} \omega_t g_t$ with $\omega_t - \omega_{t+1} \geq 0$ and $\sum_{t=1}^{\infty} \omega_t = 1$. Then u is a normalized linear real-valued function on the space ℓ_{∞} with $u(\mathbf{1}_{>n}) = \sum_{t>n} \omega_t \rightarrow_{n \rightarrow \infty} 0$. Since $u(g) = \sum_{t=1}^{\infty} \omega_t g_t = \sum_{t=1}^{\infty} (\omega_t - \omega_{t+1}) t \bar{g}_t$, it follows that if $\bar{g}_t \geq \bar{h}_t \forall t$ then $u(g) \geq u(h)$. This completes the proof of the “if” part of the theorem. \square

3.2. Proof of Theorem 2.

Let u be a patient valuation. Note that if u is a patient valuation then $u(\mathbf{e}_t) = 0$. Indeed, by additivity $u(\mathbf{0}) = 0$, and by monotonicity of a valuation $u(\mathbf{e}_t) \geq 0$. As u is a patient valuation, $1 = u(\mathbf{1}) = u(\sum_{t=1}^n \mathbf{e}_t + \mathbf{1}_{>n}) = \sum_{t=1}^n u(\mathbf{e}_t) + 1$. Therefore, $u(\mathbf{e}_t) = 0 \forall t$.

For $g \in \ell_\infty$ we set $\bar{g} := \limsup_{k \rightarrow \infty} \bar{g}_k$ and $\underline{g} := \liminf_{k \rightarrow \infty} \bar{g}_k$.

Let u be a patient valuation and $g \in \ell_\infty$. Fix $\varepsilon > 0$ and let n be sufficiently large so that $\underline{g} - \varepsilon < \bar{g}_k < \bar{g} + \varepsilon \forall k \geq n$.

Let h be defined by $h = \sum_{t=1}^n (\|g\| + \varepsilon) \mathbf{e}_t + (\bar{g} + \varepsilon) \mathbf{1}_{>n}$. Note that for every positive integer s , $\bar{h}_s \geq \bar{g}_s$ and therefore by the time-value of money property, $u(h) \geq u(g)$.

By the linearity and patience of u , $u(h) = (\bar{g} + \varepsilon)u(\mathbf{1}) = \bar{g} + \varepsilon$. Therefore, $u(g) \leq \bar{g} + \varepsilon$. As this last inequality holds for every $\varepsilon > 0$ we deduce that the right-hand side inequality of (4) holds for every patient valuation u and every $g \in \ell_\infty$.

Note that the left-hand side inequality of (4) holds for $g \in \ell_\infty$ if (and only if) the right-hand side inequality of (4) holds for $-g$. Indeed, $-u(g) = u(-g) \leq \limsup_{n \rightarrow \infty} -\bar{g}_n = -\liminf_{n \rightarrow \infty} \bar{g}_n$. Therefore, $\underline{g} \leq u(g)$ for every $g \in \ell_\infty$.

Assume that u is a linear function that is defined on ℓ_∞ and satisfies (4). Obviously, $u(\mathbf{1}) = 1$; thus u is normalized. It remains to show that u satisfies the time-value of money property. Assume that $g, h \in \ell_\infty$ with $\sum_{t=1}^n g_t \geq \sum_{t=1}^n h_t \forall n$. Then, $\underline{g} - \underline{h} \geq 0$ and therefore $u(g-h) \geq 0$ by the left-hand side inequality of (4), and therefore, as u is linear, $u(g) = u(g-h) + u(h) \geq u(h)$. \square

3.3. Proof of Theorem 3.

Obviously, a convex combination of valuations is a valuation. This proves the straightforward “if” part of the theorem. We proceed in proving the “only if” part.

Let u be a valuation, and let $\omega_t := u(\mathbf{e}_t)$. As u is a valuation, $\omega_t \geq \omega_{t+1} \geq 0 \forall t$.

As u is additive, $u(\mathbf{0}) = 0$. As u is additive, normalized, and monotonic, $u(\mathbf{1}_{>n})$ is nonincreasing in n and $0 \leq u(\mathbf{1}_{>n}) \leq 1$.

Let β be the limit of the nonincreasing sequence $u(\mathbf{1}_{>n}) = u(\mathbf{1}) - \sum_{t=1}^n \omega_t$. As $0 \leq u(\mathbf{1}_{>n}) = 1 - \sum_{t=1}^n \omega_t \leq u(\mathbf{1}) = 1$, $0 \leq \beta = 1 - \sum_{t=1}^\infty \omega_t \leq 1$.

If $\beta = 0$ then u is an impatient valuation.

If $\beta = 1$ then u is a patient valuation.

Assume that $0 < \beta < 1$. Define $w : \ell_\infty \rightarrow \mathbb{R}$ by $w(g) = \sum_{t=1}^{\infty} \frac{\omega_t}{1-\beta} g_t$, and define the function $v : \ell_\infty \rightarrow \mathbb{R}$ by $v(g) := \frac{1}{\beta}(u(g) - \sum_{t=1}^{\infty} \omega_t g_t)$.

As $\omega_t \geq \omega_{t+1} \geq 0$ and $\sum_{t=1}^{\infty} \omega_t = 1 - \beta$, w is an impatient valuation.

Obviously, $u = (1 - \beta)w + \beta v$. Therefore, it remains to prove that v is an impatient valuation.

As $u(\mathbf{1}) - \sum_{t=1}^{\infty} \omega_t = \beta$, $v(\mathbf{1}) = 1$. Therefore, the function v is normalized.

By the linearity of the function $g \mapsto u(g) - \sum_{t=1}^{\infty} \omega_t g_t$, the function v is linear.

As $v(\mathbf{1}_{>n}) = \frac{1}{\beta}(u(\mathbf{1}_{>n}) - \sum_{t>n} \omega_t) \rightarrow_{n \rightarrow \infty} 1$, the function v is patient.

Therefore, the function $v : \ell_\infty \rightarrow \mathbb{R}$ is normalized, linear, and patient.

In order to prove that v is a patient valuation, it remains to prove that v satisfies the time-value of money principle.

By the linearity of v it suffices to prove that if $g \in \ell_\infty$ with $\sum_1^s g_t \geq 0 \forall s$, then $v(g) \geq 0$.

For $g \in \ell_\infty$ and an integer n we denote by $g_{>n}$ the element of ℓ_∞ whose t -th coordinate equals g_t if $t > n$ and equals 0 if $t \leq n$.

Assume that $\sum_1^s g_t \geq 0 \forall s$. Fix $\varepsilon > 0$. As $\omega_t \geq 0$ and $\sum_{t=1}^{\infty} \omega_t < \infty$, there is a positive integer k such that $(k\omega_k + \sum_{t=k+1}^{\infty} \omega_t)\|g\| < \varepsilon$.

As v is patient, $v(g) = v(g_{>k}) = v(\sum_{t=1}^k g_t \mathbf{e}_k + g_{>k})$. Using the definition of v along the time-value of money property of u , $\beta v(\sum_{t=1}^k g_t \mathbf{e}_k + g_{>k}) = u(\sum_{t=1}^k g_t \mathbf{e}_k + g_{>k}) - (\sum_{t=1}^k g_t \omega_k + \sum_{t=k+1}^{\infty} g_t \omega_t) \geq 0 - \varepsilon \geq -\varepsilon$.

As the inequality $\beta v(g) \geq -\varepsilon$ holds for every $\varepsilon > 0$, and $\beta > 0$, we conclude that $v(g) \geq 0$. \square

3.4. Proof of Theorems 4. Let $F \subset \ell_\infty$ be a set of feasible streams of bounded payoffs and $\varepsilon > 0$.

Assume that for every valuation v there is a stream g^v in F that is a robust ε -optimizer at v with respect to F . Let $W_v \in \mathcal{N}(v)$ be a neighborhood of v such that

$$(8) \quad u(g^v) \geq w(F) - \varepsilon \quad \forall u, w \in W_v.$$

As the topological space V of all valuations is compact and the set of neighborhoods W_v covers V (i.e., $\cup_{v \in V} W_v = V$) there is a finite subcover. Namely, there are finitely many distinct valuations v_1, \dots, v_k such that $\cup_{i=1}^k W_{v_i} = V$. Set $f^i = g^{v_i}$ and let v be a valuation. As $\cup_{i=1}^k W_{v_i} = V$, there is an index $1 \leq i \leq k$ such that $v \in W_{v_i}$.

By setting $v = v_i$ and $g^v = f^i$ in inequality (8), we deduce that f^i is a robust ε -optimizer at v with respect to F .

This completes the proof of the first part of the theorem.

Let $\alpha_i : V \rightarrow \mathbb{R}_+$ be a continuous function such that $\alpha_i(v) = 0$ iff $v \notin W_{v_i}$. The existence of such a function α_i follows from the fact that V is a semi-metrizable space. (E.g., $\alpha_i(v)$ can be the distance of v from the complement $W_{v_i}^c$.) Note that $\sum_{i=1}^k \alpha_i(v) > 0 \forall v \in V$.

Next, we define the stream f^v in F by

$$f^v = \frac{\sum_{i=1}^k \alpha_i(v) f^i}{\sum_{i=1}^k \alpha_i(v)}.$$

As the functions α_i are continuous and $\sum_{i=1}^k \alpha_i(v) > 0$, the function $v \mapsto f^v$ is continuous.

Let U be the neighborhood of v consisting of all valuation u such that for all $1 \leq i \leq k$, $\alpha_i(u) > 0$ iff $\alpha_i(v) > 0$. I.e., $U = \bigcap_{i:\alpha_i(v)>0} W_{v_i} = \bigcap_{i:u \in W_{v_i}} W_{v_i}$.

Let u and w be two valuations in U . For any $1 \leq i \leq k$ such that $\alpha_i(v) > 0$, $u(f^i) \geq w(F) - \varepsilon \forall u, w \in W_{v_i}$; hence, $u(f^i) \geq w(F) - \varepsilon \forall u, w \in U \subset W_{v_i}$. As u is a linear function of the stream of payoffs, we deduce that $u(f^v) \geq w(F) - \varepsilon \forall u, w \in U$.

This completes the proof of Theorem 4.

3.5. Proof of Theorem 6. Let $\Gamma = (S, A, r, p)$ be a discrete-time MDP and let $v(s)$, $s \in S$, be the undiscounted value of the MDP with initial state s .

Set $g_t = r(s_t, a_t)$ and $\bar{g}_n = \frac{1}{n} \sum_{t=1}^n g_t$.

Let π be a stationary uniformly optimal policy⁵ of the decision maker in Γ . Thus,⁶ for every state $s \in S$ and every policy η ,

$$(9) \quad E_\pi^s \liminf_{n \rightarrow \infty} \bar{g}_n \geq v(s) \geq E_\eta^s \limsup_{n \rightarrow \infty} \bar{g}_n,$$

and for every $\varepsilon > 0$ there is n_ε such that for every state $s \in S$, every $n \geq n_\varepsilon$, and every policy η ,

$$(10) \quad \varepsilon + E_\pi^s \bar{g}_n \geq v(s) \geq E_\eta^s \bar{g}_n - \varepsilon.$$

Fix a valuation u and let $\omega_t = u(e_t)$, $t \geq 1$, be the weights of the valuation u .

⁵A uniformly optimal policy is a policy π that is optimal in all the discounted MDP with a sufficiently small discount rate. The existence of a stationary uniformly optimal policy in a finite MDP is due to [1].

⁶Properties (9) and (10) are easily derived from the fact that π is a stationary uniformly optimal policy. Alternatively, by the construction of an ε -optimal policy in [7] it follows that the policy π is, for every $\varepsilon > 0$, an ε -optimal policy in the undiscounted MDP. Alternatively, see [8, part 4] of Proposition 3].

In order to prove the theorem, it suffices to define, for every $\varepsilon > 0$, a neighborhood U of u and a policy τ , such that for every policy η and every $u^* \in U$,

$$(11) \quad 7\varepsilon + \underline{u}^*(P_\tau^s) \geq v(s) \geq \bar{u}^*(P_\eta^s) - 7\varepsilon.$$

Recall that $\sum_{t=1}^{\infty} \omega_t \leq 1$. Set $\omega_\infty = 1 - \sum_{t=1}^{\infty} \omega_t$, and let t_ε be a sufficiently large positive integer such that $(1 + \|r\|) \sum_{t=t_\varepsilon}^{\infty} \omega_t < \varepsilon$, where $\|r\| = \max_{s,a} |r(s,a)|$.

Fix $\varepsilon > 0$.

Let Γ_* be the multi-stage decision problem (N, Σ, r_*) , where the set of policies Σ coincides with those of the MDP and the payoff function r_* , as a function of the initial state s and the policy σ , is defined by

$$r_*(s, \sigma) = E_\sigma^s \sum_{1 \leq t < t_\varepsilon} \omega_t g_t + (1 - \sum_{1 \leq t < t_\varepsilon} \omega_t) E_\sigma^s v(s_{t_\varepsilon}).$$

The payoff r_* depends only on finitely many coordinates of the play of Γ . Therefore, Γ_* is equivalent to a decision problem with finitely many pure policies; thus Γ_* has an optimal pure policy.

Let σ be an optimal policy of Γ_* with payoff vector v_* . Namely,

$$(12) \quad r_*(s, \sigma) = E_\sigma^s \sum_{1 \leq t < t_\varepsilon} \omega_t g_t + (1 - \sum_{1 \leq t < t_\varepsilon} \omega_t) E_\sigma^s v(z_{t_\varepsilon}) = v_*(s),$$

and for every policy η ,

$$(13) \quad r_*(s, \eta) = E_\eta^s \sum_{1 \leq t < t_\varepsilon} \omega_t g_t + (1 - \sum_{1 \leq t < t_\varepsilon} \omega_t) E_\eta^s v(s_{t_\varepsilon}) \leq v_*(s).$$

Define the policy profile τ as follows. At stage $t < t_\varepsilon$, $\tau_t(s_1, a_1, \dots, s_t) = \sigma(s_1, a_1, \dots, s_t)$ and at stage $t \geq t_\varepsilon$, $\tau_t(s_1, a_1, \dots, s_{t_\varepsilon}, \dots, s_t) = \pi(s_{t_\varepsilon}, \dots, s_t)$.

The definition of the policy τ along inequality (9) implies that

$$(14) \quad E_\tau^s \liminf_{n \rightarrow \infty} \bar{g}_n \geq E_\tau^s v(s_{t_\varepsilon}) = E_\sigma^s v(s_{t_\varepsilon}).$$

Let U be the set of all valuations u^* whose valuation weights $\omega_t^* := u^*(e_t)$ are such that

$$(15) \quad \|r\| \sum_{t=1}^{t_\varepsilon + n_\varepsilon} |\omega_t^* - \omega_t| < \varepsilon.$$

Note that U is a neighborhood of u .

Fix a valuation $u^* \in \mathcal{U}$. By the choice of t_ε , inequality (15) implies that

$$(16) \quad \sum_{t=t_\varepsilon}^{t_\varepsilon+n_\varepsilon} \omega_t^* \|r\| < 2\varepsilon.$$

By equality (12), the definition of τ , the inequality $\omega_t^* g_t \geq \omega_t g_t - \|r\| |\omega_t^* - \omega_t|$, and inequality (15), we have

$$(17) \quad \begin{aligned} E_\tau^s \sum_{1 \leq t < t_\varepsilon} \omega_t^* g_t &= E_\tau^s \sum_{1 \leq t < t_\varepsilon} \omega_t^* g_t = E_\sigma^s \sum_{1 \leq t < t_\varepsilon} \omega_t^* g_t \\ &\geq E_\sigma^s \sum_{1 \leq t < t_\varepsilon} \omega_t g_t - \|r\| \sum_{1 \leq t < t_\varepsilon} |\omega_t - \omega_t^*| \\ &\geq v_*(s) - (1 - \sum_{1 \leq t < t_\varepsilon} \omega_t) E_\sigma^s v(s_{t_\varepsilon}) - \varepsilon. \end{aligned}$$

Let $t \geq t_\varepsilon + n_\varepsilon$. Then, using inequality (10) and the definition of τ ,

$$(18) \quad E_\tau(g_{t_\varepsilon} + \dots + g_t \mid \mathcal{H}_{t_\varepsilon}) \geq (t - t_\varepsilon + 1)(v(s_{t_\varepsilon}) - \varepsilon),$$

where \mathcal{H}_t is the algebra (of subsets of plays) that is generated by s_1, a_1, \dots, s_t .

By summation by parts, and using the inequality $\omega_t^* \geq \omega_{t+1}^* \forall t \geq t_\varepsilon$,

$$(19) \quad \sum_{t=t_\varepsilon}^{\infty} \omega_t^* g_t = \sum_{t=t_\varepsilon}^{\infty} (\omega_t^* - \omega_{t+1}^*) \sum_{s=t_\varepsilon}^t g_s$$

and

$$(20) \quad \sum_{t=t_\varepsilon}^{\infty} \omega_t^* = \sum_{t=t_\varepsilon}^{\infty} (\omega_t^* - \omega_{t+1}^*) (t - t_\varepsilon + 1).$$

Therefore, using (19), the triangle inequality, (16), (10), and (20), we have

$$\begin{aligned}
& E_\tau \left(\sum_{t=t_\varepsilon}^{\infty} \omega_t^* g_t \mid \mathcal{H}_{t_\varepsilon} \right) = E_\tau \left(\sum_{t=t_\varepsilon}^{\infty} (\omega_t^* - \omega_{t+1}^*) \sum_{s=t_\varepsilon}^t g_s \mid \mathcal{H}_{t_\varepsilon} \right) \\
& \geq E_\tau \left(\sum_{t=t_\varepsilon+n_\varepsilon}^{\infty} (\omega_t^* - \omega_{t+1}^*) \sum_{s=t_\varepsilon}^t g_s \mid \mathcal{H}_{t_\varepsilon} \right) - \sum_{t=t_\varepsilon}^{t_\varepsilon+n_\varepsilon-1} \omega_t^* \|r\| \\
& \geq \sum_{t=t_\varepsilon+n_\varepsilon}^{\infty} (\omega_t^* - \omega_{t+1}^*) (t - t_\varepsilon + 1) (v(s_{t_\varepsilon}) - \varepsilon) - 2\varepsilon \\
& \geq \sum_{t=t_\varepsilon}^{\infty} (\omega_t^* - \omega_{t+1}^*) (t - t_\varepsilon + 1) (v(s_{t_\varepsilon}) - \varepsilon) - 4\varepsilon \\
& = \sum_{t=t_\varepsilon}^{\infty} \omega_t^* (v(s_{t_\varepsilon}) - \varepsilon) - 4\varepsilon \geq \sum_{t=t_\varepsilon}^{\infty} \omega_t^* v(s_{t_\varepsilon}) - 5\varepsilon.
\end{aligned}$$

By taking the expectation, we deduce that

$$(21) \quad E_\tau^s \sum_{t=t_\varepsilon}^{\infty} \omega_t^* g_t \geq \sum_{t=t_\varepsilon}^{\infty} \omega_t^* E_\tau^s v(s_{t_\varepsilon}) - 5\varepsilon.$$

Multiplying inequality (9) by $\omega_\infty^* := 1 - \sum_{t=1}^{\infty} \omega_t^*$ and adding inequality (21), we have

$$\begin{aligned}
(22) \quad & E_\tau^s \omega_\infty^* \liminf_{n \rightarrow \infty} \bar{g}_n + E_\tau^s \sum_{t=t_\varepsilon}^{\infty} \omega_t^* g_t \geq (\omega_\infty^* + \sum_{t=t_\varepsilon}^{\infty} \omega_t^*) E_\tau^s v(s_{t_\varepsilon}) - 5\varepsilon \\
& = (1 - \sum_{1 \leq t < t_\varepsilon} \omega_t^*) E_\tau^s v(s_{t_\varepsilon}) - 5\varepsilon \geq (1 - \sum_{1 \leq t < t_\varepsilon} \omega_t) E_\tau^s v(s_{t_\varepsilon}) - 6\varepsilon.
\end{aligned}$$

By summing inequalities (17) and (22), we have

$$E_\tau^s \omega_\infty^* \liminf_{n \rightarrow \infty} \bar{g}_n + E_\tau^s \sum_{t=1}^{\infty} \omega_t^* g_t \geq v_*(s) - 7\varepsilon.$$

For any stream of bounded payoffs g , $u^*(g) \geq \omega_\infty^* \liminf_{n \rightarrow \infty} \bar{g}_n + \sum_{t=1}^{\infty} \omega_t^* g_t$ by the characterization of valuations (Theorems 1, 2, and 3), and the map $g \mapsto \omega_\infty^* \liminf_{n \rightarrow \infty} \bar{g}_n + \sum_{t=1}^{\infty} \omega_t^* g_t$ is measurable. Therefore,

$$\underline{u}^*(P_\tau^s) \geq E_\tau^s \omega_\infty^* \liminf_{n \rightarrow \infty} \bar{g}_n + E_\tau^s \sum_{t=1}^{\infty} \omega_t^* g_t \geq v_*(s) - 7\varepsilon,$$

which proves the left-hand side inequality of (11).

Fix a policy η of the decision maker. By replacing, in the above equations and inequalities, E_τ^s by E_η^s , $=$ and \geq by \leq , ε by $-\varepsilon$, and \liminf by \limsup , we have

$$\bar{u}^*(P_\eta^s) \leq E_\eta^s \omega_\infty^* \limsup_{n \rightarrow \infty} \bar{g}_n + \sum_{t=1}^{\infty} \omega_t^* g_t^{s,\eta} \leq v_*(s) + 7\varepsilon,$$

which proves the right-hand side inequality of (11).

Explicitly, using inequalities (15), (13), and the inequality $\omega_t^* g_t \leq \omega_t g_t + \|g\| |\omega_t^* - \omega_t|$, we have

$$(23) \quad E_\eta^s \sum_{1 \leq t < t_\varepsilon} \omega_t^* g_t = E_\eta^s \sum_{1 \leq t < t_\varepsilon} \omega_t^* g_t \leq v_*(s) - \left(1 - \sum_{t=1}^{\infty} \omega_t\right) E_\eta^s v(s_{t_\varepsilon}) + \varepsilon.$$

By using (19), the triangle inequality, the right-hand side inequality of (9), and (20), we have

$$\begin{aligned} & E_\eta \left(\sum_{t=t_\varepsilon}^{\infty} \omega_t^* g_t \mid \mathcal{H}_{t_\varepsilon} \right) \\ &= E_\eta \left(\sum_{t=t_\varepsilon}^{\infty} (\omega_t^* - \omega_{t+1}^*) \sum_{s=t_\varepsilon}^t g_s \mid \mathcal{H}_{t_\varepsilon} \right) \\ &\leq E_\eta \left(\sum_{t=t_\varepsilon+n_\varepsilon}^{\infty} (\omega_t^* - \omega_{t+1}^*) \sum_{s=t_\varepsilon}^t g_s \mid \mathcal{H}_{t_\varepsilon} \right) + \sum_{t=t_\varepsilon}^{t_\varepsilon+n_\varepsilon-1} \|r\| \omega_t^* \\ &\leq \sum_{t=t_\varepsilon+n_\varepsilon}^{\infty} (\omega_t^* - \omega_{t+1}^*) (t - t_\varepsilon + 1) (v(s_{t_\varepsilon}) + \varepsilon) + 2\varepsilon \\ &\leq \sum_{t=t_\varepsilon}^{\infty} (\omega_t^* - \omega_{t+1}^*) (t - t_\varepsilon + 1) (v(s_{t_\varepsilon}) + \varepsilon) + 4\varepsilon \\ &= \sum_{t=t_\varepsilon}^{\infty} \omega_t^* (v(s_{t_\varepsilon}) + \varepsilon) + \varepsilon \leq \sum_{t=t_\varepsilon}^{\infty} \omega_t^* v(s_{t_\varepsilon}) + 5\varepsilon. \end{aligned}$$

By taking the expectation, we deduce that

$$(24) \quad E_\eta^s \sum_{t=t_\varepsilon}^{\infty} \omega_t^* g_t \leq \sum_{t=t_\varepsilon}^{\infty} \omega_t^* E_\eta^s v(s_{t_\varepsilon}) + 5\varepsilon.$$

The uniform optimality of π implies that for every policy η ,

$$(25) \quad E_\eta^s \limsup_{n \rightarrow \infty} \bar{g}_n \leq E_\eta^s v(s_{t_\varepsilon}).$$

Multiplying inequality (25) by $\omega_\infty^* = 1 - \sum_{t=1}^\infty \omega_t^*$ and adding inequality (24), we have

$$(26) E_\eta^s \omega_\infty^* \liminf_{n \rightarrow \infty} \bar{g}_n + E_\eta^s \sum_{t=t_\varepsilon}^\infty \omega_t^* g_t \leq (\omega_\infty^* + \sum_{t=t_\varepsilon}^\infty \omega_t^*) E_\eta^s v(s_{t_\varepsilon}) + 5\varepsilon$$

$$= (1 - \sum_{1 \leq t < t_\varepsilon} \omega_t^*) E_\eta^s v(s_{t_\varepsilon}) + 5\varepsilon \leq (1 - \sum_{1 \leq t < t_\varepsilon} \omega_t) E_\eta^s v(s_{t_\varepsilon}) + 6\varepsilon.$$

Inequalities (23) and (26) imply that

$$\bar{u}^*(P_\eta^s) \leq (1 - \sum_{t=1}^\infty \omega_t^*) \limsup_{n \rightarrow \infty} \bar{g}_n^{s,\eta} + \sum_{t=1}^\infty \omega_t^* g_t^{s,\eta} \leq v_*(s) + 7\varepsilon,$$

which proves the right-hand side inequality of (11). \square

Any valuation u is a mixture of a patient valuation v and an impatient valuation w . If w is impatient then for any policy π we have $w(g^{s,\pi}) = w(P_\pi^s)$. If v is a patient valuation then for any policy π we have $\underline{v}(P_\pi^s) \leq v(g^{s,\pi}) \leq \bar{v}(P_\pi^s)$. Therefore, for any valuation u we have $\underline{u}(P_\pi^s) \leq u(g^{s,\pi}) \leq \bar{u}(P_\pi^s)$.

Therefore, Theorem 6 implies Theorem 5, i.e., that the set $\{g^{s,\pi} : \pi \text{ a policy}\}$ has, for every $\varepsilon > 0$ and valuation v , a robust ε -optimizer at v .

Note that the inequalities $\omega_t \geq \omega_{t+1} \geq 0$, $1 \leq t < t_\varepsilon$, were not used in the proof. Therefore, the proof demonstrates that for every finite MDP and a finite sequence of real numbers $\omega_1, \dots, \omega_N$, there is a policy π and neighborhoods U_ε , $\varepsilon > 0$, of the patient valuations such that for any policy η ,

$$E_\pi^s \sum_{t=1}^N \omega_t g_t + E_\pi^s \liminf_{t \rightarrow \infty} \bar{g}_t \geq E_\eta^s \sum_{t=1}^N \omega_t g_t + E_\eta^s \limsup_{t \rightarrow \infty} \bar{g}_t,$$

and for every $u \in U_\varepsilon$,

$$E_\pi^s \sum_{t=1}^N \omega_t g_t + E_\pi^s \underline{u}(g^{s,\pi}) \geq E_\eta^s \sum_{t=1}^N \omega_t g_t + E_\eta^s \bar{u}(g^{s,\pi}) - \varepsilon.$$

4. PROOFS OF THE PROPOSITIONS

4.1. Properties of the set F_1 in Example 2. Let u be a non-impatient valuation. Then, $u = (1 - \beta)w + \beta v$ where w is an impatient valuation, v is a patient one, and $\beta > 0$.

The impatient valuations $(1 - \beta)w + \beta u_n$, where u_n is the n -th Cesàro average valuation, converge, as $n \rightarrow \infty$, to the valuation u .

Recall that $F_1 = \{f\}$ and $\liminf_{n \rightarrow \infty} \bar{f}_n + 2\varepsilon = \liminf u_n(f) + 2\varepsilon < \limsup_{n \rightarrow \infty} \bar{f}_n = \limsup u_n(f)$.

Then, $\liminf_{n \rightarrow \infty} ((1-\beta)w + \beta u_n)(f) = (1-\beta)w(f) + \beta \liminf_{n \rightarrow \infty} \bar{f}_n < (1-\beta)w(f) + \beta \limsup_{n \rightarrow \infty} \bar{f}_n - 2\beta\varepsilon = \limsup_{n \rightarrow \infty} ((1-\beta)w + \beta u_n)(f) - 2\beta\varepsilon$. Therefore, f is not a robust $\beta\varepsilon$ - v -optimizer in F_1 . \square

4.2. Properties of the set F_3 in Example 1. Let v be a patient valuation. We will prove⁷ that $v(F_3) = 1$.

Let $n_k > 0$, $k \geq 0$, be an increasing sequence of positive integers such that $\lim n_k/n_{k+1} = 0$. Let j be a positive integer and let f^i , $0 \leq i < j$, be the stream of payoffs with $f_t^i = 1$ if $n_k < t \leq n_{k+1}$ and $k = i \pmod j$ and $f_t^i = 0$ otherwise.

Not that $\sum_{0 \leq i < j} f^i = \mathbf{1}_{>n_0}$, $\mathbf{1} - 2f^i \in F_3$, and $v(f^i) \geq 0$. Therefore, as $v(\sum_{0 \leq i < j} f^i) = v(\mathbf{1}_{>n_0}) = 1$, there is i such that $v(f^i) \leq 1/j$ and therefore $v(\mathbf{1} - 2f^i) \geq 1 - 2/j$. Hence, $v(F_3) = 1$.

Obviously, by the definitions of the n -th Cesáro average u_n and the set F_3 , for any $f \in F_3$ we have $\liminf_{n \rightarrow \infty} u_n(f) = \liminf_{n \rightarrow \infty} \bar{f}_n = -1$. Therefore, no $f \in F_3$ is a robust 1- v -optimizer.

Similarly, if v is a non-impatient valuation, then, by choosing n_0 sufficiently large, we deduce that $v(F_3) = 1$, and that F_3 does not have a robust ε -optimizer at v whenever $\varepsilon < \lim_{n \rightarrow \infty} v(\mathbf{1}_{>n})$.

4.3. Proof of Proposition 1. First, we derive an inequality that does not depend on F having a robust ε -optimizer at v .

Note that for every neighborhood W of v , $\inf_{u \in W} u(F) \leq v(F) \leq \sup_{u \in W} u(F)$. Therefore,

$$\sup_{W \in \mathcal{N}(v)} \inf_{u \in W} u(F) \leq v(F) \leq \inf_{W \in \mathcal{N}(v)} \sup_{u \in W} u(F).$$

As

$$\sup_{h \in F, W \in \mathcal{N}(v)} \inf_{u \in W} u(h) \leq \sup_{W \in \mathcal{N}(v)} \inf_{u \in W} u(F),$$

we conclude that

$$\sup_{h \in F, W \in \mathcal{N}(v)} \inf_{u \in W} u(h) \leq v(F) \leq \inf_{W \in \mathcal{N}(v)} \sup_{u \in W} u(F) = \inf_{W \in \mathcal{N}(v)} \sup_{h \in F, u \in W} u(h).$$

Second, assume that f is a robust ε -optimizer at v with respect to F . Then, there is a neighborhood U of v such that for every $u \in U$

⁷We thank Bruno Ziliotto for the proof.

we have $u(f) \geq u(F) - \varepsilon$ and $|u(f) - v(F)| \leq \varepsilon$ (and thus $u(F) \leq u(f) + \varepsilon \leq v(F) + 2\varepsilon$). Therefore,

$$v(F) - \varepsilon \leq \inf_{u \in U} u(f) \leq \sup_{h \in F, W \in \mathcal{N}(v)} \inf_{u \in W} u(h).$$

Also,

$$\inf_{W \in \mathcal{N}(v)} \sup_{h \in F, u \in W} u(h) \leq \sup_{h \in F, u \in U} u(h) \leq \sup_{u \in U} u(F) \leq v(F) + 2\varepsilon.$$

Therefore,

$$v(F) - \varepsilon \leq \sup_{h \in F, W \in \mathcal{N}(v)} \inf_{u \in W} u(h) \leq \inf_{W \in \mathcal{N}(v)} \sup_{h \in F, u \in W} u(h) \leq v(F) + 2\varepsilon.$$

If F has a robust ε -optimizer at v for every $\varepsilon > 0$ we conclude that

$$\sup_{h \in F, W \in \mathcal{N}(v)} \inf_{u \in W} u(h) = v(F) = \inf_{W \in \mathcal{N}(v)} \sup_{h \in F, u \in W} u(h).$$

In the other direction, assume that

$$\sup_{h \in F, W \in \mathcal{N}(v)} \inf_{u \in W} u(h) = a = \inf_{W \in \mathcal{N}(v)} \sup_{h \in F, u \in W} u(h).$$

The left-hand side equality implies that for every $\varepsilon > 0$ there are $f \in F$ and neighborhoods $U \in \mathcal{N}(v)$ such that $u(f) \geq a - \varepsilon/2$ for every $u \in U$. In particular, $v(F) \geq a - \varepsilon/2$.

The right-hand side equality implies that for every $\varepsilon > 0$ there is a neighborhood $W \in \mathcal{N}(v)$ such that $u(F) \leq a + \varepsilon/2$ for every $u \in U$. In particular, $v(F) \leq a + \varepsilon/2$.

Therefore, $v(F) = a$, and for every $u \in U \cap W$,

$$v(F) + \varepsilon/2 \geq u(F) \geq u(f) \geq v(F) - \varepsilon/2 \geq u(F) - \varepsilon,$$

and thus f is a robust ε -optimizer at v with respect to F . \square

4.4. Proof of Proposition 3. Assume that v is an impatient valuation with $v(g) = \sum_{t=1}^{\infty} \omega_t g_t$ where $\omega_t \geq 0$ and $\sum_{t=1}^{\infty} \omega_t = 1$.

Fix $\varepsilon > 0$ and $g^\varepsilon \in F$ with $v(g^\varepsilon) > v(F) - \varepsilon$. We will prove that g^ε is a robust 10ε - v -optimizer in F .

Fix n sufficiently large such that $\sum_{t>n} \omega_t \|F\| < \varepsilon$, where $\|F\| = \sup_{f \in F} \|f\|$. Let W be the neighborhood of v of all valuations u such that $|u(\mathbf{e}_t) - \omega_t| \|F\| < \varepsilon/n \forall t \leq n$.

Then, for every $h \in F$ and $u \in W$, $u(h) \leq u(\sum_{t=1}^n h_t \mathbf{e}_t + \|F\| \mathbf{1}_{>n}) \leq \sum_{t=1}^n \omega_t h_t + \varepsilon + \|F\| (1 - \sum_{t=1}^n \omega_t) \leq \sum_{t=1}^n \omega_t h_t + 3\varepsilon \leq v(h) + 4\varepsilon$. Therefore, $u(F) \leq v(F) + 5\varepsilon$.

Similarly, $u(g^\varepsilon) \geq u(\sum_{t=1}^n g_t^\varepsilon \mathbf{e}_t - \|F\| \mathbf{1}_{>n}) \geq \sum_{t=1}^n \omega_t h_t - \varepsilon - \|F\| (1 - \sum_{t=1}^n \omega_t) \geq \sum_{t=1}^n \omega_t g_t^\varepsilon - 3\varepsilon \geq v(g^\varepsilon) - 4\varepsilon \geq v(F) - 5\varepsilon$.

Therefore, for any $u \in W$, $|u(g^\varepsilon) - v(F)| \leq 5\varepsilon$ and $u(g^\varepsilon) \geq v(F) - 5\varepsilon \geq u(F) - 10\varepsilon$. Therefore, g^ε is a robust 10ε - v -optimizer in F . \square

4.5. Proof of Proposition 4. Fix a non-impatient valuation u and a neighborhood U of u .

Let u_\perp^\perp denote the set of all $g \in c$ with $\|g\| = 1$ and $u(g) = 0$.

For every $v \in V \setminus U$, $\sup_{g \in u_\perp^\perp} v(g) > 0$. For every $\varepsilon > 0$ set $U_\varepsilon = \{v \in V : \sup_{g \in u_\perp^\perp} v(g) > \varepsilon\} = \bigcup_{g \in u_\perp^\perp} \{v \in V : v(g) > \varepsilon\}$. As a union of open sets, U_ε is an open set. Note that $U_{\varepsilon'} \supseteq U_\varepsilon$ if $\varepsilon' < \varepsilon$ and $\bigcup_{\varepsilon > 0} U_\varepsilon \supseteq V \setminus U$. Therefore, there is $\varepsilon > 0$ such that $U_{\varepsilon'} \supseteq V \setminus U$ for every $\varepsilon' \leq \varepsilon$.

Let $\varepsilon < 1$ be sufficiently small so that $U_\varepsilon \supseteq V \setminus U$. For every $v \in U_\varepsilon$ there is an element $g^v \in u_\perp^\perp$ and a neighborhood $U_\varepsilon(v)$ of v such that for every $w \in U_\varepsilon(v)$, we have $w(g^v) > \varepsilon$. As $\bigcup_{v \in U_\varepsilon} U_\varepsilon(v) \supseteq V \setminus U$, there is a finite list v^1, \dots, v^k such that $\bigcup_{1 \leq i \leq k} U_\varepsilon(v^i) \supseteq V \setminus U$.

Let $F_\varepsilon(u)$ be the finite set $\{g^{v^i} : 1 \leq i \leq k\}$.

Let $h \in \ell_\infty$ be a stream of payoffs with $\|h\| = \varepsilon$, $\limsup_{t \rightarrow \infty} \bar{h}_t = \varepsilon$, and $\liminf_{t \rightarrow \infty} \bar{h}_t = -\varepsilon$.

Define $u_t = u(\mathbf{e}_t)$ if $t \geq 1$ and $u_0 = 1 - \sum_{t=1}^\infty u_t$. As u is a non-impatient valuation, $0 < u_0 \leq 1$.

Let n_ε be sufficiently large so that $\sum_{t > n_\varepsilon} |u_t| < u_0\varepsilon/4$.

Let H be the set of all streams of payoffs h^n , $n > n_\varepsilon$, where $h_t^n = h_t$ if $n_\varepsilon < t \leq n$ and $h_t^n = 0$ otherwise.

Let g be the stream of payoffs where $g_t = u_t / \sum_{t=0}^\infty u_t^2$ if $t \leq n_\varepsilon$ and $g_t = (u_0 - \sum_{t=n+1}^\infty u_t) / \sum_{t=0}^\infty u_t^2$ if $t > n_\varepsilon$.

Let $F = (g + H) \cup (g + F_\varepsilon(u))$.

In order to prove (1), (2), and (3), it suffices to construct a sequence of impatient valuations w_n that converges to u and a positive number $\eta > 0$, such that for any finite subset G of F , $\limsup_{n \rightarrow \infty} \sup_{f \in F} \omega_n(f) > \eta + \limsup_{n \rightarrow \infty} \max_{g \in G} \omega_n(g)$.

As u is a non-impatient valuation, $u = (1 - u_0)w + u_0v$ where v is a patient valuation. Therefore, the sequence of impatient valuations $w_n := (1 - u_0)w + u_0\gamma_n$, where $\gamma_n(g) = \bar{g}_n$, converges to u .

By the definition of g , it follows that

$$1 \geq u(g) = \sum_{t=0}^n u_t^2 / \sum_{t=0}^\infty u_t^2 - \left(\sum_{t=n+1}^\infty u_t \right)^2 / \sum_{t=0}^\infty u_t^2 > 1 - u_0\varepsilon/16.$$

By the properties of h , there are sequences of integers n_m that converge to infinity such that $\lim_{m \rightarrow \infty} w_{n_m}(g + h^{n_m}) = u(g) + u_0\varepsilon$. Therefore, $w_{n_m}(g + h^{n_m}) \geq 1 + u_0\varepsilon/2$ for all sufficiently large m .

For every $n_1 \in \mathbb{N}$ and $f \in F_\varepsilon(u)$, we have $\lim_{n \rightarrow \infty} w_n(g + h^{n_1}) = u(g + h^{n_1}) \leq 1 + u_0\varepsilon/4$ and $\lim_{n \rightarrow \infty} w_n(g + f) \leq 1$. Therefore, for every finite subset G of F , $\limsup_{n \rightarrow \infty} \max_{g \in G} w_n(g) \leq 1 + u_0\varepsilon/4$. This completes the proof of properties (1) and (3) of the set F .

Let $v \in U_\varepsilon$. Define $\alpha(F, v) := v(g) + \max_{f \in F_\varepsilon(u)} v(f)$. By the definitions of U_ε and the finite set $F_\varepsilon(u)$, there is $f^* \in F_\varepsilon(u)$ such that $\alpha(F, v) = v(g + f^*) > v(g) + \varepsilon \geq v(g + h) \forall h \in H$.

Fix $0 < \eta < v(f^*) - \varepsilon$. Let $U(v)$ be the set of all valuations w such that $|w(g) - v(g)| + |w(f) - v(f)| < \eta$ for all $f \in F_\varepsilon(u)$. As $F_\varepsilon(u)$ is a finite subset of c and g is a fixed element of c , $U(v)$ is a neighborhood of v .

The definitions of $U(v)$ and f^* imply that if $w \in U(v)$, then, $w(g + f) \leq v(g + f^*) + \eta = \alpha(F, v) + \eta$ for all $f \in F_\varepsilon(u)$.

As $\|h\| \leq \varepsilon$ for every $h \in H$, the properties of f^* and η imply that $w(g + h) \leq w(g) + \varepsilon \leq v(g) + \varepsilon + \eta < \alpha(F, v)$ for all $h \in H$.

The definitions of F , $U(v)$, and f^* imply that $w(F) \geq w(g + f^*) \geq v(g + f^*) - \eta = \alpha(F, v) - \eta$.

As F is the union of $g + H$ and $g + F_\varepsilon(u)$, and $g + f^* \in F$, we conclude that $\alpha(F, v) - \eta \leq w(F) \leq \alpha(F, v) + \eta$. This completes the proof of property (2) of the set F . \square

5. REMARKS

5.1. Related literature. Fishburn [3] presented axioms, called additivity and convergence, that, for an ordinal utility u over a denumerable product set $X = \prod_{t=1}^{\infty} X_t$, imply an additive representation of u : for every vector $x = (x_1, x_2, \dots) \in X$,

$$u(x_1, x_2, \dots) = \sum_{t=1}^{\infty} g_t(x_t),$$

where u, g_1, g_2, \dots are real-valued functions defined on the sets X, X_1, X_2, \dots , respectively.

Fishburn's additivity axiom states that for any $x, y, x^*, y^* \in X$, if $\{x_i, y_i\} = \{x_i^*, y_i^*\} \forall i$, then $u(x) + u(y) = u(x^*) + u(y^*)$. It is related to our additivity assumption. It is easy to see that in our axiomatization results the additivity can be replaced by a similar additivity axiom: $\forall g, h, g^*, h^* \in \ell_\infty$, if $g + h = g^* + h^*$, the $u(g) + u(h) = u(g^*) + u(h^*)$.

Fishburn's convergence axiom states that for any $x, y \in X$,

$$u(x_1, \dots, x_n, y_{n+1}, y_{n+2}, \dots) \rightarrow_{n \rightarrow \infty} u(x).$$

It is related to our impatience property. Indeed, an additive and monotonic valuation u is impatient iff for any $g, h \in \ell_\infty$,

$$u(g_1, \dots, g_n, h_{n+1}, h_{n+2}, \dots) \rightarrow_{n \rightarrow \infty} u(g).$$

In Fishburn's setup, the analogous patience property is: u is patient if for any $x, y \in X$,

$$u(x_1, \dots, x_n, y_{n+1}, y_{n+2}, \dots) \rightarrow_{n \rightarrow \infty} u(y).$$

In Fishburn's setup, the following is analogous to our time-value of money property. For any $x, y, z \in X$, if $u(x_1, \dots, x_n, z_{n+1}, z_{n+2}, \dots) \geq u(y_1, \dots, y_n, z_{n+1}, z_{n+2}, \dots) \forall n$, then $u(x) \geq u(y)$. This analogous axiom follows from Fishburn's convergence axiom. However, without the convergence axiom it is an independent axiom.

5.2. The time-value of money property and impatience.

Lemma 1. *A monotonic, impatient, and additive function $u : \ell_\infty \rightarrow \mathbb{R}$ that obeys $u(\mathbf{e}_t) \geq u(\mathbf{e}_{t+1})$ satisfies the time-value of money property.*

Proof. Assume that $g, h \in \ell_\infty$ with $\sum_{t=1}^s h_t \geq \sum_{t=1}^s g_t \forall s$ and let $u : \ell_\infty \rightarrow \mathbb{R}$ be a monotonic, impatient, and additive function that obeys $w_t := u(\mathbf{e}_t) \geq u(\mathbf{e}_{t+1})$. Then, as in the proof of Theorem 1, $u(g) = \sum_{t=1}^\infty w_t g_t = \sum_{t=1}^\infty (w_t - w_{t+1}) t \bar{g}_t \leq \sum_{t=1}^\infty (w_t - w_{t+1}) t \bar{h}_t = \sum_{t=1}^\infty h_t w_t = u(h)$. \square

5.3. Properties of patient valuations. Forges [5] labels a linear functional v on ℓ_∞ "time-neutral" if v satisfies (4), and Lauwers [6] proves that a linear functional u on ℓ_∞ is time-neutral iff it is monotonic and $u(g) = u(\pi g)$ for every permutation π such that $\lim_n \pi(n)/n = 1$ (where (πg) is defined by $(\pi g)_t = g_{\pi(t)}$).

The next results shows that the lower and upper bound in Theorem 2 are tight.

Lemma 2. *For every bounded g there are patient valuations u and v such that $v(g) = \underline{g}$ ($= \liminf_{n \rightarrow \infty} \bar{g}_n$) and $u(g) = \bar{g}$ ($= \limsup_{n \rightarrow \infty} \bar{g}_n$).*

Proof. Fix $g \in \ell_\infty$.

Let U be the one-dimensional subspace of ℓ_∞ that is spanned by g . Define the linear functional φ on U by $\varphi(\theta g) = \theta \bar{g}$.

Define the function $p : \ell_\infty \rightarrow \mathbb{R}$ by $p(h) = \bar{h}$. Then, p is sublinear (i.e., $p(g+h) \leq p(g) + p(h)$ and $p(\theta g) = \theta p(g)$ for all $g, h \in \ell_\infty, \theta \in \mathbb{R}_+$) and $\varphi(h) \leq p(h) = \bar{h}$ for all $h \in U$.

Therefore, by the Hahn–Banach theorem, there is a linear functional u on ℓ_∞ such that $u(h) \leq p(h) = \bar{h} \forall h \in \ell_\infty$ and $u(g) = \varphi(g) = \bar{g}$.

It remains to show that $\underline{h} \leq u(h)$ for all $h \in \ell_\infty$, which follows from $\underline{h} = \liminf_{n \rightarrow \infty} \bar{h}_n = -\limsup_{n \rightarrow \infty} -\bar{h}_n = -\limsup_{n \rightarrow \infty} \overline{(-h)}_n = -p(-h) \leq -u(-h) = u(h)$.

Applying the above-proved part to the element $-g$ of ℓ_∞ shows that there is a linear functional v on ℓ_∞ such that $v(-g) = \limsup_{n \rightarrow \infty} -\bar{g}_n$; thus $v(g) = \liminf_{n \rightarrow \infty} \bar{g}_n$, and $v(-h) \geq \liminf_{n \rightarrow \infty} -\bar{h}_n$; thus $v(h) \leq \bar{h} \forall h \in \ell_\infty$ and $v(h) = -v(-h) \geq -(\bar{-h}) = \underline{h}$. \square

Lemma 3. *There is a normalized linear function $w : \ell_\infty \rightarrow \mathbb{R}$ that is monotonic and satisfies $w(\mathbf{e}_t) = 0 \forall t$; thus $w(\mathbf{e}_t) \geq w(\mathbf{e}_{t+1})$, but does not satisfy the time-value of money property.*

Proof. Define the following two linear operators on ℓ_∞ . The linear operator $O : \ell_\infty \rightarrow \ell_\infty$ is defined by $Oh = (h_1, h_3, h_5, \dots)$, i.e., $(Oh)_t = h_{2t-1}$, and $Eh = (h_2, h_4, h_6, \dots)$, i.e., $(Eh)_t = h_{2t}$.

Let $\mathbf{0} \leq g \in \ell_\infty$ with $\underline{g} < \bar{g}$. Let u and v be two patient valuations such that $u(g) = \underline{g}$ and $v(g) = \bar{g}$.

Therefore, $u(g) - v(g) < 0$ and $u(\mathbf{e}_t) = v(\mathbf{e}_t) = 0 \forall t$.

Define the function $w : \ell_\infty \rightarrow \mathbb{R}$ by $w(h) = u(Oh)/2 + u(Eh)/2$. We claim that w is normalized, linear, monotonic, and satisfies $w(\mathbf{e}_t) \geq w(\mathbf{e}_{t+1})$, but w does not satisfy the time-value of money property.

First, note that as $u \circ O$ and $u \circ E$ are normalized, linear, and monotonic, and, therefore, so is their average w . As $w(\mathbf{e}_t) = 0 \forall t$, $0 = w(\mathbf{e}_t) \geq w(\mathbf{e}_{t+1}) = 0 \forall t$.

Next, define h by $Oh = g$ and $Eh = -g$, i.e., $h = (g_1, -g_1, g_2, -g_2, \dots)$. Note that $\sum_{t=1}^{2n} h_t = 0$ and that $\sum_{t=1}^{2n-1} h_t = g_n \geq 0$. But, $2w(h) = u(Oh) + u(Eh) = u(g) - v(g) < 0$. Therefore, w does not satisfy the time-value of money property. \square

5.4. Impatient robust optimization. Confining the theory of robust optimization to impatient valuations leads to the following modification of the definition of a robust optimizer.

For any set U we denote by U^* the set of all impatient valuations in U . Let v be a valuation and $\varepsilon \geq 0$. A small imprecision in the specification of an impatient valuation is modeled as the set of impatient valuations in a small neighborhood of a valuation v , and v need not be an impatient valuation.

A stream f in F is an *impatient-robust ε -optimizer at v with respect to F* if there is a neighborhood U of v such that

$$u(f) \geq w(F) - \varepsilon \quad \forall u, w \in U^*.$$

A robust ε -optimizer at v with respect to F is obviously an impatient-robust ε -optimizer at v with respect to F . We now show that the converse holds as well.

Fix a stream f and a neighborhood U of a valuation v . The infimum of $u(f)$ over all $u \in U^*$ equals the infimum of $u(f)$ over all $u \in U$, and the supremum of $w(f)$ over all $w \in U^*$ equals the supremum of $w(f)$ over all $w \in U$. Therefore, if f is an impatient-robust ε -optimizer at v with respect to F , then $u(f) \geq w(F) - \varepsilon$ for all $u, w \in U$; hence, f is a robust ε -optimizer at v with respect to F .

In the robustness result for a finite MDP we alluded to stringent robustness conditions that are called for when the decision maker chooses between different possible distribution over streams of payoffs. We introduce the formal definition.

Let \mathcal{P} be a set of distribution P over streams of payoffs. For every valuation u we denote by $\underline{u}(P)$ the expectation of $\underline{u}(f)$ with respect to the distribution P , and we denote by $\bar{u}(P)$ the expectation of $\bar{u}(f)$ with respect to the distribution P . The supremum of $\bar{u}(P)$ over all $P \in \mathcal{P}$ is denoted by $\bar{u}(\mathcal{P})$. Let v be a valuation.

A distribution P in \mathcal{P} is a *robust ε -optimizer at v with respect to \mathcal{P}* if there is a neighborhood U of v such that

$$\underline{u}(P) \geq \bar{w}(\mathcal{P}) - \varepsilon \quad \forall u, w \in U.$$

A distribution P in \mathcal{P} is an *impatient-robust ε -optimizer at v with respect to \mathcal{P}* if there is a neighborhood U of v such that

$$\underline{u}(P) \geq \bar{w}(\mathcal{P}) - \varepsilon \quad \forall u, w \in U^*.$$

A robust ε -optimizer at v with respect to \mathcal{P} is obviously an impatient-robust ε -optimizer at v with respect to \mathcal{P} . The converse need not hold.

For example, let g be a stream of payoffs with $\liminf_{n \rightarrow \infty} \bar{g}_n = -1 < \limsup_{n \rightarrow \infty} \bar{g}_n = 1$. Let \mathcal{P} be the set consisting of the single distribution P with $P(g) = 1/2 = P(-g)$. For any impatient valuation u , $u(P) = 0 = \underline{u}(P)$. Therefore P is a V^* -robust ε -optimizer in \mathcal{P} . In particular, for any valuation v , P is an impatient-robust v - ε -optimizer in \mathcal{P} . However, if w is a patient valuation then $\bar{w}(P) = 1$. Therefore, if v is a patient valuation, P is not a robust v - ε -optimizer P in \mathcal{P} .

5.5. The continuous-time theory. In the continuous time theory a bounded stream of payoffs is a bounded measurable function $[0, \infty) \ni t \mapsto g_t \in \mathbb{R}$. The linear space of bounded streams of payoffs is denoted by L_∞ , and $\mathbf{1}_{\leq T}$ is the stream g with $g_t = 1$ if $t \leq T$ and $g_t = 0$ if $t > T$. Similarly, one defines $\mathbf{1}$ and $\mathbf{1}_{>T}$ in analogy to the definitions in the discrete-time case.

A *valuation* is an additive function $v : L_\infty \rightarrow \mathbb{R}$ that is normalized, i.e., $v(\mathbf{1}) = 1$, and satisfies the time-value of money property: if $\int_0^T g_t dt \geq \int_0^T h_t dt \forall T \geq 0$, then $v(g) \geq v(h)$. A valuation v is *impatient* if $v(\mathbf{1}_{>T}) \rightarrow_{T \rightarrow \infty} 0$; it is *patient* if $v(\mathbf{1}_{>T}) = 1 \forall T$ (equivalently, $v(\mathbf{1}_{>T}) \rightarrow_{T \rightarrow \infty} 1$).

The characterizations of impatient valuations, patient valuations, and valuations are analogous to those in the discrete-time case.

A real-valued function u that is defined on L_∞ is an impatient valuation iff there is a function $[0, \infty) \ni t \rightarrow w_t \in \mathbb{R}$, with $\int_0^\infty w_t dt = 1$, that is nonincreasing on $(0, \infty)$ and such that

$$u(g) = \int_0^\infty g_t w_t dt.$$

A real-valued function u that is defined on L_∞ is a patient valuation iff it is a linear function on L_∞ such that

$$\liminf_{T \rightarrow \infty} \frac{1}{T} \int_0^T g_t dt \leq u(g) \leq \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T g_t dt.$$

A real-valued function u that is defined on L_∞ is a valuation iff it is a convex combination of an impatient valuation and a patient one.

Similarly, the analogous results of the other theorems and propositions hold also in the continuous-time case.

The topology on the valuation in the continuous-time case is the minimal one where for every $g \in C$, where C consists of all elements $g \in L_\infty$ such that the limit $\lim_{t \rightarrow \infty} g_t$ exists, the function $v \mapsto v(g)$ is continuous.

REFERENCES

- [1] Blackwell, D. (1962), Discrete dynamic programming, *The Annals of Mathematical Statistics*, 2, 724–738.
- [2] Debreu, G. (1960), Topological methods in cardinal utility theory, in *Mathematical Methods in the Social Sciences*, 1959. K. J. Arrow, S. Karlin, and P. Suppes (eds.), Stanford: Stanford University Press, pp. 16–26.
- [3] Fishburn, P. C. (1966), Additivity in utility with denumerable product sets, *Econometrica*, 34, 500–503.

-
- [4] Fishburn, P. C. and W. Edwards (1997), Discount-neutral utility models for denumerable time streams, *Theory and Decision*, 43, 139–166.
 - [5] Forges, F. (1986), Correlated equilibria in repeated games with lack of information on one side: A model with verifiable types, *International Journal of Game Theory*, 15, 65–82.
 - [6] Lauwers L. (1995), Time-neutrality and linearity, *Journal of Mathematical Economics*, 24, 347–351.
 - [7] Mertens, J.-F. and A. Neyman (1981), Stochastic games, *International Journal of Game Theory*, 10, 53–66.
 - [8] Neyman, A. (2003), From Markov chains to stochastic games, in *Stochastic Games and Applications*, A. Neyman and S. Sorin (eds.), NATO ASI Series, Kluwer Academic Publishers, pp. 9–25.