



**HAL**  
open science

## On multiple discount rates and present bias

Bach Dong Xuan, Philippe Bich, Bertrand Wigniolle

► **To cite this version:**

Bach Dong Xuan, Philippe Bich, Bertrand Wigniolle. On multiple discount rates and present bias. 2022. halshs-03884664

**HAL Id: halshs-03884664**

**<https://shs.hal.science/halshs-03884664v1>**

Preprint submitted on 5 Dec 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**WORKING PAPER N° 2022 – 39**

**On multiple discount rates and present bias**

**Bach Dong Xuan  
Philippe Bich  
Bertrand Wigniolle**

**JEL Codes:**

**Keywords: discounting, discount rate, exponential model, hyperbolic model, maxmin expected utility.**



# On multiple discount rates and present bias.

Bach Dong Xuan\*, Philippe Bich†, Bertrand Wigniolle‡

December 2, 2022

## Abstract

In this paper, we axiomatize the following maxmin version of the quasi-hyperbolic model

$$I(x) = \min_{(p_0, \delta) \in Q} \{p_0 x_0 + (1 - p_0)(1 - \delta)x_1 + (1 - p_0)(1 - \delta)\delta x_2 + (1 - p_0)(1 - \delta)\delta^2 x_3 \dots\},$$

where  $(x_0, x_1, \dots)$  is an infinite utility stream, and where  $Q$  is a set of possible pairs of discount rates  $\delta$  and present bias parameters  $p_0$ . When there is no present bias, we recover Chambers and Echenique's axiomatization of maxmin exponential preferences, and when  $Q$  reduces to a singleton, we get Montiel Olea and Strzalecki's axiomatization of quasi-hyperbolic preferences. To prove our main result, we provide some intertemporal variational representation results of interest for its own sake.

**Keywords:** discounting, discount rate, exponential model, hyperbolic model, maxmin expected utility.

## 1 Introduction

This paper axiomatizes the following maxmin version of the quasi-hyperbolic model

$$I(x) = \min_{(p_0, \delta) \in Q} \{p_0 x_0 + (1 - p_0)(1 - \delta)x_1 + (1 - p_0)(1 - \delta)\delta x_2 + (1 - p_0)(1 - \delta)\delta^2 x_3 \dots\}, \quad (1)$$

where  $x = (x_n)_{n \geq 0}$  is an infinite utility stream, and where  $Q$  is a set of possible pairs of discount rates  $\delta$  and present bias parameters  $p_0$ .

Intertemporal choices involve situations in which individuals have to choose between immediate-small reward and delayed-larger rewards. In practice, a large part of the literature has concentrated on the exponential model (Samuelson (1937)) to solve this trade-off, mainly because of its simplicity. Yet, most of its implications (time-consistence, time-invariance, time-separability) are contradicted by a large number of studies.

---

\*Université Paris 1 Panthéon-Sorbonne UMR 8074 Centre d'Economie de la Sorbonne, Paris School of Economics. E-mail: Xuan-Bach.Dong@univ-paris1.fr

†Université Paris 1 Panthéon-Sorbonne UMR 8074 Centre d'Economie de la Sorbonne, Paris School of Economics. E-mail: bich@univ-paris1.fr

‡Paris School of Economics, Université Paris 1 Panthéon-Sorbonne. E-mail: Bertrand.Wigniolle@univ-paris1.fr

First, many evidences suggest that when individuals make intertemporal choices, they exhibit time-inconsistent and time-dependent preferences: in general, people tend to be patient for long-term gains, but impatient for short-term gains (see [Thaler and Shefrin \(1981\)](#), [Ainslie and Herrnstein \(1981\)](#), [Ainslie and Haslam \(1992\)](#), [Loewenstein and Prelec \(1992\)](#), [Kirby and Herrnstein \(1995\)](#), [Myerson and Green \(1995\)](#), [McClure et al. \(2004\)](#), and [DellaVigna and Malmendier \(2006\)](#)). The quasi-hyperbolic discounting model (see [Phelps and Pollak \(1968\)](#) and [Laibson \(1997\)](#)) is an efficient and simple way to model dynamic inconsistency in economics (see, for example, [Barro \(1999\)](#), [Grenadier and Wang \(2007\)](#), [Marín-Solano and Navas \(2010\)](#), [Harris and Laibson \(2013\)](#), and [Zou et al. \(2014\)](#)). According to the quasi-hyperbolic model, an agent evaluates an infinite utility stream  $x$  at time  $t = 0$  by computing the hyperbolic discounted sum

$$I(x) = p_0x_0 + (1 - p_0)(1 - \delta)x_1 + (1 - p_0)(1 - \delta)\delta x_2 + (1 - p_0)(1 - \delta)\delta^2x_3 \dots,$$

where  $\delta$  is an exponential discount rate, and  $\beta := \frac{(1-p_0)(1-\delta)}{p_0\delta}$  measures *present bias* of the preferences:  $\beta = 1$  corresponds to the standard exponential discounting (i.e., there is no present bias), and for  $\beta = 0$  we have a maximal present bias.

Second, time-separability, stated in its simplest form, says that tradeoffs between today and tomorrow are not affected by future consumption. [Koopmans \(1960\)](#), in his seminal paper, provides an axiomatization of the separable exponential model, and connects separability to some independence axiom. Separability assumption is very popular, mostly because of its simplicity, but it is often violated in practice; several non-separable models have been proposed after Koopmans (see, for example, [Kreps and Porteus \(1978\)](#), [Lucas and Stokey \(1984\)](#), [Epstein and Zin \(1987\)](#), [Gilboa and Schmeidler \(1989\)](#), etc.) Remark that non-separability is deeply connected to non-additivity of preferences in static models, and one of the most convincing (static) model allowing non-additive preferences is Gilboa and Schmeidler maxmin model ([Gilboa and Schmeidler \(1989\)](#)), which will play an important role in our paper.

Our axiomatization is a simple and direct attempt to take into account the two limitations above, i.e., to encompass present biased preferences and allow non-separability, by using some maxmin approach à la Gilboa and Schmeidler. This implies, for example, two recent results: first, when there is no present bias (i.e., when  $p_0 = 1 - \delta$ ), we recover the axiomatization of the exponential maxmin criterion of [Chambers and Echenique \(2018\)](#)

$$I(x) = \min_{\delta \in E} \left\{ (1 - \delta) \sum_{t=0}^{+\infty} \delta^t x_t \right\}$$

for some set  $E$  of discount rates uniquely determined by the preference relation. Chambers and Echenique interpret their model as a way of “aggregating a decision criterion from a set of experts with different discount rates”, and thus our result opens the door to considering present biased experts.<sup>1</sup> Second, when the set  $Q$  of present bias and discount rate parameters reduces to a singleton, we get an axiomatization of the quasi-hyperbolic criterion (see [Hayashi \(2003\)](#) and [Montiel Olea and Strzalecki \(2014\)](#)).<sup>2</sup>

<sup>1</sup>There are many controversies on what would be a “good” value of  $\delta$ . For example, when dealing with climate change, Stern report ([Stern \(2006\)](#)) recommends to choose discount rates between 1.6% and 2.1%, but some economists consider these discount rates too low. In a survey, [Weitzman \(2001\)](#) asked to many specialists the discount rate that they would use to evaluate long-term projects. The mean of the answers he finds is 3.96%, with a standard deviation of 2.94%. The maxmin criterion of [Chambers and Echenique \(2018\)](#) can be seen as a possible answer to the problem generated by this multiplicity of discount rates.

<sup>2</sup>[Druegon and Ha-Huy \(2018\)](#) and [Druegon et al. \(2019\)](#) also propose to take into account multiple discount factors, but their approaches differ from the axiomatized approach presented here or in [Chambers and Echenique \(2018\)](#).

Our quasi-hyperbolic maxmin criterion has several properties in common with both the exponential maxmin model and the quasi-hyperbolic model: aside from the standard order-theoretic properties (*Completeness* and *Transitivity*) and some classical technical assumptions (*Continuity* and *Monotone continuity*), it satisfies three “standard” properties, *Convexity*, *Co-cardinality* and *Monotonicity*. First, Convexity can be seen as a fairness axiom which says roughly that the decision maker prefers to smooth payoffs across generations. Second, Co-cardinality is an invariance property of the preferences of the decision maker which requires that the order between two streams should not be modified by a same positive affine transformations of the streams. Third, Monotonicity is a kind of unanimity property among generations: this amounts to saying that if each generation  $t \geq 0$  prefers  $x_t$  to  $y_t$ , then the decision maker should prefer stream  $x$  to stream  $y$ .

On the other hand, our model relaxes two important properties: more precisely, we will not assume *Invariance to stationary relabeling (ISTAT)*, the main invariance property in [Chambers and Echenique \(2018\)](#), and will not assume separability of preferences across time, which is assumed in [Montiel Olea and Strzalecki \(2014\)](#). ISTAT assumes that every stream  $x$  is equivalent to its “smooth” delayed version  $(\theta, x)$ , where  $\theta$  is a constant equivalent of  $x$ . Here, the equivalence should be understood in a strong sense, i.e., not only is the decision maker indifferent between  $x$  and  $(\theta, x)$ , but also, quoting the authors, “there can be no hedging or smoothing value to mixing stream  $x$  with the delayed version of itself  $(\theta, x)$ ”.<sup>3</sup> Roughly, ISTAT says that “some particular relabeling of agents’ names does not matter”. ISTAT is violated in the following simple and natural model: assume that the decision maker maximizes the worst case between the current and the next generation, i.e.,  $I(x) = \min\{x_0, x_1\}$ . Taking  $x = (1, -1, 0, \dots)$ , we get  $I(x) = -1$ , so that  $\theta = -1$  is a constant equivalent of  $x$ , and  $I(\theta, x) = -1$ . But mixing  $(\theta, x) = (-1, 1, -1, 0, \dots)$  and  $x$ , we strictly improves the payoff  $-1$  since  $I(\frac{(\theta, x) + x}{2}) = I(0, 0, -1, 0, \dots) = 0$ . In this example, a simple translation of generations (while the first generation receives the constant equivalent) creates smoothing opportunities with  $x$ , which explains the failure of ISTAT.

Our quasi-hyperbolic maxmin criterion still requires some (weak) invariance properties. Our first novel axiom, we term *Delayed ISTAT*, is similar to ISTAT, but delayed at period 1, i.e., it says that some particular relabeling of all agents’ names – except the first one – does not matter. Our second novel axiom, *Existence of a future reference plan (EFREP)*, says that, for every stream  $x$ , there should exist some future reference plan  $y$ , such that  $(\theta, y)$  is equivalent to  $x$ , and satisfying the following hedging stationarity property: for every stream of the form  $(\theta, z)$  that is equivalent to  $x$  and is a hedging against  $x$ , the delayed version  $(\theta, \theta, z)$  of  $(\theta, z)$  should be a hedging against  $(\theta, y)$ , the “smooth equivalent” of  $x$ . Remark that in the simple example above where the decision maker maximizes the worst case between the two first generations,  $(\theta, x) = (-1, 1, -1, 0, \dots)$  is a hedging against  $x = (1, -1, 0, \dots)$ , but  $(\theta, \theta, x) = (-1, -1, 1, -1, 0, \dots)$  is not a hedging against  $(\theta, x) = (-1, 1, -1, 0, \dots)$ , which illustrates that, in general, we cannot take  $y = x$  in axiom EFREP. Yet, it can be proved that (1) when the preferences are assumed to be additive, then  $y$  can be taken equal to  $(\theta, \theta, \dots)$ , (2) when ISTAT is satisfied, then  $y$  can be taken equal to  $x$ . Thus, EFREP property can be seen as a weak stationarity property, up to some distortion of the delayed version of  $x$ .

Remark that we can get rid of the existential condition EFREP and get a simpler axiom if we

---

<sup>3</sup>The notion of hedging is standard in decision theory: recall that given a utility function  $I$  of a decision maker,  $z$  is a hedging against some equivalent act  $x$ , if  $I(\frac{x+z}{2}) > \frac{I(x)+I(z)}{2}$ . In the intertemporal setting, the interpretation is that the decision maker has some smoothing opportunities across time between  $x$  and  $z$ .

strengthen the maxmin quasi-hyperbolic representation result as follows:

$$I(x) = \min_{p_0 \in A, \delta \in B} \{p_0 x_0 + (1 - p_0)(1 - \delta)x_1 + (1 - p_0)(1 - \delta)\delta x_2 + (1 - p_0)(1 - \delta)\delta^2 x_3 \dots\} \quad (2)$$

The above alternative maxmin criterion is a reinforcement of our first maxmin quasi-hyperbolic criterion (1) since it requires additionally that the set  $Q$  of present bias parameters and discount factors is a product set  $A \times B$ . In this case, we can prove that in the axiomatization of  $I$ , axiom EFREP can be replaced by *Weak separability*: this axiom states that when  $(\theta, x)$  is equivalent of  $\theta$  (the constant equivalent of  $x$ ), then  $(\alpha, x)$  has to be equivalent to  $(\alpha, \theta)$  for every real  $\alpha$ , that is to say the equivalence does not depend on of the first agent's utility. Thus, the above criterion has a rather simple axiomatization, but contrarily to our more general Criterion (1), it does not provide a generalization of [Chambers and Echenique \(2018\)](#) representation result.

Last, our paper can be seen as a contribution on the literature on multiple discount rates and multiple present bias coefficients. The multiplicity of discount rates has been intensively studied (see, for example, [Chambers and Echenique \(2018\)](#), [Weitzman \(2001\)](#), [Zuber et al. \(2011\)](#), and [Jackson and Yariv \(2015\)](#)), but the multiplicity of present bias coefficients has been more recently addressed in the literature. For instance, a recent meta analysis of more than 60 papers ([Cheung et al. \(2021\)](#)) find a confidence interval for  $\beta$  equal to  $[0.74; 0, 9]$  for monetary rewards, and a confidence interval equal to  $[0.51; 0, 85]$  for non monetary rewards. Thus, our two criteria above, (1) and (2), could be seen as a means of solving the multiplicity of discount rates and present bias coefficients.

Our paper is organized as follows: in Section 2, we prove some general variational preference representation theorems (see Theorem 1, Theorem 2 and Theorem 3), which are not only useful for our main results but also of interest for its own sake. In Section 3, we provide our last axioms and our main quasi-hyperbolic discounting representation theorems (see Theorem 4 and Theorem 5), and then we give two byproducts, first the axiomatization of Chamber and Echenique maxmin exponential discounting model, second the axiomatization of the quasi-hyperbolic discounting model.

## 2 The main variational representation theorem

In this section, we prove our main variational representation theorem, with some byproducts. In particular, our axiomatization of the maxmin quasi-hyperbolic criterion will be a consequence of this representation theorem.

### 2.1 Definitions and notations

We consider a decision maker dealing with countable streams of utility. The preferences of the decision maker are represented by a binary relation  $\succeq$  on  $l_\infty$ , the space of real-valued bounded sequences.<sup>4</sup> Each index  $n = 0, 1, \dots$  represents a generation, and for every  $x = (x_n)_{n \geq 0} \in l_\infty$ ,  $x_n$  represents the utility that generation  $n$  obtains from  $x$ . As usual, we define the asymmetric part  $\succ$  and symmetric part  $\sim$  of  $\succeq$ , that is  $x \succ y$  if  $x \succeq y$  and not  $[y \succeq x]$ , and  $x \sim y$  if  $x \succeq y$  and  $y \succeq x$ . Throughout this paper, we will

<sup>4</sup>The set  $l_\infty$  is endowed with the sup norm  $\|\cdot\|_\infty$ , defined by  $\|x\|_\infty = \sup_{n \in \mathbb{N}} |x_n|$  for every  $x = (x_n)_{n \geq 0} \in l_\infty$ .

use the following standard order relations on  $l_\infty$ : for every  $(x, y) \in l_\infty^2$ ,  $x \geq y$  if and only if  $x_n \geq y_n$  for every  $n \in \mathbb{N}$ ,  $x > y$  if  $x \geq y$  and  $x \neq y$ , and  $x \gg y$  if  $x_n > y_n$  for every  $n \in \mathbb{N}$ .

For every  $\theta \in \mathbb{R}$ ,  $\boldsymbol{\theta}$  denotes the constant sequence  $(\theta, \theta, \dots)$ . For every  $x \in l_\infty$ ,  $I(x) \in \mathbb{R}$  is said to be a *constant equivalent* of  $x$  if  $x \sim \mathbf{I}(x)$ . We also need the following notation:

$$\underbrace{(\theta, \dots, \theta, x)}_{T \text{ times}}$$

which denotes the sequence that takes the value  $\theta$  for  $n = 0, \dots, T-1$  and  $x_{n-T}$  for  $n \geq T$ .

## 2.2 Standard axioms

Let us introduce the following (standard) axioms (see Section 3.3 for a discussion):

**A.1 (Weak order):** The relation  $\succeq$  is a *weak order*, i.e., it is *complete* (for every  $(x, y) \in l_\infty^2$ , either  $x \succeq y$  or  $y \succeq x$ ) and it is *transitive* (for every  $(x, y, z) \in l_\infty^3$ ,  $x \succeq y$  and  $y \succeq z$  implies  $x \succeq z$ ).

**A.2 (Monotonicity):** (1) For every  $(x, y) \in l_\infty^2$ ,  $x \geq y$  implies  $x \succeq y$ , and (2)  $\mathbf{1} \succ \mathbf{0}$ .

**A.3 (Continuity):** For every  $(x, y, z) \in l_\infty^3$  with  $x \succeq y \succeq z$ , the sets  $\{\alpha \in [0, 1] : \alpha x + (1 - \alpha)z \succeq y\}$  and  $\{\alpha \in [0, 1] : y \succeq \alpha x + (1 - \alpha)z\}$  are closed in  $[0, 1]$ .

**A.4 (Constant translation invariance):** For every  $(x, y) \in l_\infty^2$ ,  $x \succeq y$  implies  $x + \boldsymbol{\theta} \succeq y + \boldsymbol{\theta}$  for every  $\boldsymbol{\theta} \in \mathbb{R}$ .

**A.5 (Convexity):** For every  $(x, y) \in l_\infty^2$  and  $\boldsymbol{\theta} \in \mathbb{R}$ ,  $x \succeq \boldsymbol{\theta}$  and  $y \succeq \boldsymbol{\theta}$  implies  $\lambda x + (1 - \lambda)y \succeq \boldsymbol{\theta}$  for every  $\lambda \in [0, 1]$ .

The following proposition recalls under which conditions a constant equivalent of  $\succeq$  exists:

**Proposition 1.** *A binary relation  $\succeq$  on  $l_\infty$  satisfies A.1-A.5 if and only if there exists a unique weakly-increasing, concave, 1-lipschitz, normalized, and translation invariant<sup>5</sup> constant equivalent  $I : l_\infty \rightarrow \mathbb{R}$  that represents  $\succeq$ .*

Throughout this paper, to simplify, we will call the function  $I$  in Proposition 1 the “unique constant equivalent of  $\succeq$ ” (instead of the unique weakly-increasing, concave, 1-lipschitz, normalized, and translation invariant constant equivalent).

## 2.3 Results

We first state an important result of this paper which characterizes the constant equivalent  $I$  above in terms of *variational representation* (introduced by Maccheroni et al. (2006)). The norm dual of the space of bounded sequences  $(l_\infty, \|\cdot\|_\infty)$  is  $ba(\mathbb{N})$ , the space of all signed finitely additive measures<sup>6</sup> of bounded variation on the  $\sigma$ -algebra  $2^{\mathbb{N}}$  of all subsets of  $\mathbb{N}$ ; the duality is  $\langle x, \mu \rangle = \int_{\mathbb{N}} x d\mu$  for every  $x \in l_\infty$  and  $\mu \in ba(\mathbb{N})$ . When  $\mu$  is a countably additive probability, we can represent  $\mu$  by  $(\mu_n)_{n \geq 0}$ , where

<sup>5</sup> $I$  is normalized if  $I(\mathbf{1}) = 1$ , and  $I$  is translation invariant if for every  $x \in l_\infty$  and  $\boldsymbol{\theta} \in \mathbb{R}$ ,  $I(x + \boldsymbol{\theta}) = I(x) + \boldsymbol{\theta}$ .

<sup>6</sup>Such measures are called charges. A charge  $\mu$  on the  $\sigma$ -algebra  $2^{\mathbb{N}}$  of all subsets of  $\mathbb{N}$  has a bounded variation if its total variation  $V_\mu = \sup\{\sum_{i=1}^n |\mu(A_i)| : \{A_1, \dots, A_n\} \text{ is a partition of } \mathbb{N}\}$  is finite. We endow, as usual,  $ba(\mathbb{N})$  with the weak\* topology  $\sigma(ba(\mathbb{N}), l_\infty)$ : a net  $(\mu_d)$  in  $ba(\mathbb{N})$  converges to  $\mu$  for this topology if and only if for every  $x \in l_\infty$ , the net  $\langle x, \mu_d \rangle$  converges to  $\langle x, \mu \rangle$ .

$\mu_n = \mu(\{n\})$ ; and we simply have  $\langle x, \mu \rangle = \int_{\mathbb{N}} x_n d\mu = \sum_{n=0}^{+\infty} x_n \mu_n$ . Let  $\Delta$  (resp.  $\Delta^\sigma$ ) denote the set of all finitely additive probability measures (resp. the set of all countably additive probability measures) on  $2^{\mathbb{N}}$ ; we endow  $\Delta$  and  $\Delta^\sigma$  with the relative topology induced by the weak\* topology on  $ba(\mathbb{N})$ . As in [Maccheroni et al. \(2006\)](#), we say that a function  $c : \Delta \rightarrow [0, \infty]$  is grounded if  $\min_{p \in \Delta} \{c(p)\} = 0$ .

**Theorem 1.** *A binary relation  $\succeq$  on  $l_\infty$  satisfies **A.1-A.5** if and only if there exists a grounded, convex, and lower semicontinuous function  $c : \Delta \rightarrow [0, \infty]$  such that the unique constant equivalent  $I : l_\infty \rightarrow \mathbb{R}$  representing  $\succeq$  can be written as follows:*

$$I(x) = \min_{p \in \Delta} \{ \langle x, p \rangle + c(p) \}.$$

Moreover, the function  $c$  above is uniquely defined.

The main differences between Theorem 1 and the variational representation result in [Maccheroni et al. \(2006\)](#) are that (1) our framework is discrete (i.e., here, an act is a sequence  $x : \mathbb{N} \rightarrow \mathbb{R}$ ), and (2) an act  $x$  is not required to have a finite number of values as it is the case in [Maccheroni et al. \(2006\)](#). Remark that since  $l_\infty$  is an Archimedean Riesz spaces with unit and  $I : l_\infty \rightarrow \mathbb{R}$  is a convex niveloid,<sup>7</sup> Theorem 1 can be seen as a byproduct of Theorem 3 in [Cerrei-Vioglio et al. \(2014\)](#).

If we add the following Monotone continuity assumption, then we obtain a representation theorem in which the set of priors in Theorem 1, the set of finitely additive probability measures, can be reduced to the set of countably additive probability measures. Below, if  $(E_n)_{n \in \mathbb{N}}$  denotes a sequence of subsets of  $\mathbb{N}$ , then  $(E_n) \downarrow \emptyset$  means that  $E_n \supset E_{n+1}$  for every  $n \geq 0$  and  $\bigcap_{n=0}^{\infty} E_n = \emptyset$ . For every  $(x, y) \in l_\infty^2$  and  $A \subset \mathbb{N}$ ,  $xAy$  denotes the sequence defined as follows: denoting  $xAy = (z_0, z_1, \dots)$ , then  $z_n = x_n$  if  $n \in A$  and  $z_n = y_n$  if  $n \in A^c$ . Last, we let  $\mathbf{1}_A$  denote the sequence defined by  $\mathbf{1}_A = xAy$  in which  $x = \mathbf{1}$  and  $y = \mathbf{0}$ .

**A.6 (Monotone continuity):** Let  $x \in l_\infty$  and  $\theta \in \mathbb{R}$  such that  $x \succ \theta$ , and let  $(E_n)_{n \in \mathbb{N}}$  be a sequence of subsets of  $\mathbb{N}$  with  $\{E_n\} \downarrow \emptyset$ . Then for every  $k \in \mathbb{R}$ , there exists  $n_0 \in \mathbb{N}$  such that  $kE_{n_0}x \succ \theta$ .

Roughly, Monotone continuity requires that any strict inequality  $x \succ \theta$  is preserved if we modify  $x$  by some arbitrary constant  $k$  on a set  $E_{n_0}$  for  $n_0$  large enough. We now state a reinforced version<sup>8</sup> of Theorem 1 where  $\Delta$ , the set of all finitely additive probability measures, is replaced by  $\Delta^\sigma$ , the set of all countably additive probability measures on  $2^{\mathbb{N}}$ .

**Theorem 2.** *A binary relation  $\succeq$  on  $l_\infty$  satisfies **A.1-A.6** if and only if there exists a grounded,<sup>9</sup> convex, and lower semicontinuous function  $c : \Delta \rightarrow [0, \infty]$ , with  $\{p \in \Delta : c(p) \leq t\} \subset \Delta^\sigma$  for every  $t \geq 0$ , such that the unique constant equivalent  $I : l_\infty \rightarrow \mathbb{R}$  representing  $\succeq$  can be written as follows:*

$$I(x) = \min_{p \in \Delta^\sigma} \left\{ \sum_{k=0}^{+\infty} x_k p_k + c(p) \right\}.$$

<sup>7</sup>See [Cerrei-Vioglio et al. \(2014\)](#).

<sup>8</sup>Theorem 13 in [Maccheroni et al. \(2006\)](#) is a similar result but for simple acts (i.e., finite-valued functions) on a general state space  $S$ .

<sup>9</sup>In this theorem, we require a stronger definition of a grounded function  $c$  that is  $\min_{p \in \Delta^\sigma} \{c(p)\} = 0$ .



Moreover, the function  $c$  above is uniquely defined.

Notice that the additional condition  $\{p \in \Delta : c(p) \leq t\} \subset \Delta^\sigma$  guarantees that the minimum in Theorem 2 is attained, and the following example explains why the condition is necessary in the theorem above:

**Example 1.** Define a relation  $\succeq$  on  $l_\infty$  by  $x \succeq y$  if and only if  $\inf_{n \geq 0} x_n \geq \inf_{n \geq 0} y_n$ . Then  $\succeq$  satisfies axioms **A.1-A.5** but does not satisfy axiom **A.6**. To see it, we consider  $\mathbf{1} \succ \mathbf{0}$ , and let  $E_n = \{n, n+1, \dots\}$  for every  $n \in \mathbb{N}$ , then  $\mathbf{0}E_n\mathbf{1} \succ \mathbf{0}$  is not true for every  $n \in \mathbb{N}$ . Remark that  $I(x) = \inf_{p \in \Delta} \{\langle x, p \rangle\} = \inf_{p \in \Delta^\sigma} \{\sum_{k=0}^{+\infty} x_k p_k\}$  represents  $\succeq$ , so that  $c = 0$ . However, the last assumption in Theorem 2 is not satisfied because  $\{p \in \Delta : c(p) \leq t\} = \Delta$  is not included in  $\Delta^\sigma$  for every  $t \geq 0$ .

If we add the following axiom, called Homotheticity, then we obtain the maxmin model of Gilboa and Schmeidler (1989) as a particular case<sup>10</sup> of Theorem 2:

**A.7 (Homotheticity):** For every  $(x, y) \in l_\infty^2$  and  $\alpha \geq 0$ ,  $x \succeq y$  implies  $\alpha x \succeq \alpha y$ .

**Theorem 3.** A binary relation  $\succeq$  on  $l_\infty$  satisfies **A.1-A.7** if and only if there exists  $\mathbb{D}$ , a convex and weakly\* compact subset of  $\Delta^\sigma$ , such that the unique constant equivalent  $I : l_\infty \rightarrow \mathbb{R}$  representing  $\succeq$  can be written as follows:

$$I(x) = \min_{p \in \mathbb{D}} \sum_{n=0}^{\infty} x_n p_n.$$

Moreover, the set  $\mathbb{D}$  above is uniquely defined.

### 3 Maxmin representation results with quasi-hyperbolic discount factors

#### 3.1 Additional axioms

For the remaining results, we need to strengthen slightly axiom **A.2** as follows:

**A.2' (Strong monotonicity):** For every  $(x, y) \in l_\infty^2$ ,  $x \geq y$  implies  $x \succeq y$ , and for every  $(x, y) \in l_\infty^2$ ,  $x > y$  implies  $x \succ y$ .

*In the sequel, we will use the name “Maxmin axioms” for the set of axioms: **A.1, A.2', A.3-A.7**.*

We now complete the list of axioms of the previous section:

**A.8 (Invariance to stationary relabeling (ISTAT)):** For every  $\theta \in \mathbb{R}$  and  $x \in l_\infty$  such that  $x \sim \theta$ ,  $\lambda x + (1 - \lambda)(\theta, x) \sim \theta$  for every  $\lambda \in [0, 1]$ .

<sup>10</sup>We could also get from Theorem 1 the weaker representation theorem, which says that a binary relation  $\succeq$  on  $l_\infty$  satisfies **A.1-A.5** and **A.7** if and only if the unique constant equivalent  $I : l_\infty \rightarrow \mathbb{R}$  representing  $\succeq$  can be written as  $I(x) = \min_{p \in \mathbb{D}} \{\langle x, p \rangle\}$ , where  $\mathbb{D}$  is a convex weakly\* compact subset of  $\Delta$ .

**A.9 (Weak stationarity<sup>11</sup>):** For every  $\theta \in \mathbb{R}$  and  $x \in l_\infty$  such that  $x \sim \theta$ ,  $(\theta, x) \sim \theta$

**A.10 (Delayed ISTAT):** For every  $\theta \in \mathbb{R}$  and  $x \in l_\infty$  such that  $(\theta, x) \sim \theta$ ,  $\lambda(\theta, x) + (1 - \lambda)(\theta, \theta, x) \sim \theta$  for every  $\lambda \in [0, 1]$ .

**A.11 (Existence of a future reference plan (EFREP)):** For every  $\theta \in \mathbb{R}$  and  $x \in l_\infty$  with  $x \sim \theta$ , there exists a future reference  $y \in l_\infty$  of  $x$  such that (i)  $(\theta, y) \sim x$ ; (ii) for every  $z \in l_\infty$  with  $(\theta, z) \sim \theta$ , if  $\lambda x + (1 - \lambda)(\theta, z) \succ \theta$  for every  $\lambda \in (0, 1)$ , then  $\lambda(\theta, y) + (1 - \lambda)(\theta, \theta, z) \succ \theta$  for every  $\lambda \in (0, 1)$ .

**A.12 (Additivity):** For every  $(x, y) \in l_\infty^2$  such that  $x \sim y$ , then for every  $z \in l_\infty$ ,  $x + z \sim y + z$ .

**A.13 (Weak separability):** For every  $\theta \in \mathbb{R}$  and  $x \in l_\infty$  such that  $(\theta, x) \sim \theta$ , then for every  $\alpha \in \mathbb{R}$ ,  $(\alpha, x) \sim (\alpha, \theta)$ .

### 3.2 The main results

The following theorem shows that under Maxmin axioms, Delayed ISTAT and EFREP, we obtain the maxmin representation result with quasi-hyperbolic discounting.

**Theorem 4.** *A binary relation  $\succeq$  on  $l_\infty$  satisfies Maxmin axioms, Delayed ISTAT and EFREP if and only if the unique constant equivalent  $I : l_\infty \rightarrow \mathbb{R}$  representing  $\succeq$  can be written as follows:*

$$I(x) = \min_{(p_0, \delta) \in Q} \{p_0 x_0 + (1 - p_0)(1 - \delta)x_1 + (1 - p_0)(1 - \delta)\delta x_2 + (1 - p_0)(1 - \delta)\delta^2 x_3 \dots\}$$

where  $Q \subset (0, 1)^2$  is nonempty and closed in  $[0, 1]^2$ . Moreover, the set of all long term discount factors,  $E = \{\delta \in (0, 1) : \text{there exists } p_0 \in (0, 1) \text{ such that } (p_0, \delta) \in Q\}$ , is uniquely defined by  $\succeq$ .

The following example shows that we cannot have a unique set  $Q$  in Theorem 4.

**Example 2.** Consider the case where the social planner is sure about the long run discount factor,  $\delta$ , and the weight she puts on generation 0 varies in the interval  $[\underline{p}_0, \overline{p}_0]$ . Assume that she evaluates streams of utility according to the criterion stated in Theorem 4.

$$I(x) = \min_{p_0 \in [\underline{p}_0, \overline{p}_0]} \{p_0 x_0 + (1 - p_0)(1 - \delta)x_1 + (1 - p_0)(1 - \delta)\delta x_2 + (1 - p_0)(1 - \delta)\delta^2 x_3 \dots\}.$$

It is clear that we can also write

$$I(x) = \min_{p_0 \in \{\underline{p}_0, \overline{p}_0\}} \{p_0 x_0 + (1 - p_0)(1 - \delta)x_1 + (1 - p_0)(1 - \delta)\delta x_2 + (1 - p_0)(1 - \delta)\delta^2 x_3 \dots\}.$$

The evaluation depends only on the lowest and the highest short run discount factors.

In Theorem 4, if one imposes that the set of present bias parameters and discount factors is a product set, then we can get rid of the existential axiom EFREP, by replacing it by Weak separability axiom.

**Theorem 5.** *A binary relation  $\succeq$  on  $l_\infty$  satisfies Maxmin axioms, Delayed ISTAT, and Weak separability if and only if the unique constant equivalent  $I : l_\infty \rightarrow \mathbb{R}$  representing  $\succeq$  can be written as follows:*

$$I(x) = \min_{p_0 \in A, \delta \in B} \{p_0 x_0 + (1 - p_0)(1 - \delta)x_1 + (1 - p_0)(1 - \delta)\delta x_2 + (1 - p_0)(1 - \delta)\delta^2 x_3 \dots\},$$

---

<sup>11</sup>Stationarity means that, for every  $\theta \in \mathbb{R}$  and  $(x, y) \in l_\infty^2$ ,  $(\theta, x) \sim (\theta, y)$  if and only if  $x \sim y$ .

where  $A = [p_0, \bar{p}_0] \subset (0, 1)$ , and  $B \subset (0, 1)$  is nonempty and closed in  $[0, 1]$ . Moreover,  $p_0$ ,  $\bar{p}_0$  and the set  $B$  are uniquely defined by  $\succeq$ .

As explained in the introduction, the ratio  $\beta = \frac{(1-p_0)(1-\delta)}{p_0\delta}$  measures *present bias* of the preferences. Remark that in Theorem 4 and Theorem 5, we allow not only  $\beta \leq 1$  (case of present bias) but also  $\beta \geq 1$  (case of future bias). For example, when  $\beta$  converges to infinity, present can be neglected with respect to future. It is important to note that the Chambers and Echenique's exponential maxmin theorem requires to take  $\beta = 1$  (no present bias), i.e., with our notations to take  $p_0 = 1 - \delta$ . This particular case is possible with Theorem 4 (because the general form of the constraint in this theorem authorizes  $p_0$  to be a function of  $\delta$ ), but it is not compatible with Theorem 5 (where  $p_0$  and  $\delta$  are independent).

More generally, it is easy to see that Theorem 4 implies Theorem 5, and in particular, Maxmin axioms, Weak separability, and Delayed ISTAT implies EFREP. Indeed, both theorems provide axioms under which a preorder can be represented by a functional  $I(x) = \min_{p \in \mathbb{D}} \langle p, x \rangle$ , where  $\mathbb{D}$  is a uniquely defined convex and weakly\* compact subset of  $\Delta^\sigma$  (see Theorem 3). Denoting  $g(p_0, \delta) = (p_0, (1-p_0)(1-\delta), (1-p_0)(1-\delta)\delta, (1-p_0)(1-\delta)\delta^2, \dots)$ , then in Theorem 5,  $\mathbb{D} = \overline{\text{co}(\{g(p_0, \delta) : (p_0, \delta) \in A \times B\})}$ , and in Theorem 4,  $\mathbb{D} = \overline{\text{co}(\{g(p_0, \delta) : (p_0, \delta) \in Q\})}$  (where  $Q \subset (0, 1)^2$  is nonempty and closed in  $[0, 1]^2$ ). The latter is more general than the former: for example, if we take  $Q = \{(p_0, \delta), (p'_0, \delta')\}$  in Theorem 4 (with  $p_0 \neq p'_0$  and  $\delta \neq \delta'$ ), then there does not exist any sets  $A$  and  $B$  such that the functional  $I$  can be written as in Theorem 5.

### 3.3 Discussion on the main axioms

#### 3.3.1 Monotonicity

The first statement in *Monotonicity* can be seen as a unanimity property among generations: if each generation  $t \geq 0$  prefers weakly  $x_t$  to  $y_t$ , then the decision maker should prefer weakly stream  $x$  to stream  $y$ . The second statement,  $\mathbf{1} \succ \mathbf{0}$ , avoids the case where all sequences are equivalent. A similar assumption is usually studied in the literature, and is sometimes called *Nondegeneracy* assumption. *Strong Monotonicity* is just a reinforcement of Monotonicity which prevents the decision maker from being indifferent between the two streams  $x$  and  $y$  while  $x$  is strictly preferred to  $y$  by all generations. In our maxmin representation results, Strong monotonicity guarantees that each probability distribution in the set of priors  $\mathbb{D}$  puts a strictly positive weight on each generation.

#### 3.3.2 Continuity

*Continuity and Monotone continuity* are technical assumptions which ensure some kind of stability of preferences. In Appendix 5.9, we provide several equivalent reformulations of our Continuity assumption, and most of them can be found in the literature. Regarding Monotone continuity, a similar but stronger axiom was introduced by Villegas (1964) and Arrow (1970), which requires the same property as ours but replacing the constant act  $\theta$  by any act  $y$ .

#### 3.3.3 Constant translation invariance and Homotheticity

*Constant translation invariance* is equivalent to Weak certainty independence (see the proof in Appendix 5.7), the axiom introduced by Maccheroni et al. (2006) which requires that for every  $(x, y) \in l_\infty^2$ ,  $(\theta, \theta') \in$

$\mathbb{R}^2$ , and  $\lambda \in (0, 1)$ ,  $\lambda x + (1 - \lambda)\theta \succeq \lambda y + (1 - \lambda)\theta$  implies that  $\lambda x + (1 - \lambda)\theta' \succeq \lambda y + (1 - \lambda)\theta'$ . [Homotheticity and Constant translation invariance] is actually equivalent, under axioms **A.1-A.3**, to *Certainty independence* (see the proof in Appendix 5.8), an axiom introduced by Gilboa and Schmeidler (1989) in their axiomatization of maxmin criterion; Certainty independence means that for every  $(x, y) \in l_\infty^2$ ,  $\theta \in \mathbb{R}$ , and  $\lambda \in (0, 1)$ ,  $x \succeq y$  implies that  $\lambda x + (1 - \lambda)\theta \succeq \lambda y + (1 - \lambda)\theta$ . Homotheticity plus Constant translation invariance is equivalent to *Co-Cardinality* axiom (see Chambers and Echenique (2018)), which requires that the order between two streams is not modified by any positive affine transformations of the streams.

### 3.3.4 Convexity

In our intertemporal setting, *Convexity* can be seen as a fairness assumption: for example, assume that  $(1, -1, 1, -1, \dots)$  and  $(-1, 1, -1, 1, \dots)$  are equivalent for the decision maker, then the mixed distribution  $\frac{1}{2}(1, -1, 1, -1, \dots) + \frac{1}{2}(-1, 1, -1, 1, \dots) = (0, 0, \dots)$  should be preferable (in a weak sense). This corresponds to the intuition that the decision maker prefers to equalize (or “smooth”) the utilities across generations. Remark that Convexity is weaker than *Uncertainty aversion*, introduced by Schmeidler (1989), and which requires that for every  $(x, y, z) \in l_\infty^3$  with  $x \succeq z$  and  $y \succeq z$ , and for every  $\lambda \in [0, 1]$ , we have  $\lambda x + (1 - \lambda)y \succeq z$ . Under axioms **A.1-A.4**, it is equivalent to Convexity.

### 3.3.5 Weak stationarity and ISTAT

*ISTAT* requires that the decision maker is indifferent between  $x$  and  $(\theta, x)$ , the delayed version of  $x$  (where  $\theta$  is a constant equivalent of  $x$ ), and also that “there can be no hedging or smoothing value to mixing stream  $x$  with the delayed version of itself  $(\theta, x)$ ”. *Weak stationarity* just requires the first property of ISTAT (indifference between  $x$  and  $(\theta, x)$ ). The interpretation of ISTAT is discussed in Chambers and Echenique (2018) as a version of anonymity. The first part of ISTAT (Weak stationarity) means that “some particular relabeling of agents’ names does not matter”, and the second part of ISTAT means that “there is no additional smoothing motive in the comparisons of  $x$  with  $\theta$ , before or after a stationary relabeling”.

### 3.3.6 Delayed ISTAT

*Delayed ISTAT* is similar to ISTAT, but delayed at period 1, i.e., it says that some particular relabeling of all agents’ names – except the first one – does not matter. It is an intuitive property in our framework, since the quasi-hyperbolic criterion is similar to a geometric discounting when starting from period 1; as a consequence, we could expect the maxmin quasi-hyperbolic criterion to satisfy some “delayed” stationarity property. As an illustration, if we consider  $I(x) = \frac{x_0 + x_1}{2}$  for every  $x \in l_\infty$ , then Delayed ISTAT is satisfied: indeed, if  $(\theta, x) \sim \theta$ , then an immediate computation gives  $x_0 = \theta$  and  $I(\lambda(\theta, x) + (1 - \lambda)(\theta, \theta, x)) = \frac{\theta + \lambda x_0 + (1 - \lambda)\theta}{2} = \theta$ . Yet, ISTAT is not satisfied: indeed, consider  $x = (0, 2\theta, 0, 0, \dots)$  with  $\theta \neq 0$ , so that  $x \sim \theta$ . We have  $(\theta, x) = (\theta, 0, 2\theta, 0, 0, \dots)$ , so that  $I(\theta, x) = \frac{\theta}{2} \neq \theta$ . More generally, we have the following important relationship between ISTAT, Delayed ISTAT and Weak stationarity:

**Proposition 2.** *Under Maxmin axioms, ISTAT is equivalent to [Delayed ISTAT and Weak stationarity].*

*Proof.* Assume that  $\succeq$  satisfies Maxmin axiom. Obviously, ISTAT implies [Delayed ISTAT and Weak stationarity]. Now, let us prove the other implication. Suppose that Delayed ISTAT and Weak stationarity are satisfied. Let  $\theta \in \mathbb{R}$  and  $x \in l_\infty$  such that  $x \sim \theta$ . Weak stationarity implies that  $(\theta, x) \sim \theta$ , and by Convexity, we get  $\lambda x + (1 - \lambda)(\theta, x) \succeq \theta$  for every  $\lambda \in (0, 1)$ .

Now, we assume that there exists  $\lambda \in (0, 1)$  such that  $\lambda x + (1 - \lambda)(\theta, x) \succ \theta$ . Let  $u = \lambda x + (1 - \lambda)(\theta, x)$ , and let  $\theta' \in \mathbb{R}$  such that  $u \sim \theta'$ . Then we have  $\theta' > \theta$  by Strong monotonicity. Weak stationarity implies that  $(\theta', u) \sim \theta'$ , which is equivalent to  $(0, u - \theta') \sim \mathbf{0}$  by Constant translation invariance. Thus, we get that, by Strong monotonicity,  $(0, u - \theta) \succ (0, u - \theta') \sim \mathbf{0}$ ; it follows by Constant translation invariance that  $(\theta, u) \succ \theta$ . Let us replace  $u$  by  $\lambda x + (1 - \lambda)(\theta, x)$  to get  $\lambda(\theta, x) + (1 - \lambda)(\theta, \theta, x) \succ \theta$ , which violates Delayed ISTAT. Therefore, we must have  $\lambda x + (1 - \lambda)(\theta, x) \sim \theta$  for every  $\lambda \in (0, 1)$ , and the proof is finished.  $\square$

This proposition shows that under Maxmin axioms, we can recover ISTAT from Delayed ISTAT and Weak stationarity.

### 3.3.7 EFREP

*EFREP* says that for every stream  $x$ , there should exist some future reference plan  $y$ , such that (i)  $x$  and  $(\theta, y)$  are indifferent from the decision maker, and (ii) for every stream  $(\theta, z) \sim \theta$  which is a strict hedging against  $x$ , the delayed version  $(\theta, \theta, z)$  of  $(\theta, z)$  should be a strict hedging against  $(\theta, y)$ , the “smooth equivalent” of  $x$ . Condition (i) is exactly Weak stationarity when  $y = x$ , so that we can see it as a “distorted” Weak stationarity property. Remark that EFREP with (i) alone is always true since we can take  $y = \theta$ . About Condition (ii), first note that the existence of strict hedging against  $x$  captures the non-additivity of the model (since when the preferences are additive, there can be no strict hedging); and Condition (ii) above requires some invariance through time of this hedging property, still considering that the delayed version of  $x$  is distorted into  $(\theta, y)$  (instead of  $(\theta, x)$ ). The distortion comes from present bias since no present bias means no distortion ( $x = y$ ). More precisely, we can explicit  $y$  in some simple cases:

**Proposition 3.** *Under Additivity and Homotheticity, for every  $x \in l_\infty$  and  $\theta \in \mathbb{R}$  such that  $x \sim \theta$ ,  $y = \theta$  is a future reference plan associated to  $x$ . Thus, in this case, EFREP is true.*

*Proof.* Assuming Additivity and Homotheticity, it is immediate that, for every  $(x, z) \in l_\infty^2$  and  $\theta \in \mathbb{R}$  for which  $x \sim (\theta, z) \sim \theta$ , we have  $\lambda x + (1 - \lambda)(\theta, z) \sim \theta$  for every  $\lambda \in [0, 1]$ . It means that it is impossible to have  $\lambda x + (1 - \lambda)(\theta, z) \succ \theta$  for every  $\lambda \in (0, 1)$ . Consequently, Condition (ii) of EFREP is true, and it is straightforward that Condition (i) is also true when  $y = \theta$ .  $\square$

**Proposition 4.** *Under Maxmin axioms, Weak stationarity implies that for every  $x \in l_\infty$ ,  $x$  is a future reference plan associated to itself. Thus, in this case, EFREP is true.*

*Proof.* Assume that  $\succeq$  satisfies Maxmin axioms and Weak stationarity, let  $\theta \in \mathbb{R}$  and  $x \in l_\infty$  such that  $x \sim \theta$ . From Weak stationarity, Condition (i) of EFREP is true for  $y = x$ , i.e.,  $(\theta, x) \sim x$ . To prove Condition (ii), let  $z \in l_\infty$  such that  $(\theta, z) \sim \theta$  and  $u := \lambda x + (1 - \lambda)(\theta, z) \succ \theta$  for every  $\lambda \in (0, 1)$ . We want to prove that  $(\theta, u) \succ \theta$ . Let  $\theta' \in \mathbb{R}$  such that  $u \sim \theta'$ , then by Strong monotonicity, we must have  $\theta' > \theta$ . Weak stationarity implies that  $(\theta', u) \sim \theta'$ , this is equivalent to  $(0, u - \theta') \sim \mathbf{0}$  by Constant translation

invariance. Then by Strong monotonicity, we get  $(0, u - \theta) \succ (0, u - \theta') \sim \mathbf{0}$ , which finally implies that  $(\theta, u) \succ \theta$  by Constant translation invariance.  $\square$

The following example illustrates some axioms above: take  $I(x_0, x_1, \dots) = \min\{x_0, x_1\}$  and  $x = (x_0, x_1, \dots) = (1, -1, 0, \dots)$ . Then we get  $I(x) = -1$ , so that  $\theta = -1$  is a constant equivalent of  $x$ , and  $I(\theta, x) = -1$ . But  $I(\frac{(\theta, x) + x}{2}) = I(0, 0, -1, 0, \dots) = 0 > -1$ , which proves that  $I$  does not satisfy ISTAT. Now, take  $x = (1, 0, 0, \dots)$ , then  $x$  is a strict future hedging against  $x$ . Indeed,  $I(x) = 0$ ,  $I(0, x) = 0$ , and  $I(\lambda x + (1 - \lambda)(0, x)) = \min\{\lambda, 1 - \lambda\} > 0$  for every  $\lambda \in (0, 1)$ . Here, we can interpret this hedging property in terms of inequality aversion: the decision maker prefers to mix between agent 0 and agent 1, rather than giving everything to one of these agents. Remark that this hedging property cannot be translated in time, i.e.,  $(0, x)$  is no more a strict future hedging against  $(0, x)$  (since  $I(\lambda(0, x) + (1 - \lambda)(0, 0, x)) = 0$ ). Our main axiom EFREP assumes the existence of a future equivalent of  $x$  which allows to recover that hedging can be translated in time.

### 3.3.8 Additivity

In many models like Gilboa and Schmeidler maxmin model, *Additivity* corresponds to the case where the set of priors is a singleton, modeling the situation where there is no ambiguity. We will see, in the next section, that it is also the case in our main representation theorem. Recall that from Proposition 3, Additivity together with Homotheticity imply EFREP, but Theorem 4 proves that the converse implication is false since the quasi-hyperbolic maxmin criterion does not satisfy, in general, Additivity.

### 3.3.9 Weak separability

If ISTAT is interpreted as “some particular relabeling of agents’ names does not matter”, then Weak separability allows this to be false because there can exist  $\theta \in \mathbb{R}$  and  $x \in l_\infty$ , with  $x \sim \theta$ , such that  $(\theta, x) \sim \theta$  is false. But when it is true, Weak separability strengthens ISTAT at stream  $x$ , requiring that for every  $\alpha \in \mathbb{R}$ ,  $(\alpha, x) \sim (\alpha, \theta)$ ; this means that the level of utility of the first agent does not matter in the previous equivalence. Thus, Weak separability is neither less general, nor more general, than ISTAT axiom. Weak separability can also be seen as a Weak additivity axiom: it says that if the equivalence  $(\theta, x) \sim \theta$  is true, then it should remain true if we add any stream  $(y_0, 0, \dots)$  to the equivalence. In particular, Additivity implies Weak separability.

## 3.4 Two consequences of Theorem 4

In this subsection we study two important results which are special cases of Theorem 4. The first one is the axiomatization of the exponential maxmin criterion by Chambers and Echenique (see Theorem 3 in Chambers and Echenique (2017)), and the second one is the axiomatization of the quasi-hyperbolic discounting model (see Montiel Olea and Strzalecki (2014)).

### 3.4.1 The maxmin representation result with geometric discounting

If we reinforce [Delayed ISTAT and EFREP] into ISTAT in Theorem 4, then we get the axiomatization of the exponential maxmin criterion introduced in Chambers and Echenique (2017).

**Corollary 1.** *A binary relation  $\succeq$  on  $l_\infty$  satisfies Maxmin axioms and ISTAT if and only if the unique constant equivalent  $I : l_\infty \rightarrow \mathbb{R}$  representing  $\succeq$  can be written as follows:*

$$I(x) = \min_{\delta \in E} (1 - \delta) \sum_{t=0}^{\infty} \delta^t x_t,$$

where  $E \subset (0, 1)$  is non-empty and closed in  $[0, 1]$ . Moreover, the set  $E$  is uniquely defined by  $\succeq$ .

*Proof.* We only prove the *only-if* part. Assume that  $\succeq$  satisfies all axioms in the Corollary. From Proposition 2, Proposition 4, and Theorem 4, there exists  $Q \subset (0, 1)^2$  nonempty and closed in  $[0, 1]^2$  such that

$$I(x) = \inf_{(p_0, \delta) \in Q} \{p_0 x_0 + (1 - p_0)(1 - \delta)x_1 + (1 - p_0)(1 - \delta)\delta x_2 + (1 - p_0)(1 - \delta)\delta^2 x_3 \dots\}.$$

Applying Weak stationarity with  $\theta := I(x)$ , we get  $\theta = I((\theta, x))$ . From the definition of  $I$ , we have  $\langle (\theta, x), g(p_0, \delta) \rangle \geq \theta$  for every  $(p_0, \delta) \in Q$ , where  $g(p_0, \delta) = (p_0, (1 - p_0)(1 - \delta), (1 - p_0)(1 - \delta)\delta, \dots)$ , with equality for at least one  $(p_0, \delta) \in Q$ . An easy computation implies that  $\langle x, \boldsymbol{\delta} \rangle \geq \theta$  for every  $(p_0, \delta) \in Q$ , where  $\boldsymbol{\delta} = (1 - \delta)(1, \delta, \delta^2, \dots)$ , with equality for at least such one  $(p_0, \delta) \in Q$ . Thus,

$$I(x) = \min_{\delta \in E} (1 - \delta) \sum_{t=0}^{\infty} \delta^t x_t,$$

where  $E = \{\delta \in (0, 1) : \text{there exists } p_0 \in (0, 1) \text{ such that } (p_0, \delta) \in Q\}$ . Last, the uniqueness of  $E$  is immediate from the uniqueness result of Theorem 4.  $\square$

From Proposition 2, we know that ISTAT is equivalent to [Delayed ISTAT and Weak stationarity]. To obtain our axiomatization of maxmin result with quasi-hyperbolic discounting, we keep Delayed ISTAT and introduce a strictly weakening version of Weak stationarity, namely EFREP (see Proposition 4). Therefore, Weak stationarity is the key axiom which allows us to differentiate the Chambers and Echenique's exponential maxmin behavior from our quasi-hyperbolic maxmin behavior.

### 3.4.2 The representation of quasi-hyperbolic preferences

We can easily get the representation theorem of quasi-hyperbolic preferences by imposing Additivity in Theorem 4.

**Corollary 2.** *A binary relation  $\succeq$  on  $l_\infty$  satisfies Maxmin axioms, Delayed ISTAT, and Additivity if and only if there exists  $(p_0, \delta) \in (0, 1)^2$  such that the unique constant equivalent  $I : l_\infty \rightarrow \mathbb{R}$  representing  $\succeq$  can be written as follows:*

$$I(x) = p_0 x_0 + (1 - p_0)(1 - \delta)x_1 + (1 - p_0)(1 - \delta)\delta x_2 + (1 - p_0)(1 - \delta)\delta^2 x_3 \dots$$

*Proof.* We only prove the *only-if* part. Assume that  $\succeq$  satisfies all axioms in the Corollary. We know, from Theorem 4 and Proposition 3, that the unique constant equivalent representing  $\succeq$ ,  $I : l_\infty \rightarrow \mathbb{R}$ , can

be written as follows:

$$I(x) = \min_{(p_0, \delta) \in Q} \{p_0 x_0 + (1 - p_0)(1 - \delta)x_1 + (1 - p_0)(1 - \delta)\delta x_2 + (1 - p_0)(1 - \delta)\delta^2 x_3 \dots\},$$

where  $Q \subset (0, 1)^2$  is nonempty and closed in  $[0, 1]^2$ . But it is easy to see, from Additivity, that  $I$  is additive (i.e.,  $I(x + y) = I(x) + I(y)$  for every  $(x, y) \in l_\infty^2$ ); thus, from Homotheticity,  $I$  is a linear function. Since the topological dual of  $l_\infty$  is isomorphic to the set of signed charges, there exists some  $\mu \in ba(\mathbb{N})$  such that  $I(x) = \int x d\mu$ . Now, Monotone continuity implies that  $\mu$  has to be countably additive; and since  $I$  is normalized (i.e.,  $I(1) = 1$ ), we finally get  $\mu \in \Delta^\sigma$ . Thus, we can write

$$I(x) = \min_{p \in \{\mu\}} \int x dp.$$

From the uniqueness of the set of probability distributions in Theorem 3, the set  $Q$  must be a singleton, which ends the proof.  $\square$

## 4 Conclusion

Our work can be seen as a relaxation of additivity in quasi-hyperbolic models, by incorporating some ambiguity in the present bias and the discount rate parameter. We also relax the stationarity assumption in Chambers and Echenique’s maxmin model, namely, ISTAT: if this assumption is interpreted as “delaying the start of a stream should leave the society indifferent”, then our weakening axiom consists in delaying this requirement for one period. We also require some kind of “distorted” ISTAT axiom, which means that delaying the start of a stream  $x$  should leave the society indifferent if the future translated stream  $x$  can be replaced by some equivalent stream  $y$ , interpreted as a distortion of  $x$  due to present bias. Last, we have proved that if we allow the present bias parameters and the discount factors to be independent, then the axiomatization is largely simplified since ISTAT can be replaced by Weak separability.

## 5 Appendix

### 5.1 Proof of Proposition 1

In the proof, we will use the following lemma:

**Lemma 1.** *For every  $(\theta', \theta) \in \mathbb{R}^2$  for which  $\theta' > \theta$ , we have  $\theta' \succ \theta$ .*

*Proof.* We prove this lemma by contradiction: assume  $\theta' > \theta$  and  $\theta \succeq \theta'$ . Then by using **A.4**, we have  $\mathbf{0} \succeq \varepsilon$ , where  $\varepsilon = \theta' - \theta > 0$ . From **A.1** and **A.4**, we can easily get  $\mathbf{0} \succeq n\varepsilon$  for every  $n \in \mathbb{N}$ . However, Monotonicity implies  $n\varepsilon \succeq \mathbf{1}$  for some  $n$  large enough. Thus, we get  $\mathbf{0} \succeq \mathbf{1}$  which is a contradiction with Monotonicity (part (2)).  $\square$

Now, to prove Proposition 1, assume that  $\succeq$  satisfies **A.1-A.5**, and let us prove that there exists  $I : l_\infty \rightarrow \mathbb{R}$  such that for every  $x$  in  $l_\infty$ ,  $x \sim \mathbf{I}(x)$ . First, we show that, for every  $(x, y, z) \in l_\infty^3$  (now fixed) for which  $x \succeq y \succeq z$ , there exists  $\alpha \in [0, 1]$  such that  $y \sim \alpha x + (1 - \alpha)z$ . By **A.3**, the sets



$\{\beta \in [0, 1] : \beta x + (1 - \beta)z \succeq y\}$  and  $\{\beta \in [0, 1] : y \succeq \beta x + (1 - \beta)z\}$  are closed, they are non-empty (1 belongs to the first one, 0 to the second one), and their union is the whole set  $[0, 1]$ . Since  $[0, 1]$  is connected, their intersection is non-empty, it follows that there exists  $\alpha \in [0, 1]$  such that  $y \sim \alpha x + (1 - \alpha)z$ . Now, let  $x \in l_\infty$ . By Monotonicity,  $\bar{\mathbf{x}} \succeq x \succeq \underline{\mathbf{x}}$ , in which  $\bar{x} = \sup_{t \geq 0} \{x_t\}$  and  $\underline{x} = \inf_{t \geq 0} \{x_t\}$ . Thus, there exists  $\alpha \in [0, 1]$  such that  $x \sim \alpha \bar{\mathbf{x}} + (1 - \alpha)\underline{\mathbf{x}}$ , which proves that there exists  $I : l_\infty \rightarrow \mathbb{R}$  such that for every  $x$  in  $l_\infty$ ,  $x \sim \mathbf{I}(\mathbf{x})$  (by simply defining  $I(x) = \alpha \bar{x} + (1 - \alpha)\underline{x}$ ).

Let us prove that  $I$  is uniquely defined. If there exists  $J$  such that  $x \sim \mathbf{J}(\mathbf{x})$ , then  $\mathbf{I}(\mathbf{x}) \sim \mathbf{J}(\mathbf{x})$ . It must be that  $I(x) = J(x)$ ; otherwise, we would have, without any loss of generality,  $I(x) > J(x)$ , implying that  $\mathbf{I}(\mathbf{x}) \succ \mathbf{J}(\mathbf{x})$  by Lemma 1, a contradiction. Remark that to prove the existence of the unique constant equivalent  $I$ , we only need axioms **A.1-A.3**.

We now prove that for every  $(x, y) \in l_\infty^2$ ,  $x \succeq y$  if and only if  $I(x) \geq I(y)$  (i.e.,  $I$  represents  $\succeq$ ). For the first implication, consider  $(x, y) \in l_\infty^2$  for which  $x \succeq y$ . From **A.1**, we have  $\mathbf{I}(\mathbf{x}) \succeq \mathbf{I}(\mathbf{y})$ , then it follows from Lemma 1 that  $I(x) \geq I(y)$ . Conversely, let  $(x, y) \in l_\infty^2$  with  $I(x) \geq I(y)$ . By Monotonicity,  $\mathbf{I}(\mathbf{x}) \succeq \mathbf{I}(\mathbf{y})$ , which implies, from **A.1**, that  $x \succeq y$ . Therefore,  $I$  represents  $\succeq$ .

To prove that  $I$  is weakly increasing, let  $(x, y) \in l_\infty^2$  for which  $x \geq y$ , then **A.1** and **A.2** imply that  $\mathbf{I}(\mathbf{x}) \succeq \mathbf{I}(\mathbf{y})$ . From Lemma 1, we can conclude that  $I(x) \geq I(y)$ .

Let us prove translation invariance by contradiction: without any loss of generality, we can assume that there exists  $x \in l_\infty$  and  $\theta \in \mathbb{R}$  such that  $I(x + \theta) > I(x) + \theta$ . But since  $I$  represents  $\succeq$ , we have  $x + \theta \succ \mathbf{I}(\mathbf{x}) + \theta$ , that is  $x \succ \mathbf{I}(\mathbf{x})$  (from **A.4**), a contradiction.

Now, since  $I$  is increasing and translation invariant, and for every  $(x, y) \in l_\infty^2$ , we have  $x \leq y + \|x - y\|_\infty$ , we get  $I(x) - I(y) \leq \|x - y\|_\infty$ . Permuting  $x$  and  $y$ , we obtain  $|I(x) - I(y)| \leq \|x - y\|_\infty$ , i.e.,  $I$  is 1-lipschitz.

To prove that  $I$  is normalized, simply remark that from the definition of the constant equivalent  $I$ , we have  $\mathbf{I}(\mathbf{1}) \sim \mathbf{1}$ , and finally Lemma 1 implies  $I(\mathbf{1}) = 1$ .

Last, to prove concavity, let  $(x, y) \in l_\infty^2$  and  $\lambda \in [0, 1]$ . Since  $x \sim \mathbf{I}(\mathbf{x})$  and  $x \sim y + \mathbf{I}(\mathbf{x}) - \mathbf{I}(\mathbf{y})$ , we get  $x \succeq \mathbf{I}(\mathbf{x})$  and  $y + \mathbf{I}(\mathbf{x}) - \mathbf{I}(\mathbf{y}) \succeq \mathbf{I}(\mathbf{x})$ . Consequently, **A.5** implies that

$$\lambda x + (1 - \lambda)(y + \mathbf{I}(\mathbf{x}) - \mathbf{I}(\mathbf{y})) \succeq \mathbf{I}(\mathbf{x}),$$

or equivalently

$$I[\lambda x + (1 - \lambda)(y + \mathbf{I}(\mathbf{x}) - \mathbf{I}(\mathbf{y}))] \geq I(x).$$

From the translation invariance of  $I$ , a simple computation finally gives  $I(\lambda x + (1 - \lambda)y) \geq \lambda I(x) + (1 - \lambda)I(y)$ , i.e.,  $I$  is concave.

The other implication (if there exists a weakly increasing, concave, 1-lipschitz, normalized, and translation invariant constant equivalent  $I : l_\infty \rightarrow \mathbb{R}$  that represents  $\succeq$ , then  $\succeq$  satisfies **A.1-A.5**) is left to the reader.

## 5.2 Proof of Theorem 1

Assume that  $\succeq$  satisfies **A.1-A.5**. Let  $I : l_\infty \rightarrow \mathbb{R}$  be the unique constant equivalent of  $\succeq$  defined by Proposition 1.

First, we start with an application of the generalized Moreau duality. Let  $X$  be a topological vector space,  $F$  be a nonempty subset of  $\mathbb{R}^X$ , and  $\varphi : X \rightarrow \mathbb{R}$  be a function. We then recall that the  $F$ -superdifferential at  $x \in X$  of  $\varphi$  is

$$\partial^F \varphi(x) = \{f \in F : \varphi(y) - \varphi(x) \leq \langle y, f \rangle - \langle x, f \rangle \text{ for every } y \in X\},$$

where  $\langle x, f \rangle = f(x)$ . The  $F$ -conjugate of  $\varphi$ ,  $\varphi^F : F \rightarrow [-\infty, \infty]$ , is defined by

$$\varphi^F(f) = \inf_{x \in X} \{\langle x, f \rangle - \varphi(x)\}.$$

When  $F = X'$  (the topological dual of  $X$ ), then  $\partial^F \varphi$  is equal to  $\partial \varphi$ , the superdifferential of  $\varphi$ , and  $\varphi^F$  is equal to  $\varphi^*$ , the Fenchel conjugate<sup>12</sup> of  $\varphi$ . The following lemma plays a central role in the proof of Theorem 1 (for a proof, see Lemma 4 in [Cerreia-Vioglio et al. \(2014\)](#)):

**Lemma 2.** *If  $\partial^F \varphi(x) \neq \emptyset$  for every  $x \in X$ , then*

$$\begin{aligned} \varphi(x) &= \min_{f \in F} \{\langle x, f \rangle - \varphi^F(f)\}, \text{ and} \\ \partial^F \varphi(x) &= \arg \min_{f \in F} \{\langle x, f \rangle - \varphi^F(f)\} \end{aligned}$$

for every  $x \in X$ . Furthermore,  $\varphi^F$  is the maximal function among all functions  $R : F \rightarrow \mathbb{R} \cup \{-\infty\}$  satisfying

$$\varphi(x) = \min_{f \in F} \{\langle x, f \rangle - R(f)\}.$$

To prove Theorem 1, we define  $X = l_\infty$ ,  $F = \Delta$ ,<sup>13</sup> and  $\varphi = I$ , and we apply Lemma 2 which requires to prove that  $\partial^\Delta I(x)$  is non-empty for every  $x \in l_\infty$ . Let  $x \in l_\infty$ , we have  $\partial^\Delta I(x) = \partial I(x) \cap \Delta$ , in which

$$\partial I(x) = \{\mu \in (l_\infty)' = ba(\mathbb{N}) : I(y) - I(x) \leq \langle y - x, \mu \rangle \text{ for every } y \in l_\infty\}$$

is the superdifferential of  $I$  at  $x$ . Since  $x$  is in the interior of the domain of the proper<sup>14</sup> concave continuous function  $I$ , we have  $\partial I(x) \neq \emptyset$  (see [Aliprantis and Border \(2006\)](#), Theorem 7.12). Let us now prove that  $\partial I(x) \subseteq \Delta$ . If  $\mu \in \partial I(x)$ , then we have  $I(x + \mathbf{1}) - I(x) \leq \langle \mathbf{1}, \mu \rangle$  and  $I(x - \mathbf{1}) - I(x) \leq \langle -\mathbf{1}, \mu \rangle$ . From the translation invariance of  $I$ ,  $I(x + \mathbf{1}) - I(x) = 1$  and  $I(x - \mathbf{1}) - I(x) = -1$ , thus, we get  $\langle \mathbf{1}, \mu \rangle = 1$ , i.e.,  $\mu$  has a total mass equal to 1. Second, let  $(y, z) \in l_\infty^2$  such that  $y \geq z$ . Since  $I$  is monotone,  $0 \leq I(y - z + x) - I(x) \leq \langle y - z, \mu \rangle$ , so that  $\mu$  is monotone. We conclude that  $\mu \in \Delta$ . Hence,  $\partial^\Delta I(x) = \partial I(x) \neq \emptyset$ .

From Lemma 2, we can write

$$I(x) = \min_{p \in \Delta} \{\langle x, p \rangle - I^\Delta(p)\},$$

<sup>12</sup>The Fenchel conjugate can also be defined by using an supremum rather than an infimum. Transforming one version of the Fenchel conjugate into the other is just a matter of introducing minus signs at the right place.

<sup>13</sup>Actually,  $F$  has to be a subset of  $\mathbb{R}^X$ , but an element  $p \in \Delta$  can be identified with the function  $p : x \in X \mapsto \langle x, p \rangle$ .

<sup>14</sup> $I$  is proper by definition because it does not take infinite values.

where

$$I^\Delta(p) = \inf_{x \in l_\infty} \{\langle x, p \rangle - I(x)\}$$

for every  $p \in \Delta$ . The pointwise infimum of a family of (weakly<sup>\*</sup>) upper semicontinuous affine functions is upper semicontinuous and concave. Thus,  $I^\Delta$  is weakly<sup>\*</sup> upper semicontinuous and concave on  $\Delta$ . It follows that  $c(p) := -I^\Delta(p) = \sup_{x \in l_\infty} \{I(x) - \langle x, p \rangle\}$  is convex and weakly<sup>\*</sup> lower semicontinuous, and we get

$$I(x) = \min_{p \in \Delta} \{\langle x, p \rangle + c(p)\}.$$

In addition, from the definition of  $c$ , we have  $c(p) \geq I(\mathbf{1}) - \langle \mathbf{1}, p \rangle = 0$  for every  $p \in \Delta$ , which proves that  $c$  is non-negative. Thus, from  $1 = I(\mathbf{1}) = \min_{p \in \Delta} \{\langle \mathbf{1}, p \rangle + c(p)\} = 1 + \min_{p \in \Delta} \{c(p)\}$ , we get  $\min_{p \in \Delta} \{c(p)\} = 0$ , i.e.,  $c$  is grounded.

Last, we will prove the uniqueness result. Let  $\tilde{c} : \Delta \rightarrow [0, \infty]$  is a convex and weakly<sup>\*</sup> lower semicontinuous function such that

$$I(x) = \min_{p \in \Delta} \{\langle x, p \rangle + \tilde{c}(p)\}.$$

Then by defining  $R = -\tilde{c}$ , we get that  $R : \Delta \rightarrow \mathbb{R} \cup \{-\infty\}$  is a weakly<sup>\*</sup> upper semicontinuous and concave function such that  $I(x) = \min_{p \in \Delta} \{\langle x, p \rangle - R(p)\}$ , i.e.,  $I = R^\Delta$  is the  $\Delta$ -conjugate of  $R$ .

Define  $\bar{R} : ba(\mathbb{N}) \rightarrow \mathbb{R} \cup \{-\infty\}$  by  $\bar{R}(\mu) = R(\mu)$  for every  $\mu \in \Delta$  and  $\bar{R}(\mu) = -\infty$  for every  $\mu \notin \Delta$ . We have  $I(x) = \min_{\mu \in ba(\mathbb{N})} \{\langle x, \mu \rangle - \bar{R}(\mu)\} = \bar{R}^*(x)$ , implying

$$-c(\mu) = I^\Delta(\mu) = \inf_{x \in l_\infty} \{\langle x, \mu \rangle - I(x)\} = \inf_{x \in l_\infty} \{\langle x, \mu \rangle - \bar{R}^*(x)\} = \bar{R}^{**}(\mu) = \bar{R}(\mu) = -\tilde{c}(\mu)$$

for every  $\mu \in \Delta$  (the equality  $\bar{R}^{**} = \bar{R}$  being true because  $\bar{R}$  is weakly<sup>\*</sup> upper semicontinuous and concave), which finally proves the uniqueness of  $c$ .

The other implication is left to the reader.

### 5.3 Proof of Theorem 2

**If part.** Assume that the constant equivalent representing  $\succeq$  is

$$I(x) = \min_{p \in \Delta^\sigma} \{\langle x, p \rangle + c(p)\}$$

for every  $x \in l_\infty$ , where  $c : \Delta \rightarrow [0, \infty]$  is a grounded, convex, and lower semicontinuous function with  $\{p \in \Delta : c(p) \leq t\} \subset \Delta^\sigma$  for every  $t \geq 0$ . Axioms **A.1-A.5** are obvious. We only have to prove **A.6**. If it is false, then there exists  $x \in l_\infty$  and  $(\theta, k) \in \mathbb{R}^2$  such that  $x \succ \theta$  (thus  $I(x) > \theta$ ), and there exists  $(E_n)_{n \in \mathbb{N}}$  which is a sequence of subsets of  $\mathbb{N}$ , with  $(E_n) \downarrow \emptyset$ , such that  $I(\mathbf{k}E_n x) \leq \theta$  for every  $n \in \mathbb{N}$ . By the definition of  $I$ , for every  $n \in \mathbb{N}$ , there exists  $p^n \in \Delta^\sigma$  such that

$$\langle \mathbf{k}E_n x, p^n \rangle + c(p^n) \leq \theta. \tag{3}$$

In particular, the sequence  $\{c(p^n)\}_{n \geq 0}$  is bounded, that is there exists a real  $t \geq 0$  such that for every  $n \in \mathbb{N}$ ,  $p^n \in \{q \in \Delta : c(q) \leq t\}$ , which is a compact subset of  $\Delta$  for the weak\* topology. Thus, up to a subsequence, we can assume that  $\{p^n\}_{n \geq 0}$  converges to some  $p \in \{q \in \Delta : c(q) \leq t\} \subset \Delta^\sigma$ . Now, passing to the limit in Inequality (3), we get

$$\langle x, p \rangle + c(p) \leq \theta.$$

Thus,  $I(x) \leq \theta$ , a contradiction. Therefore, **A.6** is proved.

**Only-if part.** From Theorem 1, for  $x \in l_\infty$  fixed, we can write

$$I(x) = \min_{p \in \Delta} \{\langle x, p \rangle + c(p)\},$$

and from the proof of Theorem 1, we know that

$$c(p) = \sup_{x \in l_\infty} \{I(x) - \langle x, p \rangle\}.$$

The proof that  $c$  is grounded, convex, and lower semicontinuous is the same as in the proof of Theorem 1.

Let  $p \in \arg \min_{p \in \Delta} \{\langle x, p \rangle + c(p)\} = \partial^\Delta I(x) = \partial I(x)$  (see Lemma 2). From the definition of superdifferential of  $I$  at  $x$ , we get

$$I(y) - I(x) \leq \langle y - x, p \rangle$$

for every  $y \in l_\infty$ . We will prove  $p \in \Delta^\sigma$ . Equivalently, we need to show that if  $(E_n)_{n \in \mathbb{N}}$  is a sequence of subsets of  $\mathbb{N}$  for which  $(E_n) \downarrow \emptyset$ , then  $\langle \mathbf{1}_{E_n}, p \rangle$  converges to 0 as  $n \rightarrow \infty$ .

Let  $k = \inf\{x_n - 1 : n \in \mathbb{N}\}$ . Then for every  $n \in \mathbb{N}$ , we have (from the above inequality)

$$I(\mathbf{k}E_n x) - I(x) \leq I(x - \mathbf{1}_{E_n}) - I(x) \leq -\langle \mathbf{1}_{E_n}, p \rangle. \quad (4)$$

Let us prove that  $\lim_{n \rightarrow \infty} I(\mathbf{k}E_n x) = I(x)$ . For every  $m \in \mathbb{N}/\{0\}$ ,  $x \succ \mathbf{I}(\mathbf{x}) - \frac{1}{m}$ . Consequently, from **A.6**, there exists an integer  $k_m$  such that  $\mathbf{k}E_{k_m} x \succ \mathbf{I}(\mathbf{x}) - \frac{1}{m}$ , or  $I(\mathbf{k}E_{k_m} x) > I(x) - \frac{1}{m}$ . Notice that  $\mathbf{k}E_n x \leq \mathbf{k}E_{n+1} x < x$  for every  $n \in \mathbb{N}$  since  $k < \inf_{n \in \mathbb{N}} \{x_n\}$ , it follows that  $I(x) \geq \lim_{n \rightarrow \infty} I(\mathbf{k}E_n x) \geq I(x) - \frac{1}{m}$  for every  $m \in \mathbb{N}$ . Thus  $\lim_{n \rightarrow \infty} I(\mathbf{k}E_n x) = I(x)$ .

Consequently, from (4), we obtain that  $\lim_{n \rightarrow \infty} \langle \mathbf{1}_{E_n}, p \rangle = 0$ . Therefore,  $p \in \Delta^\sigma$ , and we can conclude that  $\partial^{\Delta^\sigma} I(x) = \partial I(x) \neq \emptyset$ . It follows from Lemma 2 that

$$I(x) = \min_{p \in \Delta^\sigma} \{\langle x, p \rangle + c(p)\}.$$

Now, we will prove that for every  $t \geq 0$  (now fixed),  $\{q \in \Delta : c(q) \leq t\}$  is a subset of  $\Delta^\sigma$ . By contradiction, assume there exists  $p \in \Delta$  such that  $c(p) \leq t$  and  $p \notin \Delta^\sigma$ . Then, there exists a sequence  $(E_n)_{n \in \mathbb{N}}$  of subsets of  $\mathbb{N}$  with  $(E_n) \downarrow \emptyset$  such that  $(\langle \mathbf{1}_{E_n}, p \rangle)_{n \in \mathbb{N}}$  does not converges to 0. It means that there exists  $\epsilon > 0$  such that  $\langle \mathbf{1}_{E_n}, p \rangle > \epsilon$  for every  $n \geq 0$ . Recall  $c(p) = \sup_{x \in l_\infty} \{I(x) - \langle x, p \rangle\}$ , we get  $t \geq c(p) \geq I(-k\mathbf{1}_{E_n}) + \langle k\mathbf{1}_{E_n}, p \rangle$  for every integer  $n$  and every real  $k > 0$ . Let us prove that  $\lim_{n \rightarrow \infty} I(-k\mathbf{1}_{E_n}) = I(\mathbf{0}) = 0$ . Indeed,  $I(-k\mathbf{1}_{E_n})$  is an increasing sequence, majorized by  $I(\mathbf{0}) = 0$ . If

it converges to  $\alpha < 0$ , then from **A.6** applied to  $\alpha < \mathbf{0}$ , there exists  $n_0$  such that for every  $n \geq n_0$ ,  $I(-k\mathbf{E}_n) = I(-k\mathbf{1}_{E_n}) > \alpha$ , a contradiction, thus we must have  $\lim_{n \rightarrow \infty} I(-k\mathbf{1}_{E_n}) = 0$ . Hence, we get  $t \geq c(p) \geq \lim_{n \rightarrow \infty} k\langle \mathbf{1}_{E_n}, p \rangle \geq k\epsilon$  for every real  $k > 0$ , a contradiction. This finally proves that  $\{q \in \Delta : c(q) \leq t\} \subset \Delta^\sigma$ .

## 5.4 Proof of Theorem 3

Assume that  $\succeq$  satisfies **A.1-A.7**. Let  $x \in l_\infty$ , by Theorem 2, we can write

$$I(x) = \min_{p \in \Delta^\sigma} \{\langle x, p \rangle + c(p)\},$$

where we know from the proof of Theorem 1 that  $c : \Delta \rightarrow [0, \infty]$  is defined by

$$c(p) = \sup_{x \in l_\infty} \{I(x) - \langle x, p \rangle\},$$

and is such that  $\{p \in \Delta : c(p) \leq t\} \subset \Delta^\sigma$  for every  $t \geq 0$ .

Let  $\alpha > 0$ , and let us prove that  $I(\alpha x) = \alpha I(x)$ . By contradiction, we assume  $I(\alpha x) > \alpha I(x)$ . This implies that  $\alpha x \succ \alpha \mathbf{I}(x)$ . By **A.7**, we get  $x \succ \mathbf{I}(x)$ , a contradiction. Similarly, if we assume  $I(\alpha x) < \alpha I(x)$ , then we get  $\mathbf{I}(x) \succ x$ , a contradiction. Therefore, we must have  $I(\alpha x) = \alpha I(x)$ .

Now, let  $p^* \in \arg \min_{p \in \Delta^\sigma} \{\langle x, p \rangle + c(p)\}$ , we will prove that  $c(p^*) = 0$ . By contradiction, assume  $c(p^*) > 0$ , then from the definition of  $c$  above, there exist  $x^* \in l_\infty$  such that  $I(x^*) - \langle x^*, p^* \rangle > 0$ . This implies that  $I(\alpha x^*) - \langle \alpha x^*, p^* \rangle = \alpha [I(x^*) - \langle x^*, p^* \rangle]$ , which converges to  $\infty$  when  $\alpha$  goes to  $\infty$ . From the definition of  $c$ , we get  $c(p^*) = \infty$ , which is a contradiction as  $I(x) = \langle x, p^* \rangle + c(p^*)$ . Therefore,  $c(p^*) = 0$ , and since this is true for every  $p^* \in \arg \min_{p \in \Delta^\sigma} \{\langle x, p \rangle + c(p)\}$ , we finally get

$$I(x) = \min_{p \in \{q \in \Delta^\sigma : c(q) = 0\}} \{\langle x, p \rangle\}.$$

Let  $\mathbb{D} = \{q \in \Delta : c(q) = 0\} = \{q \in \Delta : c(q) \leq 0\}$ , this set is weakly\* closed and convex since  $c : \Delta \rightarrow [0, \infty]$  is weakly\* lower semicontinuous and convex. Moreover, it is clear that  $\mathbb{D}$  is pointwise bounded, i.e., for every  $x \in l_\infty$  the set  $\{\langle x, p \rangle : p \in \mathbb{D}\}$  is bounded in  $\mathbb{R}$ , so it is weakly\* compact. We have proved that  $\mathbb{D} = \{q \in \Delta : c(q) \leq 0\} \subset \Delta^\sigma$ , which implies that  $\mathbb{D} = \{q \in \Delta^\sigma : c(q) \leq 0\} = \{q \in \Delta^\sigma : c(q) = 0\}$ . This ends the proof of the first implication. To prove the uniqueness of the set  $\mathbb{D}$ , see Theorem 1 in [Gilboa and Schmeidler \(1989\)](#).

The other implication is left to the reader.

## 5.5 Proof of Theorem 4

Define the function  $s$  by

$$s(p) = \left( \frac{p_1}{1-p_0}, \frac{p_2}{1-p_0}, \frac{p_3}{1-p_0}, \dots \right)$$

for every  $p \in \Delta^\sigma$  such that  $p_t > 0$  for every  $t \in \mathbb{N}$ . In the proof, we will need the following two lemmas:

**Lemma 3.** For every  $x \in l_\infty$ ,  $\theta \in \mathbb{R}$ , and  $p \in \Delta^\sigma$  such that  $p_t > 0$  for every  $t \in \mathbb{N}$ , we have

$$\langle (\theta, x), p \rangle = \theta \iff \langle x, s(p) \rangle = \theta,$$

and this holds true if we replace the two equations by similar inequalities.

*Proof.* We have the following equivalent equations:

$$\begin{aligned} \langle (\theta, x), p \rangle &= \theta \\ \iff \theta p_0 + \langle x, (p_1, p_2, \dots) \rangle &= \theta \\ \iff \langle x, (p_1, p_2, \dots) \rangle &= (1 - p_0)\theta \\ \iff \langle x, s(p) \rangle &= \theta. \end{aligned}$$

The proof for weak and strict inequalities are similar.  $\square$

Let  $D$  be a subset of  $l_1$ , a point  $p \in D$  is an exposed point of  $D$  if there exists some  $x \in l_\infty$  such that for every  $p' \in \mathbb{D} \setminus \{p\}$ ,  $\langle x, p \rangle < \langle x, p' \rangle$ . We then say that  $p$  is exposed in  $D$  by  $x$ . Throughout this paper, we use  $\mathcal{E}(D)$  to denote the set of all exposed points of  $D$ .

**Lemma 4.** Let  $T \subset (0, 1)$ , closed in  $[0, 1]$ . Let  $F = \{(1 - \delta)(1, \delta, \delta^2, \dots) : \delta \in T\}$  and  $C$  be the closure of the convex hull of  $F$ . Then for every  $\delta \in T$ ,  $(1 - \delta)(1, \delta, \delta^2, \dots)$  is an exposed point of  $C$ .

*Proof.* Let  $\delta \in T$  and  $q = (1 - \delta)(1, \delta, \delta^2, \dots)$ . For every  $z \in l_\infty$ , we denote the discounted value of  $z$  corresponding with the discount factor  $\delta' \in (0, 1)$  by  $D_{\delta'}(z)$ , that is  $D_{\delta'}(z) = (1 - \delta') \sum_{n=0}^{\infty} (\delta')^n z_n$ . We then define  $y \in l_\infty$  by  $y_0 = \delta^2$ ,  $y_1 = \delta^2 - 2\delta$ , and  $y_n = (1 - \delta)^2$  for every  $n > 1$ . By an easy computation, we have

$$D_{\delta'}(y) = (\delta' - \delta)^2 \text{ for every } \delta' \in T. \quad (5)$$

Thus, we get that

$$0 = D_\delta(y) < D_{\delta'}(y) \text{ for every } \delta' \neq \delta. \quad (6)$$

We will prove that  $q$  is exposed in  $C$  by  $y$ , or equivalently that for every  $q' \in C \setminus \{q\}$ ,

$$\langle y, q \rangle < \langle y, q' \rangle. \quad (7)$$

To prove Inequality (7), we have to prove that  $\langle y, q \rangle = D_\delta(y) = 0 < \langle y, q' \rangle$  for every  $q' \in C \setminus \{q\}$ . Clearly,  $\langle y, q' \rangle \geq 0$  for every  $q' \in C$ . By contradiction, assume that there exists  $r \in C \setminus \{q\}$  such that  $\langle y, r \rangle = 0$ . From the definition of  $C$ ,  $r = \lim_{n \rightarrow \infty} r^n$  in which<sup>15</sup>  $r^n \in \text{co}(F)$ .

To simplify notation, we denote  $\delta' = (1 - \delta')(1, \delta', (\delta')^2, \dots)$  for every  $\delta' \in T$ . By the assumption  $r \neq \delta$ , there exist  $\varepsilon > 0$  such that  $r \notin B(\delta, \varepsilon)$ , where  $B(\delta, \varepsilon)$  is the closed ball centered  $\delta$  and of radius  $\varepsilon$ . Since  $r^n$  converges to  $r$ , there exists  $N \in \mathbb{N}$  such that  $r^n \notin B(\delta, \varepsilon)$  for every  $n > N$ . Let  $A = \{\delta' \in T : \delta' \in B(\delta, \varepsilon)\}$

<sup>15</sup>Norm convergence and weakly convergence are equivalent in  $l_1$ , thus, we can understand the limit for the norm topology.

and  $B = T \setminus A$ . The set  $B$  must be non-empty; otherwise, if  $B = \emptyset$ , then  $F \subset B(\boldsymbol{\delta}, \varepsilon)$ , and in particular,  $r \in \overline{\text{co}(F)} \subset B(\boldsymbol{\delta}, \varepsilon)$ , a contradiction.

By the fact that  $r^n \in \text{co}(F)$ , we can write  $r^n = \sum_{i=1}^{m_n} \lambda_i^n \boldsymbol{\delta}_i^n$ , where  $\lambda_i^n > 0$ ,  $\sum_{i=1}^{m_n} \lambda_i^n = 1$ , and  $\boldsymbol{\delta}_i^n \in T$  for every  $n \in \mathbb{N}$  and  $i \in \{1, \dots, m_n\}$ . For every  $n > N$ , let

$$\bar{\lambda}^n = \sum_{i=1, \dots, m_n: \boldsymbol{\delta}_i^n \in B} \lambda_i^n,$$

which is well defined because for every  $n > N$ , there exists  $i \in \{1, 2, \dots, m_n\}$  such that  $\boldsymbol{\delta}_i^n \in B$  (otherwise,  $r^n$  would be in  $B(\boldsymbol{\delta}, \varepsilon)$  as a convex combination of elements in this convex set). The sequence  $\{\bar{\lambda}^n\}_{n > N}$  admits a subsequence that converges to a point  $k \in [0, +\infty]$ . We still denote this subsequence by the initial sequence to simplify notations.

If  $k = 0$ , then since  $r^n = \sum_{i=1}^{m_n} \lambda_i^n \boldsymbol{\delta}_i^n = \sum_{\boldsymbol{\delta}_i^n \in A} \lambda_i^n \boldsymbol{\delta}_i^n + \sum_{\boldsymbol{\delta}_i^n \in B} \lambda_i^n \boldsymbol{\delta}_i^n$ , we get that the pointwise limit,  $r$ , of  $\{r^n\}_{n \geq 0}$  must belong to  $B(\boldsymbol{\delta}, \varepsilon)$ , a contradiction. Therefore, we must have  $k \in (0, +\infty]$ .

For every  $\boldsymbol{\delta}' \in T$ , we define the function  $f(\boldsymbol{\delta}') = \boldsymbol{\delta}'$ . The norm continuity of  $f$  implies that there exists  $\gamma > 0$  such that  $|\boldsymbol{\delta}' - \boldsymbol{\delta}| > \gamma$  for every  $\boldsymbol{\delta}' \in B$  (indeed, recall that  $\boldsymbol{\delta}' \in B$  if and only if  $\boldsymbol{\delta}' \in T$  and  $\|\boldsymbol{\delta}' - \boldsymbol{\delta}\| = \|f(\boldsymbol{\delta}') - f(\boldsymbol{\delta})\| > \varepsilon$ ). Thus, from Equation (5), we get  $D_{\boldsymbol{\delta}'}(y) = (\boldsymbol{\delta}' - \boldsymbol{\delta})^2 > \gamma^2$  for every  $\boldsymbol{\delta}' \in B$ . Thus, for every  $n > N$ , we get

$$\langle y, r^n \rangle = \sum_{\boldsymbol{\delta}_i^n \in A} \lambda_i^n D_{\boldsymbol{\delta}_i^n}(y) + \sum_{\boldsymbol{\delta}_i^n \in B} \lambda_i^n D_{\boldsymbol{\delta}_i^n}(y) \geq \sum_{\boldsymbol{\delta}_i^n \in B} \lambda_i^n D_{\boldsymbol{\delta}_i^n}(y) > \sum_{\boldsymbol{\delta}_i^n \in B} \lambda_i^n \gamma^2 = \bar{\lambda}^n \gamma^2,$$

where the first inequality follows from Inequality (6), and the second inequality follows from the fact that for every  $\boldsymbol{\delta}' \in B$ ,  $D_{\boldsymbol{\delta}'}(y) = (\boldsymbol{\delta}' - \boldsymbol{\delta})^2 > \gamma^2$ . Hence,  $\langle y, r \rangle = \lim_{n \rightarrow \infty} \langle y, r^n \rangle \geq k \gamma^2 > 0$ . It contradicts the assumption  $\langle y, r \rangle = 0$ . Thus, by contradiction, Inequality (7) is true, and the proof of Lemma 4 is finished.  $\square$

**If part.** Assume that

$$I(x) = \min_{(p_0, \delta) \in Q} \{p_0 x_0 + (1 - p_0)(1 - \delta)x_1 + (1 - p_0)(1 - \delta)\delta x_2 + (1 - p_0)(1 - \delta)\delta^2 x_3 \dots\},$$

where  $Q \subset (0, 1) \times (0, 1)$  is a closed set in  $[0, 1]^2$ . The proof that  $\succeq$  satisfies Maxmin axioms and Delayed ISTAT are left to the reader. We will only prove EFREP.

Let  $\mathbb{D} = g(Q)$ , where  $g(p_0, \delta) = (p_0, (1 - p_0)(1 - \delta), (1 - p_0)(1 - \delta)\delta, \dots)$  for every  $(p_0, \delta) \in Q$ . Let  $x \in l_\infty$  and  $\theta \in \mathbb{R}$  with  $x \sim \theta$ , and let  $(\bar{p}_0, \bar{\delta})$  be a solution of the minimization problem defining  $I(x) = \theta$ . In particular,

$$\langle x, g(\bar{p}_0, \bar{\delta}) \rangle = \theta, \tag{8}$$

and for every  $p \in \mathbb{D}$ ,

$$\langle x, p \rangle \geq \theta.$$

Let  $E = \{\boldsymbol{\delta} : \text{there exists } p_0 \in (0, 1) \text{ such that } (p_0, \delta) \in Q\}$ . Since  $\{\boldsymbol{\delta} : \delta \in E\} = \mathcal{E}(s(\mathbb{D}))$  (see Lemma 4) and  $\bar{\boldsymbol{\delta}} \in E$ , we know that  $\bar{\boldsymbol{\delta}}$  is exposed in  $s(\mathbb{D})$  by some  $y \in l_\infty$ , and translating  $y$  by a constant, we

can choose  $y$  such that  $\langle y, \bar{\delta} \rangle = \theta$ . Equivalently,

$$\langle y, \bar{\delta} \rangle = \theta < \langle y, s(p) \rangle$$

for every  $p \in \mathbb{D}$  such that  $s(p) \neq \bar{\delta}$ .

We now prove the following properties of  $y$ :

1.  $(\theta, y) \sim \theta$ .
2. If  $z \in l_\infty$  with  $(\theta, z) \sim \theta$  and  $\lambda x + (1-\lambda)(\theta, z) \succ \theta$  for every  $\lambda \in (0, 1)$ , then  $\lambda(\theta, y) + (1-\lambda)(\theta, \theta, z) \succ \theta$  for every  $\lambda \in (0, 1)$ .

For the first property, the definition of  $y$  implies that  $\theta \leq \langle y, s(p) \rangle$  for every  $p \in \mathbb{D}$ , this is equivalent to  $\theta \leq \langle (\theta, y), p \rangle$  for every  $p \in \mathbb{D}$  by Lemma 3. The equality is attained since  $\bar{\delta} \in s(\mathbb{D})$ . Thus,  $(\theta, y) \sim \theta$  by the definition of  $I$ .

For the second property, let  $z \in l_\infty$  such that

$$(\theta, z) \sim \theta, \text{ and for every } \lambda \in (0, 1), \lambda x + (1-\lambda)(\theta, z) \succ \theta. \quad (9)$$

Let us prove that  $\lambda(\theta, y) + (1-\lambda)(\theta, \theta, z) \succ \theta$  for every  $\lambda \in (0, 1)$ . It is equivalent to show that, for every  $p \in \mathbb{D}$ ,

$$\langle \lambda(\theta, y) + (1-\lambda)(\theta, \theta, z), p \rangle > \theta,$$

or (from Lemma 3)

$$\langle \lambda y + (1-\lambda)(\theta, z), s(p) \rangle > \theta. \quad (10)$$

We first prove that this inequality is true when  $s(p) = \bar{\delta}$ . We already have  $\langle y, \bar{\delta} \rangle = \theta$ . Now, if we assume  $\langle (\theta, z), \bar{\delta} \rangle = \theta$  (from  $(\theta, z) \sim \theta$ , we cannot have  $\langle (\theta, z), \bar{\delta} \rangle < \theta$ ), then  $\langle z, \bar{\delta} \rangle = \theta$  (by using Lemma 3 and  $s(\bar{\delta}) = \bar{\delta}$ ). Since  $s(g(\bar{p}_0, \bar{\delta})) = \bar{\delta}$ , by using Lemma 3 again, we get  $\langle (\theta, z), g(\bar{p}_0, \bar{\delta}) \rangle = \theta$ . This, together with Equation (8), implies that  $\langle \lambda x + (1-\lambda)(\theta, z), g(\bar{p}_0, \bar{\delta}) \rangle = \theta$  for every  $\lambda \in [0, 1]$ , a contradiction to (9). Thus, we obtain  $\langle (\theta, z), \bar{\delta} \rangle > \theta$  by contradiction, and this implies that Inequality (10) is true when  $s(p) = \bar{\delta}$ .

Next, we prove that Inequality (10) holds when  $p$  (fixed) satisfies  $s(p) \neq \bar{\delta}$ . First, we have  $\langle y, s(p) \rangle > \theta$  (because  $\bar{\delta}$  is exposed in  $s(\mathbb{D})$  by  $y$ ). Second, we get  $\langle (\theta, z), p \rangle \geq \theta$  from (9), which implies that  $\langle z, s(p) \rangle \geq \theta$  by Lemma 3. Since  $s(p)$  is a geometric distribution, we get  $s(p) = s(s(p))$ . Then it follows from Lemma 3 that  $\langle (\theta, z), s(p) \rangle \geq \theta$ . Hence, Inequality (10) holds because it is a convex combination of one strict and one weak inequality, which finally ends the proof.

**Only-if part.** Assume that  $\succeq$  satisfies Maxmin axiom, Delayed ISTAT, and EFREP. From Theorem 3, and more precisely from the proof of Theorem 1, we have

$$I(x) = \min_{p \in \mathbb{D}} \{ \langle x, p \rangle \}, \quad (11)$$



where  $\mathbb{D} = \{q \in \Delta^\sigma : c(q) = 0\}$ , and where  $c : \Delta \rightarrow [0, \infty]$  is defined by

$$c(p) = \sup_{x \in l_\infty} \{I(x) - \langle x, p \rangle\}.$$

Remark that we can identify the weak topology and the weak\* topology on  $l_1$ . Now, the set  $\mathbb{D}$  is a convex and weakly compact subset of  $l_1$  (it is convex and weakly\* compact from Theorem 3, thus weakly compact). A theorem from Benyamini and Lindenstrauss (see Corollary 5.18 in [Benyamini and Lindenstrauss \(1998\)](#)) ensures that  $\mathbb{D}$  is the weakly closed convex hull of its exposed points.<sup>16</sup> Thus, the main idea of the proof is to characterize all exposed points of the set  $\mathbb{D}$ .

*Step 1: For every element  $p = (p_0, p_1, \dots)$  in  $\mathbb{D}$ , we have  $p_t \in (0, 1)$  for every  $t \in \mathbb{N}$ .*

We will prove the statement by contradiction. Assume that  $p_t = 0$  for some  $t \geq 0$ , then for every  $\varepsilon > 0$ ,  $\langle \theta + \varepsilon \mathbf{e}_t, p \rangle = \langle \theta, p \rangle = \theta$ , where  $\mathbf{e}_t$  is the sequence whose  $t^{\text{th}}$ -term is one and all other terms are zero. It follows by Strong monotonicity that  $I(\theta + \varepsilon \mathbf{e}_t) - \langle \theta + \varepsilon \mathbf{e}_t, p \rangle > I(\theta) - \langle \theta, p \rangle = 0$ . From the definition of  $c$ , we get

$$c(p) = \sup_{x \in l_\infty} \{I(x) - \langle x, p \rangle\} \geq I(\theta + \varepsilon \mathbf{e}_t) - \langle \theta + \varepsilon \mathbf{e}_t, p \rangle > 0,$$

a contradiction to  $c(p) = 0$  (from  $p \in \mathbb{D}$ ). Thus, by contradiction, we must have  $p_t > 0$  for every  $t \in \mathbb{N}$ .

*Step 2:  $s(\mathbb{D})$  is a convex and weakly compact subset of  $\Delta^\sigma$ .*

From Step 1, the function  $s$  is well defined in  $\mathbb{D}$ . If  $p \in \mathbb{D}$ , then clearly  $s(p) \in \Delta^\sigma$ . To prove that  $s(\mathbb{D})$  is convex, let  $(p, p') \in \mathbb{D}^2$  and  $\lambda \in (0, 1)$ , we need to show that  $\lambda s(p) + (1 - \lambda)s(p') \in s(\mathbb{D})$ . We have

$$\lambda s(p) + (1 - \lambda)s(p') = \left( \frac{\lambda p_1}{1 - p_0} + \frac{(1 - \lambda)p'_1}{1 - p'_0}, \frac{\lambda p_2}{1 - p_0} + \frac{(1 - \lambda)p'_2}{1 - p'_0}, \dots \right).$$

For every  $i = 1, 2, \dots$ , we can write

$$\frac{\lambda p_i}{1 - p_0} + \frac{(1 - \lambda)p'_i}{1 - p'_0} = \frac{\alpha p_i + (1 - \alpha)p'_i}{1 - \alpha p_0 - (1 - \alpha)p'_0}, \quad (12)$$

where  $\alpha = \frac{\lambda(1 - p'_0)}{\lambda(1 - p'_0) + (1 - \lambda)(1 - p_0)} \in (0, 1)$ . Let  $p'' = \alpha p + (1 - \alpha)p'$ , we have  $p'' \in \mathbb{D}$  since  $\mathbb{D}$  is convex. Then Equation (12) implies that  $\lambda s(p) + (1 - \lambda)s(p') = s(p'') \in s(\mathbb{D})$ . Thus,  $s(\mathbb{D})$  is convex.

The weakly compactness of  $s(\mathbb{D})$  is a consequence of the compactness of  $\mathbb{D}$ , the continuity of  $s$  for the norm topology, and the equivalence of norm compactness and weakly compactness in  $l_1$  (see Theorem 16.24 in [Aliprantis and Border \(2006\)](#)).

*Step 3: There exists some set  $T \subset (0, 1)$ , closed in  $[0, 1]$ , such that  $\mathcal{E}(s(\mathbb{D})) = \{(1 - \delta)(1, \delta, \delta^2, \dots) : \delta \in T\}$ ,*

<sup>16</sup>Actually, Benyamini and Lindenstrauss theorem implies that  $\mathbb{D}$  is the weakly closed convex hull of its *strongly exposed points*, but strongly exposed points of convex subsets of Hausdorff topological vector spaces are automatically exposed points.

where  $\mathcal{E}(s(\mathbb{D}))$  is the set of all exposed points of  $s(\mathbb{D})$ .

We first remark that, for every  $x \in l_\infty$ , there exists a unique  $J(x) \in \mathbb{R}$  such that  $(J(x), x) \sim \mathbf{J}(\mathbf{x})$ . Indeed, for every  $a \in \mathbb{R}$ , consider  $h(a) = I(a, x) - I(\mathbf{a}) = I(a, x) - a$ . From Strong monotonicity, there exists  $k > 0$  such that  $\mathbf{k} > (k, x)$ , i.e.,  $h(k) < 0$ , and  $K < 0$  such that  $\mathbf{K} < (K, x)$ , i.e.,  $h(K) > 0$ . Since  $I$  is continuous,  $h$  is continuous also. Thus, there exists  $J(x) \in \mathbb{R}$  such that  $h(J(x)) = 0$ , i.e.,  $(J(x), x) \sim \mathbf{J}(\mathbf{x})$ . Now, the uniqueness of such a value is straightforward. Otherwise, if there exist two reals  $a > b$  with  $(a, x) \sim \mathbf{a}$  and  $(b, x) \sim \mathbf{b}$ , then from Constant translation invariance, we get  $(0, x - \mathbf{a}) \sim (0, x - \mathbf{b})$ . However,  $(0, x - \mathbf{b}) > (0, x - \mathbf{a})$ , a contradiction to Strong monotonicity.

We define a new preference relation  $\succeq^*$  on  $l_\infty$  as follows: “ $x \succeq^* y$  if and only if  $J(x) \geq J(y)$ ”. The symmetric and asymmetric parts of  $\succeq^*$  are defined as usual. By definition, we have  $x \sim^* \mathbf{J}(\mathbf{x})$ , thus  $J : l_\infty \rightarrow \mathbb{R}$  is the unique constant equivalent that represents  $\succeq^*$ . Now, we prove that  $J(x) = \min_{q \in s(\mathbb{D})} \{\langle x, q \rangle\}$  for every  $x \in l_\infty$ . Let  $x \in l_\infty$ , and let  $\theta = J(x)$ . From the definition of  $J$ , we get  $(\theta, x) \sim \theta$ , then from the definition of  $I$ , we get

$$I((\theta, x)) = \min_{q \in \mathbb{D}} \{\langle (\theta, x), q \rangle\} = \theta.$$

Thus, we have  $\langle (\theta, x), q \rangle \geq \theta$  for every  $q \in \mathbb{D}$ , and there exists  $q^* \in \mathbb{D}$  such that  $\langle (\theta, x), q^* \rangle = \theta$ . It follows from Lemma 3 that  $\langle x, s(q) \rangle \geq \theta$  for every  $q \in \mathbb{D}$ , and  $\langle x, s(q^*) \rangle = \theta$ . Hence,

$$J(x) = \theta = \min_{q \in s(\mathbb{D})} \{\langle x, q \rangle\}. \quad (13)$$

Since  $s(\mathbb{D})$  is convex and weakly compact (from Step 2), Theorem 3 implies that the preference relation  $\succeq^*$  satisfies axioms **A.1-A.7**. Moreover, we can also represent  $J(x) = \min_{q \in \mathbb{D}^*} \{\langle x, q \rangle\}$ , where  $\mathbb{D}^* = \{q \in \Delta^\sigma : c^*(q) = 0\}$  and  $c^*(q) = \sup_{x \in l_\infty} \{J(x) - \langle x, q \rangle\}$  (see the proof of Theorem 3). The uniqueness result of Theorem 3 implies that  $s(\mathbb{D}) = \mathbb{D}^*$ . Now, let us prove that  $\succeq^*$  satisfies ISTAT. Consider  $x \sim^* \theta$  and  $\lambda \in [0, 1]$ , from the definition of  $\succeq^*$ , we have  $(\theta, x) \sim \theta$ , and by using Delayed ISTAT, we get  $(\theta, \lambda x + (1 - \lambda)(\theta, x)) = \lambda(\theta, x) + (1 - \lambda)(\theta, \theta, x) \sim \theta$ , which is equivalent to  $\lambda x + (1 - \lambda)(\theta, x) \sim^* \theta$ . Thus,  $\succeq^*$  satisfies ISTAT.

Next, we prove that every exposed point of the set  $s(\mathbb{D}) = \mathbb{D}^*$  is a geometric distribution. Let  $q \in \mathbb{D}^*$  be exposed by some  $x \in l_\infty$ . Let  $\theta \in \mathbb{R}$  such that  $x \sim^* \theta$ . Observe that  $q$  is a geometric distribution if and only if  $q = s(q)$ ; thus, it is sufficient to show that  $s(q)$  is also exposed in  $\mathbb{D}^*$  by  $x$  which is equivalent to <sup>17</sup>

$$\langle x, s(q) \rangle = \langle x, q \rangle \quad (14)$$

and

$$c^*(s(q)) = 0, \text{ i.e., } s(q) \in \mathbb{D}^*. \quad (15)$$

First, we prove Equation (14). From the definition of  $q$ , we have

$$\langle x, q \rangle = \theta, \quad (16)$$

---

<sup>17</sup>Step 1 implies that for every  $q \in s(\mathbb{D})$ ,  $q \gg \mathbf{0}$ . Thus,  $s(q)$  is well defined.

and for every  $q' \in \mathbb{D}^* \setminus \{q\}$ ,

$$\langle x, q' \rangle > \theta. \quad (17)$$

Since  $\succeq^*$  satisfied ISTAT, we get that  $J(\lambda x + (1 - \lambda)(\theta, x)) = \theta$  for every  $\lambda \in [0, 1]$ . It follows from the definition of  $J$  that

$$\theta = J(\lambda x + (1 - \lambda)(\theta, x)) = \min_{q' \in \mathbb{D}^*} \{\langle \lambda x + (1 - \lambda)(\theta, x), q' \rangle\}. \quad (18)$$

In particular, by taking  $\lambda = 0$ , we get  $\theta = J(\theta, x) = \min_{q' \in \mathbb{D}^*} \{\langle (\theta, x), q' \rangle\}$ , this implies that for every  $q' \in \mathbb{D}^*$ ,

$$\langle (\theta, x), q' \rangle \geq \theta. \quad (19)$$

Now, consider  $\lambda \in (0, 1)$ . From Inequalities (17) and (19), we get  $\langle \lambda x + (1 - \lambda)(\theta, x), q' \rangle > \theta$  for every  $q' \in \mathbb{D}^* \setminus \{q\}$ . Then Equation (18) implies that we must have  $\lambda \langle x, q \rangle + (1 - \lambda) \langle (\theta, x), q \rangle = \theta$ , it follows from Equation (16) that  $\langle (\theta, x), q \rangle = \langle x, q \rangle = \theta$ . By Lemma 3, we finally get that  $\langle x, s(q) \rangle = \langle x, q \rangle = \theta$ .

Now, let us prove Equation (15). By the definition of  $c^*$ , we have

$$c^*(q) = \sup_{x' \in l_\infty} \{J(x') - \langle x', q \rangle\} = 0,$$

which implies that  $J(x') - \langle x', q \rangle \leq 0$  for every  $x' \in l_\infty$ . Applying this inequality to  $x' := (J(y), y)$  for  $y \in l_\infty$ , we get

$$J(J(y), y) \leq \langle (J(y), y), q \rangle.$$

But from  $J(J(y), y) = J(y)$  ( $\succeq^*$  satisfies ISTAT), it is equivalent to

$$J(y) \leq J(y)q_0 + \sum_{k=0}^{+\infty} y_k q_{k+1}.$$

A simple computation entails that, for every  $y \in l_\infty$ ,  $J(y) \leq \langle y, s(q) \rangle$ . In particular, if we take  $y$  a constant sequence, then the previous inequality is an equality. Thus, from the definition of  $c^*$ , we finally get

$$c^*(s(q)) = \sup_{y \in l_\infty} \{J(y) - \langle y, s(q) \rangle\} = 0.$$

Since  $q$  is exposed in  $s(\mathbb{D})$  by  $x$ , Equations (14) and (15) imply  $q = s(q)$ . Solving the equation explicitly, we easily get

$$q = (1 - \delta)(1, \delta, \delta^2, \dots),$$

where  $\delta \in (0, 1)$ , i.e.,  $q$  is a geometric probability. Thus, we have proved that all exposed points of  $s(\mathbb{D})$ ,  $\mathcal{E}(s(\mathbb{D}))$ , are geometric distributions.

Let  $U = f(\mathcal{E}(s(\mathbb{D})))$ , where  $f : \Delta^\sigma \rightarrow [0, 1]$  is defined by  $f(x_0, x_1, \dots, x_n, \dots) = 1 - x_0$ . The function  $f$  is clearly normed continuous (thus weakly continuous). In particular,  $f(s(\mathbb{D}))$  is closed (for the topology on  $[0, 1]$ ), as an image of a compact set by a continuous function) and is included in  $(0, 1)$  (by Step 1). Now, if we define  $T = \overline{U}$ , then  $T \subset f(s(\mathbb{D}))$ , which implies that  $T$  is closed (in  $[0, 1]$ ) and is included in

$(0, 1)$ . Obviously, the continuity of the function  $f$  deduces that

$$\overline{\mathcal{E}(s(\mathbb{D}))} = \{(1 - \delta)(1, \delta, \delta^2, \dots) : \delta \in T\}.$$

Now, Corollary 5.18 in [Benyamini and Lindenstrauss \(1998\)](#) implies that  $s(\mathbb{D})$  is the weakly closed convex hull of the set  $\mathcal{E}(s(\mathbb{D}))$ , equivalently, it is the weakly closed convex hull of the set  $\overline{\mathcal{E}(s(\mathbb{D}))} = \{(1 - \delta)(1, \delta, \delta^2, \dots) : \delta \in T\}$ . But from Lemma 4, we know that, for every  $\delta \in T$ ,  $(1 - \delta)(1, \delta, \delta^2, \dots)$  is an exposed point of  $s(\mathbb{D})$ , which finally proves that  $\mathcal{E}(s(\mathbb{D})) = \{(1 - \delta)(1, \delta, \delta^2, \dots) : \delta \in T\}$ .

*Step 4:* For every  $p \in \mathcal{E}(\mathbb{D})$ ,  $s(p) \in \mathcal{E}(s(\mathbb{D})) = \{(1 - \delta)(1, \delta, \delta^2, \dots) : \delta \in T\}$ .

Let  $p \in \mathbb{D}$  be exposed by some  $x \sim \theta$ . Let  $y \in l_\infty$  be a stream that satisfies both conditions in EFREP.

1.  $(\theta, y) \sim \theta$ .
2. If  $z \in l_\infty$  with  $(\theta, z) \sim \theta$  and  $\lambda x + (1 - \lambda)(\theta, z) \succ \theta$  for every  $\lambda \in (0, 1)$ , then  $\lambda(\theta, y) + (1 - \lambda)(\theta, \theta, z) \succ \theta$  for every  $\lambda \in (0, 1)$ .

The first condition above implies that  $y \sim^* \theta$ , or  $J(y) = \theta$ . Thus, from the formula defining  $J$  (see Equation (13)), we get that  $\theta$  is the value of the following minimization problem (P):

$$\min_{q' \in s(\mathbb{D})} \{\langle y, q' \rangle\}.$$

Let  $K \subset s(\mathbb{D})$  be the set of all solutions of (P). Since  $s(\mathbb{D})$  is equal to the weakly closed convex hull of its set of exposed points, we have  $\min_{q' \in s(\mathbb{D})} \{\langle y, q' \rangle\} = \min_{q' \in \mathcal{E}(s(\mathbb{D}))} \{\langle y, q' \rangle\}$ .<sup>18</sup> Thus,  $K \cap \mathcal{E}(s(\mathbb{D})) \neq \emptyset$ , and from Step 2, there exists  $\delta \in T$  such that  $\delta \in K \cap \mathcal{E}(s(\mathbb{D}))$ . Since the value of (P) is equal to  $\theta$ , we have

$$\langle y, \delta \rangle = \theta. \tag{20}$$

We will prove that  $\delta = s(p)$  by contradiction: assume  $\delta \neq s(p)$ . From  $\delta \in \mathcal{E}(s(\mathbb{D}))$ , there exists  $z \in l_\infty$  such that  $\delta$  is exposed by  $z$  in  $s(\mathbb{D})$ . Notice that, for every  $\alpha \in \mathbb{R}$ ,  $\delta$  is also exposed by  $z + \alpha$  in  $s(\mathbb{D})$ , and we can take  $\alpha$  such that  $\langle z + \alpha, \delta \rangle = \theta$  because  $\delta \neq 1$ . Thus, we can assume, without any loss of generality, that  $\delta$  is exposed by  $z$  in  $s(\mathbb{D})$  with

$$\langle z, \delta \rangle = \theta. \tag{21}$$

Now, we will prove that  $(\theta, z) \sim \theta$  and  $\lambda x + (1 - \lambda)(\theta, z) \succ \theta$  for every  $\lambda \in (0, 1)$ . From the definition of  $z$ , we have

$$\min_{q' \in s(\mathbb{D})} \{\langle z, q' \rangle\} = \langle z, \delta \rangle = \theta. \tag{22}$$

It follows from the definition of  $J$  that  $J(z) = \theta$ , which is equivalent to  $(\theta, z) \sim \theta$ . As a consequence of Delayed ISTAT, we also get  $(\theta, \theta, z) \sim \theta$ .

<sup>18</sup>Remark that from Step 2,  $\mathcal{E}(s(\mathbb{D}))$  is weakly compact, and the function to be minimized is continuous, which proves that this minimization problem is well defined.

Next, let us show that  $\lambda x + (1 - \lambda)(\theta, z) \succ \theta$  for every  $\lambda \in (0, 1)$ . Since  $I(x) = \theta$ , from the definition of  $I$ ,  $\langle x, q \rangle \geq \theta$  for every  $q \in \mathbb{D}$ . Moreover, the definition of  $z$  implies that  $\langle z, q' \rangle > \theta$  for every  $q' \in s(\mathbb{D}) \setminus \{\delta\}$ , and Lemma 3 in turn implies that  $\langle (\theta, z), q \rangle > \theta$  for every  $q \in \mathbb{D}$  such that  $s(q) \neq \delta$ . Thus, we can deduce that for every  $\lambda \in (0, 1)$  and  $q \in \mathbb{D}$  with  $s(q) \neq \delta$ ,  $\langle \lambda x + (1 - \lambda)(\theta, z), q \rangle > \theta$ . Now, if  $s(q) = \delta$ , then  $q \neq p$  (because we have assumed  $s(p) \neq \delta$ ); and since  $p$  is exposed by  $x$ , we finally have  $\langle x, q \rangle > \theta$ . This proves that, in this case, we still have  $\langle \lambda x + (1 - \lambda)(\theta, z), q \rangle > \theta$ . The definition of  $I$  implies that for every  $\lambda \in (0, 1)$ ,  $\lambda x + (1 - \lambda)(\theta, z) \succ \theta$ .

From the second condition of EFREP, we must have  $\lambda(\theta, y) + (1 - \lambda)(\theta, \theta, z) \succ (\theta, y)$  for every  $\lambda \in (0, 1)$ . Since  $\langle z, s(\delta) \rangle = \langle z, \delta \rangle = \theta$  (from  $s(\delta) = \delta$ ), Lemma 3 implies that  $\langle (\theta, z), \delta \rangle = \theta$ . Thus, by using Equation (20), we have the following equation:

$$\langle \lambda y + (1 - \lambda)(\theta, z), \delta \rangle = \theta. \quad (23)$$

We also get that, for every  $q' \in s(\mathbb{D})$ ,

$$\langle \lambda y + (1 - \lambda)(\theta, z), q' \rangle \geq \theta. \quad (24)$$

Indeed,  $\langle y, q' \rangle \geq \theta$  for every  $q' \in s(\mathbb{D})$  (because  $\theta$  is value of the minimization problem  $(P)$ ), and for every  $q \in \mathbb{D}$ ,  $\langle (\theta, z), s(q) \rangle \geq \theta$  is equivalent to  $\langle (\theta, \theta, z), q \rangle \geq \theta$  (From Lemma 3); and the last inequality follows from the definition of  $I(\theta, \theta, z) = \min_{q \in \mathbb{D}} \langle (\theta, \theta, z), q \rangle$  and  $(\theta, \theta, z) \sim \theta$ .

Next, by applying Lemma 3 to Equation (23) and Inequality (24), we get

$$\langle \lambda(\theta, y) + (1 - \lambda)(\theta, \theta, z), \delta \rangle = \theta \quad (25)$$

since  $s(\delta) = \delta$ , and for every  $q \in \mathbb{D}$ ,

$$\langle \lambda(\theta, y) + (1 - \lambda)(\theta, \theta, z), q \rangle \geq \theta. \quad (26)$$

Last, by definition,  $I(\lambda(\theta, y) + (1 - \lambda)(\theta, \theta, z)) = \min_{q \in \mathbb{D}} \langle \lambda(\theta, y) + (1 - \lambda)(\theta, \theta, z), q \rangle$ . Thus, we finally get  $\lambda(\theta, y) + (1 - \lambda)(\theta, \theta, z) \sim \theta$ , which violates the second condition of EFREP. Therefore, by contradiction, we get  $\delta = s(p)$ , which proves that  $s(p)$  is exposed in  $s(\mathbb{D})$ .

*Step 5: End of the proof.*

In Step 4, we have proved that for every  $p$  exposed in  $\mathbb{D}$ ,  $s(p) = (1 - \delta)(1, \delta, \delta^2, \dots)$  for some  $\delta \in T$ . An easy computation yields that  $p = (p_0, (1 - p_0)(1 - \delta), (1 - p_0)(1 - \delta)\delta, \dots)$  for some  $p_0 \in (0, 1)$ . Let  $Q$  be the set of all  $(p_0, \delta)$  corresponding to all exposed points of  $\mathbb{D}$ . From Corollary 5.18 in [Benyamini and Lindenstrauss \(1998\)](#),  $\mathbb{D}$  is the weakly closed convex hull of the set  $\mathcal{E}(\mathbb{D})$ ; equivalently, it is the weakly closed convex hull of the set  $\{(p_0, (1 - p_0)(1 - \delta), (1 - p_0)(1 - \delta)\delta, \dots) : (p_0, \delta) \in Q\}$ .

Thus, from Theorem 3, we can write

$$I(x) = \inf_{(p_0, \delta) \in Q} \{p_0 x_0 + (1 - p_0)(1 - \delta)x_1 + (1 - p_0)(1 - \delta)\delta x_2 + (1 - p_0)(1 - \delta)\delta^2 x_3 \dots\}.$$

Since the function  $g(p_0, \delta) = (p_0, (1-p_0)(1-\delta), (1-p_0)(1-\delta)\delta, \dots)$  is norm continuous, we can write

$$I(x) = \min_{(p_0, \delta) \in \overline{Q}} \{p_0 x_0 + (1-p_0)(1-\delta)x_1 + (1-p_0)(1-\delta)\delta x_2 + (1-p_0)(1-\delta)\delta^2 x_3 \dots\},$$

where  $\overline{Q}$  denotes the closure of  $Q$ . Remark that the function  $(p_0, p_1, \dots, p_n, \dots) \mapsto (p_0, 1 - \frac{p_1}{1-p_0})$  is continuous on  $\Delta^\sigma \cap (0, 1)^\mathbb{N}$ , and  $\mathbb{D}$  is compact in  $\Delta^\sigma \cap (0, 1)^\mathbb{N}$ , which implies that  $f(\mathbb{D}) \subset (0, 1)^2$  is closed in  $[0, 1]^2$ . In particular,  $\overline{Q} \subset f(\mathbb{D}) \subset (0, 1)^2$  is closed in  $[0, 1]^2$ .

*Step 6. The uniqueness of the set of long term discount factors,  $E = \{\delta \in (0, 1) : \text{there exists } p_0 \in (0, 1) \text{ such that } (p_0, \delta) \in Q\}$ .*

Assume  $I(x) = \min_{(p_0, \delta) \in Q} \{p_0 x_0 + (1-p_0)(1-\delta)x_1 + (1-p_0)(1-\delta)\delta x_2 + (1-p_0)(1-\delta)\delta^2 x_3 \dots\} = \min_{(p_0, \delta) \in \overline{Q}} \{p_0 x_0 + (1-p_0)(1-\delta)x_1 + (1-p_0)(1-\delta)\delta x_2 + (1-p_0)(1-\delta)\delta^2 x_3 \dots\}$ . Define  $E = \{\delta \in (0, 1) : \text{there exists } p_0 \in (0, 1) \text{ such that } (p_0, \delta) \in Q\}$  and  $E' = \{\delta \in (0, 1) : \text{there exists } p_0 \in (0, 1) \text{ such that } (p_0, \delta) \in \overline{Q}\}$ , and let us prove that  $E = E'$ .

We only prove that  $E \subset E'$ . By contradiction, assume that there exists  $\bar{\delta} \in E$ , but  $\bar{\delta} \notin E'$ . By choosing  $z = (\bar{\delta}^2, \bar{\delta}^2 - 2\bar{\delta}, (1-\bar{\delta})^2, (1-\bar{\delta})^2, \dots)$ , we get  $0 = \langle z, \bar{\delta} \rangle < \langle z, \delta \rangle$  for every  $\delta \in E'$ . Now, we have

$$\min_{(p_0, \delta) \in Q} \{p_0 0 + (1-p_0)(1-\delta)z_0 + (1-p_0)(1-\delta)\delta z_1 + (1-p_0)(1-\delta)\delta^2 z_2 \dots\} \leq (1-p_0)D_{\bar{\delta}}(z) = 0,$$

for some  $p_0 \in (0, 1)$  such that  $(p_0, \bar{\delta}) \in Q$ , which implies that  $I(0, z) \leq 0$ . However, for every  $(p_0, \delta) \in Q$ , we have  $(1-p_0)D_{\delta}(z) > 0$ . Thus,

$$\min_{(p_0, \delta) \in Q} \{p_0 0 + (1-p_0)(1-\delta)z_0 + (1-p_0)(1-\delta)\delta z_1 + (1-p_0)(1-\delta)\delta^2 z_2 \dots\} > 0,$$

which implies that  $I(0, z) > 0$ , a contraction.

Therefore, we must have  $E = E'$ .

## 5.6 Proof of Theorem 5

The *if part* of Theorem 5 is left to the reader. We will only prove the *only-if* part. As in the proof of Theorem 4, assuming that  $\succeq$  satisfies Maxmin axioms, Delayed ISTAT and Weak separability, from Theorem 3, we get

$$I(x) = \min_{p \in \mathbb{D}} \{\langle x, p \rangle\}, \tag{27}$$

where  $\mathbb{D} = \{q \in \Delta^\sigma : c(q) = 0\}$ , and where  $c : \Delta \rightarrow [0, \infty]$  is defined by

$$c(p) = \sup_{x \in l_\infty} \{I(x) - \langle x, p \rangle\}.$$

The three first steps in the proof of Theorem 4 remain true: for every element  $p = (p_0, p_1, \dots)$  in  $\mathbb{D}$ , we have  $p_t \in (0, 1)$  for every  $t \in \mathbb{N}$ ,  $s(\mathbb{D})$  is a convex and weakly compact subset of  $\Delta^\sigma$ , and there exists some set  $B \subset (0, 1)$ , closed in  $[0, 1]$ , such that  $\mathcal{E}(s(\mathbb{D})) = \{(1-\delta)(1, \delta, \delta^2, \dots) : \delta \in B\}$ . Next, we will

prove the following lemma, where we denote  $g(p_0, \delta) = (p_0, (1 - p_0)(1 - \delta), (1 - p_0)(1 - \delta)\delta, \dots)$ :

**Lemma 5.** *Let  $A = \{p_0 \in (0, 1) : \text{there exists } p = (p_0, p_1, \dots) \in \mathbb{D}\}$ , then  $\mathbb{D} = \overline{\text{co}\{g(p_0, \delta) : p_0 \in A, \delta \in B\}}$ .*

To prove the first inclusion  $\mathbb{D} \subset \overline{\text{co}\{g(p_0, \delta) : p_0 \in A, \delta \in B\}}$ , take  $p \in \mathbb{D}$  (in particular  $p_0 \in A$  by definition).  $s(\mathbb{D})$  is the closure of the convex hull of its exposed points (see Corollary 5.18 in [Benyamini and Lindenstrauss \(1998\)](#)). Since  $\mathcal{E}(s(\mathbb{D})) = \{(1 - \delta)(1, \delta, \delta^2, \dots) : \delta \in B\}$ ,  $s(p)$  can be written as a limit of convex combinations of geometric probabilities as follows:  $s(p) = \lim_{n \rightarrow +\infty} \sum_{i=1}^{M_n} \lambda_i(n) \boldsymbol{\delta}(n)$ , where  $M_n > 0$ ,  $\delta(n) \in B$ , and  $\boldsymbol{\delta}(n)$  denotes  $(1 - \delta(n))(1, \delta(n), \delta(n)^2, \dots)$ . This implies that  $p = \lim_{n \rightarrow +\infty} \sum_{i=1}^{M_n} \lambda_i(n) (p_0, (1 - p_0)(1 - \delta(n)), (1 - p_0)(1 - \delta(n))\delta(n), \dots)$ , and we get the first inclusion.

To prove the second inclusion  $\overline{\text{co}\{g(p_0, \delta) : p_0 \in A, \delta \in B\}} \subset \mathbb{D}$ , since  $\mathbb{D}$  is convex and closed, we only have to prove that  $\{g(p_0, \delta) : p_0 \in A, \delta \in B\} \subset \mathbb{D}$ . Take  $p_0 \in A$  and  $\delta \in B$ . The condition  $g(p_0, \delta) \in \mathbb{D}$  is equivalent to  $c(g(p_0, \delta)) \leq 0$ , thus, from the above formula defining  $c$ , to  $I(x) \leq \langle x, g(p_0, \delta) \rangle$  for every  $x \in l_\infty$ . Subtracting  $x_0 = I(x_0, x_0, \dots)$ , this is also equivalent to

$$I(0, x) \leq \langle (0, x), g(p_0, \delta) \rangle = (1 - p_0)\langle x, \boldsymbol{\delta} \rangle \text{ for every } x \in l_\infty, \quad (28)$$

where  $\boldsymbol{\delta} = (1 - \delta)(1, \delta, \delta^2, \dots)$ . Take  $x \in l_\infty$  and let  $\theta \in \mathbf{R}$  such that  $(\theta, x) \sim \boldsymbol{\theta}$  (i.e.,  $I(\theta, x) = \theta$ ). In particular, from the definition of  $I$ , we obtain  $\theta \leq \langle (\theta, x), p \rangle$  for every  $p \in \mathbb{D}$ . From Lemma 3, this is equivalent to  $\theta \leq \langle x, s(p) \rangle$  for every  $p \in \mathbb{D}$ . From  $\delta \in B$ , we have  $\boldsymbol{\delta} \in s(\mathbb{D})$ , which implies that  $\theta \leq \langle x, \boldsymbol{\delta} \rangle$ . Then, Inequality (28) follows from

$$(1 - p_0)\langle x, \boldsymbol{\delta} \rangle \geq (1 - p_0)\theta \geq I(0, \boldsymbol{\theta}) = I(0, x).$$

Indeed:

1. The inequality  $(1 - p_0)\theta \geq I(0, \boldsymbol{\theta})$  is a consequence of  $p_0 \in A$  and the definition of  $I$ .
2. The equality  $I(0, \boldsymbol{\theta}) = I(0, x)$  is a consequence of  $I(\boldsymbol{\theta}) = I(\theta, x)$  and Weak separability.

This ends the proof of Lemma 5. It is immediate that  $A$  is a compact subset of  $(0, 1)$ , and thus we can write

$$\begin{aligned} I(x) &= \min_{p_0 \in A, \delta \in B} \{p_0 x_0 + (1 - p_0)(1 - \delta)x_1 + (1 - p_0)(1 - \delta)\delta x_2 + (1 - p_0)(1 - \delta)\delta^2 x_3 \dots\} \\ &= \min_{p_0 \in [\underline{p}_0, \overline{p}_0], \delta \in B} \{p_0 x_0 + (1 - p_0)(1 - \delta)x_1 + (1 - p_0)(1 - \delta)\delta x_2 + (1 - p_0)(1 - \delta)\delta^2 x_3 \dots\}, \end{aligned}$$

where  $\underline{p}_0 = \min_{p_0 \in A} \{p_0\} \in (0, 1)$ , and  $\overline{p}_0 = \max_{p_0 \in A} \{p_0\} \in (0, 1)$ . The second equation follows from the linearity of the objective function in  $p_0$ . Finally, the uniqueness of  $\underline{p}_0$  and  $\overline{p}_0$  follows from the uniqueness of  $\mathbb{D}$ , and the uniqueness of  $B$  follows from the uniqueness result in Theorem 4.

## 5.7 The equivalence of Constant translation invariance and Weak certainty independence

First, we prove that Constant translation invariance implies Weak certainty independence. Let  $(x, y) \in l_\infty^2$ ,  $(\theta, \theta') \in \mathbf{R}^2$ , and  $\lambda \in (0, 1)$  such that  $\lambda x + (1 - \lambda)\boldsymbol{\theta} \succeq \lambda y + (1 - \lambda)\boldsymbol{\theta}$ . Using Constant translation

invariance, we get  $\lambda x + (1 - \lambda)\boldsymbol{\theta} + (1 - \lambda)(\boldsymbol{\theta}' - \boldsymbol{\theta}) \succeq \lambda y + (1 - \lambda)\boldsymbol{\theta} + (1 - \lambda)(\boldsymbol{\theta}' - \boldsymbol{\theta})$ , which is  $\lambda x + (1 - \lambda)\boldsymbol{\theta}' \succeq \lambda y + (1 - \lambda)\boldsymbol{\theta}'$ . Thus, Constant translation invariance implies Weak certainty independence.

Now, let us prove the inverse direction. Let  $(x, y) \in l_\infty^2$  with  $x \succeq y$ , and let  $\theta \in \mathbb{R}$ . We can write  $x \succeq y$  as  $\lambda \frac{x}{\lambda} + (1 - \lambda)\mathbf{0} \succeq \lambda \frac{y}{\lambda} + (1 - \lambda)\mathbf{0}$  for some  $\lambda \in (0, 1)$ . By Weak certainty independence, we get  $\lambda \frac{x}{\lambda} + (1 - \lambda) \frac{\boldsymbol{\theta}}{1 - \lambda} \succeq \lambda \frac{y}{\lambda} + (1 - \lambda) \frac{\boldsymbol{\theta}}{1 - \lambda}$ , which is  $x + \boldsymbol{\theta} \succeq y + \boldsymbol{\theta}$ . Thus, Weak certainty independence implies Constant translation invariance.

Therefore, Constant translation invariance and Weak certainty independence are equivalent.

## 5.8 [Constant translation invariance and Homotheticity] is equivalent to Certainty independence

Let  $\succeq$  be a binary relation on  $l_\infty$  that satisfies **A.1-A.3**. Assume that Constant translation invariance and Homotheticity hold. Let  $(x, y) \in l_\infty^2$  with  $x \succeq y$ , and let  $\lambda \in (0, 1)$ . By Homotheticity, we get  $\lambda x \succeq \lambda y$ , then it follows by Constant translation invariance that  $\lambda x + (1 - \lambda)\boldsymbol{\theta} \succeq \lambda y + (1 - \lambda)\boldsymbol{\theta}$ . Thus, [Constant translation invariance and Homotheticity] imply Certainty independence.

Assume that Certainty independence holds. It is clear that Certainty independence is stronger than Weak certainty independence, then Constant translation invariance is true by Appendix 5.7. Now, we prove that Homotheticity holds. Let  $(x, y) \in l_\infty^2$  with  $x \succeq y$ , we need to show that  $\alpha x \succeq \alpha y$  for every  $\alpha \geq 0$ . From Certainty independence, we have  $\alpha x \succeq \alpha y$  for  $\alpha \in (0, 1)$  (by taking  $\boldsymbol{\theta} = \mathbf{0}$ ). When  $\alpha = 0$  and  $\alpha = 1$ , Homotheticity is straightforward. Last, we consider  $\alpha > 1$ . Since **A.1-A.3** are satisfied, there exists a unique constant equivalent  $I$  that represents  $\succeq$  (see the proof of Proposition 1). Observe that  $\alpha x \succeq \mathbf{I}(\alpha \mathbf{x})$  (actually this is an equivalence), then Certainty independence implies that  $\frac{1}{\alpha}\alpha x + (1 - \frac{1}{\alpha})\mathbf{0} \succeq \frac{1}{\alpha}\mathbf{I}(\alpha \mathbf{x}) + (1 - \frac{1}{\alpha})\mathbf{0}$ , which is  $x \succeq \frac{1}{\alpha}\mathbf{I}(\alpha \mathbf{x})$ . Similarly, from  $\mathbf{I}(\alpha \mathbf{x}) \succeq \alpha x$ , we get  $\frac{1}{\alpha}\mathbf{I}(\alpha \mathbf{x}) \succeq x$ . Thus,  $x \sim \frac{1}{\alpha}\mathbf{I}(\alpha \mathbf{x})$ , which implies that  $\alpha I(x) = I(\alpha x)$ . It follows that  $\alpha I(x) = I(\alpha x) \geq I(\alpha y) = \alpha I(y)$  since  $x \succeq y$  gives  $I(x) \geq I(y)$ , from which we obtain  $\alpha x \succeq \alpha y$ . Thus, Homotheticity holds, and the proof is completed.

## 5.9 Equivalent Continuity assumptions

There exist several notions of continuity in the literature. The following lemma proves that, in  $l_\infty$ , most of them are equivalent under Weak order and Monotonicity:

**Lemma 6.** *Assume that  $\succeq$  satisfies **A.1** and **A.2**. Then each of the following continuity assumptions are equivalent to **A.3**:*

**A.3.1:** *For every  $(x, y) \in l_\infty^2$ , the sets  $\{z \in l_\infty : z \succ y\}$  and  $\{t \in l_\infty : y \succ t\}$  are open.*

**A.3.2:** *For every  $(x, y, z) \in l_\infty^3$  with  $x \succ y \succ z$ , there exists  $(\alpha, \beta) \in (0, 1)^2$  such that  $\alpha x + (1 - \alpha)z \succ y$  and  $y \succ \beta x + (1 - \beta)z$ .*

**A.3.3:** *For every  $(x, y) \in l_\infty^2$  with  $x \succ y$ , there exists  $\theta, \theta' > 0$  such that  $x - \boldsymbol{\theta} \succ y$  and  $x \succ y + \boldsymbol{\theta}'$ .*

*Proof.* Let us prove that **A.3.1**  $\implies$  **A.3**  $\implies$  **A.3.2**  $\implies$  **A.3.3**  $\implies$  **A.3.1**.

**A.3.1**  $\implies$  **A.3**: let  $(x, y, z) \in l_\infty^3$  with  $x \succeq y \succeq z$ , and let us prove that  $\{\alpha \in [0, 1] : \alpha x + (1 - \alpha)z \succeq y\}$  is closed in  $[0, 1]$ . Consider the continuous function  $f : [0, 1] \rightarrow l_\infty$  defined by  $f(\alpha) = \alpha x + (1 - \alpha)z$ . Then

$$\{\alpha \in [0, 1] : \alpha x + (1 - \alpha)z \succeq y\} = f^{-1}(\{t \in l_\infty : t \succeq y\}),$$



which is open as the complement of a preimage of an open subset by a continuous function. The proof is similar for the second statement of **A.3**.

**A.3**  $\implies$  **A.3.2**: let  $(x, y, z) \in l_\infty^3$  with  $x \succeq y \succeq z$ . Suppose on the contrary that there does not exist  $\beta \in (0, 1)$  such that  $y \succ \beta x + (1 - \beta)z$ , then we must have  $\beta x + (1 - \beta)z \succeq y$  for every  $\beta \in (0, 1)$ . From **A.3**, we get  $z \succeq y$  by letting  $\beta \rightarrow 0$ , a contradiction. It is thus concluded that there exists  $\beta \in (0, 1)$  such that  $y \succ \beta x + (1 - \beta)z$ . The existence of  $\alpha \in (0, 1)$ , for which  $\alpha x + (1 - \alpha)z \succ y$ , follows by a similar argument.

**A.3.2**  $\implies$  **A.3.3**: let  $(x, y) \in l_\infty^2$  with  $x \succ y$ . From Monotonicity, there exists two reals  $m$  and  $n$  large enough for which  $x \succ y \succ x - \mathbf{m}$  and  $y + \mathbf{n} \succ x \succ y$ . Axiom **A.3.2** guarantees the existence of  $(\alpha, \beta) \in (0, 1)^2$  such that  $\alpha x + (1 - \alpha)(x - \mathbf{m}) \succ y$  and  $x \succ \beta(y + \mathbf{n}) + (1 - \beta)y$ . Thus, defining  $\theta = (1 - \alpha)m > 0$  and  $\theta' = \beta n$ , we get  $x - \theta \succ y$  and  $x \succ y + \theta'$ .

**A.3.3**  $\implies$  **A.3.1**: let  $y \in l_\infty$ , and let us prove that the set  $\{z \in l_\infty : z \succ y\}$  is open. Take  $\bar{z} \in l_\infty$  such that  $\bar{z} \succ y$ . From **A.3.3**, there exists  $\theta > 0$  such that  $\bar{z} - \theta \succ y$ . In particular, from Monotonicity and Transitivity, we get that, for every  $z' \gg \bar{z}$ ,  $z' - \theta \succ y$ ; thus, the set  $\{z' - \theta, z' \gg \bar{z}\}$  is an open neighborhood of  $\bar{z}$  in  $\{z \in l_\infty : z \succ y\}$ , which finally proves that  $\{z \in l_\infty : z \succ y\}$  is open. Analogously, we can prove that the set  $\{t \in l_\infty : y \succ t\}$  is open.  $\square$

## References

- AINSLIE, G., AND N. HASLAM (1992): “Hyperbolic discounting,” In G. Loewenstein & J. Elster (Eds.), *Choice over time*, 57-92. Russell Sage Foundation.
- AINSLIE, G., AND R. J. HERRNSTEIN (1981): “Preference reversal and delayed reinforcement,” *Animal Learning & Behavior*, 9, 476-482.
- ALIPRANTIS, C. D., AND K. C. BORDER (2006): *Infinite Dimensional Analysis: A Hitchhiker’s Guide* (Third Ed.). Berlin: Springer.
- ARROW, K. (1970): *Essays in the Theory of Risk-Bearing*. Amsterdam: North-Holland.
- BARRO, R. J. (1999): “Ramsey Meets Laibson in the Neoclassical Growth Model,” *The Quarterly Journal of Economics*, 114 (4), 1125-1152.
- BENYAMINI, Y., AND J. LINDENSTRAUSS (1998): *Geometric Nonlinear Functional Analysis*, Vol. 1. American Mathematical Society Colloquium Publications, Vol. 48. Providence, RI: American Mathematical Society.
- CERREIA-VIOGLIO, S., F. MACCHERONI, M. MARINACCI, AND A. RUSTICHINI (2014): “Niveloids and their extensions: Risk measures on small domains,” *Journal of Mathematical Analysis and Applications*, 413 (1), 343-360.
- CHAMBERS, C., AND F. ECHENIQUE (2017): “On Multiple Discount Rates,” Working paper.
- CHAMBERS, C., AND F. ECHENIQUE (2018): “On Multiple Discount Rates,” *Econometrica*, 86 (4), 1325-1346.

- CHEUNG, S. L., A. TYMULA, AND X. WANG (2021): “Quasi-Hyperbolic Present Bias: A Meta-Analysis,” IZA Discussion Papers 14625, Institute of Labor.
- DRUGEON, J. P., AND T. HA-HUY (2018): “A Not so Myopic Axiomatization of Discounting,” PSE Working Papers halshs-01761962, HAL.
- DRUGEON, J. P., T. HA-HUY, AND T. D. H. NGUYEN (2019): “On maximin dynamic programming and the rate of discount,” *Economic Theory*, 67 (3), 703-729.
- DELLAVIGNA, S., AND U. MALMENDIER (2006): “Paying Not to Go to the Gym,” *American Economic Review*, 96 (3), 694-719.
- EPSTEIN, L. G., AND S. E. ZIN (1989): “Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework,” *Econometrica*, 57 (4), 937-969.
- GILBOA, I., AND D. SCHMEIDLER (1989): “Maxmin expected utility with non-unique prior,” *Journal of Mathematical Economics*, 18 (2), 141-153.
- GRENADIER, S. R., AND N. WANG (2007): “Investment under uncertainty and time-inconsistent preferences,” *Journal of Financial Economics*, 84 (1), 2-39.
- HARRIS, C., AND D. LAIBSON (2013): “INSTANTANEOUS GRATIFICATION,” *The Quarterly Journal of Economics*, 128 (1), 205-248.
- HAYASHI, T. (2003): “Quasi-stationary cardinal utility and present bias,” *Journal of Economic Theory*, 112, 343-352.
- JACKSON, M. O., AND L. YARIV (2015): “Collective Dynamic Choice: The Necessity of Time Inconsistency,” *American Economic Journal: Microeconomics*, 7 (4), 150-178.
- KIRBY, K. N., AND R. J. HERRNSTEIN (1995): “Preference Reversals Due to Myopic Discounting of Delayed Reward,” *Psychological Science*, 6 (2), 83-89.
- KOOPMANS, T. C. (1960): “Stationary Ordinal Utility and Impatience,” *Econometrica*, 28 (2), 287-309.
- KREPS, D. M., AND E. L. PORTEUS (1978): “Temporal Resolution of Uncertainty and Dynamic Choice Theory,” *Econometrica*, 46 (1), 185-200.
- LAIBSON, D. (1997): “Golden Eggs and Hyperbolic Discounting,” *The Quarterly Journal of Economics*, 112 (2), 443-478.
- LOEWENSTEIN, G., AND D. PRELEC (1992): “Anomalies in Intertemporal Choice: Evidence and an Interpretation,” *The Quarterly Journal of Economics*, 107 (2), 573-597.
- LUCAS, R. E., AND N. L. STOKEY (1984): “Optimal growth with many consumers,” *Journal of Economic Theory*, 32 (1), 139-171.
- MACCHERONI, F., M. MARINACCI, AND A. RUSTICHINI (2006): “Ambiguity Aversion, Robustness, and the Variational Representation of Preferences,” *Econometrica*, 74 (6), 1447-1498.

- MARÍN-SOLANO, J., AND J. NAVAS (2010): “Consumption and Portfolio Rules for Time-Inconsistent Investors,” *European Journal of Operational Research*, 201 (3), 860-872.
- MCCLURE, S. M., D. LAIBSON, G. LOEWENSTEIN AND J.D. COHEN (1992): “Separate Neural Systems Value Immediate and Delayed Monetary Rewards,” *Science*, 306, 503-507.
- MONTIEL OLEA, J. L., AND T. STRZALECKI (2014): “Axiomatization and Measurement of Quasi-Hyperbolic Discounting,” *The Quarterly Journal of Economics*, 129 (3), 1449-1499.
- MYERSON, J., AND L. GREEN (1995): “Discounting of delayed rewards: Models of individual choice,” *Journal of the Experimental Analysis of Behavior*, 64 (3), 263-276.
- PHELPS, E., AND R. POLLAK (1968): “On Second-Best National Saving and Game-Equilibrium Growth,” *Review of Economic Studies*, 35 (2), 185-199.
- SAMUELSON, P. A. (1937): “A Note on Measurement of Utility,” *Review of Economic Studies*, 4, 155-161.
- SCHMEIDLER, D. (1989): “Subjective Probability and Expected Utility without Additivity,” *Econometrica*, 57 (3), 571-587.
- STERN, N. (2006): “Stern Review on the economics of climate change,” HM Treasury, London.
- THALER, R., AND H. M. SHEFRIN. (1981): “An economic theory of self-control,” *The Journal of Political Economy*, 89, 392-406.
- VILLEGAS, C. (1964): “On Qualitative Probability Sigma-Algebras,” *The Annals of Mathematical Statistics*, 35, 1787-1796.
- WEITZMAN, M. L. (2011): “Gamma Discounting,” *American Economic Review*, 91 (1), 260-271.
- ZOU, Z., S. CHEN, AND L. WEDGE (2014): “Finite horizon consumption and portfolio decisions with stochastic hyperbolic discounting,” *Journal of Mathematical Economics*, 52, 70-80.
- ZUBER, S. (2011): “The Aggregation of Preferences: Can We Ignore the Past?,” *Theory and Decision*, 70 (3), 367-384.