DECREASING IMPATIENCE

CHRISTOPHER P. CHAMBERS, FEDERICO ECHENIQUE, AND ALAN D. MILLER

Abstract. Decreasing impatience, a common behavioral phenomenon in intertemporal choice, and a property with certain normative support in the literature on project evaluation, is characterized in several different ways. Discount factors that display decreasing impatience are characterized through a convexity axiom for investments at fixed interest rates. Then we show that they are equivalent to a geometric average of generalized quasi-hyperbolic discount rates. Finally, they emerge as prices in parimutuel betting markets, that is as the outcome of parimutuel aggregation.

1. Introduction

Decreasing impatience, a property of intertemporal preferences, has been the intense focus of positive studies in behavioral economics, and of a normative literature (and practice) in project evaluation. Suppose that a unit payoff, or reward, in period t is valued the same as a payoff of $f(t)$ today. Decreasing impatience is the property that $f(t+1)/f(t)$ is weakly monotone increasing. For example, an agent may prefer to receive a check for \$ 110 in 31 days to \$ 100 in 30 days, but \$ 100 today over \$ 110 tomorrow. Such choices reflect decreasing impatience, and are usually modeled by means of parametric models of hyperbolic, or β-δ (also called quasi-hyperbolic), discounting. See Loewenstein and Prelec (1992), or Frederick, Loewenstein, and O'Donoghue (2002) for a review of the experimental literature. Many behavioral economists have used these parametric models of discounting in studies of consumption, savings, and retirement (Laibson, 1997; O'Donoghue and Rabin, 1999; Diamond and Köszegi, 2003, for

(Chambers) Department of Economics, Georgetown University
(Echenique) Division of the Humanities and Social Sciences, California Institute of Technology
(Miller) Faculty of Law, Western University. Faculty of Law and Department of Economics, University of Haifa
Echenique thanks the National Science Foundation for financial support (Grants SES-1558757 and CNS-518941). This research was undertaken, in part, thanks to funding from the Canada Research Chairs program.

example). Our paper provides new characterizations of decreasing impatience as a “non parametric” property of discount factors.

Aside from its role in behavioral economics, decreasing impatience is an important property in the normative literature, and practice, of project evaluation: specifically in evaluating very long term projects and policies. The idea can be traced to Weitzman (1998; 2001), who observed that different policy makers, or different members of a team of policy advisers, propose to use different exponential discount rates. To use one rate over another can have drastically different implications for the evaluation of monetary streams. Weitzman used surveys to document the widely different discount rates proposed by professional economists for use in the evaluation of long-term projects, and argued that aggregating discount rates is crucial for policy making. The solution advocated by Weitzman is to use a discount factor that is an average of individual agents' discount factors. Such an average will *necessarily* display decreasing impatience.

The idea of decreasing impatience is a common practice in project evaluation, and Weitzman's rule has been adopted by practitioners. As an example, the British government's official guide to project evaluation (H.M. Treasury's “Green book,” 2020) recommends a discount rate that decreases from 3.50% to 2.50% as the time horizon increases from 30 to 75 years. Such differences may seem small, but compounding implies that they have dramatic implications over long horizons.

Our contribution is to provide characterizations of decreasing impatience as a general property of discounting. Our main result proposes three different characterizations. One is axiomatic, showing that, within models of intertemporal choice that rely on discounting, a simple convexity axiom captures decreasing impatience. The axiom relies on a single choice problem that can be implemented in the lab; thus avoiding incentive-compatibility issues that arise with multiple incentivized decisions (Azrieli et al., 2018). The other two characterizations show that decreasing impatience emerges as the property of an aggregate preference from different models of multiple selves.

Our first characterization using multiple selves shows that any discount factor that exhibits decreasing impatience is the result of a geometric average of generalized β-δ preferences. We provide a justification for using the geometric average by means of a dynamic consistency axiom, together with other standard normative axioms.[1] Standard β-δ preferences treat the present period as special, and worthy of a premium in intertemporal tradeoffs. Generalized β-δ preferences simply extend the special treatment to all periods before some cutoff date. Our characterization means that β-δ discounting is, in a sense, the canonical decreasing impatience discount factor: All other discount factors can be understood as a dynamically consistent aggregation of such generalized β-δ discount factors. There is then a sense in which any dynamic inconsistency due to decreasing impatience can be traced back to the β-δ model.

The second characterization using multiple selves shows that any discount factor that exhibits decreasing impatience is the equilibrium price of a parimutuel market. Parimutuel markets were first proposed as an information-aggregation mechanism by Eisenberg and Gale (1959), and have been the focus of an extensive empirical and experimental literature (see, for example, Plott et al. (2003)). Aside from its traditional use in horse races, they have been implemented inside large corporations as an information aggregation mechanism (Gillen et al., 2017). Here we use them as preference aggregation mechanisms. Our result means that decreasing impatience is always the aggregate of a collection of agents with traditional exponential discount factors, where the aggregation takes the parimutuel form. Note that instead of a representation in terms of generalized β-δ preferences, we obtain any decreasing impatience discount factor by means of aggregating exponential (i.e $\beta = 1$) discount factors.

Related literature: Decreasing impatience is a well known behavioral phenomenon: see for example Loewenstein and Prelec (1992), Prelec (2004), and Frederick et al. (2002). It is commonly modeled using hyperbolic or quasi-hyperbolic discount factors, and has been incorporated in multiple theoretical studies (Laibson, 1997; O'Donoghue and Rabin, 1999; Diamond and Köszegi, 2003). Nonparametric treatments such as ours are less common. Halevy (2015) presents theoretical results disentangling various deviations from stationary discounting, together with an experimental implementation. Chakraborty (2021) considers the phenomenon of present bias in isolation. His weak present bias axiom is satisfied by a host of models that have been introduced to relax the stationarity assumption of exponential discounting. Chakraborty characterizes the utility

[1] This justification of geometric averaging of discount factors is, we believe, of some interest independently of the present application.

representation (within a certain family) that satisfies the axiom of weak present bias.

The axiom we use to characterize decreasing impatience might be called a form of risk-seeking for time lotteries in DeJarnette et al. (2020), but the similarity is simply mathematical – our model is fully deterministic, and no lotteries are considered.

The problem of aggregating discount rates has received a lot of attention. The seminal paper by Weitzman (2001) documents disagreements about the discount rate, and proposes a solution that implies decreasing impatience. More recently, Zuber (2011) and Jackson and Yariv (2015) show that linear aggregation of exponential discounting preferences and dynamic consistency are incompatible.[2] Feng and Ke (2018) and Hayashi and Lombardi (2019) discuss ways of avoiding this impossibility by weakening the assumed Pareto criteria. Chambers and Echenique (2018) and Chambers and Echenique (2020) introduce and axiomatize decision criteria for environments with multiple discount rates. These papers focus on desirable properties of the resulting aggregate criterion for making intertemporal choices. Millner (2020) assumes agents who, each having their own (distinct) discount rates, entertain the possibility that alternative discount rates are possible. He shows an agreement about long term discounting on the smaller discount rates (an instance of decreasing impatience).

Our result in Section 4 connects with the literature on multiplicative aggregation. Our result is probably most similar to Hayashi (2016), but there are several differences. The first, and most obvious difference, is that our framework involves no social disagreement over period rewards. All disagreement is due to the form of discounting. A second main difference is that we envision this result as being most relevant when applied to dated rewards. One of our main properties, indeed, is a Pareto condition applied to dated rewards. Were we to postulate a form of Pareto for streams, we would be back to the framework of Harsanyi (1955) and Jackson and Yariv (2015). So, intertemporal tradeoffs for consumption streams should be viewed as "irrelevant" here. Finally, the point of Hayashi (2016) is that, while dynamic consistency may be interesting, we should not necessarily invoke it in an environment in which social preference is independent of history. By contrast, our framework has no language for allowing us to condition a ranking on history. So we implicitly rule out his aggregation functions. The interesting

[2]Linear aggregation is discussed in Section 6.

examples motivating his study involve intertemporal tradeoffs across individuals, a phenomenon that does not obtain here.

As a mathematical result, our aggregation result is not particularly novel, and indeed was motivated by the log-opinion pool of statistics. This is a method of aggregating Bayesian priors, by taking a geometric mean of the density functions. Versions of this aggregator were characterized by Genest (1984) and West (1984) using axioms very similar to the ones we describe here. It is worth noting that the failure of the log-opinion pool in probability aggregation to commute with respect to marginal distributions (a property used by McConway (1981) to characterize linear aggregation) in the case of probability aggregation does not pose a problem for us. Discount factors over finer or shorter lengths of time are not additive, but multiplicative by their very nature.

2. The Model

2.1. Notational conventions. .

A preference relation over a set is a complete and transitive binary relation (also called a weak order). A function $f : A \subseteq \mathbf{R} \to \mathbf{R}$ is ***weakly monotone increasing***, or ***non-decreasing***, if $f(x) \geq f(y)$ when $x \geq y$; and ***strictly monotone increasing***, if $f(x) > f(y)$ when $x > y$. It is ***weakly monotone decreasing***, or ***non-increasing***, if $-f$ is weakly monotone increasing; and ***strictly monotone decreasing*** if $-f$ is strictly monotone decreasing.

The set of bounded real sequences is denoted by ℓ^∞, and the subset of non-negative sequences by ℓ^∞_+. A sequence $\{x_t\} \in \ell^\infty$ is (absolutely) ***summable*** if $\sum_{t=0}^\infty |x_t|$ converges. The set of summable real sequences is denoted by ℓ^1, and the subset of nonnegative summable sequences by ℓ^1_+.

2.2. Discount factors.

We consider a model of intertemporal choice in which time is discrete, the horizon is infinite, and the objects of choice are bounded real sequences: $\{x_t : t = 0, 1, \dots,\}$. One may interpret each x_t as a monetary payoff, or as the value in "utils" of some underlying physical outcome. It is worth emphasizing that all our results hold if we assume a finite, instead of an infinite, time horizon.

We suppose throughout that the sequences in question are non-negative and bounded, and restrict attention to preferences that are represented by means of a monotone weakly decreasing ***discount factor*** $f : \mathbf{N} \to \mathbf{R}_+$. So a sequence x

is ranked above y for the discount factor f if $\sum_{t=0}^{\infty} f(t)x_t \geq \sum_{t=0}^{\infty} f(t)y_t$. In fact we shall take f to be a summable sequence and have values in $(0,1]$.

Of course, these assumptions are not without loss. We restrict attention to preferences with a linear utility representation, $x \mapsto \sum_t x_t f(t)$. The linear representation presumes a form of independence, or separability; but these assumptions are well understood and merit no further discussion here.[3] Future payoffs are discounted, as the values of f are weakly decreasing. Moreover the assumption that discount factors are summable expresses a particular form of impatience (no weight is placed "at infinity."). The linear structure encapsulates the idea that our sequences represent utils, as the marginal rate of intertemporal substitution depends on only on dates and not consumption.

Formally, then, the objects of choice are bounded non-negative sequences: elements of ℓ^∞_+. We consider preferences $\succeq$ on ℓ^∞_+ for which there is a monotone weakly decreasing $f \in \ell^1_+$ where $x \succeq y$ iff $\sum_{t=0}^{\infty} f(t)x(t) \geq \sum_{t=0}^{\infty} f(t)y(t)$. The class of such preferences is denoted by $\mathcal{P}$.

A ***dated reward*** is a sequence that is identically zero, except for at most one value t. Dated rewards are identified as pairs (x,t), which denotes a sequence that is zero everywhere and equal to $x \geq 0$ at time t. Let $\mathcal{D}$ denote the set of all dated rewards. If $\succeq \in \mathcal{P}$ is one of the preferences under consideration, we have that $(x,t) \succeq (y,s)$ if and only if $xf(t) \geq yf(s)$. So we can say that an agent with preferences $\succeq$ is happy to delay consumption of x at period t in exchange for $y \geq x$ at $t+1$ if and only if $\frac{x}{y} \geq \frac{f(t+1)}{f(t)}$. In particular, the agent is indifferent between consuming or delaying when $\frac{x}{y} = \frac{f(t+1)}{f(t)}$: so the magnitude $f(t+1)/f(t)$ is an expression of how ***impatient*** the agent is when it comes to consumption in periods t and $t+1$.

The main focus of our paper are preferences for which the ratio $f(t+1)/f(t)$ is monotone weakly increasing. Such preferences, and their associated discount factors, are said to satisfy ***decreasing impatience***.

A preference $\succeq \in \mathcal{P}$ with associated discount factor f is ***stationary*** if the ratio $f(t+1)/f(t)$ is constant; independent of t. It is well known, and easy to see, that this case corresponds to the existence of $\delta \in (0,1]$ and a scalar A for which $f(t) = A\delta^t$ (indeed, $A = f(0)$). The discount factor is then associated with a

[3]Loewenstein and Prelec (1992) advocate for $\mathcal{P}$ as a model of intertemporal choice, and then impose a version of decreasing impatience that they show implies hyperbolic discounting.

constant ***exponential discount rate*** δ. The subclass of stationary preferences, also called exponential discounting preferences, is denoted by $\mathcal{P}^S$.

In our discussion, stationarity and decreasing impatience are defined as properties of f. They can also be defined as behavioral properties of $\succeq$: ***decreasing impatience*** says that if $x > y$, $s < t$, and $(x, t) \succeq (y, s)$ then $(x, t+r) \succeq (y, s+r)$ for all $r > 0$. Stationarity strengthens this to be an "if and only if" statement, holding for all x, y (see Chakraborty (2021) for an eloquent discussion of these properties). It is easy to see that the behavioral definitions are equivalent to our definitions, within the class $\mathcal{P}$, see *e.g.* Prelec (2004).

In applications, it is common to model decreasing impatience through a β-δ, or quasi-hyperbolic, discount factor (Laibson, 1997): these take the form $f(t) = \beta^{\min\{t,1\}}\delta^t$, with $\delta, \beta \in (0, 1]$, so that $f(t+1)/f(t)$ goes from $\beta\delta$ when $t = 0$ to δ for all $t > 0$. The idea is that period $t = 0$ plays a special role. We are interested in a generalization of this model that extends this special role to all initial periods: $t = 0, \dots, t^* - 1$ for some $t^* \geq 1$.

Specifically, say that a discount factor f is ***generalized*** β**-**δ if there are $\beta, \delta \in (0, 1]$ and t^* such that

$$f(t) = \begin{cases} f(0)(\beta\delta)^t \text{ if } t \leq t^* \\ f(0)\beta^{t^*}\delta^t \text{ if } t > t^*. \end{cases}$$

In a generalized β-δ discount factor, the measure of impatience $f(t + 1)/f(t)$ goes from $\beta\delta$ in periods $t = 0, \dots t^* - 1$ to δ in periods $t \geq t^*$. The standard quasi-hyperbolic model obtains when $t^* = 1$, and exponential discounting when $\beta = 1$.

3. Main results

Before we state our main results, we introduce a few preliminary ideas. The first is a behavioral axiom: a pattern of intertemporal choice which says that for any principal k and rate of return r, investing half of k at maturity $t - 1$, and half at maturity $t + 1$, is always preferred to investing all of k at maturity t. The pattern is called ***Compound-interest convexity.*** Formally, the statement of the axiom is:

Axiom (Compound-interest convexity)**.** *For all $k > 0$, all $t \geq 1$ and all $r > 0$,*

$$\left(\frac{k}{2}(1 + r)^{t-1}, t - 1\right) + \left(\frac{k}{2}(1 + r)^{t+1}, t + 1\right) \succeq (k(1 + r)^t, t).$$

Recall that the dated reward notation (x,t) refers to a sequence in ℓ^∞, making the addition of dated rewards meaningful. As we shall see, within the class $\mathcal{P}$, compound-interest convexity characterizes decreasing impatience.

The other two notion are related to aggregating discount factors. The point will be that a discount factor satisfies decreasing impatience if and only if it is the aggregate of some basic parametric models of discounting.

The first method of aggregation is the geometric mean. Given a finite or countable collection of discount factors f_s, a ***weighted geometric mean*** is $\prod_s f_s(t)^{\eta_s}$, for some $\eta_s > 0$ with $\sum_s \eta_s = 1$. Importantly, in Section 4 we show that the weighted geometric mean of a finite number of discount factors is the unique aggregation rule that uniquely satisfies a notion of dynamic consistency, together with some standard normative axioms.

A second, perhaps unexpected, connection to decreasing impatience comes from the method of "parimutuel aggregation" introduced by Eisenberg and Gale (1959).[4] The idea is to use a market mechanism (or a pseudomarket, where agents use exogenously given incomes to purchase goods) and have the discount factor arise as an equilibrium price.

A ***parimutuel economy*** is a collection $(\succeq_i, w_i)_{i\in I}$, where I is finite or countable, each $\succeq_i$ being a preference relation in $\mathcal{P}^S$ (meaning a stationary preference over streams in ℓ^∞_+) and $w_i > 0$ satisfying $\sum_i w_i = 1$. In words, a parimutuel economy consists of a set I of agents with exponential preferences and strictly positive income w_i. Aggregate income sums to one.

We restrict attention to parimutuel economies with a finite, or countable, set of agents and a unit supply of "good," or money, per period. An ***allocation*** in a parimutuel economy is a collection $x = (x_i)_{i\in I}$ of sequences in ℓ^∞_+ with the property that

$$\sum_{i\in I} x_i(t) = 1$$

for all t.

A ***parimutuel equilibrium*** in $(\succeq_i, w_i)$ is a pair (x^*, p^*) in which x^* is an allocation and $p^* \in \ell^1_+$ is a sequence of prices, for which x^*_i is maximal for $\succeq_i$ in the budget set

$$\{x \in \ell^\infty_+ : \sum_t p(t)x(t) \leq w_i\}.$$

[4]The idea in this paper is generalized in Eisenberg (1961).

Finally, for any two sequences f and g, $f \propto g$ means that there is some $\alpha > 0$ for which $f = \alpha g$. Observe if f and g are discount factors, then $f \propto g$ means that they represent the same $\succeq$.

3.1. **Characterization of decreasing impatience.**

Theorem 1. *Let $\succeq$ be a preference in $\mathcal{P}$, with associated discount factor f, and suppose that $f(t+1)/f(t)$ is bounded away from 1. Normalize f so that $\sum_t f(t) = 1$. The following statements are equivalent:*

(1) *$\succeq$ satisfies decreasing impatience.*
(2) *$\succeq$ satisfies compound-interest convexity.*
(3) *f is proportional to the (finite or countable) geometric mean of generalized β-δ discount factors. That is, there are β-δ discount factors f_s, and $\eta_s > 0$ with $\sum_s \eta_s = 1$, such that $f(t) \propto \prod_s f_s(t)^{\eta_s}$.*
(4) *There exists a parimutuel economy, and a parimutuel equilibrium (x^*, p^*) for which $p^*_t = f(t)$ for all t.*

Remark. Observe that in order to falsify compound-interest convexity, it is sufficient find a single observation

$$\left(\frac{k}{2}(1+r)^{t-1}, t-1\right) + \left(\frac{k}{2}(1+r)^{t+1}, t+1\right) \prec (k(1+r)^t, t).$$

The significance of this observation is that, in a framework of dated rewards, decreasing impatience requires at least two observations to falsify. Multiple observations of choices in experimental economics usually require resorting to some type of random problem selection, see *e.g.* Azrieli et al. (2018), and thus committing to a theory of behavior over random outcomes. The tradeoff is that we must commit to a theory over consumption streams for compound-interest convexity to be meaningful.

Remark. Behavioral economists (see for example Laibson (1997) or Diamond and Köszegi (2003)) have used the β-δ model for its analytical tractability, as an approximation to the hyperbolic discount factor model.[5] The equivalence between (1) and (3) means, in contrast, that generalized β-δ preferences are, in a sense, the canonical model of decreasing impatience. One can always think of a discount factor satisfying decreasing impatience as an aggregate of β-δ discount

[5]See Loewenstein and Prelec (1992) and the references therein for a discussion of hyperbolic discounting. A form of the β-δ model was axiomatized by Hayashi (2003).

factors. Moreover, as emphasized by Theorem 2 below, the geometric mean as an aggregator of discount rates satisfies a notion of dynamic consistency. (As will become apparent below, this aggregator makes the most sense when restricting to $\mathcal{D}$.)

There is then a sense in which any dynamic inconsistency displayed by a preference with decreasing impatience can be traced to the β-δ model.

Remark. An inspection of the proof of Theorem 1 establishes that decreasing impatience implies something slightly stronger than compound-interest convexity. In particular, decreasing impatience implies that for all $k > 0$, all $t \geq 1$ and all $\beta > 0$,

$$\left(\frac{k}{2}\beta^{t-1}, t-1\right) + \left(\frac{k}{2}\beta^{t+1}, t+1\right) \succeq (k\beta^t, t).$$

This stronger hypothesis provides further simple tests for refuting the hypothesis of decreasing impatience.

Remark. The hypothesis that $f(t+1)/f(t)$ is bounded away from 1 guarantees that each of the β-δ preferences obtained in 3 has exponents that are strictly smaller than 1. If we only assume that f is strictly decreasing, the remaining equivalences in the theorem continue to hold.

Remark. The equilibrium we construct to prove the equivalence of parimutuel prices for stationary agents features a countable number of agents, each of whom consumes a dated reward in equilibrium.

4. Multiplicative aggregators

In this section, we discuss a formal model of preference aggregation, and establish the class of multiplicative aggregators as the unique ones satisfying a collection of properties. The idea here is that utility is common, and consumption is public. That is, agents are asked to rank consumption; the only disagreements are about discount factors. We envision the exercise here as making the most sense for the domain $\mathcal{D}$ of dated rewards. We imagine that the goal is to aggregate a group of individual discount factors into a social one, and impose several properties on how this aggregation takes place. Key amongst our assumptions are a Pareto property *for dated rewards only* and a dynamic consistency property.

Given two discount factors f and f', recall that the notation $f \propto f'$ means that there is $\alpha > 0$ for which $f = \alpha f'$. And, for any time period t, the t-translation of f is the discount factor denoted by f^t defined as $f^t(s) \equiv f(t+s)$. The set of all discount factors is denoted by $\mathcal{NI}$.[6]

Given is a finite set of agents $M \equiv \{1, \ldots, m\}$, indexed by $i \in M$. An ***aggregator*** is a function $\varphi : \mathcal{NI}^M \to \mathcal{NI}$. The idea behind an aggregator is that there is a "social preference," which takes the same form as individual preference. In general, the aggregate sequence of discount factors may depend on the entire sequence of discount factors for every individual agent.

An aggregator φ is a ***geometric mean*** if there exists $\eta_i > 0$ for each $i \in M$ such that $\sum_i \eta_i = 1$ and

$$\varphi(f_1, \ldots, f_m) \propto \phi_\alpha(f_1, \ldots, f_m)(t) = \prod_{i \in M} f_i(t)^{\alpha_i},$$

for all $(f_1, \ldots, f_m) \in \mathcal{NM}^M$.

We postulate the following axioms:

- (Ordinality) If for all $i \in M$, $f_i \propto f_i'$, then $\varphi(f_1, \ldots, f_m) \propto \varphi(f_1', \ldots, f_m')$.
- (Pareto) If for all $i \in M$, $f_i(t)x \geq f_i(s)y$, then $\varphi(f_1, \ldots, f_n)(t)x \geq \varphi(f_1, \ldots, f_n)(s)y$, with a strict inequality if any individual inequality is strict.
- (Independence of Irrelevant Alternatives) For any t. For all $f, f' \in \mathcal{NI}^M$, if for all $i \in M$ and all $x, y \in \mathbf{R}_{++}$: $f_i(t)x \geq f_i(0)y$ iff $f_i'(t)x \geq f_i'(0)y$, then $\varphi(f)(t)x \geq \varphi(f)(0)y$ iff $\varphi(f')(t)x \geq \varphi(f')(0)y$.
- (Dynamic Consistency) For all $t \geq 0$, $\varphi(f_1^t, \ldots, f_n^t) \propto \varphi(f_1, \ldots, f_n)^t$.

Given the role of our next result in Theorem 1, we want to emphasize the Dynamic Consistency axiom.

Suppose that we ask individuals about their preference between two dated rewards when t periods have passed, each of them treating time t as if it were the new period 0. And suppose that the aggregator then judges (x, s) to be preferred to (x', s'). Dynamic consistency requires that a the aggregator at the original time 0 should judge $(x, s+t)$ to be preferred to $(x', s'+t)$. Otherwise a plan for choosing $(x', s'+t)$ over $(x, s+t)$ would be reversed when time t arrives.

The remaining axioms should be familiar. Ordinality requires that the aggregator only uses information about how a discount factor ranks dated rewards. Pareto is the usual Pareto efficiency axiom restricted to dated rewards. Finally Independence of Irrelevant Alternatives (IIA) is a version of Arrow's IIA: it says

[6]The notation is a nod to the fact that our result holds if we restrict attention to all monotone nonincreasing and positive discount factors.

that given a period t, the aggregator should only use information about the sets of agents that ranks a date t reward against a period 0 reward.

Theorem 2. *An aggregator satisfies Ordinality, Pareto, IIA, and Dynamic Consistency if and only if it is a geometric mean.*

Remark. The geometric mean of summable discount factors may not be summable, but if the discount factors in question are generalized β-δ with $\delta < 1$ then their geometric mean is guaranteed to be summable. In any case, summability is not needed to rank dated rewards, which has been our focus in this section.

5. General parimutuel markets

Equilibria in parimutuel economies may be viewed as solutions to a particular kind of social welfare maximization problem. Indeed, Samuelson (1956) proposed a general aggregation procedure whereby a representative consumer arises from the maximization of a social welfare functional. With the right prices, these solutions can be decentralized, as in the second welfare theorem, to provide individual optimizing behavior. For Eisenberg-Gale aggregation in our context, the social welfare function in question is the so-called Nash welfare (Nash, 1950)

$$W((u_i)_{i\in I}) = \prod_{i\in I} u_i^{w_i},$$

where $u_i(x) = \sum_t x_t \delta_i^t$ represents $\succeq_i$. The social welfare maximization program is then

$$\begin{array}{cc} \max_{x_i \in \ell_+^\infty} & W((u_i)_{i\in I}) \\ \text{s.t} & \sum_i x_{i,t} = 1 \text{ for all } t. \end{array}$$

The equilibrium allocations identified in Theorem 1 solve this maximization problem, and equilibrium prices take the form of the upper envelope of "weighted" versions of the agents discount factors. An illustration is provided in Figure 1.

In the figure, there are three agents with exponential discount factors. The equilibrium prices are indicated in black, as the pointwise maximum of the agents weighted discount factors. It should be clear from the picture that the price exhibits decreasing impatience. Theorem 1 says that any discount factor with this property can be interpreted as such an equilibrium price.

Now, in Theorem 1, the discount factor was taken as the primitive starting point. In contrast, in this section we take a population of agents N as the starting point. Each of the agents $i \in N$ have preferences in $\mathcal{P}$, and we consider

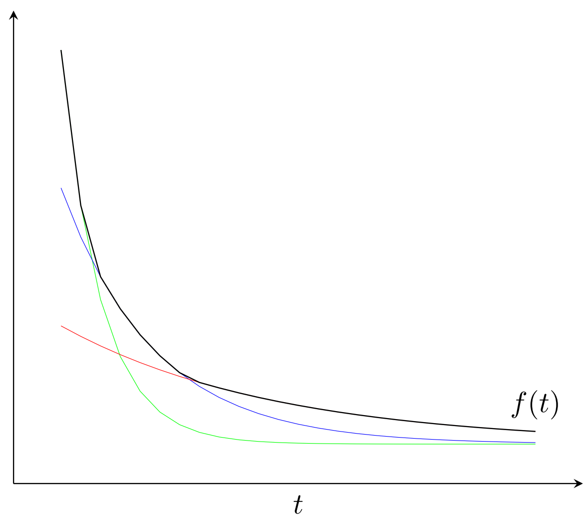

FIGURE 1. Parimutuel equilibrium price with three exponential discount factors: $\delta_1 < \delta_2 < \delta_3$.

an aggregate discount factor obtained through parimutuel aggregation. The next result concerns the structure of the set of possible prices for parimutuel equilibria with a given set of preferences. We study the set of possible prices as incomes vary.

The result we present is very general, covering environments with both continuous and discrete time. The reason for bothering with this level of generality is that it is usually much easier to compute examples in continuous time, so we want to have a result that can be applied to a continuous time model. On the other hand, the main results of the paper were stated for an environment with discrete time, and we want a result that applies to the same environments as Theorem 1.[7] In the end, it turns out that there is a common structure that works quite generally.

Let (Ω, Σ) be a measurable space, and for each $i \in N$, let δ_i be countably additive probability measure on (X, Σ). We assume that the set $\{\delta_i\}_{i \in N}$ is mutually absolutely continuous. The discrete time model is obtained when $(\Omega, \Sigma) = (\mathbf{N}, 2^{\mathbf{N}})$ and δ_i is identified with an exponential measure on $2^{\mathbf{N}}$ (that is, with a number $\hat{\delta}_i \in (0, 1)$ so that $\delta_i(A) = \sum_{t \in A}(1 - \hat{\delta}_i)\hat{\delta}_i^t$). The continuous-time model is obtained when $(\Omega, \Sigma) = (\mathbf{R}_+, \mathbf{B})$, where $\mathbf{B}$ is the Borel σ-algebra on $\mathbf{R}_+$ and δ_i is an exponential probability measure on $\mathbf{B}$.

[7]It is, however, not too difficult to obtain a version of Theorem 1 for continuous time. At least versions of the equivalence between statements (1), (2) and (4).

In a parimutuel market, $w_i \geq 0$ denotes i's wealth. Here we assume that $\sum_{i \in N} w_i > 0$. An ***economy*** then consists of probabilities and wealth. An ***allocation*** consists of, for each $i \in N$, $x_i \in Ł^{\infty}_{+}(\Omega, \Sigma)$, for which $\sum_{i \in N} x_i = \mathbf{1}$.[8]

A ***parimutuel equilibrium*** is a pair (p^*, x^*), consisting of a finite non-negative measure[9] p^* and an allocation $x^* = \{x_i^*\}_{i \in N}$ for which for all $i \in N$:

For all $g \in L^{\infty}_{+}(X, \Sigma)$, $\int g dp^* \leq w_i$ implies $\int g d\delta_i \leq \int x_i^* d\delta_i$.

Now, for any measure δ_i and any scalar α_i, $\alpha_i \delta_i$ denotes the scalar multiple of the measure. Then $\bigvee_{i \in N} \alpha_i \delta_i$ denotes the join of the measures in the pointwise dominance order. See, for example, Aliprantis and Border (2006), Theorem 10.56.

In particular, $p = \bigvee_i \alpha_i \delta_i$ exactly when there is a measurable partition $\{E_1, \ldots, E_n\}$ of Ω for which

(1) For all $E \in \Sigma$ and all $i \in N$, $p(E) \geq \alpha_i \delta_i(E)$.
(2) For all $E \in \Sigma$, $p(E) = \sum_{i \in N} \alpha_i \delta_i(E \cap E_i)$.

Proposition 1. *Suppose given $\{\delta_i\}_{i \in N}$ and $\{w_i\}_{i \in N}$ for which $\sum_{i \in N} w_i > 0$. For any equilibrium (x^*, p^*) of the corresponding economy, there is $\alpha_i \geq 0$, with $\sum_{i \in N} \alpha_i > 0$ for which $p = \bigvee_{i \in N} \alpha_i \delta_i$. Conversely, if there are $\alpha_i \geq 0$ for which $\sum_{i \in N} \alpha_i > 0$ and $p = \bigvee_{i \in N} \alpha_i \delta_i$, then there for all $i \in N$, there is $w_i \geq 0$ with $\sum_i w_i > 0$ for which p constitutes an equilibrium price in the resulting economy.*

The following establishes uniqueness of Eisenberg-Gale aggregation, supposing mutual absolute continuity of p_i. The proof essentially replicates the argument found in Eisenberg and Gale (1959).

Proposition 2. *In the framework of Proposition 1, if (x, p) and $(\bar{x}, \bar{p})$ are equilibria, where each of p and $\bar{p}$ is a probability measure, then $p = \bar{p}$.*

6. Discussion and conclusion

6.1. Linear aggregation.

Many previous studies have focused on linear aggregation of exponential discount rates (Zuber, 2011; Jackson and Yariv, 2015). Our Theorem 1 provides some alternative representations, but it turns out that it is possible to use related ideas to obtain a linear representation for any discount factor that displays decreasing impatience. That is, a statement analogous to the equivalence between (1) and (3) in the theorem, but with a linear function

[8]Of course, we could also describe the preferences induced by the probability measures δ_i as we did in the preceding sections.

[9]We do not impose countable additivity. In fact, this countable additivity will be shown to be a consequence of equilibrium.

instead of a multiplicative one. The representation is not, however, in terms of exponential or β-δ discount factors.

A very basic insight behind Theorem 1 is that the set of discount factors that satisfy decreasing impatience is convex, and so can be represented in terms of its extreme elements. Indeed, the proof of Theorem 1 reveals that f and g satisfy decreasing impatience if and only if $\beta^t f(t)$ and $\beta^t g(t)$ are convex functions, for $\beta > 1$. But then when $\lambda \in (0,1)$, $\beta^t(\lambda f(t)+(1-\lambda)g(t))$ is convex; establishing the convexity of the set of discount factors with the decreasing impatience property.

Now, using results from Langberg et al. (1980), one can show that for each of these extreme elements, there is a (potentially infinite) increasing sequence of discount factors, $(\beta_1, \beta_2, \ldots)$. Each discount factor, except possibly the first one, is used for *at least two consecutive periods*, meaning that for each β_l, there are two periods $t, t+1$ for which $\frac{f(t+2)}{f(t+1)} = \frac{f(t+1)}{f(t)} = \beta_l$. As these form extreme rays of the relevant class of discount factors, classical Paretian aggregation assuming linearity (as in (Harsanyi, 1955)) would mean that these discount factors form "canonical" ones from which all others can be built linearly. This establishes another kind of multiple selves representation. The extremal discount factors figuring in this representation are of course a proper superset of the ones invoked in part (3) of Theorem 1. We have chosen to emphasize the generalized β-δ discount factors because of their connection to popular models in behavioral economics.

Langberg et al. (1980) is devoted to decreasing failure life rate distributions. A decreasing failure rate in their paper is determined by log-convexity of the decumulative distribution function. The authors in that paper characterize the extreme points of the log-convex decumulative distribution functions. Our paper instead focuses on decreasing sequences of discount factors, but up to scale the mathematics behind the two concepts are identical: a decreasing nonnegative sequence that satisfies log-convexity. In the context of Langberg et al. (1980), summability is not a focus, but otherwise the concepts are the same; and a close inspection of their arguments establishes that summability poses no special issue.

6.2. **Sequences of transformations.** Condition (2) of Theorem 1 would allow us to provide a characterization of pairs $x, y \in \ell^\infty$ for which $x \succeq y$ for every $\succeq \in \mathcal{P}$ (we would additionally need to introduce linear inequalities asserting that discount rates are nonincreasing). Such a result would claim that $x \succeq y$ for all $\succeq \in \mathcal{P}$ iff x arises from y from a sequence of transformations; analogous to mean-preserving spreads as in Rothschild and Stiglitz (1970). A similar exercise appears in Chambers and Echenique (2020).

7. Proofs

7.1. **Proof of Theorem 1.** First note that (1) holds iff $\log f(t)$ is (discretely) convex in t: log convexity means that $2\log(f(t+1)) \leq \log(f(t)) + \log(f(t+2))$, or $\frac{f(t+1)}{f(t)} \leq \frac{f(t+2)}{f(t+1)}$. We claim that log convexity is equivalent to the convexity of $\beta^t f(t)$ for any $\beta > 1$, a property that is equivalent to Statement (2) in the theorem.[10]

Convexity of $\beta^t f(t)$ in t means that for every $t \geq 0$, $\beta^t f(t) + \beta^{t+2} f(t+2) \geq 2\beta^{t+1} f(t+1)$, hence

$$h(\beta) = \beta^2 f(t+2) - 2\beta f(t+1) + f(t) \geq 0.$$

So fix $t \geq 1$, and observe that h is convex in β as $f(t+2) > 0$. We solve for the minimum value of h over β: the first-order condition gives $2\beta f(t+2) - 2f(t+1) = 0$. Now, since $\beta = f(t+1)/f(t+2) \geq 1$ (as f is monotone decreasing), the minimum value of h is

$$\left(\frac{f(t+1)}{f(t+2)}\right)^2 f(t+2) - 2\left(\frac{f(t+1)}{f(t+2)}\right) f(t+1) + f(t) \geq 0.$$

Thus $f(t) \geq \frac{f(t+1)^2}{f(t+2)}$, which is log-convexity.

The converse implication is obtained by reversing the steps in the proof we just finished.

Now we show that (1) is equivalent to (4). We have seen that (1) is equivalent to log convexity of f. That $\log f(t)$ is monotone decreasing and convex holds if and only if it is the pointwise maximum of a collection A of decreasing affine functions. We may take A to be at most countable.

Each element of A is of the form $t \mapsto a - dt$, and hence identified with a pair (a, d) of scalars with $d > 0$. Each t can be associated with a member of A.

Consider then a parimutuel economy with $N = \mathbf{N}$, which is countable, and for which each $i \in N$ is associated with $(a_i, d_i) \in A$ for which $\log f(i) = a_i - id_i$ and $\log f(t) \geq a_i - td_i$. Then, $i \in N$ has preferences $\succeq_{(a_i,d_i)}$ associated with the

[10]The equivalence between decreasing impatience and log convexity is emphasized by Prelec (2004). See his Corollary 1. The equivalence between log convexity and the convexity of $\beta^t f(t)$ is essentially an idea from Montel (1928).

stationary discount factor $f_{(a_i,d_i)}(t) = (e^{-d_i})^t$. Let $x_i^*(i) = 1$ and zero otherwise. Let $w_i = f(i)$.

Observe that for each i, $\sum_t f(t)x_i(t) = f(i) = w_i$. Next, let $y \in \ell_+^\infty$ so that $\sum_t f(t)y(t) \leq w_i$. Then since $f(t) \geq e^{a_i}(e^{-d_i})^t$, it follows that $\sum_t e^{a_i}(e^{-d_i})^t y(t) \leq \sum_t f(t)y(t) \leq w_i$. Finally, $\sum_t e^{a_i}(e^{-d_i})^t x_i^*(t) = e^{a_i}(e^{-d_i})^i = f(i) = w_i$. So x_i^* is feasible and maximizes agent i's utility.

The converse, that any parimutuel equilibrium prices display log-convexity, proceeds as follows. Consider any parimutuel equilibrium (x^*, p^*) for an economy of agents with stationary discount factors. Let $f = p^*$, and let $t, t+1, t+2$ and let $j \in N$ for which $x_j^*(t+1) > 0$. Then $\frac{\delta_j^{t+1}}{f(t+1)} \geq \frac{\delta_j^t}{f(t)}$, which implies $\delta_j \geq \frac{f(t+1)}{f(t)}$. And $\frac{\delta_j^{t+1}}{f(t+1)} \geq \frac{\delta_j^{t+2}}{f(t+2)}$, which implies $\delta_j \leq \frac{f(t+2)}{f(t+1)}$. Conclude $\frac{f(t+2)}{f(t+1)} \geq \frac{f(t+1)}{f(t)}$.

Now we turn to the equivalence between (1) and (3). Observe that if f is positive, decreasing and satisfies non-decreasing impatience, then there exists $\gamma \in (0,1]$ with $f(t+1)/f(t) \uparrow \gamma$. Let $g(t) \equiv \gamma^{-t} f(t)$. Note that g is decreasing and satisfies log-convexity. To see that it is decreasing, observe that $\frac{g(t+1)}{g(t)} = \frac{f(t+1)}{\gamma f(t)} \leq 1$ as $f(t+1) \leq \gamma f(t)$. To see that it is log-convex, recall that log-convexity is the same as non-decreasing impatience, and observe that $\frac{g(t+2)}{g(t+1)} = \frac{f(t+2)}{\gamma f(t+1)} \geq \frac{f(t+1)}{\gamma f(t)} = \frac{g(t+1)}{g(t)}$, as f is log-convex.

This means that the sequence $h(t) = \log g(0) - \log g(t)$ is increasing, concave, and equals 0 at $t = 0$. We also have that $h(t+1) - h(t) = \log(g(t+1)/g(t)) \to 0$, as $g(t+1)/g(t) \to 1$ by definition of g. By Lemma 1 there exists $\alpha \in \ell_+^1$ with $h(t) = \sum_s \alpha(s) \min\{s, t\}$.

This tells us that

$$g(t) = g(0) \prod_{s=0}^{\infty} \max\{e^{-s\alpha(s)}, e^{-t\alpha(s)}\}.$$

Thus,

$$f(t) = \gamma^t g(t) = g(0) \prod_{s=0}^{\infty} \max\{\beta(s)^s \gamma(s)^t, (\beta(s)\gamma(s))^t\},$$

where

$$\gamma(s) = \gamma^{\frac{1}{2^{s+1}}}$$
$$\beta(s) = e^{-\alpha(s)}.$$

Fix any sequence $\eta_s > 0$ with $\sum_s \eta_s = 1$. For each $s = 0, \ldots$ the discount factor $t \mapsto \max\{\beta(s)^s\gamma(s)^t, (\beta(s)\gamma(s))^t\}$ is generalized $\beta - \delta$ with $\beta = (\beta(s)\gamma(s))^{1/\eta_s}$ and

$\delta = (\gamma(s))^{1/\eta_s}$, where the switch point is at s. Let f_s denote this discount factor. Then we have that $f(t) = g(0)\prod_s f_s(t)^{\eta_s}$.

Conversely, it is basic algebra to see that positive and log-convex functions are preserved under both products and powers:

- If $f, g > 0$ are log-convex, then so is $(fg)(t) = f(t)g(t)$.
- If $f > 0$ is log-convex and $\alpha > 0$, then so is $f^\alpha(t) = (f(t))^\alpha$.

Clearly each generalized $\beta - \delta$ discount factor is positive and log-convex. The result then follows for countable geometric means by taking limits.

In the proof we have used the following lemma, which is an analogue of a result of Blaschke and Pick (1916).

Lemma 1. *Suppose that f satisfies*

(1) $f(0) = 0$
(2) $f(t) \geq 0$ *for all* $t > 0$
(3) f *concave, increasing, satisfies* $\lim_{t\to\infty} f(t+1) - f(t) = 0$.

Then there exists $\alpha \in \ell^1_+$ for which for all t, $f(t) = \sum_{s=0}^{\infty} \alpha(s)\min\{s,t\}$.

Proof. Observe that if it holds that $f(t) = \sum_{s=0}^{\infty} \alpha(s)\min\{s,t\}$, then $f(t+1) - f(t) = \sum_{s\geq t+1} \alpha(s)$. So starting from f we may define, for $t \geq 1$, $\alpha(t) = -f(t+1) + 2f(t) - f(t-1) = 2[f(t) - (\frac{1}{2}f(t+1) + \frac{1}{2}f(t-1))] \geq 0$ as f is concave. Let $\alpha(0)$ be arbitrary.

Observe that $f(t)-f(t-1) = f(t+1)-f(t)+\alpha(t)$ and by induction $f(t)-f(t-1) = f(t+k+1)-f(t+k)+\sum_{s=0}^{k} \alpha(t+s)$. Since $\lim_{t\to\infty} f(t+k+1)-f(t+k) = 0$, we can conclude that α is summable. Further, this implies that

$$f(t+1) - f(t) = \sum_{s=0}^{\infty} \alpha(t+1+s) = \sum_{s=0}^{\infty} \alpha(s)[\min\{s, t+1\} - \min\{s,t\}]. \tag{1}$$

Finally, the function $f^*(t) \equiv \sum_{s=0}^{\infty} \alpha(s)\min\{s,t\}$ is well defined because $f^*(t) = \sum_{s=0}^{t} \alpha(s)s + \sum_{s=t+1}^{\infty} t\alpha(s)$, and we have already established that α is summable. Then Equation (1) establishes that for all t, $f(t) - f(t-1) = f^*(t) - f^*(t-1)$, and since $f(0) = 0 = f^*(0)$, we know that $f = f^*$. □

7.2. **Proof of Theorem 2.** The necessity of the axioms is for the most part immediate. To see that the Pareto axiom holds, let x, y for which for all i, $f_i(t)x \geq f_i(s)y$. Then if $x \geq 0 > y$ or $x > 0 \geq y$, the result is obvious. Otherwise, if $x, y > 0$, then $\frac{f_i(t)}{f_i(s)} \geq \frac{y}{x}$ so that $\frac{\prod_i f_i(t)^{\alpha_i}}{\prod_i f_i(s)^{\alpha_i}} = \prod_i \left(\frac{f_i(t)}{f_i(s)}\right)^{\alpha_i} \geq \frac{y}{x}$, so that $\prod_i f_i(t)^{\alpha_i} x \geq \prod_i f_i(s)^{\alpha_i} y$; with a strict inequality if any individual inequality is strict (since each $\alpha_i > 0$). A similar argument establishes the result when $x, y < 0$.

We turn them to showing that the axioms are sufficient. Let φ be an aggregator that satisfies the axioms. We shall prove that it is a geometric mean. Let $\mathcal{NNI}$ (for normalized non-increasing) be defined by $f \in \mathcal{NI}$ and $f(0) = 1$. Define the map $\hat{\varphi} : \mathcal{NNI}^M \to \mathcal{NNI}$ by $\hat{\varphi}(f) \equiv \frac{\varphi(f)}{\varphi(f)(0)}$. Observe that by the first axiom, $\hat{\varphi}(\frac{f_1}{f_1(0)}, \ldots, \frac{f_m}{f_m(0)}) = \frac{\varphi(f_1,\ldots,f_m)}{\varphi(f_1,\ldots,f_m)(0)}$ for any $f \in \mathcal{NNI}^M$.

By the independence property we may define a map $\varphi_t : [0,1]^M \to [0,1]$ via $\varphi_t(f_1(t), \ldots, f_m(t)) = \hat{\varphi}(f)(t)$, where each $f_i \in \mathcal{NNI}$.

By the Pareto property, for all t, s, $\varphi_t = \varphi_s$: suppose that $f_i(t) = f_i(s)$ for all $i \in M$. Then $(1, t)$ is ranked the same as $(1, s)$ for all agents, and therefore must be for the social ranking; so that $\varphi_t(f_1(t), \ldots, f_m(t)) = \varphi_s(f_1(s), \ldots, f_m(s))$.Write φ^* for φ_t. Observe similarly by Pareto that φ^* is strictly increasing in all coordinates, and that for any $x \in [0, 1]$, $\varphi^*(x, \ldots, x) = x$.

Now, we want to claim that for all $a, b \in [0,1]^M$ with $a \leq b$, we have $\frac{\varphi^*(a)}{\varphi^*(b)} = \varphi^*\left(\frac{a_1}{b_1}, \ldots, \frac{a_m}{b_m}\right)$.

To this end, let $f_1, \ldots, f_m \in \mathcal{NNI}$ for which $f_i(1) = b_i$ and $f_i(2) = a_i$. Observe that $\frac{f_i^1(1)}{f_i^1(0)} = \frac{a_i}{b_i}$, and that for each $i \in M$, $\frac{f_i^1}{f_i^1(0)} \in \mathcal{NNI}$. Then $\varphi^*\left(\frac{a_1}{b_1}, \ldots, \frac{a_m}{b_m}\right)$ is the social discount factor at period 1 relative to period 0 ascribed to $(f_1^1, \ldots, f_m^1)$. By dynamic consistency, this must be the same as the social discount factor at period 2 relative to period 0 ascribed to $(f_1, \ldots, f_m)$, which is $\frac{\varphi^*(f_1(2),\ldots,f_m(2))}{\varphi^*(f_1(1),\ldots,f_m(1))}$, or $\frac{\varphi^*(a_1,\ldots,a_m)}{\varphi^*(b_1,\ldots,b_m)}$. So indeed for all $a, b \in [0,1]^M$ with $a \leq b$, we have $\frac{\varphi^*(a)}{\varphi^*(b)} = \varphi^*\left(\frac{a_1}{b_1}, \ldots, \frac{a_m}{b_m}\right)$.

Observe that this is a form of the Cauchy functional equation. For $a, b \in [0,1]^M$ with $a \leq b$, we have

$$\frac{\varphi^*(a)}{\varphi^*(b)} = \varphi^*\left(\frac{a_1}{b_1}, \ldots, \frac{a_m}{b_m}\right).$$

We can define $\psi : (-\infty, 0]^M \to (-\infty, 0]$ as $\psi(x_1, \ldots, x_m) = \log \varphi^*(\exp(x_1), \ldots, \exp(x_m))$. Clearly $\psi(0, \ldots, 0) = 0$. Observe then that ψ satisfies $\psi(x-y) = \psi(x) - \psi(y)$ whenever $x \leq y$. Analogously, $\psi(x-y) + \psi(y) = \psi(x)$, when $x \leq y$, which can equivalently be written as $\psi(x) + \psi(y) = \psi(x+y)$ for any $x, y \leq 0$.

The result now follows from a standard Cauchy argument: observe that for any $x \leq 0$ and any $q \in \mathbb{Q}_+$, we get $\psi(qx) = q\psi(x)$. The monotonicity of ψ then implies that for any $c \in \mathbb{R}_+$, $\psi(cx) = c\psi(x)$.

Define $\eta_i \equiv -\psi(-\mathbf{1}_i) > 0$. Then $\psi(x) = \psi(\sum_i(-x_i)(-\mathbf{1}_i)) = \sum_i x_i\eta_i$, and $\sum_i \eta_i = -\psi(-\sum_i \mathbf{1}_i) = 1$ as $\varphi^*(x,\dots,x) = x$. Thus $\log\varphi^*(e^{x_1},\dots,e^{x_M}) = \sum_i x_i\eta_i$, and hence $\varphi^*(e^{x_1},\dots,e^{x_M}) = \prod_i (e^{x_i})^{\eta_i}$.

7.3. **Proof of Proposition 1. Establishing the direction given w_i**

First, fix $w_i \geq 0$ for which $\sum_i w_i > 0$. Without loss suppose that all $w_i > 0$; otherwise, we may discard agents for which $w_i = 0$ and proceed. Let (x^*, p^*) be a parimutuel equilibrium, and for each $i \in N$ define $E_i \equiv \{\omega : x_i(\omega) > 0\}$. Note that $\delta_i(E_i) > 0$.

Absolute continuity of δ_i with respect to p^*.

If $\delta_i(E) > 0$ for some $E \in \Sigma$ and $p^*(E) = 0$, then $\int x_i + \mathbf{1}_E d\delta_i > \int x_i d\delta_i$ yet $\int x_i + \mathbf{1}_E dp^* = \int x_i dp^*$, contradicting that (x^*, p^*) is an equilibrium.

Equilibrium prices are countably additive

We first show that for any $E \in \Sigma$, if $E \subseteq E_i$ and $p^*(E) > 0$, then $\int_E x_i^* dp^* > 0$. This follows as $p^*(E) > 0$ implies $\delta_i(E) > 0$ (otherwise i could increase wealth by selling her consumption on E), and $\int_E x_i d\delta_i > 0$ by countable additivity of δ_i. So $\int_E x_i dp^* = 0$ is not possible when (x^*, p^*) is an equilibrium, again because it would mean that i can raise her utility for free.

Now let us suppose by means of contradiction that p^* is not countably additive. Then there is some $y > 0$ and sequence $\{F_n\}_{n\in\mathbb{N}}$ for which $F_{n+1} \subseteq F_n$ and $\bigcap_n F_n = \varnothing$, but $p^*(F_n) \geq y$.

We may assume without loss that there is some $i \in N$ for which for all $n \in \mathbf{N}$, $\int_{F_n} x_i^* dp^* \geq y/|N|$. This follows as equilibrium implies that $p^*(F_n) = \sum_i \int_{F_n} x_i^* dp^*$ (x_i^* is an allocation), so for each n there is i for which $\int_{F_n} x_i^* dp^* \geq \frac{y}{|N|}$. We can just take an i that appears infinitely often.

We know by Lebesgue dominated convergence, and the countable additivity of δ_i, that $\int_{F_n} x_i^* d\delta_i \to 0$. So $\frac{\int_{F_n} x_i^* d\delta_i}{\int_{F_n} x_i^* dp^*} \to 0$. Pick n large so that $\int_{F_n} x_i^* d\delta_i < \frac{\int_{F_n} x_i^* dp^*}{p^*(\Omega)}$. Then $x_i^* - x_i^*|_{F_n} + \frac{\int_{F_n} x_i^* dp^*}{p^*(\Omega)}\mathbf{1}_\Omega$ is strictly preferred to x_i^* for agent i, and costs the same as x_i^*.

Establishing a property of ratios of measures

Second, we show that if $E \subseteq E_i$ and $F \in \Sigma$ then

$$\delta_i(F)p^*(E) \leq \delta_i(E)p^*(F). \tag{2}$$

Note that (2) is immediate if $p^*(E)=0$ or (by absolute continuity) if $p^*(F)=0$. Then to prove (2) suppose, towards a contradiction, that

$$\frac{\delta_i(F)}{p^*(F)} > \frac{\delta_i(E)}{p^*(E)}.$$

For $y>0$, let $E^y=\{\omega\in E : x_i^*(\omega)\geq y\}$. Note that $\cup_{y>0}E^y=E$, so the countable additivity of δ_i and p^* imply that there is $y>0$ with $\frac{\delta_i(F)}{p^*(F)} > \frac{\delta_i(E^y)}{p^*(E^y)}$. Now observe that

$$\int [x_i^* - y\mathbf{1}_{E^y} + y\frac{p^*(E^y)}{p^*(F)}\mathbf{1}_F]d\delta_i = \int x_i^* d\delta_i + y[\frac{p^*(E^y)}{p^*(F)}\delta_i(F)-\delta_i(E^y)] > \int x_i^* d\delta_i,$$

where $x_i^* - y\mathbf{1}_{E^y} + y\frac{p^*(E^y)}{p^*(F)}\mathbf{1}_F \geq 0$, while

$$\int [x_i^* - y\mathbf{1}_{E^y} + y\frac{p^*(E^y)}{p^*(F)}\mathbf{1}_F]dp^* = \int x_i^* dp^* + y[\frac{p^*(E^y)}{p^*(F)}p^*(F)-p^*(E^y)] = \int x_i^* dp^*;$$

a contradiction.

Establishing absolute continuity of p^* with respect to each δ_i

For any $G\in\Sigma$ with $p^*(G)>0$, $\sum_i x_i^* = \mathbf{1}$ implies that there is $G_j\subseteq E_j\cap G$ with $p^*(G_j)>0$. Then (2) with $F=\Omega$ implies that $\delta_j^*(G_j)>0$ which, by mutual absolute continuity of the (δ_i), implies that $0<\delta_i(G_j)\leq\delta_i(G)$.

Concluding this direction

Next, define $\alpha_i=\frac{p(E_i)}{\delta_i(E_i)}>0$. Then (2) implies that, for any $F\in\Sigma$, $p^*(F)\geq \alpha_i\delta_i(F)$. It also implies that for any $F_i\subseteq E_i$ $p^*(F)=\alpha_i\delta_i(F)$.

Finally, by $\sum_i x_i=\mathbf{1}$ we can find a collection $F_i\subseteq E_i$, for $i\in N$, pairwise disjoint, and with $F=\cup F_i$. Then

$$p^*(F)=\sum_{i\in N}\alpha_i\delta_i(F_i).$$

It follows that $p^*=\bigvee\alpha_i\delta_i$.

Establishing the converse direction, given α_i

Conversely, suppose that $p=\bigvee\alpha_i\delta_i$, for a collection $\alpha_i\geq 0$ and $\sum_i\alpha_i>0$. Let $\{E_i\}$ be a measurable partition of X with the property that $p(F)=\alpha_i\delta_i(F)$ for all $F\subseteq E_i$. Choose $E_i=\varnothing$ when $\alpha_i=0$. Set $w_i=\alpha_i\delta_i(E_i)$ and $x_i=\mathbf{1}_{E_i}$, so we have that $\int x_i dp = p(E_i)=\alpha_i\delta_i(E_i)=w_i$ and $\sum_i x_i=\mathbf{1}$.

Finally, suppose that g_i is such that $\int g_i dp \leq w_i = \alpha_i\delta_i(E_i)$.

First suppose that $\alpha_i>0$. Then since $p\geq\alpha_i\delta_i$, we have $\int g_i dp\geq\int g_i d(\alpha_i\delta_i) = \alpha_i\int g_i d\delta_i$. Conclude that $\int g_i d\delta_i\leq\delta_i(E_i)=\int x_i d\delta_i$.

Suppose now that $\alpha_j = 0$, and suppose there is g_j for which $\int g_j dp_j > 0$, but $\int g_j dp = 0$. We know that for any $E \in \Sigma$, $p(E) = \sum_{i \in N} \alpha_i \delta_i(E_i \cap E)$. So $\int g_j dp = \sum_i \alpha_i \int_{E_i} g_j d\delta_i$. Conclude that for every $i \in N$ for which $\alpha_i > 0$, $\int_{E_i} g_j d\delta_i = 0$. By mutual absolute continuity, this implies that $\int_{E_i} g_j d\delta_j = 0$, and in particular $\int g_j d\delta_j = 0$, a contradiction.

7.4. **Proof of Proposition 2.** First, let us suppose the economy is given, and that all δ_i are mutually absolutely continuous. As a consequence of the proof of Proposition 1, all of $\{\delta_i\}$ and $p, \bar{p}$ are mutually absolutely continuous. From here, we pick a probability measure μ with respect to which all measures are mutually absolutely continuous, and with a slight abuse of notation, refer to the Radon Nikodym derivative of any measure ν with respect to μ as $\nu \in L^1(\Omega, \mu)$.

All relevant statements below are understood to hold μ-almost everywhere, without further mention.

Now, as a first point, by Proposition 1, we have the existence of α_i and $\bar{\alpha}_i$, for each equilibrium.

It is easy to see that for any $\omega \in \Omega$ and any i, if $x_i(\omega) > 0$, then $p(\omega) = \alpha_i \delta_i(\omega)$, so that $p(\omega) x_i(\omega) \frac{1}{\alpha_i} p(\omega) = p(\omega) x_i(\omega) \delta_i(\omega)$. And since $\bar{\alpha}_i \delta_i(\omega) \leq \bar{p}(\omega)$, we conclude that $p(\omega) x_i(\omega) \delta_i(\omega) \leq p(\omega) x_i(\omega) \frac{1}{\bar{\alpha}_i} \bar{p}(\omega)$. Consequently:

$$p(\omega) x_i(\omega) \frac{1}{\alpha_i} p(\omega) \leq p(\omega) x_i(\omega) \frac{1}{\bar{\alpha}_i} \bar{p}(\omega).$$

Symmetrically,

$$\bar{p}(\omega) \bar{x}_i(\omega) \frac{1}{\bar{\alpha}_i} \bar{p}(\omega) \leq \bar{p}(\omega) \bar{x}_i(\omega) \frac{1}{\alpha_i} p(\omega).$$

By mutual absolute continuity, and by multiplying the two inequalities pointwise, we have that for every $(\omega, \omega') \in \Omega \times \Omega$, $p(\omega) x_i(\omega) \bar{p}(\omega') \bar{x}_i(\omega') p(\omega) \bar{p}(\omega') \leq p(\omega) x_i(\omega) \bar{p}(\omega') \bar{x}_i(\omega') \bar{p}(\omega) p(\omega')$.

Hence, since these densities are μ-almost everywhere strictly positive,

$$p(\omega) x_i(\omega) \bar{p}(\omega') \bar{x}_i(\omega') \frac{\bar{p}(\omega')}{p(\omega')} \leq p(\omega) x_i(\omega) \bar{p}(\omega') \bar{x}_i(\omega') \frac{\bar{p}(\omega)}{p(\omega)}$$

So, integrating with respect to the product measure $\mu \times \mu$ on $\Omega \times \Omega$ we obtain that

$$w_i \int \bar{p}(\omega') \bar{x}_i(\omega') \frac{\bar{p}(\omega')}{p(\omega')} d\mu(\omega') \leq w_i \int p(\omega) x_i(\omega) \frac{\bar{p}(\omega)}{p(\omega)} d\mu(\omega)$$

Since $w_i > 0$ and adding over $i \in N$ (which is finite), we may pass the sum inside the integral to obtain

$$\int \bar{p}(\omega)(\sum_i \bar{x}_i(\omega))\frac{\bar{p}(\omega)}{p(\omega)}d\mu(\omega) \leq \int p(\omega)(\sum_i x_i(\omega))\frac{\bar{p}(\omega)}{p(\omega)}d\mu(\omega).$$

Each of x_i and $\bar{x}_i$ is an allocation, so

$$\int \bar{p}(\omega)\frac{\bar{p}(\omega)}{p(\omega)}d\mu(\omega) \leq \int p(\omega)\frac{\bar{p}(\omega)}{p(\omega)}d\mu(\omega) = 1. \tag{3}$$

Observe that this inequality establishes that the function $g : \Omega \to \mathbf{R}$ defined by $g(\omega) = \frac{\bar{p}(\omega)}{\sqrt{p(\omega)}}$ satisfies $g \in L^2(\Omega, \mu)$. Further, the function $h : \Omega \to \mathbf{R}$ defined by $h(\omega) = \sqrt{p(\omega)}$ satisfies $h \in L^2(\Omega, \mu)$ as $\int h^2 d\mu = \int p d\mu = 1$.

By the Cauchy-Schwarz inequality:

$$\left(\int g(\omega)h(\omega)d\mu(\omega)\right)^2 \leq \int (g(\omega))^2 d\mu(\omega) \int (h(\omega))^2 d\mu(\omega)$$

Observe that since $\int (h(\omega))^2 d\mu(\omega) = 1$, the right hand side of this inequality is given by $\int \frac{\bar{p}(\nu)\bar{p}(\nu)}{p(\nu)} d\mu(\nu)$, which we know by equation (3) is bounded by 1. On the other hand, we also know that $\int ghd\mu = \int \bar{p} d\mu = 1$. Conclude that $\left(\int ghd\mu\right)^2 = \int g^2 d\mu \int h^2 d\mu$. The Cauchy-Schwarz inequality, however, only holds with equality for collinear vectors. So we may conclude that that $g = \beta h$ almost everywhere, for some $\beta > 0$, from which we conclude using the definitions of g and h, that $\bar{p}(\omega) = \beta p(\omega)$, which implies that $p = \bar{p}$ almost everywhere as each of them are densities of probability measures. So $\bar{p} = p$.

References

Aliprantis, C. D. and K. C. Border (2006): *Infinite Dimensional Analysis: A Hitchhiker's Guide, third edition*, Springer.

Azrieli, Y., C. P. Chambers, and P. J. Healy (2018): "Incentives in experiments: A theoretical analysis," *Journal of Political Economy*, 126, 1472–1503.

Blaschke, W. and G. Pick (1916): "Distanzschätzungen im Funktionenraum II," *Mathematische Annalen*, 77, 277–300.

Chakraborty, A. (2021): "Present bias," Forthcoming, Econometrica.

Chambers, C. P. and F. Echenique (2018): "On multiple discount rates," *Econometrica*, 86, 1325–1346.

——— (2020): "The Pareto Comparisons of a Group of Exponential Discounters," *Mathematics of Operations Research*, 45, 622–640.

DeJarnette, P., D. Dillenberger, D. Gottlieb, and P. Ortoleva (2020): "Time lotteries and stochastic impatience," *Econometrica*, 88, 619–656.

Diamond, P. and B. Köszegi (2003): "Quasi-hyperbolic discounting and retirement," *Journal of Public Economics*, 87, 1839–1872.

Eisenberg, E. (1961): "Aggregation of utility functions," *Management Science*, 7, 337–350.

Eisenberg, E. and D. Gale (1959): "Consensus of subjective probabilities: The pari-mutuel method," *The Annals of Mathematical Statistics*, 30, 165–168.

Feng, T. and S. Ke (2018): "Social discounting and intergenerational Pareto," *Econometrica*, 86, 1537–1567.

Frederick, S., G. Loewenstein, and T. O'Donoghue (2002): "Time discounting and time preference: A critical review," *Journal of economic literature*, 40, 351–401.

Genest, C. (1984): "A characterization theorem for externally Bayesian groups," *Annals of Statistics*, 12, 1100–1105.

Gillen, B. J., C. R. Plott, and M. Shum (2017): "A pari-mutuel-like mechanism for information aggregation: A field test inside Intel," *Journal of Political Economy*, 125, 1075–1099.

Halevy, Y. (2015): "Time consistency: Stationarity and time invariance," *Econometrica*, 83, 335–352.

Harsanyi, J. C. (1955): "Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility," *Journal of Political Economy*, 63, 309–321.

Hayashi, T. (2003): "Quasi-stationary cardinal utility and present bias," *Journal of Economic Theory*, 112, 343–352.

——— (2016): "Consistent updating of social welfare functions," *Social Choice and Welfare*, 46, 569–608.

Hayashi, T. and M. Lombardi (2019): "Social discount rate: spaces for agreement," Working paper.

Jackson, M. O. and L. Yariv (2015): "Collective dynamic choice: the necessity of time inconsistency," *American Economic Journal: Microeconomics*, 7, 150–78.

Laibson, D. (1997): "Golden eggs and hyperbolic discounting," *Quarterly Journal of Economics*, 112, 443–478.

Langberg, N. A., R. V. León, J. Lynch, and F. Proschan (1980): "Extreme points of the class of discrete decreasing failure rate life distributions," *Mathematics of Operations Research*, 5, 35–42.

Loewenstein, G. and D. Prelec (1992): "Anomalies in intertemporal choice: Evidence and an interpretation," *The Quarterly Journal of Economics*, 107, 573–597.

McConway, K. J. (1981): "Marginalization and linear opinion pools," *Journal of the American Statistical Association*, 76, 410–414.

Millner, A. (2020): "Nondogmatic Social Discounting," *American Economic Review*, 110, 760–75.

Montel, P. (1928): "Sur les fonctions convexes et les fonctions sousharmoniques," *Journal de Mathématiques Pures et Appliquées*, 7, 29–60.

Nash, J. F. (1950): "The bargaining problem," *Econometrica*, 18, 155–162.

O'Donoghue, T. and M. Rabin (1999): "Doing it now or later," *American Economic Review*, 89, 103–124.

Plott, C. R., J. Wit, and W. C. Yang (2003): "Parimutuel betting markets as information aggregation devices: experimental results," *Economic Theory*, 22, 311–351.

Prelec, D. (2004): "Decreasing impatience: a criterion for Non-stationary time preference and "hyperbolic" discounting," *Scandinavian Journal of Economics*, 106, 511–532.

Rothschild, M. and J. E. Stiglitz (1970): "Increasing risk: I. A definition," *Journal of Economic theory*, 2, 225–243.

Samuelson, P. A. (1956): "Social Indifference Curves," *The Quarterly Journal of Economics*, 70, 1–22.

Treasury, H. M. (2020): "The green book," *Appraisal and evaluation in central government*.

Weitzman, M. L. (1998): "Why the far-distant future should be discounted at its lowest possible rate," *Journal of environmental economics and management*, 36, 201–208.

——— (2001): "Gamma discounting," *American Economic Review*, 91, 260–271.

West, M. (1984): "Bayesian aggregation," *Journal of the Royal Statistical Society: Series A (General)*, 147, 600–607.

Zuber, S. (2011): "The aggregation of preferences: can we ignore the past?" *Theory and decision*, 70, 367–384.