JOURNAL OF ECONOMIC THEORY **83**, 105–144 (1998)
ARTICLE NO. ET972451

An Axiomatic Approach to Complete Patience and Time Invariance

Massimo Marinacci*

Department of Economics, University of Toronto, Toronto M5S 3G7, Canada
massimo@chass.utoronto.ca

Received October 24, 1997; revised May 4, 1998

The standard criterion used to compare streams of payoffs in the undiscounted case is $\liminf_{T\to\infty} 1/T\sum_{t=1}^{T} u(x_t)$. In this paper we approach the problem axiomatically. This sheds light on the behavioral underpinnings of such a rule and leads to a novel choice criterion, the Polya Index. *Journal of Economic Literature* Classification Numbers: C72, C73, D90. © 1998 Academic Press

1. INTRODUCTION

In some intertemporal problems it is important to consider agents who do not discount future utilities but instead attach the same importance to all periods, no matter how far apart they are. This is the case for a social planner who allocates resources among different generations, or for players who greatly value a long time horizon in a repeated strategic interaction. For instance, the celebrated folk theorems of Aumann and Shapley [1] and Rubinstein [20] consider complete patient players, as well as earlier works on infinitely repeated stochastic games (see, e.g., Blackwell and Ferguson [3]). Complete patient social planners have been considered in growth theory by the literature pioneered by Ramsey [19].

In the discounted case, the standard criterion used to compare infinite streams of payoffs $\{x_1, ..., x_n, ...\}$ is $(1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} u(x_t)$ for $0<\delta<1$. Without discounting, it seems natural to focus on the limit of the time averages $\lim_{T\to\infty} 1/T\sum_{t=1}^{T} u(x_t)$. In particular, the following classic result shows that this criterion can be thought of as the limiting case of discounting.[1]

* I thank Larry Epstein, Itzhak Gilboa, Marco Scarsini, an anonymous referee, and participants at the Chantilly Workshop on Decision Theory, June, 1997, for some very useful discussions. The financial support of the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged.

[1] The "if" part is due to Littlewood [13], while the more famous "only if" part is due to Frobenius [8].

0022-0531/98 $25.00

THEOREM 1. *Let* $\{u(x_1), ..., u(x_n), ...\}$ *be a bounded stream of utilities. Then* $\lim_{T\to\infty} 1/T \sum_{t=1}^{T} u(x_t)$ *exists if and only if* $\lim_{\delta\to 1}(1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} u(x_t)$ *exists. In this case*

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} u(x_t) = \lim_{\delta\to 1} (1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} u(x_t).$$

However, these limits often do not exist, as the following simple example shows.

EXAMPLE. Let $u(x_t)$ be the following sequence of zeros and ones:

$$1, 1, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 0, ...$$

It is easy to check that $\lim_{T\to\infty} 1/T \sum_{t=1}^{T} u(x_t)$ does not exist, and, in particular, we have[2]

$$\liminf_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} u(x_t) = \frac{1}{3} \qquad \text{and} \qquad \limsup_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} u(x_t) = \frac{2}{3}.$$

Because of this existence problem, the limit of time averages criterion is often replaced by the more general

$$\liminf_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} u(x_t). \tag{1}$$

However, no behavioral underpinning for this more general criterion has been provided. As there are many other methods which can be used to rank non-convergent sequences (see, e.g., Hardy [11]), it is not clear why (1), a rather crude alternative, should be preferred. Moreover, the limit of time averages itself lacks a clear behavioral underpinning; Theorem 1, however interesting, falls short of providing one.

1.1. *Our approach*

In this paper we approach the problem axiomatically. In particular, we look at complete patience and time invariance, the two main features of these time averaging criteria that have been discussed in the literature (see, e.g., Fudenberg and Tirole [9] pp. 148–149). For convenience, we briefly summarize the intuitive meanings of these two properties:

1. A time preference reflects complete patience if all periods of time are equally weighted.

[2] For a more general result, see Proposition 22 in the appendix.

2. A time preference is time invariant if the payoffs obtained in any finite number of periods do not matter.

In the paper we first study the second property. In particular, we prove that time invariance per se would deliver the following criterion:

$$\lim_{T\to\infty}\left[\inf_{j\geqslant 1}\left(\frac{1}{T}\sum_{t=1}^{T} u(x_{j+t})\right)\right].$$

Notice that such a limit always exists.

After having established this result, we focus on patience. Unlike time invariance, complete patience is much trickier to axiomatize. In the finite case there is a natural definition of complete patience: we have complete patience when the ordering of any two payoff streams does not change by taking arbitrary permutations of their respective time indexes. However, we show that a literal translation of this definition to the infinite case is highly unsatisfactory. In particular, it delivers the following criterion:

$$\lim_{t\to\infty}\inf\ u(x_t),$$

where only instantaneous utilities are considered.

To provide a more interesting definition of complete patience we use natural densities. For a given subset of points of time A, its natural density $\delta(A)$ is defined by

$$\lim_{t\to\infty}\frac{|A\cap\{1,...,t\}|}{t}$$

whenever the limit exists, which is not always the case. Loosely speaking, a permutation x^{π} of a payoff stream $x=\{x_t\}_{t\geqslant 1}$ preserves the upper sets' densities if $\delta(\{t\colon u(x_t)\geqslant\alpha\})=\delta(\{t\colon u(x_t^{\pi})\geqslant\alpha\})$ for all real numbers α. These permutations do not change the relative frequencies with which the different payoffs come up in the stream (in section 6 a simple example is provided).

We say that an agent is completely patient when the ordering of two payoff streams does not change by taking permutations that preserve the upper sets' densities. This more compelling definition of patience delivers, up to a technical condition, the following criterion:

$$\lim_{\varepsilon\to 0}\left[\lim_{T\to\infty}\inf\frac{1}{\varepsilon T}\sum_{t=(1-\varepsilon)T}^{T} u(x_t)\right]. \qquad (2)$$

We call this criterion the Polya Index. It is well defined for every possible bounded payoff stream. In particular,

$$\lim_{\varepsilon \to 0} \left[\lim_{T \to \infty} \inf \frac{1}{\varepsilon T} \sum_{t=(1-\varepsilon) T}^{T} u(x_t) \right] = \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} u(x_t)$$

when the limit of time averages exists, so that the Polya Index extends the standard limit criterion. This implies that our representation theorem provides a foundation for the limit criterion as well.

The Polya Index has an interesting characterization. Let $\mathscr{F}$ be the set of all bounded payoff streams, $\succsim$ a time preference on $\mathscr{F}$ that can be represented by the Polya Index, and $u(x_t)$ its corresponding instantaneous utility. Let

$$\mathscr{F}_a = \left\{ x = \{x_t\}_{t \geqslant 1} \in \mathscr{F} \colon \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} u(x_t) \text{ exists} \right\},$$

that is, $\mathscr{F}_a$ is the set of all payoff streams that have a well defined limit of time averages. For any given stream $x = \{x_t\}_{t \geqslant 1}$, it holds that

$$P(x) = \sup \left\{ \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} u(x'_t) \colon x' \in \mathscr{F}_a \text{ and } x \succsim x' \right\}, \tag{3}$$

where $P(x)$ denotes the Polya Index of the stream x. That is, the Polya Index of x is the supremum of the time average limits taken over all streams x' for which such a limit is well defined, and such that $x \succsim x'$.

Besides its intrinsic interest, this characterization, together with the original form (2), seems to provide the Polya Index with an interesting analytic tractability.[3]

1.2. *Lim inf*

Instead of deriving the standard lim inf criterion, our axiomatic approach led us to the Polya Index. However, our analysis also sheds new light on the lim inf criterion. To see why this is the case we have to make a short digression. This work started as a dividend of the analysis of Marinacci [15]. In that paper it was shown that for any bounded sequence

[3] As we will prove, it also holds that

$$P(x) = \sup \left\{ \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} u(x'_t) \colon x' \in \mathscr{F}_a \text{ and } u(x'_t) \leqslant u(x_t) \text{ for all } t \geqslant 1 \right\}.$$

This representation seems especially useful in terms of analytical tractability.

$f\colon \mathbb{N} \to \mathbb{R}$ there exists a non-additive normalized measure $v\colon 2^{\mathbb{N}} \to [0, 1]$ such that

$$\liminf_{n} f(n) = \int f(n)\, dv,$$

where an appropriate notion of integral, due to Choquet [5], is used (see the appendix for details).

As Choquet integrals have been used to model vague subjective beliefs in Schmeidler [21], this observation suggested the possibility of using that framework to study time preferences. It turned out, however, that the most appropriate framework was the closely connected multiple priors model of Gilboa and Schmeidler [10].

We now illustrate this point. In our temporal context we have weights instead of priors, which represent how much the agents value the different points of time. Combined with utilities, different weights deliver different rankings of the payoff streams. In particular, a natural weight for complete patience would be the natural density defined above. However, this density fails to exist for many sets, and so we cannot use it as a weight. Lacking this "ideal" weight, we assume that agents replace it with sets of weights, in particular with those weights that coincide with the natural density whenever it exists. This is why the multiple priors model is useful for our purposes.[4]

By using this model as our set-up we derive the Polya Index. Moreover, we show that there exists a strict subset $\mathscr{C}_l$ of the set of weights just described such that

$$\liminf_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} u(x_t) = \min_{\mu \in \mathscr{C}_l} \int u(x_t)\, d\mu. \tag{4}$$

Even though we have not been able to determine which further requirements on preferences are needed to move to the strict subset $\mathscr{C}_l$, the equality (4) sheds light on the nature of the lim inf criterion, and on the way in which it combines weights and utilities. In particular, it provides a novel behavioral perspective on the lim inf criterion: an agent who uses such a criterion can be viewed as using a set of weights, all coinciding with the natural density when this "ideal" weight exists. The set is then summarized through the minimum $\min_{\mu \in \mathscr{C}_l} \int u(x_t)\, d\mu$.

[4] Of course, an alternative justification for this model in our temporal context is that the agents are not sure which weight to use and, instead of a single one, use a set of them. This justification is, *mutatis mutandis*, more in line with the original argument of Gilboa and Schmeidler [10].

In sum, our axiomatic approach to time invariance and complete patience led to some new criteria to rank streams of payoffs, notably the Polya Index, and shed new light on the lim inf criterion, the most used in the non discounted case. As a secondary contribution, we provide a connection between the two apparently unrelated issues of modelling vagueness in subjective beliefs and complete patience in time preferences. Finally, in a companion paper, Marinacci [16], we show that our axiomatic approach leads to a considerable generalization of the undiscounted Folk Theorems. Specifically, we show that they can be proved by imposing only conditions on preferences, without relying on any particular evaluation functionals. In so doing, we generalize and unify several important results obtained in the case of complete patience, included the classic results of Aumann and Shapley [1] and Rubinstein [20]. Moreover, our results are based only on properties of preferences and this makes transparent their behavioral foundation.

The rest of the paper is organized as follows. Section 2 describes the set-up, and reports the representation theorem of Gilboa and Schmeidler [10]. Section 3 examines time invariance, and proves a representation result for this property. Section 4 considers the "naive" definition of complete patience, and shows what kind of representation result it entails. Section 5 shows how unsatisfactory the naive definition is and argues that natural densities have to be considered. It also shows that a form of patience is already incorporated in the time invariance axiom. Section 6 formally defines patience by means of densities, and proves the relative representation theorem, where the Polya Index first occurs. Section 7 provides two interesting characterizations of the Polya Index. Finally, section 8 considers the lim inf criterion, and shows that our analysis sheds new light on this criterion as well. All the proofs, and the most technical analysis, are relegated to the appendix. A glossary of the more relevant notation is provided at the beginning of the appendix.

2. SET-UP

We use the generalization of the Anscombe–Aumann model introduced in Gilboa and Schmeidler [10] and Schmeidler [21].

Let $\mathscr{X}$ be a nonempty set of consequences and $\mathscr{P}$ the set of all probability distributions with finite support on $\mathscr{X}$, i.e.,

$$\mathscr{P}=\left\{p\colon \mathscr{X}\to[0,1] : p(c)\neq 0 \text{ for finitely many } c\text{'s in } \mathscr{X} \text{ and } \sum_{c\in\mathscr{X}} p(c)=1\right\}.$$

Let $\mathscr{T} = \{1, ..., t, ...\}$ be the set of points of time, and $2^{\mathscr{T}}$ its power set. An act f is a function from $\mathscr{T}$ into $\mathscr{P}$. For $p \in \mathscr{P}$, p^* denotes the constant act $p^*(t) = p$ for all $t \in \mathscr{T}$.

The set of all acts is endowed with a preference ordering $\succsim$. In particular, $\mathscr{F}$ denotes the set of all bounded acts, i.e., $f \in \mathscr{F}$ if there are $p_1, p_2 \in \mathscr{P}$ such that $p_1^* \precsim f \precsim p_2^*$ for all $t \in \mathscr{T}$.

We now present several axioms on $\succsim$.

A.1. *Weak Order.* $\succsim$ is complete and transitive.

A.2. *Monotonicity.* For all acts $f, g \in \mathscr{F}$ we have $f \succsim g$ whenever $f(t) \succsim g(t)$ for all $t \in \mathscr{T}$.

A.3. *Continuity.* For all $f, g, h \in \mathscr{F}$, if $f \succ g$ and $g \succ h$, then there are $\alpha, \beta \in (0, 1)$ such that $\alpha f + (1-\alpha) h \succ g \succ \beta f + (1-\beta) h$.

A.4. *Nondegeneracy.* There exist $f, g \in \mathscr{F}$ such that $f \succ g$.

All the above axioms are standard, and have a simple interpretation. The next axiom is a weak version of the Independence Axiom, which only requires independence with respect to constant acts.

A.5. *Certainty Independence.* For all acts f and g, for all constant acts c, and all $0 < \alpha < 1$, $f \succsim g$ if and only if $\alpha f + (1-\alpha) c \succsim \alpha g + (1-\alpha) c$.

Next we introduce a smoothing axiom: the agent always weakly prefers to smooth his payoff stream by mixing two indifferent acts rather than have only one of them all the time.

A.6. *Intertemporal Smoothing.* For all f and g in $\mathscr{F}$, $f \sim g$ implies $\alpha f + (1-\alpha) g \succsim f$ for all $0 < \alpha < 1$.

Finally, a key ingredient in a temporal decision is how the agents weight the different points of time. In this set-up, where infinite points of time are considered, formally a weight is a set function $\mu: 2^{\mathscr{T}} \to [0, 1]$ that satisfies the following conditions:

(i) $\mu(\emptyset) = 0$;

(ii) $\mu(A \cup B) = \mu(A) + \mu(B)$ whenever $A \cap B = \emptyset$;

(iii) $\mu(\mathscr{T}) = 1$.

We can now report the Gilboa and Schmeidler Theorem for our set-up (Chateauneuf [4] proved independently a similar result).

THEOREM 2. *The following two statements are equivalent*:

(i) *The preference relation* $\succsim$ *on* $\mathscr{F}$ *satisfies the axioms* A.1–A.6.

(ii) *There exists an affine real valued function* u *on* $\mathscr{P}$ *and a unique non-empty weak*-compact and convex set* $\mathscr{C}$ *of weights on* $2^{\mathscr{T}}$ *such that, for all* f *and* g *in* $\mathscr{F}$,

$$f \succsim g \qquad \text{if and only if} \quad \min_{\mu \in \mathscr{C}} \int_{\mathscr{T}} u(f(t))\, d\mu \geqslant \min_{\mu \in \mathscr{C}} \int_{\mathscr{T}} u(g(t))\, d\mu.$$

Finally, the function u *is unique up to a positive linear transformation.*

Interpreted in our temporal context, this representation means that the agent does not evaluate the payoff streams through a single weight, but instead uses a set of weights, summarized by the minimum $\min_{\mu \in \mathscr{C}} \int_{\mathscr{T}} u(f(t))\, d\mu$.

By Theorem 2, every preference relation $\succsim$ that satisfies axioms A.1-A.6 is associated with a pair $(u, \mathscr{C})$, the utility function on $\mathscr{P}$ and the set of multiple weights. Using these pairs it is possible to introduce a natural partial order R on the set of preference relations satisfying axioms A.1–A.6: We write $\succsim R \succsim'$ if the two following conditions are satisfied:

(i) the utility functions u and u' on $\mathscr{P}$ are equal, up to positive linear transformations;

(ii) the set $\mathscr{C}'$ is contained in $\mathscr{C}$, i.e., $\mathscr{C}' \subseteq \mathscr{C}$.

The partial order R is reflexive, transitive, and antisymmetric. It is easy to see that for two preferences $\succsim$ and $\succsim'$ satisfying axioms A.1–A.6, the following two statements are equivalent:

(i) $\succsim R \succsim'$

(ii) for all $f \in \mathscr{F}$ and all constant acts p^* it holds that

$$\begin{aligned} p^* \succsim' f \qquad &\text{implies} \quad p^* \succsim f \\ p^* \succ' f \qquad &\text{implies} \quad p^* \succ f. \end{aligned}$$

DEFINITION 3. Let $\succsim$ be a preference relation that satisfies a given set of axioms, which includes A.1–A.6. We call $\succsim$ canonical if $\succsim R \succsim'$ for all other preferences $\succsim'$ that satisfy the same set of axioms.

In other words, a preference relation $\succsim$ that satisfies a given set of axioms is canonical if it holds that $\mathscr{C}' \subseteq \mathscr{C}$ for all preferences $\succsim'$ which satisfy the same set of axioms and have the same utility function on $\mathscr{P}$.

It is important to keep in mind that $\succsim$ is canonical with respect to a given set of axioms. Indeed, different sets of axioms may be associated with different canonical preference relations.

2.1. *Comonotonic Independence*

In the sequel we will need another axiom, due to Schmeidler [21]. Two acts f and g in $\mathscr{F}$ are comonotonic if for no $t, t' \in \mathscr{T}$ it holds $f(t) \succ f(t')$ and $g(t) \prec g(t')$. In other words, $f(t) \succ f(t')$ implies $g(t) \succsim g(t')$, i.e., two comonotonic acts have the same kind of monotonicity. Consequently, their intertemporal payoff profile has a similar shape and the mixture of two comonotonic acts does not alter the shape. It therefore seems natural to require that this mixture does not change the original preference ordering between the two acts. This motivates the next axiom, a stronger version of A.5.

A.7. *Comonotonic Independence.* For all pairwise comonotonic acts $f, g, h \in \mathscr{F}$ and all $\alpha \in (0, 1)$, $f \succsim g$ implies $\alpha f + (1-\alpha) h \succsim \alpha g + (1-\alpha) h$.

This axiom plays an important role in the representation theorem for non-additive measures proved in Schmeidler [21], which is reported in the appendix.

3. TIME INVARIANCE

We first study time invariance. Given an act $f \in \mathscr{F}$, define

$$f^k(t) = f(t+k) \qquad \text{for all} \quad t \geqslant 1.$$

A.8*. *Time Invariance.* For every $f \in \mathscr{F}$ and every $k > 0$ it holds that $f \sim f^k$.

According to this axiom, the agent puts zero weight on the consequences obtained on all past and present periods, and full weight to the future periods. In other words, it does not matter what happens in any finite set of points of time.

A crucial implication of Theorem 1 is that Time Invariance must be satisfied by any preference relation that aims to model the undiscounted case as a limit case of discounting as δ goes to 1, when such a limit exists. This is a very important feature of Time Invariance.

In the sequel we will sometimes need a very weak independence axiom related to time invariance. A bit of notation: For a set of points of time

$A \subseteq \mathscr{T}$, and for a pair $p_1, p_2 \in \mathscr{P}$, with $p_2 \succ p_1$, let f_A be the act defined by:

$$f_A(t) = \begin{cases} p_2 & \text{if } \quad t \in A \\ p_1 & \text{if } \quad t \notin A. \end{cases}$$

In other words, f_A is any binary act which gives a higher payoff on A than on A^c. In the notation f_A we omit explicit reference to p_1 and p_2 since what matters is only their relative order, not their specific values.

A.8. *Time Invariance Independence.* For all $A \subseteq \mathscr{T}$ there exist a $f_A \in \mathscr{F}$ such that

$$\tfrac{1}{2} f_A + \tfrac{1}{2} f_{A^c} \sim \tfrac{1}{2} f_A^k + \tfrac{1}{2} f_{A^c}$$

for all $k \geqslant 1$.

This is a very weak notion of independence, and it only involves time invariance. Interestingly, it turns out that A.8 implies A.8* (this is why we have used the star in A.8*).

PROPOSITION 4. *Suppose the preference relation* $\succsim$ *on* $\mathscr{F}$ *satisfies axioms* A.1–A.6 *and* A.8. *Then, it satisfies* A.8*. *The converse is false* (*that is, there exist preference relations* $\succsim$ *on* $\mathscr{F}$ *that satisfy axioms* A.1–A.6 *and* A.8*, *but not* A.8).

An example of a utility functional that satisfies A.8* but not A.8 is

$$\liminf_{t \to \infty} u(f(t)),$$

i.e., the lim inf of instantaneous utilities.

3.1. *A Representation*

As will be proved in the appendix, A.8 implies in terms of multiple weights that all the weights have to be invariant.[5] As there is no a priori reason to exclude any of these invariant weights, we focus on the maximal set of invariant weights. In other words, we focus on the canonical preference relation that satisfies the set of axioms A.1–A.6 and A.8.

For the canonical preference relation we obtain the following representation result, which provides a complete characterization of time invariance for this natural case.

[5] Let $x = \{x_n\}_{n \geqslant 1}$ be a bounded sequence, and τ the shift transformation defined by

$$(\tau(x))_n = x_{n+1}$$

for all $n \geqslant 1$. A weight μ: $2^{\mathscr{T}} \to [0, 1]$ is invariant if $\mu(A) = \mu(\tau(A))$ for all $A \subseteq \mathscr{T}$.

THEOREM 5. *The following two statements are equivalent*:

(i) *The preference relation* $\succsim$ *on* $\mathcal{F}$ *is canonical and satisfies axioms* A.1–A.6 *and* A.8.

(ii) *There exists an affine real valued function u on* $\mathcal{P}$ *such that, for all f and g in* $\mathcal{F}$, *we have*, $f \succsim g$ *if and only if*

$$\lim_{T\to\infty}\left[\inf_{j\geqslant 1}\left(\frac{1}{T}\sum_{t=1}^{T} u(f(j+t))\right)\right] \geqslant \lim_{T\to\infty}\left[\inf_{j\geqslant 1}\left(\frac{1}{T}\sum_{t=1}^{T} u(g(j+t))\right)\right].$$

Finally, the function u is unique up to a positive linear transformation.

3.2. *Time Averages*

For some payoff streams, the above representation reduces to the comparison of the limits of time averages. Indeed, suppose the time average $\lim_{T\to\infty} 1/T\sum_{t=1}^{T} u(f(t))$ converges to some $l \in \mathbb{R}$. This implies $\lim_{T\to\infty} 1/T\sum_{t=1}^{T} u(f(j+t)) = l$ for all $j \geqslant 1$. In particular,

$$\lim_{T\to\infty}\ \inf_{j\geqslant 1}\left[\frac{1}{T}\sum_{t=1}^{T} u(f(j+t))\right] = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} u(f(t))$$

whenever $\lim_{T\to\infty} 1/T\sum_{t=1}^{T} u(f(j+t)) = l$ uniformly in j. This proves the following corollary.[6]

COROLLARY 6. *Suppose that the preference relation* $\succsim$ *on* $\mathcal{F}$ *satisfies axioms* A.1–A.6 *and* A.8. *Then*

$$f \succsim g \qquad \text{if and only if} \quad \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} u(f(t)) \geqslant \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} u(g(t))$$

whenever both $\lim_{T\to\infty} 1/T\sum_{t=1}^{T} u(f(j+t))$ *and* $\lim_{T\to\infty} 1/T\sum_{t=1}^{T} u(g(j+t))$ *converge uniformly in j.*

[6] Recall that it always holds that

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} u(f(t)) = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} u(f(j+t))$$

for all $j \geqslant 1$.

4. PATIENCE

We now move to the analysis of patience. As was mentioned in the introduction, there is a natural definition of complete patience in the finite case: we have complete patience when the ordering of any two acts does not change by taking arbitrary permutations of their respective time indexes.

It is therefore natural to first look at the direct counterpart in the infinite case of this notion, which is so compelling in the finite case. To do so, let Π be the set of all one-to-one and onto maps $\pi: \mathscr{T} \to \mathscr{T}$. Given an act $f \in \mathscr{F}$ and a map $\pi \in \Pi$, define

$$f^{\pi}(t) = f(\pi(t)) \qquad \text{for all} \quad t \geqslant 1.$$

The act f^{π} is obtained from f through a rearrangement of its elements $f(t)$.

A.9. *Naive Patience.* For every $f \in \mathscr{F}$ and every $\pi \in \Pi$, it holds that $f \sim f^{\pi}$.

This axioms states that the agent evaluates the consequences per se, regardless of the points of time where he gets them. This axiom characterizes an agent with "infinite" patience for whom all the points of time, no matter how remote, have the same weight. We use the adjective naive because this is the literal translation in the infinite case of the natural notion of patience for finite sets. As will be seen, axiom A.9 is not at all satisfactory. However, before moving on, we show what kind of representation it entails.

THEOREM 7. *The following two statements are equivalent*:

(i) *The preference relation* $\succsim$ *on* $\mathscr{F}$ *satisfies the axioms* A.1–A.4, A.6, A.8*, *and* A.9.

(ii) *There exists an affine utility* $u: \mathscr{P} \to \mathbb{R}$ *such that, for all* f *and* g *in* $\mathscr{F}$,

$$f \succsim g \qquad \text{if and only if} \quad \liminf_{t \to \infty} u(f(t)) \geqslant \liminf_{t \to \infty} u(g(t)). \tag{5}$$

The function u *is unique up to a positive linear transformation.*

Notice that we use the lim inf of the instantaneous utilities and not of their time averages. It is worth noting that we obtain lim sup instead of lim inf in Eq. (5) if we replace A.6 with the dual axiom in which $f \sim g$ implies $\alpha f + (1-\alpha)\, g \precsim f$ for all $0 < \alpha < 1$. Similar dual versions hold for all the representation results in the paper in which lim inf and min occur.

5. PATIENCE REVISITED

Axiom A.9, which is the literal translation of complete patience from the finite to the infinite horizon, is much stronger than it might seem at a first glance. For example, consider the two acts f and g defined as follows

$$f(t)=\begin{cases} p_2 & \text{if} \quad t \in \{2k\}_{k \geqslant 1} \\ p_1 & \text{if} \quad t \notin \{2k\}_{k \geqslant 1} \end{cases} \qquad \text{and} \qquad g(t)=\begin{cases} p_2 & \text{if} \quad t \in \{4k\}_{k \geqslant 1} \\ p_1 & \text{if} \quad t \notin \{4k\}_{k \geqslant 1} \end{cases} \tag{6}$$

where $p_2 \succ p_1$. It is easy to check that, according to A.9, it holds that $f \sim g$. However, the relative frequency $\lim_{t \to \infty}(|\{2k\}_{k \geqslant 1} \cap \{1, ..., t\}|)/t = 1/2$ with which the agent gets the higher consequence p_2 under act f, is twice than that under g, i.e., $\lim_{t \to \infty}(|\{4k\}_{k \geqslant 1} \cap \{1, ..., t\}|)/t = 1/4$. Nevertheless, by A.9, $f \sim g$ because this axiom does not take into account the relative frequencies with which payoffs come up.

As this example suggests, we must modify A.9 in order to take care of the relative frequencies. In order to do this, we have to introduce densities.

DEFINITION 8. Let $A \subseteq \mathscr{T}$. The lower natural density of A is

$$\delta_*(A) = \lim \inf_{t \to \infty} \frac{|A \cap \{1, ..., t\}|}{t},$$

while the upper natural density is

$$\delta^*(A) = \lim \sup_{t \to \infty} \frac{|A \cap \{1, ..., t\}|}{t}.$$

Finally, a set $A \subseteq \mathscr{T}$ has natural density $\delta(A)$ if $\delta_*(A) = \delta^*(A) = \delta(A)$.

The collection $\mathscr{A}_d = \{A \subseteq \mathscr{T} : \delta(A) \text{ exists}\}$ has the following properties:

1. $A \in \mathscr{A}_d$ implies $A^c \in \mathscr{A}_d$;
2. if $A, B \in \mathscr{A}_d$ and $A \cap B = \emptyset$, then $A \cup B \in \mathscr{A}_d$.

However, $\mathscr{A}_d$ is not an algebra.

5.1. *Patience and Time Invariance*

A form of patience based on frequencies is already implicit in the Time Invariance axiom A.8. To see it, we need the following definition.

DEFINITION 9. The lower Banach density of a set $A \subseteq \mathcal{T}$ is

$$\beta_*(A) = \lim_{T \to \infty} \left[\inf_{j \geqslant 1} \frac{|A \cap \{j, ..., j+T-1\}|}{T} \right],$$

while the upper Banach density is

$$\beta^*(A) = \lim_{T \to \infty} \left[\sup_{j \geqslant 1} \frac{|A \cap \{j, ..., j+T-1\}|}{T} \right].$$

Finally, a set $A \subseteq \mathcal{T}$ has Banach density $\beta(A)$ if $\beta_*(A) = \beta^*(A) = \beta(A)$.

It holds that

$$\beta_*(A) \leqslant \delta_*(A) \leqslant \delta^*(A) \leqslant \beta^*(A)$$

for all $A \subseteq \mathcal{T}$. In particular, $\beta(A) = \delta(A)$ whenever $\beta(A)$ exists.

$\mathcal{A}_b$ is the collection of all sets that have a Banach density, i.e. $\mathcal{A}_b = \{A \subseteq \mathcal{T} : \beta(A) \text{ exists}\}$, and $\mathcal{F}_b$ is the set of all acts $f \in \mathcal{F}$ such that $\{t: f(t) \succsim p\} \in \mathcal{A}_b$ for all $p \in \mathcal{P}$.

Given an act $f \in \mathcal{F}_b$, we are interested in permutations $\pi \in \Pi$ which are invariant under Banach densities, i.e., $\beta(\{t: f(t) \succsim p\}) = \beta(\{t: f^\pi(t) \succsim p\})$ for all $p \in \mathcal{P}$. We denote them by Π^f_b.

We can now state the notion of patience that comes with Time Invariance.

THEOREM 10. *Suppose the preference relation* $\succsim$ *on* $\mathcal{F}$ *satisfies axioms* A.1–A.6, *and* A.8. *Then, for every* $f \in \mathcal{F}_b$ *and every* $\pi \in \Pi^f_b$, *it holds that* $f \sim f^\pi$.

6. PATIENCE AND DENSITIES

We have seen how Time Invariance implies a form of patience based on Banach densities. A natural step is to replace Banach densities with natural densities in the notion of patience implied by Time Invariance.

To do this, we need a bit of notation. Denote by $\mathcal{F}_d$ be the set of all acts $f \in \mathcal{F}$ such that $\{t: f(t) \succsim p\} \in \mathcal{A}_d$ for all $p \in \mathcal{P}$. For a given $f \in \mathcal{F}_d$, denote by Π^f_d the set of all permutations $\pi \in \Pi$ invariant under natural densities, that is, $\delta(\{t: f(t) \succsim p\}) = \delta(\{t: f^\pi(t) \succsim p\})$ for all $p \in \mathcal{P}$.

We are now in a position to state the new patience axiom.

A.10. *Cardinal Patience.* For every $f \in \mathcal{F}_d$ and every $\pi \in \Pi^f_d$, it holds that $f \sim f^\pi$.

Of course, A.9 implies A.10, while the converse is not true. Moreover, if we replace $\mathscr{F}_d$ and Π_d^f with, respectively, their subsets $\mathscr{F}_b$ and Π_b^f, by Theorem 10 the axiom becomes a consequence of Time Invariance.

The following simple example further illustrates the nature of this axiom.

EXAMPLE. Let $f \in \mathscr{F}$ be an act such that

$$f(t) = \begin{cases} p_1 & \text{if} \quad t \text{ is even,} \\ p_2 & \text{if} \quad t \text{ is odd,} \end{cases}$$

with $p_1 \succ p_2$. Clearly, $f \in \mathscr{F}_d$. Let π be the permutation defined as follows: $\pi(2t-1) = 2t$ and $\pi(2t) = 2t-1$ for $t \geqslant 1$. Then

$$f^{\pi}(t) = \begin{cases} p_1 & \text{if} \quad t \text{ is odd,} \\ p_2 & \text{if} \quad t \text{ is even.} \end{cases}$$

This permutation preserves the upper sets' densities, i.e., $\pi \in \Pi_d^f$. In fact

$$\begin{aligned} &\delta(\{t\colon f(t) \succsim p\}) = \delta(\{t\colon f^{\pi}(t) \succsim p\}) = 1 && \text{if} \quad p \precsim p_2 \\ &\delta(\{t\colon f(t) \succsim p\}) = \delta(\{t\colon f^{\pi}(t) \succsim p\}) = \tfrac{1}{2} && \text{if} \quad p_2 \prec p \precsim p_1 \\ &\delta(\{t\colon f(t) \succsim p\}) = \delta(\{t\colon f^{\pi}(t) \succsim p\}) = 0 && \text{if} \quad p \succ p_1 \end{aligned}$$

6.1. *A Technical Condition*

For the representation result we need a technical condition, called regularity. It is introduced in this subsection, which can be skipped at a first reading.

For any $\hat{p} \in \mathscr{P}$, let $[\hat{p}] = \{p\colon p \sim \hat{p}\}$, i.e., $[\hat{p}]$ is the indifference class containing $\hat{p}$. Similarly, for $A \subseteq \mathscr{P}$ set $[A] = \{[p]\colon p \in A\}$.

DEFINITION 11. For a given $f \in \mathscr{F}$, let $A_f = \{p\colon \{t\colon f(t) \succsim p\} \notin \mathscr{A}_\delta\}$. We denote by $\mathscr{F}_\delta$ the set of all acts $f \in \mathscr{F}$ such that the set $[A_f]$ is at most countable.

In other words, if $u\colon \mathscr{P} \to \mathbb{R}$ is an affine utility associated with $\succsim$, $\mathscr{F}_\delta$ is the set of all acts f such that $\{t\colon u(f(t)) \geqslant \alpha\} \notin \mathscr{A}_d$ for an at most countable set of $\alpha \in \mathbb{R}$. Loosely speaking, $\mathscr{F}_\delta$ is the set of acts "measurable" w.r.t. $\mathscr{A}_\delta$. Notice that $\mathscr{F}_d \subseteq \mathscr{F}_\delta$.

As will be seen in the appendix, if $f \in \mathscr{F}_\delta$, and $p_0 \in \mathscr{P}$ is such that $u(p_0) = 0$, then $f \sim p_0$ whenever

$$\sum_{t=1}^{\infty} \frac{u(f(t))}{t}$$

converges (i.e., $\lim_{T\to\infty}\sum_{t=1}^{T} u(f(t))/t$) exists[7] and is finite). For our next representation theorem, a similar property must hold for all acts in $\mathcal{F}$. To this end we introduce the following technical condition, called regularity.

DEFINITION 12. Let $\succsim$ be a preference ordering that satisfies A.1–A.6, and $u: \mathcal{P}\to\mathbb{R}$ an affine utility provided by Theorem 2. Let $p_0\in\mathcal{P}$ be such that $u(p_0)=0$. We say that $\succsim$ is regular if, for all acts $f\in\mathcal{F}/\mathcal{F}_\delta$ such that

$$\sum_{t=1}^{\infty}\frac{u(f(t))}{t}$$

converges and such that $\inf_{t\geqslant 1} u(f(t))<0<\sup_{t\geqslant 1} u(f(t))$, it holds that $f\sim p_0$.[8]

We conclude by presenting an interesting class of acts in $\mathcal{F}_\delta$.

PROPOSITION 13. *Let $\succsim$ be a preference ordering that satisfies* A.1–A.6, *and $u: \mathcal{P}\to\mathbb{R}$ an affine utility provided by Theorem* 2. *Let $f\in\mathcal{F}$. If there exists an $l\in\mathbb{R}$ such that*

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}|u(f(t))-l|=0, \tag{7}$$

then $f\in\mathcal{F}_\delta$.

Remarks. (i) It is easy to check that (7) holds whenever $\lim_{t\to\infty} u(f(t))$ exists. Hence, by Proposition 13, in this special case $f\in\mathcal{F}_\delta$. (ii) If $u(f(t))\geqslant 0$ for all $t\geqslant 1$ and $\sum_{t=1}^{\infty} u(f(t))/t$ converges, then, by Kronecker's Lemma

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}u(f(t))=\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}|u(f(t))|=0.$$

Hence, $f\in\mathcal{F}_\delta$ by Proposition 13. This is why in Definition 12 we require that $\inf_{t\geqslant 1} u(f(t))<0<\sup_{t\geqslant 1} u(f(t))$. Otherwise, as just proved, f would automatically be in $\mathcal{F}_\delta$.

[7] By Kronecker's Lemma, the convergence of $\sum_{t=1}^{\infty} u(f(t))/t$ implies $\lim_{T\to\infty} 1/T \sum_{t=1}^{T} u(f(t))=0$. However, the converse is false. For example, let $u(f(t))=1/lgt$. Then $\sum_{t=1}^{\infty} 1/tlgt=\infty$, but $\lim_{T\to\infty} 1/T\sum_{t=1}^{T} 1/lgt=0$.

[8] It is important to note that, as shown in the last footnote, the convergence of $\sum_{t=1}^{\infty} u(f(t))/t$ is a much stronger requirement on f than $\lim_{T\to\infty} 1/T\sum_{t=1}^{T} u(f(t))=0$. Therefore, regularity is a much weaker condition than assuming $f\sim p_0$ when $\lim_{T\to\infty} 1/T\sum_{t=1}^{T} u(f(t))=0$.

6.2. *Representation Theorem*

We can now state the representation result.

THEOREM 14. *The following two statements are equivalent*:

(i) *The preference relation* $\succsim$ *on* $\mathcal{F}$ *is canonical, satisfies axioms* A.1–A.6, A.8–A.10, *and is regular.*

(ii) *There exists an affine utility* $u: \mathcal{P} \to \mathbb{R}$ *such that, for all* f *and* g *in* $\mathcal{F}$, *we have* $f \succsim g$ *if and only if*

$$\lim_{\varepsilon \to 0} \left[\lim_{T \to \infty} \inf \frac{1}{\varepsilon T} \sum_{t=(1-\varepsilon)T}^{T} u(f(t)) \right] \geqslant \lim_{\varepsilon \to 0} \left[\lim_{T \to \infty} \inf \frac{1}{\varepsilon T} \sum_{t=(1-\varepsilon)T}^{T} u(g(t)) \right].$$

Finally, the function u *is unique up to a positive linear transformation.*

Given its importance in this work, we now give a name to the functional that comes up in Theorem 14.

DEFINITION 15. Let $f \in \mathcal{F}$ and $u: \mathcal{P} \to \mathbb{R}$ an affine utility function. The functional $P: \mathcal{F} \to \mathbb{R}$ defined by

$$P(f) = \lim_{\varepsilon \to 0} \left[\lim_{T \to \infty} \inf \frac{1}{\varepsilon T} \sum_{t=(1-\varepsilon)T}^{T} u(f(t)) \right]$$

is called the Polya Index.

Remark. We call P the Polya Index because for the special case of acts f such that $u(f(t)) \in \{0, 1\}$ for all $t \geqslant 1$, $P(f)$ is equal to the Polya minimal density of the set $\{t: u(f(t)) = 1\}$. These densities have been introduced by Polya [18, pp. 556–568].

Notice that

$$\lim_{\varepsilon \to 0} \left[\lim_{T \to \infty} \inf \frac{1}{\varepsilon T} \sum_{t=(1-\varepsilon)T}^{T} u(f(t)) \right] = \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} u(f(t)) \tag{8}$$

whenever the limit on the r.h.s. exists. Therefore, the Polya Index is an extension of the limit of time average criterion to streams that do not have well defined time average limits.

Using the equality (8), we get the following interesting consequence of Theorem 14. It provides a behavioral underpinning for the use of time averages, provided they exist.

COROLLARY 16. *Suppose the preference relation* $\succsim$ *on* $\mathscr{F}$ *satisfies axioms* A.1–A.6, A.8, *and* A.10, *and regularity. Then there exists an affine utility* $u: \mathscr{P} \to \mathbb{R}$ *such that, for all f and g in* $\mathscr{F}$,

$$f \succsim g \qquad \text{if and only if} \quad \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} u(f(t)) \geqslant \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} u(g(t)),$$

provided the two limits exist. The function u is unique up to a positive linear transformation.

7. THE POLYA INDEX

We now provide two further characterizations of the Polya Index.

7.1. *Polya Index as an Inner Approximation*

We first characterize the Polya Index as an inner approximation. For a given $f \in \mathscr{F}$, let $\mathscr{F}_a = \{f \in \mathscr{F}: \lim_{T\to\infty} 1/T \sum_{t=1}^{T} u(f(t)) \text{ exists}\}$.

THEOREM 17. *Let* $\succsim$ *be the preference ordering of Theorem* 14, *and* $u: \mathscr{P} \to \mathbb{R}$ *its corresponding affine utility. If* $u(\mathscr{P}) = \mathbb{R}$ (*i.e., the range of u is* $\mathbb{R}$), *for all* $f \in \mathscr{F}$ *it holds that*

$$P(f) = \sup\left\{\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} u(g(t)) \colon g \in \mathscr{F}_a \text{ and } f \succsim g\right\}.$$

This characterization shows that the Polya Index can be viewed as an inner approximation taken over the less preferred acts that have well defined time average limits. In other words, comparing two acts through the Polya Index is equivalent to comparing them by taking the supremum over the set of all less preferred acts which have well defined time average limits.

The proof of Theorem 17 rests on the following Lemma, which is interesting in itself because it provides the Polya Index with further analytical tractability.

LEMMA 18. *Let* $u: \mathscr{P} \to \mathbb{R}$ *be the affine utility provided by Theorem* 14. *If* $u(\mathscr{P}) = \mathbb{R}$, *then, for all* $f \in \mathscr{F}$, *it holds that*

$$P(f) = \sup\left\{\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} u(g(t)) \colon g \in \mathscr{F}_a \text{ and } u(f(t)) \geqslant u(g(t)) \text{ for all } t \geqslant 1\right\}.$$

7.2. *Polya Index and Weights*

Next we characterize the Polya Index in terms of sets of weights. Let $\tau: \mathcal{T} \to \mathcal{T}$ be the shift transformation defined by

$$\tau(t) = t + 1.$$

We denote by $\mathcal{N}_d$ the set of all normalized finitely additive measures μ on $2^{\mathcal{T}}$ such that

1. $\mu(A) = \delta(A)$ for all $A \in \mathcal{A}_d$.
2. $\mu(A) = \mu(\tau(A))$ for all $A \subseteq \mathcal{T}$, i.e., μ is invariant w.r.t. τ.

Besides being invariant, the weights in $\mathcal{N}_d$ coincide with the natural density when it exists. They are the natural weights for complete patience (see the discussion below), and the next result shows that the Polya Index can be justified through them.

THEOREM 19. *Let $u: \mathcal{P} \to \mathbb{R}$ be the affine utility provided by Theorem* 14. *Then there exists a unique weak*-compact and convex set $\mathcal{C}_p \subseteq \mathcal{N}_d$ such that*

$$P(f) = \min_{\mu \in \mathcal{C}_p} \int u(f(t))\, d\mu$$

for all $f \in \mathcal{F}$.

Notice that up to a mild technical condition (i.e., regularity), the set $\mathcal{C}_p$ coincides with $\mathcal{N}_d$. The set $\mathcal{N}_d$ consists of all invariant weights that coincide with the natural density δ whenever it exists. The density δ would be the natural weight for complete patience, but it fails to exist for many sets and cannot be used as a weight. Lacking this "ideal" weight, we can think of an agent using the Polya Index as replacing it with the set of all invariant weights that coincide with δ whenever it exists, i.e., with the set $\mathcal{N}_d$.[9]

This interpretation of the Polya Index was already outlined in the introduction and it is important because it provides the Polya Index with a behavioral foundation. Interestingly, this interpretation is completely different from that used in the multiple priors model of Gilboa and Schmeidler, whose aim is to model vagueness in subjective beliefs.

[9] It is important to note that the fact that the agent uses all the weights compatible with δ is the reason why in Theorem 14 we have a canonical ordering.

8. LIM INF OF TIME AVERAGES

In the last Theorem we have seen how the Polya Index can be represented in terms of weights in $\mathcal{N}_d$. We now show that the same is true for the lim inf of time averages.

THEOREM 20. *Let $\succsim$ be a preference relation on $\mathcal{F}$ that satisfies axioms* A.1–A.6, A.8, *and* A.10, *and regularity. Then there exists an affine utility $u: \mathcal{P} \to \mathbb{R}$ and a unique non-empty weak*-compact and convex set $\mathcal{C}_l \subseteq \mathcal{N}_d$ of weights on $2^{\mathcal{T}}$ such that, for all f and g in $\mathcal{F}$,*

$$f \succsim g \qquad \text{if and only if} \quad \min_{\mu \in \mathcal{C}_l} \int_{\mathcal{T}} u(f(t))\, d\mu \geqslant \min_{\mu \in \mathcal{C}_l} \int_{\mathcal{T}} u(g(t))\, d\mu$$

and

$$\min_{\mu \in \mathcal{C}_l} \int_{\mathcal{T}} u(f(t))\, d\mu = \lim_{T \to \infty} \inf \frac{1}{T} \sum_{t=1}^{T} u(f(t)).$$

Finally, the function u is unique up to a positive linear transformation.

This result shows that we can justify $\liminf_{T \to \infty} 1/T \sum_{t=1}^{T} u(f(t))$ through a set of weights $\mathcal{C}_l$ contained in $\mathcal{N}_d$. In particular, $\mathcal{C}_l \subsetneqq \mathcal{C}_p$, as the following result shows.

PROPOSITION 21. $\mathcal{C}_l \subsetneqq \mathcal{C}_p$.

The set $\mathcal{C}_l$ is therefore strictly smaller than $\mathcal{C}_p$. Some weights in $\mathcal{C}_p$ have to be eliminated in order to represent $\liminf_{T \to \infty} 1/T \sum_{t=1}^{T} u(f(t))$. However, it is not clear what conditions on $\succsim$, on top of A.1–A.6, A.8, A.10 and regularity, would lead to this elimination. In other words, it is not clear which further axiom to impose on $\succsim$ in order to move from $\mathcal{C}_p$ to the smaller set $\mathcal{C}_l$.

Therefore, the behavioral underpinning of the liminf criterion is less transparent than that of the Polya Index. Nevertheless, Theorem 20 is interesting because it sheds light on the liminf criterion by providing a representation where weights and utilities are clearly separated.

9. PROOFS AND RELATED ANALYSIS

9.1. *Glossary of notation*

$\mathcal{T}$ set of points of time (sect. 2)
$\mathcal{X}$ set of consequences (sect. 2)

$\mathscr{P}$	set of all probability distributions with finite support (sect. 2)
$\mathscr{F}$	set of all bounded acts (sect. 2)
$\mathscr{C}$	set of weights (sect. 2)
$\mathscr{F}_b$	set of all bounded acts that preserve the upper sets' Banach densities (sect. 5.1)
$\mathscr{F}_d$	set of all bounded acts that preserve the upper sets' natural densities (sect. 6)
$\mathscr{F}_a$	set of all bounded acts that have a well defined limit of time averages (sect. 7.1)
f^k	shift (sect. 3)
β	Banach density (sect. 5.1)
δ	natural density (sect. 5)
Π	set of all one-to-one and onto permutations on $\mathscr{T}$ (sect. 4)
Π_b^f	set of all permutations invariant under Banach densities (sect. 5.1)
Π_d^f	set of all permutations invariant under natural densities (sect. 6)
$\mathscr{A}_b$	collection of all sets that have a Banach density (sect. 5.1)
$\mathscr{A}_d$	collection of all sets that have a natural density (sect. 6)
$\mathscr{N}_d$	invariant weights that coincide with the natural density when it exists (sect. 7.2)

9.2. *Example*

In the introduction we presented an example of a sequence whose limit average does not exist. We now give a more general result that includes the example as a special case. It shows how far apart can be the upper and lower bounds of the partial sums $1/T\sum_{t=1}^{T} x_t$.

PROPOSITION 22. *For any positive integer N there exists a sequence x such that $x_t \in \{0, 1\}$ for all $t \geqslant 1$, and*

$$\lim_{T\to\infty}\inf \frac{1}{T}\sum_{t=1}^{T} x_t = \frac{1}{1+N} \qquad \text{and} \qquad \lim_{T\to\infty}\sup \frac{1}{T}\sum_{t=1}^{T} x_t = \frac{N}{1+N}.$$

Notice that the difference

$$\lim_{T\to\infty}\sup \frac{1}{T}\sum_{t=1}^{T} x_t - \lim_{T\to\infty}\inf \frac{1}{T}\sum_{t=1}^{T} x_t = \frac{N-1}{N+1}$$

tends to 1 as N gets larger and larger. As x takes on only the values 0 and 1, this means that there are sequences for which the lim inf and lim sup are very far away.

Proof. Given N, let x be the sequence whose first N elements are 1, the second N^2 elements are 0, the third N^3 elements are 1, and so forth. It is easy to check that

$$\limsup_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} x_t = \lim_{T\to\infty} \frac{\sum_{t=0}^{T} N^{2t+1}}{\sum_{t=0}^{T} N^{2t+1} + \sum_{t=0}^{T} N^{2t} - 1} = \frac{N}{N+1}$$

$$\liminf_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} x_t = \lim_{T\to\infty} \frac{\sum_{t=0}^{T} N^{2t+1}}{\sum_{t=0}^{T} N^{2t+1} + \sum_{t=0}^{T} N^{2t} - 1 + N^{2T+2}}$$
$$= \frac{1}{N+1}$$

as wanted. ▮

9.3. *Proposition* 4 *and Theorem* 5

Let l^∞ be the set of real sequences $x = \{x_n\}_{n\geqslant 1}$ bounded w.r.t. the sup-norm $\|x\| = \sup_n |x_n|$, and let $\tau\colon l^\infty \to l^\infty$ be the shift transformation defined by

$$(\tau(x))_n = x_{n+1}.$$

A linear functional $L\colon l^\infty \to \mathbb{R}$ is a Banach limit if it satisfies the following conditions.

1. $L(x) \geqslant 0$ if $x \geqslant 0$.
2. $L(\tau(x)) = L(x)$ for all $x \in l^\infty$.
3. $L(1) = 1$, where 1 denotes the sequence $\{1, ..., 1, ...\}$.

We denote by $\mathscr{L}$ the set of all Banach limits on l^∞.

A finitely additive measure $\mu\colon 2^{\mathscr{T}} \to [0, 1]$ is invariant if $\mu(A) = \mu(\tau(A))$ for all $A \subseteq \mathscr{T}$. We denote by $\mathscr{N}$ the set of all finitely additive invariant measures. Let $F\colon \mathscr{T} \to \mathbb{R}$ be a sequence in l^∞. Set $\mathscr{M} = \{\int F\, d\mu\colon \mu \in \mathscr{N}\}$. We have

LEMMA 23. $\mathscr{L} = \mathscr{M}$.

Proof. We first notice that, by Sucheston [22], $\lim_n \sup_{j\geqslant 1} [1/n \sum_{j=1}^{n} F(i+j)]$ exists for all $F \in l^\infty$. We now prove that $\mathscr{M} \subseteq \mathscr{L}$. For all $A \subseteq \mathscr{T}$, let $\tau^{-1}(A) = \{t \in \mathscr{T}\colon \tau(t) \in A\}$. We have $\tau^{-1}(\tau(A)) = A$ for all $A \subseteq \mathscr{T}$, $\tau(\tau^{-1}(A)) = A$ if $\{1\} \subsetneq A$, and $\tau(\tau^{-1}(A)) \cup \{1\} = A$ if $\{1\} \subseteq A$ (notice that $\tau(\tau^{-1}(A)) \cap \{1\} = \varnothing$ in this last case). For every $\mu \in \mathscr{N}$ it holds that $\mu(A) = \mu(\tau(\tau^{-1}(A))) = \mu(\tau^{-1}(A))$ because $\mu(\{1\}) = 0$. We have $1_A(\tau(t)) = 1_{\tau^{-1}(A)}(t)$, so that

$$\int 1_A(\tau(t))\, d\mu = \int 1_A(t)\, d\mu \qquad \text{for all} \quad A \subseteq \mathscr{T}. \tag{9}$$

Let $F \in l^\infty$ be a simple function. Using (9), it is easy to check that for every invariant measure in $\mathcal{N}$ it holds that

$$\int F\,d\mu = \int \left[\frac{1}{n}\sum_{i=1}^{n} F(i+j)\right] d\mu \qquad \text{for all} \quad j \geqslant 1.$$

Therefore, $\int F\,d\mu \leqslant \sup_{j\geqslant 1} 1/n \sum_{j=1}^{n} F(i+j)$, so that

$$\int F\,d\mu \leqslant \lim_n \sup_{j\geqslant 1} \left[\frac{1}{n}\sum_{i=1}^{n} F(i+j)\right]. \tag{10}$$

Let $F \in l^\infty$. There exists a sequence of simple functions F_k that converges uniformly to F. We can write

$$\begin{aligned}\lim_n \sup_{j\geqslant 1} \left[\frac{1}{n}\sum_{i=1}^{n} F(i+j)\right] &= \lim_n \sup_{j\geqslant 1} \left[\frac{1}{n}\sum_{i=1}^{n} \lim_k F_k(i+j)\right]\\ &= \lim_n \sup_{j\geqslant 1} \lim_k \left[\frac{1}{n}\sum_{i=1}^{n} F_k(i+j)\right].\end{aligned}$$

For every $\varepsilon > 0$ there exists $K_\varepsilon > 0$ such that $|F(i) - F_k(i)| < \varepsilon$ for all $k \geqslant K_\varepsilon$ and all $i \geqslant 1$. Hence

$$\begin{aligned}&\left|\frac{1}{n}\sum_{i=1}^{n} F_k(j+i) - \frac{1}{n}\sum_{i=1}^{n} F(j+i)\right|\\ &\qquad \leqslant \frac{1}{n}\sum_{i=1}^{n} |F_k(j+i) - F(j+i)| = \varepsilon\end{aligned}$$

for all $k \geqslant K_\varepsilon$ and all $j \geqslant 1$. This implies that

$$\lim_n \sup_{j\geqslant 1} \lim_k \left[\frac{1}{n}\sum_{i=1}^{n} F_k(i+j)\right] = \lim_n \lim_k \sup_{j\geqslant 1} \left[\frac{1}{n}\sum_{i=1}^{n} F_k(i+j)\right].$$

On the other hand, assume $\sup_{j\geqslant 1} 1/n \sum_{i=1}^{n} F(i+j) \geqslant \sup_{j\geqslant 1} 1/n \sum_{i=1}^{n} F_k(i+j)$. Then

$$\begin{aligned}&\sup_{j\geqslant 1} \frac{1}{n}\sum_{i=1}^{n} F(i+j) - \sup_{j\geqslant 1} \frac{1}{n}\sum_{i=1}^{n} F_k(i+j)\\ &\quad = \sup_{j\geqslant 1} \frac{1}{n}\sum_{i=1}^{n} [F(i+j) - F_k(i+j) + F_k(i+j)] - \sup_{j\geqslant 1} \frac{1}{n}\sum_{i=1}^{n} F_k(i+j)\\ &\quad \leqslant \sup_{j\geqslant 1} \frac{1}{n}\sum_{i=1}^{n} [F(i+j) - F_k(i+j)] < \varepsilon\end{aligned}$$

for all $k \geqslant K_\varepsilon$ and all $n \geqslant 1$. The same holds if $\sup_{j \geqslant 1} 1/n \sum_{i=1}^n F(i+j) < \sup_{j \geqslant 1} 1/n \sum_{i=1}^n F_k(i+j)$, so that

$$\left| \sup_{j \geqslant 1} \frac{1}{n} \sum_{i=1}^n F(i+j) - \sup_{j \geqslant 1} \frac{1}{n} \sum_{i=1}^n F_k(i+j) \right| < \varepsilon$$

for all $k \geqslant K_\varepsilon$ and all $n \geqslant 1$. In turn, this implies

$$\lim_n \lim_k \sup_{j \geqslant 1} \left[\frac{1}{n} \sum_{i=1}^n F_k(i+j) \right] = \lim_k \lim_n \sup_{j \geqslant 1} \left[\frac{1}{n} \sum_{i=1}^n F_k(i+j) \right]$$

and we conclude that

$$\lim_n \sup_{j \geqslant 1} \left[\frac{1}{n} \sum_{i=1}^n F(i+j) \right] = \lim_k \lim_n \sup_{j \geqslant 1} \left[\frac{1}{n} \sum_{i=1}^n F_k(i+j) \right]. \tag{11}$$

Putting together (10) and (11), we get

$$\int F \, d\mu = \lim_k \int F_k \, d\mu \geqslant \lim_k \lim_n \sup_{j \geqslant 1} \left[\frac{1}{n} \sum_{i=1}^n F_k(i+j) \right]$$
$$= \lim_n \sup_{j \geqslant 1} \left[\frac{1}{n} \sum_{i=1}^n F(i+j) \right].$$

Sucheston [22] proves that

$$\lim_n \sup_{j \geqslant 1} \left[\frac{1}{n} \sum_{i=1}^n F(i+j) \right] = \sup_{m_1, \ldots, m_s} \left[\lim_{j \to \infty} \inf \frac{1}{s} \sum_{k=1}^s F(m_k + j) \right].$$

As $\int F \, d\mu$ is a linear functional on l^∞, by Banach [2]

$$\int F \, d\mu \geqslant \sup_{m_1, \ldots, m_s} \left[\lim_{j \to \infty} \inf \frac{1}{s} \sum_{k=1}^s F(m_k + j) \right]$$

implies that $\int F \, d\mu$ is a Banach limit. This proves that $\mathscr{M} \subseteq \mathscr{L}$.

As to the converse, for any $L \in \mathscr{L}$ there exists a finitely additive measure on $2^{\mathscr{T}}$ such that $L(F) = \int F \, d\mu$ for all $F \in l^\infty$ (cf. Dunford and Schwartz [6, p. 258]). But $L(1_A) = L(1_{\tau(A)}(\tau(t))) = L(1_{\tau(A)})$, so that $\mu(A) = \mu(\tau(A))$, i.e. $\mu \in \mathscr{N}$. This implies $\mathscr{L} \subseteq \mathscr{M}$. ∎

LEMMA 24. *Suppose the preference relation* $\succsim$ *satisfies axioms* A.1–A.6, *and let* $\mathscr{C}$ *be the convex and compact set provided by Theorem* 2. *Then* $\mathscr{C} \subseteq \mathscr{N}$ *if and only if* $\succsim$ *satisfies* A.8.

Proof. "If" part: A.8 implies

$$\min_{\mu \in \mathscr{C}} \int \left(\tfrac{1}{2}u(f_A) + \tfrac{1}{2}u(f_{A^c})\right) d\mu = \min_{\mu \in \mathscr{C}} \int \left(u(\tfrac{1}{2} f_A + \tfrac{1}{2} f_{A^c})\right) d\mu$$

$$= \min_{\mu \in \mathscr{C}} \int \left(u(\tfrac{1}{2} f^1_A + \frac{1}{2} f_{A^c})\right) d\mu$$

$$= \min_{\mu \in \mathscr{C}} \int \left(\tfrac{1}{2}u(f^1_A) + \tfrac{1}{2} u(f_{A^c})\right) d\mu$$

that is

$$\min_{\mu \in \mathscr{C}} \int (u(f_A) + u(f_{A^c}))\, d\mu = \min_{\mu \in \mathscr{C}} \int (u(f^1_A) + u(f_{A^c}))\, d\mu. \tag{12}$$

Without loss of generality, set $u(p_1) = 0$ and $u(p_2) = 1$. Then (12) becomes:

$$\min_{\mu \in \mathscr{C}} (\mu(A) + \mu(A^c)) = \min_{\mu \in \mathscr{C}} (\mu(\tau^{-1}(A)) + \mu(A^c))$$

so that

$$\min_{\mu \in \mathscr{C}} (\mu(\tau^{-1}(A)) + \mu(A^c)) = 1,$$

that is

$$\min_{\mu \in \mathscr{C}} (\mu(\tau^{-1}(A)) - \mu(A)) = 0.$$

This implies

$$\mu(\tau^{-1}(A)) \geqslant \mu(A)$$

for all $\mu \in \mathscr{C}$.

Now, let us consider f_{A^c}. Proceeding as above we get

$$\mu(\tau^{-1}(A^c)) \geqslant \mu(A^c)$$

for all $\mu \in \mathscr{C}$. As $\mu(\tau^{-1}(A^c)) = \mu((\tau^{-1}(A))^c)$, we have

$$\mu((\tau^{-1}(A))^c) \geqslant \mu(A^c)$$

for all $\mu \in \mathscr{C}$, so that $\mu(A) \geqslant \mu(\tau^{-1}(A))$ for all $\mu \in \mathscr{C}$. We conclude that

$$\mu(A) = \mu(\tau^{-1}(A))$$

for all $\mu \in \mathscr{C}$, so that $\mathscr{C} \subseteq \mathscr{N}$.

As to the "only if" part, assume $\mathscr{C} \subseteq \mathscr{N}$. Then, by Lemma 23, $\int u(f)\, d\mu = \int u(f^k)\, d\mu$ for all $\mu \in \mathscr{C}$ and $k \geqslant 1$, so that

$$\int \left(\tfrac{1}{2} u(f_A) + \tfrac{1}{2} u(f_{A^c})\right) d\mu = \int \left(\tfrac{1}{2} u(f_A^k) + \tfrac{1}{2} u(f_{A^c})\right) d\mu$$

for all $\mu \in \mathscr{C}$ and $k \geqslant 1$. This implies

$$\min_{\mu \in \mathscr{C}} \int \left(\tfrac{1}{2} u(f_A) + \tfrac{1}{2} u(f_{A^c})\right) d\mu = \min_{\mu \in \mathscr{C}} \int \left(\tfrac{1}{2} u(f_A^k) + \tfrac{1}{2} u(f_{A^c})\right) d\mu$$

so that

$$\tfrac{1}{2} f_A + \tfrac{1}{2} f_{A^c} \sim \tfrac{1}{2} f_A^k + \tfrac{1}{2} f_{A^c}$$

for all $k \geqslant 1$. ∎

LEMMA 25. *Suppose the preference relation $\succsim$ satisfies axioms* A.1–A.4, A.8 *and* A.6, *and let $\mathscr{C}$ be the convex and compact set provided by Theorem* 2. *Then $\mathscr{C} = \mathscr{N}$ if and only if $\succsim$ is canonical.*

Proof. Suppose $\succsim$ satisfies A.9. Define a preference $\succsim_{\mathscr{N}}$ as follows

$$f \succsim_{\mathscr{N}} g \qquad \text{if and only if} \quad \min_{\mu \in \mathscr{N}} \int u(f)\, d\mu \geqslant \min_{\mu \in \mathscr{N}} \int u(g)\, d\mu$$

where u is the utility function derived from $\succsim$ using Theorem 2. It is easy to check that $\succsim_{\mathscr{N}} \in \mathscr{I}$. As $\succsim$ is canonical, $\mathscr{N} \subseteq \mathscr{C}$. However, $\mathscr{C} \subseteq \mathscr{N}$ by Lemma 24, so that $\mathscr{N} = \mathscr{C}$. The converse is obvious. ∎

9.3.1. *Proof of Proposition* 4

By Lemma 24, $\mathscr{C} \subseteq \mathscr{N}$, so that, by Lemma 23, $\min_{\mu \in \mathscr{C}} \int u(f)\, d\mu = \min_{\mu \in \mathscr{C}} \int u(f^k)\, d\mu$. In turn, this implies $f \sim f^k$.

We now show that the converse is false, i.e., there exists a preference relation $\succsim$ that satisfies the axioms A.1–A.6 and A.8*, but not A.8. Let $u: \mathscr{P} \to \mathbb{R}$ be a von Neumann-Morgenstern utility function on $\mathscr{P}$. Define $\succsim$ as follows

$$f \succsim g \qquad \text{if and only if} \quad \lim_{t \to \infty} \inf u(f(t)) \geqslant \lim_{t \to \infty} \inf u(g(t)) \tag{13}$$

for all $f, g \in \mathscr{F}$. This ordering $\succsim$ satisfies axioms A.1–A.6 and A.8. We now show that it does not satisfy A.8*. Let A be the set of odd integers. Set

$$f_A(t) = \begin{cases} p_2 & \text{if} \quad t \in A \\ p_1 & \text{if} \quad t \notin A \end{cases}$$

with $p_1, p_2 \in \mathscr{P}$, $p_2 \succ p_1$. W.l.o.g., set $u(p_2) = 1$ and $u(p_1) = 0$. Then

$$u(f_A) = \{1, 0, 1, 0, ...\}$$

$$u(f_{A^c}) = \{0, 1, 0, 1, ...\}$$

$$u(f_A^1) = \{0, 1, 0, 1, ...\}$$

so that

$$\lim_{t \to \infty} \inf u(\tfrac{1}{2} f_A + \tfrac{1}{2} f_{A^c}) = \lim_{t \to \infty} \inf (\tfrac{1}{2} u(f_A) + \tfrac{1}{2} u(f_{A^c})) = \tfrac{1}{2}$$

while

$$\lim_{t \to \infty} \inf u(\tfrac{1}{2} f_A^1 + \tfrac{1}{2} f_{A^c}) = \lim_{t \to \infty} \inf u(f_{A^c}) = 0.$$

Therefore, by (13)

$$\tfrac{1}{2} f_A + \tfrac{1}{2} f_{A^c} \succ \tfrac{1}{2} f_A^1 + \tfrac{1}{2} f_{A^c}$$

which violates A.8. ▌

9.3.2. *Proof of Theorem* 5

By Lemma 24,

$$f \succsim_{\mathscr{N}} g \qquad \text{if and only if} \quad \min_{\mu \in \mathscr{N}} \int u(f)\, d\mu \geqslant \min_{\mu \in \mathscr{N}} \int u(g)\, d\mu.$$

By Lemma 23, $\min_{\mu \in \mathscr{N}} \int u(f)\, d\mu = \min_{L \in \mathscr{L}} L(u(f))$. By Lorentz [14]

$$\min_{L \in \mathscr{L}} L(u(f)) = \sup_{m_1, ..., m_s} \left[\lim_{j \to \infty} \inf \frac{1}{S} \sum_{k=1}^{s} u(f(m_k + j)) \right]$$

and by Sucheston [22]

$$\sup_{m_1, ..., m_s} \left[\lim_{j \to \infty} \inf \frac{1}{S} \sum_{k=1}^{s} u(f(m_k + j)) \right] = \lim_{n} \sup_{j \geqslant 1} \left[\frac{1}{n} \sum_{i=1}^{n} u(f(i+j)) \right]$$

and this proves the result. ▌

9.3.3. *Proof of Corollary* 6

If $1/n \sum_{i=1}^{n} u(f(i+j))$ converges uniformly in j, then

$$\lim_n \inf_{j \geqslant 1} \left[\frac{1}{n} \sum_{i=1}^{n} u(f(i+j)) \right] = \lim_n \frac{1}{n} \sum_{i=1}^{n} u(f(i))$$

as wanted. ▮

9.4. *Theorem* 7

A non-additive set function $v: 2^{\mathcal{T}} \to [0, 1]$ is a capacity if it satisfies:

(i) $v(\emptyset) = 0$.

(ii) $v(A) \leqslant v(B)$ whenever $A \subseteq B \subseteq \mathcal{T}$.

(iii) $v(T) = 1$.

Of course, all additive measures are capacities, while the converse is false.
Let $f: \mathcal{T} \to \mathbb{R}$ be a bounded real-valued function on $\mathcal{T}$. The Choquet integral of f with respect to a capacity v is

$$\int_{\mathcal{T}} f\, dv = \int_0^{\infty} v(\{t: f(t) \geqslant \alpha\})\, d\alpha + \int_{-\infty}^{0} [v(\{t: f(t) \geqslant \alpha\}) - 1]\, d\alpha$$

where the right hand side is a Riemann integral. The integral is well defined because v is monotone. When v is additive, the Choquet integral becomes a standard additive integral.

We can now report Schmeidler's Theorem.

THEOREM 26. *The following two statements are equivalent*:

(i) *The preference relation* $\succsim$ *on* $\mathcal{F}$ *satisfies the axioms* A.1–A.4, *and* A.7.

(ii) *There exists a unique capacity* v *on* $2^{\mathcal{T}}$ *and an affine real valued function* u *on* $\mathcal{P}$ *such that for all* f *and* g *in* $\mathcal{F}$

$$f \succsim g \qquad \text{if and only if} \quad \int_{\mathcal{T}} u(f(t))\, dv \geqslant \int_{\mathcal{T}} u(g(t))\, dv.$$

Finally, the function u *is unique up to a positive linear transformation.*

We are now in a position to prove Theorem 7.

9.4.1. *Proof of Theorem 7*

The implication (ii) $\Rightarrow$ (i) is easy. As to the converse, since $\succsim$ satisfies A.1–A.4 and A.7, by Schmeidler's Theorem there exists a unique capacity v on $2^{\mathcal{T}}$ and an affine real valued function u on $\mathcal{P}$ such that for all f and g in $\mathcal{F}$

$$f \succsim g \qquad \text{if and only if} \quad \int_{\mathcal{T}} u(f(t))\, dv \geqslant \int_{\mathcal{T}} u(g(t))\, dv. \tag{14}$$

Let $p_1, p_2 \in \mathcal{P}$ with $p_2 \succ p_1$. Let $A \subseteq \mathcal{T}$ be a cofinite set, and f_A an act defined as follows

$$f_A(t) = \begin{cases} p_2 & \text{if} \quad t \in A \\ p_1 & \text{if} \quad t \notin A \end{cases}$$

Let k be a positive integer such that $t < k$ for all $t \notin A$. Then

$$f_A^k(t) = p_2 \qquad \text{for all} \quad t \in \mathcal{T}.$$

By A.8, $f_A \sim f_A^k$. By (14), this implies

$$u(p_1) + [u(p_2) - u(p_1)]\, v(A) = u(p_2)$$

so that $v(A) = 1$.

Next, let $A \subseteq \mathcal{T}$ be a finite set. By (14) and A.8,

$$u(p_1) + [u(p_2) - u(p_1)]\, v(A) = u(p_1)$$

so that $v(A) = 0$.

Let $A, B \subseteq \mathcal{T}$ be two infinite sets which are not cofinite. Define two acts f_A and f_B as follows

$$f_A(t) = \begin{cases} p_2 & \text{if} \quad t \in A \\ p_1 & \text{if} \quad t \notin A \end{cases} \qquad \text{and} \qquad f_B(t) = \begin{cases} p_2 & \text{if} \quad t \in B \\ p_1 & \text{if} \quad t \notin B \end{cases}.$$

As both A and B are not cofinite, there exists a map $\pi \in \Pi$ such that $f_A^\pi(t) = f_B(\pi(t))$ for all $t \in \mathcal{T}$. By A.9, $f_A \sim f_B$. By (14), this implies

$$u(p_1) + [u(p_2) - u(p_1)]\, v(A) = u(p_1) + [u(p_2) - u(p_1)]\, v(B)$$

so that $v(A) = v(B)$. Every infinite set A not cofinite can be decomposed in two infinite sets A_1, A_2 not cofinite. By what has been just proved, $v(A_1) = v(A_2) = v(A)$. By Schmeidler's Theorem, $f_{A_1} \sim f_{A_2}$. By A.6, $\frac{1}{2} f_{A_1} + \frac{1}{2} f_{A_2} \succsim f_{A_1}$, so that $v(A) \geqslant v(A_1) + v(A_2)$. Hence, $v(A) = 0$ for all infinite set which are not cofinite.

To summarize, the capacity v has the following form

$$v(A) = \begin{cases} 1 & \text{if } A \text{ is cofinite,} \\ 0 & \text{if } A \text{ is not cofinite} \end{cases}.$$

In other words, v is a filter game defined on the free filter of cofinite sets (see Marinacci [15]). It can be checked that for the filter game v it holds

$$\int_{\mathcal{T}} u(f(t))\, dv = \lim_{t \to \infty} \inf\, u(f(t))$$

for each $f \in \mathcal{F}$. By (14), we conclude that (i) $\Rightarrow$ (ii), as wanted. ∎

9.5. *Proof of Theorem* 10

By Lemmas 23 and 24

$$\lim_{T \to \infty} \inf_{j \geqslant 1} \left[\frac{1}{T} \sum_{t=1}^{T} u(f(j+t)) \right] \leqslant \int u(f)\, d\mu \leqslant \lim_{T \to \infty} \sup_{j \geqslant 1} \left[\frac{1}{T} \sum_{t=1}^{T} u(f(j+t)) \right]$$

for all $\mu \in \mathcal{C}$, so that

$$\beta_*(A) \leqslant \mu(A) \leqslant \beta^*(A).$$

Therefore, $\mu(A) = \beta(A)$ for all $A \in \mathcal{A}_b$. W.l.o.g., assume $u(f)$ is non-negative. Hence

$$\int u(f)\, d\mu = \int_0^\infty \mu(u(f) \geqslant \alpha)\, d\alpha = \int_0^\infty \beta(u(f) \geqslant \alpha)\, d\alpha$$

$$= \int_0^\infty \beta(u(f^\pi) \geqslant \alpha)\, d\alpha = \int u(f^\pi)\, d\mu$$

for all $\mu \in \mathcal{C}$. This implies $\min_{\mu \in \mathcal{C}} \int u(f)\, d\mu = \min_{\mu \in \mathcal{C}} \int u(f^\pi)\, d\mu$, so that $f \sim f^\pi$. ∎

9.6. *Proof of Proposition* 13

By a result due to Koopman and von Neumann [12], $\lim_{T \to \infty} 1/T \sum_{t=1}^{T} |u(f(t)) - l| = 0$ implies the existence of a set $J \subseteq \mathcal{T}$, with $\delta(J) = 0$, such that $\lim_{t \notin J} u(f(t)) = l$. We can write

$$\{t\colon u(f(t)) \geqslant \alpha\} = \{t\colon u(f(t)) \geqslant \alpha \text{ and } t \in J\} \cup \{t\colon u(f(t)) \geqslant \alpha \text{ and } t \notin J\}.$$

If $\alpha > l$, then $\{t\colon u(f(t)) \geqslant \alpha$ and $t \notin J\}$ is finite, while if $\alpha < l$, then $\{t\colon u(f(t)) \geqslant \alpha$ and $t \notin J\}$ is cofinite. In both cases, it belongs to $\mathscr{A}_d$. As $\{t\colon u(f(t)) \geqslant \alpha$ and $t \in J\} \in \mathscr{A}_d$, we conclude that $\{t\colon u(f(t)) \geqslant \alpha\} \in \mathscr{A}_d$ whenever $\alpha \neq l$. ∎

9.7. *Theorems* 14, 17, *and* 19

Let $\mathscr{L}_c$ be the set of all linear functionals L on l^∞ such that

1. $L(x) \geqslant 0$ if $x \geqslant 0$.
2. $L(x) = \lim_{T \to \infty} 1/T \sum_{t=1}^{T} x_t$ whenever this limit exists.
3. $L(1) = 1$, where 1 denotes the sequence $\{1, ..., 1, ...\}$.

In other words, $\mathscr{L}_c$ is the set of all positive functionals that coincide with the Cesaro limit of the sequence x when this limit exists. We now prove that all these functionals are Banach limits. This is a simple, but important result, for our purposes.

PROPOSITION 27. $\mathscr{L}_c \subseteq \mathscr{L}$.

Proof. Let $L \in \mathscr{L}_c$. Notice that for all $x \in l^\infty$ the sequence $x - \tau(x)$ has a Cesaro limit. For,

$$\frac{1}{T} \sum_{t=1}^{T} (x_t - \tau(x_t)) = \frac{1}{T} (x_{T+1} - x_1)$$

so that

$$0 = \lim_{T \to \infty} -\frac{1}{T} \|x\| \leqslant \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} (x_t - \tau(x_t)) \leqslant \lim_{T \to \infty} \frac{1}{T} \|x\| = 0.$$

Therefore

$$L(x) - L(\tau(x)) = L(x - \tau(x)) = 0$$

which shows that $L \in \mathscr{L}$. ∎

As $\mathscr{L}_c$ is a convex and weak*-compact set, we can define the lower envelope I_c on l^∞ as follows:

$$I_c(x) = \min\{L(x)\colon L \in \mathscr{L}_c\}.$$

Let $V = \{x\colon \lim_{T \to \infty} 1/T \sum_{t=1}^{T} x_t \text{ exists}\}$. We now prove a characterization of the envelope.

THEOREM 28. *For all $x \in l^\infty$ we have*

$$I_c(x) = \sup\{L(x'): L \in \mathscr{L}_c, x' \in V \text{ and } x' \leqslant x\}.$$

Proof. For each $x \in l^\infty$ define

$$I_*(x) = \sup\{L(x'): L \in \mathscr{L}_c, x' \in V \text{ and } x' \leqslant x\}.$$

It is easy to check that I_* is a positive homogeneous sublinear functional. Moreover, $-\infty < -\|x\| \leqslant I_*(x) \leqslant \|x\| < \infty$. I_* is a linear functional on V. For a given $\tilde{x} \in l^\infty$, let us look at the linear subspace $V \cup \{\tilde{x}\}$ generated by V and $\tilde{x}$. A typical element of $V \cup \{\tilde{x}\}$ has the form $x + t\tilde{x}$, with $x \in V$ and $t \in \mathbb{R}$. Define $\tilde{L}$ on $V \cup \{\tilde{x}\}$ by

$$\tilde{L}(x + t\tilde{x}) = I_*(x) + tI_*(\tilde{x}).$$

As I_* is a linear functional on V, $\tilde{L}$ as well is a linear functional on $V \cup \{\tilde{x}\}$. We show that it is positive. Let $x + t\tilde{x} \geqslant 0$. There are two cases to consider according to the sign of t:

1. Suppose $t \geqslant 0$. Then

$$\tilde{L}(x + t\tilde{x}) = I_*(x) + tI_*(\tilde{x}) = I_*(x) + I_*(t\tilde{x}) \\ \geqslant I_*(x + t\tilde{x}) \geqslant 0.$$

2. Suppose $t < 0$. Then $x + t\tilde{x} \geqslant 0$ implies $x \geqslant -t\tilde{x}$, so that

$$I_*(x) \geqslant I_*(-t\tilde{x}) = -tI_*(\tilde{x}).$$

In turn, this implies $I_*(x) + tI_*(\tilde{x}) \geqslant 0$. Hence $\tilde{L}(x + t\tilde{x}) \geqslant 0$.

We conclude that $\tilde{L}$ is a positive linear functional on $V \cup \{\tilde{x}\}$. By well ordering the set $l^\infty / V \cup \{\tilde{x}\}$, a similar argument proves that for every linear subspace $V \cup \{\tilde{x}\} \subseteq M \subseteq l^\infty$ there exists a positive linear functional $\tilde{L}_M$ such that $\tilde{L}_M(x) = I_*(x)$ for all $x \in V$, and $\tilde{L}_M(\tilde{x}) = I_*(\tilde{x})$. A standard application of Zorn's lemma finally shows that there exists a positive linear functional $\tilde{L}$ on l^∞ such that $\tilde{L}(x) = I_*(x)$ for all $x \in V$, and $\tilde{L}(\tilde{x}) = I_*(\tilde{x})$. Hence $\tilde{L} \in \mathscr{L}_c$, so that $\tilde{L}(x) \geqslant I_c(x) \geqslant I_*(x)$ for all $x \in l^\infty$. This implies

$$\tilde{L}(\tilde{x}) = I_c(\tilde{x}) = I_*(\tilde{x})$$

and this proves the result because $\tilde{x}$ was arbitrary. ∎

COROLLARY 29. *Let $u: \mathscr{P} \to \mathbb{R}$ be an affine utility. If $u(\mathscr{P}) = \mathbb{R}$, for every $f \in \mathscr{F}$*

$$I_c(u(f)) = \sup\left\{\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} u(g(t)): g \in \mathscr{F}, u(g) \in V, g(t) \precsim f(t) \text{ for all } t \geqslant 1\right\}.$$

For a given set of weights $\mathscr{C}$, let $I_{\mathscr{C}}: \mathscr{F} \to \mathbb{R}$ be defined by $I_{\mathscr{C}}(f) = \min_{\mu \in \mathscr{C}} \int u(f)\, d\mu$.

THEOREM 30. *The following two statements are equivalent*:

(i) *The preference relation* $\succsim$ *on* $\mathscr{F}$ *is canonical, satisfies the axioms* A.1–A.6, A.8, A.9, *and* A.10, *and is regular.*

(ii) *There exists an affine utility* $u: \mathscr{P} \to \mathbb{R}$ *and a unique weak*-compact and convex set of weights* $\mathscr{C} \subseteq \mathscr{N}_d$ *such that for all* f *and* g *in* $\mathscr{F}$ *we have* $f \succsim g$ *if and only if* $I_{\mathscr{C}}(f) \geqslant I_{\mathscr{C}}(g)$.

Proof. By Theorem 5 we know that $\mathscr{C} \subseteq \mathscr{N}$. We want to show that $\mathscr{C} \subseteq \mathscr{N}_d$. We first show that

$$\lim_{t \to \infty} \frac{|\{mk+s\}_{m \geqslant 0} \cap \{1, ..., t\}|}{t} = \frac{1}{k} \qquad \text{for all} \quad 0 \leqslant s \leqslant k-1.$$

For every $t \geqslant s$, there exists $m \geqslant 0$ such that $mk+s \leqslant t \leqslant (m+1)k+s$. Hence

$$\frac{\dfrac{mk+s}{k}}{(m+1)\,k+s} \leqslant \frac{|\{mk+s\}_{m \geqslant 0} \cap \{1, ..., t\}|}{t} \leqslant \frac{\dfrac{(m+1)\,k+s}{k}}{mk+s}$$

so that

$$\frac{1}{k} = \lim_{m \to \infty} \frac{\dfrac{mk+s}{k}}{(m+1)\,k+s} \leqslant \liminf_{t \to \infty} \frac{|\{mk+s\}_{m \geqslant 0} \cap \{1, ..., t\}|}{t}$$

$$\leqslant \limsup_{t \to \infty} \frac{|\{mk+s\}_{m \geqslant 0} \cap \{1, ..., t\}|}{t} \leqslant \lim_{m \to \infty} \frac{\dfrac{(m+1)\,k+s}{k}}{mk+s} = \frac{1}{k}$$

as wanted. In this way, for every $k \geqslant 1$ we have a partition $\{A_i^k\}_{i=1}^k$ of $\mathscr{T}$ with

$$\lim_{t \to \infty} \frac{|A_i^k \cap \{1, ..., t\}|}{t} = \frac{1}{k} \qquad \text{for all} \quad 1 \leqslant i \leqslant k.$$

It is easy to check that $\tau(A_i^k)=A_{i+1}^k$ for all $1\leqslant i\leqslant k-1$. Therefore, being $\mathscr{C}\subseteq\mathscr{N}$, for all $\mu\in\mathscr{C}$

$$\mu(A_1^k)=\mu(A_2^k)=\cdots=\mu(A_{k-1}^k)=\frac{1}{k}$$

so that

$$\mu(A_i^k)=\delta\,(A_i^k)\qquad\text{for all}\quad 1\leqslant i\leqslant k-1\qquad\text{and all}\quad \mu\in\mathscr{C}.$$

Using additivity we can construct a chain $\Gamma_q=\{A_q\}_{q\in\mathbb{Q}\cap[0,1]}$ such that $\delta(A_q)=\mu(A_q)=q$ for all $q\in\mathbb{Q}\cap[0,1]$ and all $\mu\in\mathscr{C}$.

Suppose $\{\bar{q}\}_{l\geqslant 1}$ and $\{\underline{q}_l\}_{l\geqslant 1}$ are two sequences in $\mathbb{Q}\cap[0,1]$ such that $\bar{q}_l\downarrow r$ and $\underline{q}_l\uparrow r$. It holds that

$$r=\lim_{l\to\infty}\underline{q}_l=\lim_{l\to\infty}\mu(A_{\underline{q}_l})$$

$$\leqslant\mu\left(\bigcup_{l\geqslant 1}A_{\underline{q}_l}\right)\leqslant\mu\left(\bigcap_{l\geqslant 1}A_{\bar{q}_l}\right)\leqslant\lim_{l\to\infty}\mu(A_{\bar{q}_l})=\lim_{l\to\infty}\underline{q}_l=r$$

for all $\mu\in\mathscr{C}$. Set $A_r=\bigcap_{l\geqslant 1}A_{\bar{q}_l}$, and $\Gamma=\{A_r\}_{r\in[0,1]}$. Then $\mu(A_r)=r$ for all $r\in[0,1]$ and all $\mu\in\mathscr{C}$. Proceeding in a similar way, one gets $\delta_*(A)=\delta^*(A)=r$ for all $\bigcup_{l\geqslant 1}A_{\underline{q}_l}\subseteq A\subseteq\bigcap_{l\geqslant 1}A_{\bar{q}_l}$. This implies that $\delta(A)$ exists and $\delta(A)=r$. Hence, $\delta(A_r)=\mu(A_r)$ for all $A_r\in\Gamma$.

Let $A\in\mathscr{A}_d$. Let A_δ, $A_\delta^c\in\Gamma$ be such that $\delta(A)=\delta(A_\delta)$ and $\delta(A^c)=\delta(A_\delta^c)$. By A.10, $f_A\sim f_{A_\delta}$ and $f_{A^c}\sim f_{A_\delta^c}$. Therefore, $\min_{\mu\in\mathscr{C}}\mu(A)=\min_{\mu\in\mathscr{C}}\mu(A_\delta)$ and $\min_{\mu\in\mathscr{C}}(A^c)=\min_{\mu\in\mathscr{C}}(A_\delta^c)$. Since A_δ, $A_\delta^c\in\Gamma$, $\delta(A_\delta)=\min_{\mu\in\mathscr{C}}\mu(A_\delta)$ and $\delta(A_\delta^c)=\min_{\mu\in\mathscr{C}}\mu(A_\delta^c)$. Hence, $\min_{\mu\in\mathscr{C}}\mu(A)=\delta\,(A)$ and $\min_{\mu\in\mathscr{C}}\mu(A^c)=\delta(A^c)$, so that

$$\mu(A)\geqslant\delta(A)\qquad\text{and}\qquad\mu(A^c)\geqslant\delta(A^c)$$

for all $\mu\in\mathscr{C}$. This implies $\mu(A)=\delta(A)$, and we conclude that $\mu(A)=\delta(A)$ for all $A\in\mathscr{A}_d$ and all $\mu\in\mathscr{C}$. Hence, $\mathscr{C}\subseteq\mathscr{N}_d$. ∎

THEOREM 31. *Suppose the preference relation* $\succsim$ *on* $\mathscr{F}$ *is canonical, satisfies the axioms* A.1–A.6, A.8, A.9, *and* A.10, *and is regular. Then* $I_{\mathscr{C}}(f)=I_c(u(f))$.

Proof. We first show that $I_{\mathscr{C}}(f)=\lim_{T\to\infty}1/T\sum_{t=1}^{T}u(f(t))$ whenever this limit exists. We first consider $f\in\mathscr{F}_\delta$. W.l.o.g., assume $u(f(t))\geqslant 0$ for all $t\in\mathscr{T}$. Except for a set of Lebesgue measure zero M, for every $\alpha\geqslant 0$ the natural density $\delta(\{t\colon u(f(t))\geqslant\alpha\})$ exists. For each $A\subseteq\mathscr{T}$ and each $T\in\mathscr{T}$, let $\delta_T(A) = |A \cap \{1, ..., T\}|/T$. The set function $\delta_T(A)$ is a finitely additive probability on $2^{\mathscr{T}}$, and

$$\int_{\mathscr{T}} u(f)\, d\,\delta_T = \int_0^\infty \delta_T(\{t: u(f(t)) \geqslant \alpha\})\, d\alpha = \frac{1}{T} \sum_{t=1}^{T} u(f(t)).$$

For every $0 \leqslant \alpha \notin M$, $\lim_{T \to \infty} \delta_T(\{t: u(f(t)) \geqslant \alpha\}) = \delta(\{t: u(f(t)) \geqslant \alpha\})$. As f is bounded, there exists $K > 0$ such that $0 \leqslant u(f(t)) \leqslant K$ for all $t \in \mathscr{T}$. Therefore, for every $t \geqslant 1$, we have

$$0 \leqslant \delta_T(\{t: u(f(t)) \geqslant \alpha\}) \leqslant 1 \qquad \text{for all} \quad 0 \leqslant \alpha \leqslant K,$$

$$\delta_T(\{t: u(f(t)) \geqslant \alpha\}) = 0 \qquad \text{for} \quad \alpha > K.$$

By the Arzelà Bounded Convergence Theorem, this implies

$$\begin{aligned}
\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} u(f(t)) &= \lim_{T \to \infty} \int_0^\infty \delta_T(\{t: u(f(t)) \geqslant \alpha\})\, d\alpha \\
&= \lim_{T \to \infty} \int_{M^c} \delta_T(\{t: u(f(t)) \geqslant \alpha\})\, d\alpha \\
&= \int_{M^c} \delta(\{t: u(f(t)) \geqslant \alpha\})\, d\alpha = \int_{M^c} \mu(\{t: u(f(t)) \geqslant \alpha\})\, d\alpha \\
&= \int_0^\infty \mu(\{t: u(f(t)) \geqslant \alpha\})\, d\alpha
\end{aligned}$$

for all $\mu \in \mathscr{C}$ as $\mathscr{C} \subseteq \mathscr{N}_d$. Therefore,

$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} u(f(t)) = \int u(f(t))\, d\mu \qquad \text{for all} \quad \mu \in \mathscr{C}.$$

In turn, this implies

$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} u(f(t)) = \min_{\mu \in \mathscr{C}} \int u(f(t))\, d\mu.$$

We now consider $f \in \mathscr{F}/\mathscr{F}_\delta$. W.l.o.g., assume $\inf_{t \geqslant 1} u(f(t)) < 0 < \sup_{t \geqslant 1} u(f(t))$. We first decompose $u(f(t))$ as

$$u(f(t)) = x(t) + x'(t)$$

where $x,\ x' \in l^{\infty}$, $\lim_{t\to\infty} x_t = \lim_{T\to\infty} 1/T \sum_{t=1}^{T} u(f(t))$, and $\sum_{t=1}^{\infty} x'(t)/t < \infty$. Set $x'(t) = t[1/t \sum_{k=1}^{t} u(f(k)) - 1/(t-1) \sum_{k=1}^{t-1} u(f(k))]$. Simple algebra shows that

$$u(f(t)) - x'(t) = \frac{1}{t-1} \sum_{k=1}^{t-1} u(f(k))$$

and, by setting $x(t) = u(f(t)) - x'(t)$, we have $\lim_{t\to\infty} x(t) = \lim_{T\to\infty} 1/T \sum_{t=1}^{T} u(f(t))$. On the other hand,

$$\sum_{t=1}^{\infty} \frac{x'(t)}{t} = \lim_{T\to\infty} \sum_{t=1}^{T} \left[\frac{1}{t} \sum_{k=1}^{t} u(f(k)) - \frac{1}{t-1} \sum_{k=1}^{t-1} u(f(k)) \right]$$

$$= \lim_{T\to\infty} \left(\frac{1}{T} \sum_{t=1}^{T} u(f(t)) - u(f(1)) \right) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} u(f(t))$$

so that $\sum_{t=1}^{\infty} x'(t)/t$ is a convergent series.

As x_t is such that $\inf_{t\geqslant 1} u(f(t)) \leqslant x(t) \leqslant \sup_{t\geqslant 1} u(f(t))$, there exists an act $g \in \mathscr{F}$ such that $u(g(t)) = x(t)$ for all $t \geqslant 1$. As $\lim_{T\to\infty} u(g(t)) = \lim_{T\to\infty} 1/T \sum_{t=1}^{T} u(f(t))$, $g \in \mathscr{F}_{\delta}$ (cf. the Remarks after Proposition 13). Hence, $\int u(g(t))\, d\mu = \lim_{T\to\infty} 1/T \sum_{t=1}^{T} u(f(t))$ for all $\mu \in \mathscr{C}$, so that

$$\min_{\mu\in\mathscr{C}} \int u(f(t))\, d\mu = \min_{\mu\in\mathscr{C}} \int [u(g(t)) + x'(t)]\, d\mu$$

$$= \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} u(f(t)) + \min_{\mu\in\mathscr{C}} \int x'(t)\, d\mu.$$

As $\inf_{t\geqslant 1} u(f(t)) < 0 < \sup_{\geqslant 1} u(f(t))$, there exist $p_1^*, p_2^* \in \mathscr{P}$ such that $u(p_1^*) < 0 < u(p_2^*)$. Therefore, there exists $\alpha > 0$ such that

$$\alpha u(p_1^*) \leqslant \inf_{t\geqslant 1} x'(t) \leqslant \sup_{t\geqslant 1} x'(t) \leqslant \alpha u(p_2^*).$$

As u is affine on $\mathscr{P}$, there exists a $0 \leqslant \lambda_{x,t} \leqslant 1$ such that $\alpha u(\lambda_{x,t} p_1^* + (1-\lambda_{x,t})\, p_2^*) = x'(t)$. Set $g'(t) = \lambda_{x,t} p_1^* + (1-\lambda_{x,t})\, p_2^*$ so that $g' \in \mathscr{F}$ and $\alpha u(g'(t)) = x'(t)$ for all $t \geqslant 1$. Clearly, $\sum_{t=1}^{\infty} u(g'(t))/t$ converges. If $u(g'(t)) \geqslant 0$ for all $t \geqslant 1$, $g' \in \mathscr{F}_{\delta}$ (cf. the Remarks after Proposition 13) and

$$\min_{\mu\in\mathscr{C}} \int u(g_2(t))\, d\mu = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} u(g'(t)) = 0.$$

Suppose, instead, that $\inf_{t\geqslant 1} u(g'(t)) < 0 < \sup_{t\geqslant 1} u(g'(t))$. By regularity $\min_{\mu\in\mathscr{C}} \int u(g'(t))\,d\mu = 0$. In both cases we can conclude that $\min_{\mu\in\mathscr{C}} \int x'(t)\,d\mu = 0$, so that

$$I_{\mathscr{C}}(f) = \min_{\mu\in\mathscr{C}} \int u(f(t))\,d\mu = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} u(f(t))$$

as wanted.

Let $\mathscr{C}_c$ be the set of weights associated with I_c. By regularity, $\mathscr{C}_c \subseteq \mathscr{C}$, so that

$$I_c(f) \geqslant I_{\mathscr{C}}(f) \qquad \text{for all} \quad f\in\mathscr{F}.$$

Suppose that for some f it holds $I_c(f) > I_{\mathscr{C}}(f)$. As

$$I_c(u(f)) = \sup\left\{\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} u(g(t)) \colon g\in\mathscr{F},\, u(g)\in V,\, g(t)\precsim f(t) \text{ for all } t\geqslant 1\right\}$$

for any $\varepsilon > 0$, there exists $u(g)\in V$ such that $g(t)\precsim f(t)$ for all $t\geqslant 1$, and

$$I_c(u(f)) - \varepsilon \leqslant \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} u(g(t)).$$

Put $\varepsilon < I_c(f) - I_{\mathscr{C}}(f)$. We have

$$I_c(u(f)) - \varepsilon \leqslant \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} u(g(t)) = I_c(u(g)) = I_{\mathscr{C}}(u(g)) \leqslant I_{\mathscr{C}}(u(f))$$

so that $I_c(u(f)) - I_{\mathscr{C}}(u(f)) \leqslant \varepsilon$, a contradiction. We conclude that $I_c(u(f)) = I_{\mathscr{C}}(u(f))$, and this completes the proof. ∎

THEOREM 32. *$I_c(x) = P(x)$ for all $x\in l^\infty$.*

Proof. If $x\in V$, then $I_c(x) = P(x) = \lim_{T\to\infty} 1/T\sum_{t=1}^{T} x_t$. By Theorem 28, this implies that $I_c(x)\leqslant P(x)$ for all $x\in l^\infty$. We now show that $I_c(x)\geqslant P(x)$ for all $x\in l^\infty$. It suffices to show that for all $k\in\mathbb{R}$ we have $I_c(x)\geqslant k$ whenever $P(x) > k$. Without loss, assume $k=0$ and $x\geqslant 0$. We now elaborate on an argument used in Peres [17]. Suppose $P(x) > 0$. It is easy to check that this implies $\lim_{\varepsilon\to 0} \liminf 1/\varepsilon T\sum_{t=T}^{T(1+\varepsilon)} x_t > 0$. This means that

$$\lim \inf_{N,M\to\infty} \frac{1}{N-M}\sum_{t=N+1}^{M} x_t > 0 \qquad \text{as} \quad 1\leqslant\frac{M}{N}\to 1.$$

Equivalently, for all $\delta > 0$ there exists $\gamma_\varepsilon > 0$ and N_δ such that $1/(N-M) \sum_{t=N+1}^{M} x_t > 0$ whenever $N \geqslant N_\delta$ and $N \leqslant M \leqslant (1+\delta)\, N$. As $M - N > 0$,

$$\sum_{t=N+1}^{M} x_t > 0$$

whenever $N \geqslant N_\delta$ and $N \leqslant M \leqslant (1+\delta)\, N$. Let $N_1 = 1$ and

$$N_k = \min \left\{ N > N_{k-1} \colon \sum_{N_{k-1}}^{N} x_t > 0 \right\}.$$

By (15), $\lim_{k \to \infty} N_k / N_{k-1} = 1$. Set

$$x'_t = x_t - \frac{1}{N_{k+1} - N_k} \sum_{N_k+1}^{N_{k+1}} x_t \qquad \text{for} \quad N_k \leqslant t < N_{k+1}.$$

As $1/(N_{k+1} - N_k) \sum_{N_k+1}^{N_{k+1}} x_t \geqslant 0$, $x' \leqslant x$. Moreover, it can be checked that $\lim_{T \to \infty} 1/T \sum_{t=1}^{T} x'_t = 0$. Hence, by Theorem (28), $I_c(x) \geqslant 0$, as wanted. ▮

COROLLARY 33. *Let $u \colon \mathscr{P} \to \mathbb{R}$ be an affine utility. Then for every $f \in \mathscr{F}$*

$$I_{\mathscr{C}}(f) = I_c(u(f)) = P(f).$$

All this proves Theorem 14 and Lemma 18, and the Polya Index's characterizations in Theorems 17 and 19.

10. THEOREM 20

The result is a simple consequence of Theorems 2 and 19 once one observes that the functional $I_i(x) = \liminf_{T \to \infty} 1/T \sum_{t=1}^{T} x_t$ satisfies the following properties:

(i) $I_i(x + x') \geqslant I_i(x) + I_i(x')$ for all $x, x' \in l^\infty$.

(ii) $I_i(\alpha x) = \alpha I_i(x)$ for all $\alpha \geqslant 0$ and $x \in l^\infty$.

(iii) $I_i(x) = I_i(\tau(x))$ for all $x \in l^\infty$.

(iv) $I_i(x) = I_i(x^\pi)$ for all $x \in l^\infty$. ▮

11. PROPOSITION 21

It suffices to prove that there exists $x \in l^\infty$ such that

$$\lim_{\varepsilon \to 0} \liminf \frac{1}{\varepsilon T} \sum_{t=T(1-\varepsilon)}^{T} x_t < \lim \inf_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} x_t.$$

Let x be the sequence considered in the introduction, i.e,

1, 1, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, ...

Then

$$\lim_{\varepsilon \to 0} \liminf \frac{1}{\varepsilon T} \sum_{t=T(1-\varepsilon)}^{T} x_t = 0,$$

while

$$\lim \inf_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} x_t = 1/3. \quad \blacksquare$$

REFERENCES

1. R. Aumann and L. Shapley, Long-term competition—A game theoretic analysis, mimeo, Hebrew University of Jerusalem, 1976.
2. S. Banach, "Théorie des opérations linéaires," Warsaw, 1932.
3. D. Blackwell and T. S. Ferguson, The big match, *Ann. Math. Stat.* **39** (1968), 159–163.
4. A. Chateauneuf, On the use of capacities in modeling uncertainty aversion and risk aversion, *J. Math. Econ.* **20** (1991), 343–369.
5. G. Choquet, Theory of capacities, *Ann. Inst. Fourier* **5** (1953), 131–295.
6. N. Dunford and J. T. Schwartz, "Linear Operators," Interscience, New York, 1957.
7. L. Epstein, Stationary cardinal utility and optimal growth under uncertainty, *J. Econ. Theory* **31** (1983), 133–152.
8. G. Frobenius, Über die Leibnizsche Reihe, *J. fur Math.* **89** (1880), 262–264.
9. D. Fudenberg and J. Tirole, "Game Theory," MIT Press, Cambridge, MA, 1991.
10. I. Gilboa and D. Schmeidler, Maxmin expected utility with a non-unique prior, *J. Math. Econ.* **18** (1989), 141–153.
11. G. Hardy, "Divergent Series," Oxford Univ. Press, Oxford, 1949.
12. B. Koopman and J. von Neumann, Dynamical systems of continuous spectra, *Proc. Nat. Acad. Sci. U.S.A.* **17** (1932), 255–263.
13. J. Littlewood, The converse of Abel's theorem on power series, *Proc. London Math. Soc.* **9** (1911), 434–448.
14. G. Lorentz, A contribution to the theory of divergent sequences, *Acta Math.* **80** (1948), 167–190.
15. M. Marinacci, Decomposition and representation of coalitional games, *Math. Oper. Res.* **21** (1996), 1000–1015.
16. M. Marinacci, The behavioral foundation of the undiscounted Folk Theorems, mimeo, University of Toronto, 1997.

17. Y. Peres, Application of Banach limits to the study of sets of integers, *Israel J. Math.* **62** (1988), 17–31.
18. G. Polya, Untersuchungen über Lücken und Singularitäten von Potenzreihen, *Math. Z.* **29** (1929), 549–640.
19. F. Ramsey, A mathematical theory of saving, *Econ. J.* **38** (1928), 543–559.
20. A. Rubinstein, Equilibrium in supergames, Research Memorandum 25, Center for Research in Game Theory and Mathematical Economics, Hebrew University of Jerusalem, Jerusalem, 1977.
21. D. Schmeidler, Subjective probability and expected utility without additivity, *Econometrica* **57** (1989), 571–587.
22. L. Sucheston, On the existence of finite invariant measures, *Math. Z.* **86** (1964), 327–336.