Monotone Additive Statistics*

Xiaosheng Mu[†] Luciano Pomatto[‡] Philipp Strack[§] Omer Tamuz[¶]

March 22, 2024

Abstract

The expectation is an example of a descriptive statistic that is monotone with respect to stochastic dominance, and additive for sums of independent random variables. We provide a complete characterization of such statistics, and explore a number of applications to models of individual and group decision-making. These include a representation of stationary monotone time preferences, extending the work of Fishburn and Rubinstein (1982) to time lotteries. This extension offers a new perspective on risk attitudes toward time, as well as on the aggregation of multiple discount factors. We also offer a novel class of nonexpected utility preferences over gambles which satisfy invariance to background risk as well as betweenness, but are versatile enough to capture mixed risk attitudes.

1 Introduction

How should a random quantity be summarized by a single number? In Bayesian statistics, point estimators capture an entire posterior distribution. In finance, risk measures quantify the risk of a financial position. And in economics, certainty equivalents characterize an agent's preference for uncertain outcomes.

We use the term *descriptive statistic*, or simply *statistic*, to refer to a map that assigns a number to each bounded random variable. We study statistics that are monotone with respect to first-order stochastic dominance, and additive for sums of independent random variables. An example of a monotone additive statistic is the expectation. The median is monotone but not additive, while the variance is additive, but not monotone.

*We would like to thank Kim Border, Sebastian Ebert, Giacomo Lanzani, George Mailath, Meg Meyer, Pietro Ortoleva, Doron Ravid and Weijie Zhong for helpful comments and suggestions.

[†]Princeton University. Email: xmu@princeton.edu.

[‡]Caltech. Email: luciano@caltech.edu.

[§]Yale University. Email: philipp.strack@yale.edu. Philipp Strack was supported by a Sloan fellowship.

[¶]Caltech. Email: tamuz@caltech.edu. Omer Tamuz was supported by a grant from the Simons Foundation (#419427), a Sloan fellowship, a BSF award (#2018397) and a National Science Foundation CAREER award (DMS-1944153).

Monotonicity is a well studied property of statistics (see, e.g., Bickel and Lehmann, 1975a,b), and holds, for example, for certainty equivalents of monotone preferences over lotteries. Additivity is a stronger assumption. We focus on this property because of its conceptual simplicity and because it serves as a baseline assumption in many settings. As we argue, additivity corresponds to stationarity in the context of preferences over time lotteries (see §3). In the context of choices over monetary gambles it corresponds to invariance to background risk (§4.1), or to a form of separability across independent decision problems (§4.6).

Beyond the expectation, an additional example of a monotone additive statistic is the map K_a that, given $a \in \mathbb{R}$, assigns to each random variable X the value

$$K_a(X) = \frac{1}{a} \log \mathbb{E}\left[\mathrm{e}^{aX}\right]. \tag{1}$$

In the fields of probability and statistics, this function is known as the *(normalized) cumulant generating function* evaluated at a. In finance it is called the *entropic risk measure*. In economics, it corresponds to the certainty equivalent of an expected utility maximizer who exhibits constant absolute risk aversion (CARA) over gambles. For bounded random variables, the essential minimum and maximum provide further examples of such statistics; as we explain later, they are the limits of K_a as a approaches $\pm\infty$. The expectation is equal to K_0, the limit of K_a as a approaches 0.

Our main result establishes that these examples, and their weighted averages, are the only monotone additive statistics. That is, we show that if a statistic Φ is monotone, additive and normalized so that it satisfies $\Phi(c) = c$ for every constant c, then it is of the form

$$\Phi(X) = \int K_a(X)\,\mathrm{d}\mu(a)$$

for some probability measure μ over the extended real line. This result provides a simple representation of a natural family of statistics, which one may a priori have expected to be much richer.

Our first application is to time preferences. The starting point for our analysis is the work by Fishburn and Rubinstein (1982), who study preferences over dated rewards—a monetary reward, together with the time at which it will be received. They show that exponential discounting of time arises from a set of axioms, of which the most substantial, *stationarity*, postulates that preferences between two dated rewards are unaffected by the addition of a common delay.

We extend the analysis of Fishburn and Rubinstein (1982) to *time lotteries*, which consist of a monetary reward x and a random time T at which it will be received. In this setting, we too introduce a stationarity axiom that requires preferences to be invariant with respect to random independent delays. As we argue in the main text, this stationarity axiom captures an assumption of dynamic consistency, together with the idea that preferences do not depend on calendar time.

We show that a monotone and stationary preference over time lotteries admits a representation of the form

$$u(x)\mathrm{e}^{-r\Phi(T)},$$

where Φ is a monotone additive statistic (Theorem 3). Thus, $\Phi(T)$ is the certainty equivalent of the random time T, i.e. the deterministic time that is as desirable as T. Over deterministic dated rewards, the above representation coincides with standard discounted utility. General time lotteries are reduced to deterministic ones through the certainty equivalent Φ. By our main representation theorem, it takes the form $\Phi(T) = \int K_a(T)\,\mathrm{d}\mu(a)$. In this context, each $K_a(T)$ is the certainty equivalent of T under an expected discounted preference with discount rate $-a$. The different certainty equivalents are then averaged according to the measure μ.

In this representation it is as if the decision maker had in mind not one but multiple discount factors. Thus, Φ can be interpreted as the certainty equivalent of a decision maker who is uncertain about the correct discount factor, or as the aggregated certainty equivalent of a group of agents with different discount factors.

Our representation theorem for monotone and stationary time preferences has implications for understanding the relation between stationarity and risk attitudes toward time. How people choose among prospects that involve risk over time has been studied both theoretically and experimentally (Chesson and Viscusi, 2003; Onay and Öncüler, 2007; Ebert, 2020; DeJarnette et al., 2020; Ebert, 2021). A basic paradox these papers highlight is that many subjects display risk aversion over the time dimension, even though the standard theory of expected discounted utility predicts that people are risk-seeking with respect to time lotteries.

This raises the question of whether there is a tractable and well-motivated class of preferences that allows risk-aversion, risk-seeking, or a combination of the two, as well as stationarity. As pointed out by DeJarnette, Dillenberger, Gottlieb, and Ortoleva (2020), this is not entirely obvious. The preferences we study in this paper offer a solution to this problem, as they satisfy a strong notion of stationarity while allowing for a variety of risk attitudes.

We further apply the characterization of monotone stationary preferences to the problem of aggregating heterogeneous time preferences. It is well known that when aggregating expected discounted utility preferences, a utilitarian approach that averages individual utility functions leads to present bias (see Jackson and Yariv, 2014, 2015). Based on this observation, the literature concludes that within expected discounted utility, it is impossible to aggregate individual preferences into a social preference unless the latter is dictatorial.

We show that this difficulty is not due to stationarity, but rather to an insistence on the idea that the social preference should conform to expected discounted utility. When preferences are allowed to belong to the more general class of monotone stationary preferences, then a social preference obtained by averaging the certainty equivalents of the individuals satisfies Pareto efficiency and stationarity (Proposition 3). Moreover, under some assumptions, every Paretian and stationary social preference is obtained in this way (Proposition 4).

Monotone additive statistics also find applications to models of choice over monetary gambles. It is well known that an expected utility agent whose preferences are invariant to independent background risk must have CARA preferences. This invariance property makes CARA utility functions useful modeling tools when the analyst does not observe the agents' wealth level or the additional risks they face (see, e.g., Barseghyan, Molinari, O'Donoghue, and Teitelbaum, 2018). Beyond expected utility, monotone preferences that are invariant to background risk have certainty equivalents that are monotone additive statistics. Thus, by our main theorem, any such preference can be represented by a weighted average of CARA certainty equivalents, where the mixing measure μ is now over the coefficient of absolute risk aversion. In this representation, the decision maker entertains multiple utility functions, each defining a different certainty equivalent. Every lottery is evaluated by averaging over these certainty equivalents.

An interesting feature of preferences represented by monotone additive statistics is that they can display behavior that is not uniformly risk-averse nor risk-seeking, such as that of an agent buying both lottery tickets and insurance (Friedman and Savage, 1948), all while maintaining invariance to background risk. At the same time, a potential difficulty for this class of preferences is that their defining parameter, the measure μ over the coefficient of risk aversion, is infinite-dimensional. To narrow down the parameter space, we focus on those preferences that also satisfy *betweenness*, a well-known weakening of the independence axiom that has been extensively studied in the literature (see Dekel, 1986; Gul, 1991; Cerreia-Vioglio, Dillenberger, and Ortoleva, 2020). We show that a preference represented by a monotone additive statistic Φ satisfies betweenness if and only if it is of the form

$$\Phi(X) = \beta K_{-a\beta}(X) + (1-\beta) K_{a(1-\beta)}(X).$$

The parameter $\beta \in [0, 1]$ controls the the relative weights of the risk-averse and risk-seeking components, with increased β making the decision maker more risk-averse. The parameter a is a scale parameter. This is a simple two-parameter family, but it is rich enough to accommodate preferences that are neither risk-averse nor risk-seeking, while maintaining invariance to background risk.

Our final application concerns group decision-making under risk. We consider an organization that employs multiple agents, each of whom makes decisions following an individual preference relation, which can be seen as a decision rule prescribed by the firm. We show that in order for the agents' independent choices to not violate stochastic dominance when combined, it is sufficient and necessary that their preferences are represented by the same monotone additive statistic. Thus, these are the only preferences with the property that decentralized decisions cannot result in stochastically dominated outcomes for the organization.

1.1 Related Literature

A large literature in statistics studies descriptive statistics of probability distributions. A representative example is the work of Bickel and Lehmann (1975a,b), who study location statistics using an axiomatic, non-parametric approach that is similar to ours. This literature has however focused on different properties, and, to the best of our knowledge, does not contain a similar characterization of additivity and monotonicity. The mathematics literature has studied additive statistics as homomorphisms from the convolution semigroup to the real numbers (see Ruzsa and Székely, 1988; Mattner, 1999, 2004), but without imposing monotonicity.

In finance and actuarial sciences, $-K_a(X)$ is also known as an *entropic risk measure*, and is used to assess the riskiness of a financial position X. It is a canonical example of a coherent risk measure (see Föllmer and Schied, 2002, 2011; Föllmer and Knispel, 2011). In this literature, Goovaerts, Kaas, Laeven, and Tang (2004) study additive statistics that are monotone with respect to all entropic risk measures, i.e. those with the property that $K_a(X) \geq K_a(Y)$ for all $a \in \mathbb{R}$ implies $\Phi(X) \geq \Phi(Y)$, and show that they must be weighted averages of entropic risk measures, as in our main representation theorem. In contrast, we do not assume this property of Φ, but instead show that it is implied by monotonicity and additivity of Φ.

In an earlier paper, Pomatto, Strack, and Tamuz (2020) show that on the domain of random variables that have all moments, the only monotone additive statistic is the expectation.[1] This is because the larger domain includes fat-tailed random variables, which rule out all other monotone additive statistics. In contrast, in this paper we primarily study the domain of bounded random variables, which allows for much richer preferences with a variety of risk attitudes.

Monotone additive statistics also relate to what we called *additive divergences* in Mu,

[1] The same phenomenon is studied more in depth in Fritz, Mu, and Tamuz (2020). It is shown there that the expectation remains the unique monotone additive statistic on the domain L^p, for any $p \geq 1$, while there are no monotone additive statistics on L^p with $p < 1$, or on the domain of all random variables.

Pomatto, Strack, and Tamuz (2021). An additive divergence is a map defined over Blackwell experiments that satisfies monotonicity with respect to the Blackwell order and additivity for product experiments. While some of the techniques used in the two papers are similar, the main mathematical argument is fundamentally different. The last step of the proof of Theorem 1 is an application of the Riesz Representation Theorem and is similar to the argument used in the previous paper. However, because the Blackwell order has different properties from first-order stochastic dominance, the remainder of the proof is different, with the previous paper having no analogue of Theorem 6, which is the main technical step in the proof of Theorem 1. This new technique is also needed for the proof of Theorem 2, which characterizes monotone additive statistics beyond bounded random variables.

DeJarnette, Dillenberger, Gottlieb, and Ortoleva (2020) study preferences over time lotteries that display risk aversion. One class of preferences they propose is a generalization of expected discounted utility (GEDU) that for a random prize X delivered at a random time T takes the form $\mathbb{E}\left[\phi(u(X)\mathrm{e}^{-rT})\right]$ for some strictly increasing transformation ϕ. The curvature of ϕ determines the attitude towards risk. While GEDU satisfies stationarity for deterministic X and T, stationarity does not in general hold once random times T are considered, even with respect to adding a deterministic delay. In contrast, we impose a strong form of stationarity but do not insist on expected utility. The only intersection between our model and GEDU are preferences represented by K_a, corresponding to a point-mass mixing measure μ. These preferences have the standard EDU representation, but perhaps with a negative discount rate, as we explain in §3.3.[2]

Applied to choice under risk, our representation also bears resemblance to cautious expected utility theory (Cerreia-Vioglio, Dillenberger, and Ortoleva, 2015), in which a gamble is evaluated according to the minimum certainty equivalent across a family of utility functions. The two representations are conceptually related, as both involve uncertainty over a utility function. Our axioms are however different in that we study preferences that are invariant to adding an independent gamble, while Cerreia-Vioglio, Dillenberger, and Ortoleva (2015) consider the effect of mixing with another gamble.

Decision criteria that aggregate multiple certainty equivalents have appeared before in the literature. Myerson and Zambrano (2019) advocate the maximization of a sum of certainty equivalents as an effective rule for risk sharing. Chambers and Echenique (2012) formalize and characterize this rule as a social welfare functional.

The remainder of the paper is organized as follows. In §2 we introduce monotone additive statistics and state our main result. In §3 we apply this result to time lotteries,

[2] When the prize X is held constant, a GEDU preference reduces to an expected utility preference over random times. In contrast, our prizes are always deterministic, and the stationary preferences over random times are represented by monotone additive statistics, which are not expected utility unless the mixing measure μ is a point mass (see Proposition 8 in §F of the online appendix).

and in §4 we apply it to monetary gambles. Finally, §5 provides an overview of the proof of our main result. The appendix and online appendix contain omitted proofs for the results in the main text.

2 Monotone Additive Statistics

We denote by L^∞ the collection of bounded real random variables, defined over a nonatomic probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We will identify each $c \in \mathbb{R}$ with the corresponding constant random variable taking value $X(\omega) = c$ at each $\omega \in \Omega$.

We say that a map $\Phi\colon L^\infty \to \mathbb{R}$ is a *statistic* if it satisfies (i) $\Phi(X) = \Phi(Y)$ whenever $X, Y \in L^\infty$ have the same distribution, and (ii) $\Phi(c) = c$ for every $c \in \mathbb{R}$; that is, Φ assigns c to the constant random variable c. We are interested in statistics that satisfy monotonicity with respect to first-order stochastic dominance and additivity for sums of independent random variables. Formally, Φ is

- *additive* if $\Phi(X+Y) = \Phi(X) + \Phi(Y)$ whenever X and Y are independent, and
- *monotone* if $X \geq_1 Y$ implies $\Phi(X) \geq \Phi(Y)$, where $\geq_1$ denotes first-order stochastic dominance.

Since, by assumption, the value $\Phi(X)$ depends only the distribution of the random variable X, monotonicity is equivalent to the requirement that $\Phi(X) \geq \Phi(Y)$ whenever $X \geq Y$ almost surely. This equivalence is based on the well-known fact that $X \geq_1 Y$ if and only if there are random variables $\tilde{X}, \tilde{Y}$ such that X and $\tilde{X}$ are identically distributed, Y and $\tilde{Y}$ are identically distributed, and $\tilde{X} \geq \tilde{Y}$ almost surely.[3]

We denote by $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ the extended real numbers. Given $X \in L^\infty$ and $a \in \overline{\mathbb{R}} \setminus \{0, \pm\infty\}$, we consider the statistic

$$K_a(X) = \frac{1}{a} \log \mathbb{E}\left[\mathrm{e}^{aX}\right]. \tag{2}$$

The value $K_a(X)$ is the certainty equivalent of X for a CARA utility function with coefficient of risk aversion $-a$. In probability and statistics, $K_a(X)$ is known as the *(normalized) cumulant generating function* of X, evaluated at a.

If X and Y have the same distribution, then $K_a(X) = K_a(Y)$. Moreover, $K_a(c) = c$ for every $c \in \mathbb{R}$, so K_a is a statistic. If X and Y are independent, then $\mathbb{E}\left[\mathrm{e}^{a(X+Y)}\right] = \mathbb{E}\left[\mathrm{e}^{aX}\right]\mathbb{E}\left[\mathrm{e}^{aY}\right]$, and hence K_a is additive. It is also monotone.

[3] An alternative, equivalent definition for a statistic is to let the domain of Φ be the set of probability distributions on $\mathbb{R}$ with bounded support. In this domain, additivity would be defined with respect to convolution. We choose to have the domain consist of random variables, as this approach offers some notational advantages.

We additionally define $K_0(X), K_\infty(X), K_{-\infty}(X)$ to be the expectation, the essential maximum, and the essential minimum of X, respectively.[4] This choice of notation makes $a \mapsto K_a(X)$ a continuous function from $\overline{\mathbb{R}}$ to $\mathbb{R}$, for any X. Our main result is a representation theorem for monotone additive statistics:

Theorem 1. *$\Phi\colon L^\infty \to \mathbb{R}$ is a monotone additive statistic if and only if there exists a (unique) Borel probability measure μ on $\overline{\mathbb{R}}$ such that for every $X \in L^\infty$*

$$\Phi(X) = \int_{\overline{\mathbb{R}}} K_a(X)\,\mathrm{d}\mu(a). \tag{3}$$

We refer to μ as the *mixing measure* of the statistic Φ. Each K_a satisfies monotonicity and additivity, and it is immediate that these two properties are preserved under convex combinations. Theorem 1 says that the one-parameter family $\{K_a\}$ forms the extreme points of the set of monotone additive statistics; every such statistic must be a weighted average obtained by mixing over this family. In §5 we provide an overview of the proof of Theorem 1.

Theorem 1 can be extended to other domains of random variables. We consider the set L_M of random variables X for which $K_a(X)$, as defined in (2), is finite for all $a \in \mathbb{R}$. The domain L_M contains those unbounded random variables whose distributions have sub-exponential tails, as in the case of the Gaussian distribution.

Theorem 2. *$\Phi\colon L_M \to \mathbb{R}$ is a monotone additive statistic if and only if it admits a (unique) representation of the form (3) where the measure μ has compact support in $\mathbb{R}$.*

The extension of Theorem 1 to the larger domain L_M adds to the applicability of our representation, as it includes distributions with unbounded support, such as Gaussian or Poisson, for which the function K_a has closed-form expressions. For example, Theorem 2 implies that when applied to a Gaussian random variable Z, a monotone additive statistic Φ defined on L_M takes the simple mean-variance form $\Phi(Z) = \mathbb{E}[Z] + c\mathrm{Var}[Z]/2$, where $c \in \mathbb{R}$ is the mean of the measure μ characterizing Φ.

A few additional remarks are in order. First, Theorems 1 and 2 answer an open question in the mathematical finance literature on risk measures posed by Goovaerts, Kaas, Laeven, and Tang (2004), who asked if entropic risk measures are the only additive risk measures. Second, a possible strengthening of our additivity condition requires $\Phi(X+Y) = \Phi(X) + \Phi(Y)$ to hold for all pairs of random variables, rather than just the independent ones. As is well known, the only statistic that is additive in this more restrictive sense is the expectation (see, for example, de Finetti, 1970). A different strengthening is additivity with respect to uncorrelated random variables. It follows from the analysis

[4] The essential maximum and minimum are the maximum and minimum of the support: $\max[X] = \sup\{a : \mathbb{P}[X \le a] < 1\}$ and $\min[X] = \inf\{a : \mathbb{P}[X \le a] > 0\}$.

of Chambers and Echenique (2020) that on a finite probability space the expectation is, again, the only monotone statistic that is additive for uncorrelated random variables.

3 Monotone Stationary Time Preferences

Next, we apply monotone additive statistics to the study of time preferences. We consider decision problems where an agent is asked to choose between time lotteries that pay a fixed payoff at a future random time, as in the case of a driver choosing between different routes, where some routes are more likely than others to face heavy traffic, or a company choosing between projects with different random completion times. We argue that in this context additivity is connected to a notion of stationarity, according to which a choice between future rewards is not affected by the addition of an independent delay. In this section we study preferences over time lotteries that are monotone and stationary, characterize them using monotone additive statistics, discuss the risk attitudes they can model, and apply them to the problem of aggregating heterogeneous time preferences.

3.1 Domain and Axioms

A *time lottery* is a monetary reward received by a decision maker at a future, random time. Formally, it consists of a pair (x, T), where $x \in \mathbb{R}_{++}$ is a positive payoff and $T \in L^\infty_+$ is the random time at which it realizes.[5] Thus, time is non-negative and continuous. Our primitive is a complete and transitive binary relation $\succeq$ on the domain $\mathbb{R}_{++} \times L^\infty_+$. We denote by $\sim$ the indifference relation induced by $\succeq$. To avoid notational confusion, in the rest of this section x and y always denote monetary payoffs, t, s and d denote deterministic times, and T, S, and D denote random times.

We say that a preference relation $\succeq$ on $\mathbb{R}_{++} \times L^\infty_+$ is a *monotone stationary time preference* (henceforth, MSTP) if it satisfies the following axioms:

Axiom 3.1 (More is Better). *If $x > y$ then $(x, T) \succ (y, T)$.*

Axiom 3.2 (Earlier is Better). *If $s > t$ then $(x, t) \succ (x, s)$, and if $S \geq_1 T$ then $(x, T) \succeq (x, S)$.*

Axiom 3.3 (Stochastic Stationarity). *If $(x, T) \succeq (y, S)$ then $(x, T + D) \succeq (y, S + D)$ for any D that is independent from T and S.*

Axiom 3.4 (Continuity). *For any (y, S), the sets $\{(x, t) \colon (x, t) \succeq (y, S)\}$ and $\{(x, t) \colon (x, t) \preceq (y, S)\}$ are closed in $\mathbb{R}_{++} \times \mathbb{R}_+$.*

[5] Per standard notation, L^∞_+ denotes the set of non-negative bounded random variables.

The first two Axioms 3.1 and 3.2 are standard conditions that directly generalize those in Fishburn and Rubinstein (1982), who studied preferences over dated rewards (x, t) with a deterministic time t. They require the decision maker to prefer higher payoffs, and to prefer earlier times. Axiom 3.4 is a standard continuity assumption that does not require a choice of topology over random times. The most substantive of our axioms is stochastic stationarity. In §3.4 we discuss this axiom in depth and motivate it using the notions of time invariance and dynamic consistency (Halevy, 2015).[6]

3.2 Representation

Our next result characterizes monotone stationary time preferences:

Theorem 3. *A preference relation $\succeq$ over time lotteries is an MSTP if and only if there exist a monotone additive statistic Φ, a constant $r > 0$, and a continuous and increasing function $u \colon \mathbb{R}_{++} \to \mathbb{R}_{++}$ such that $\succeq$ is represented by*

$$V(x, T) = u(x) \cdot \mathrm{e}^{-r\Phi(T)}. \tag{4}$$

As in Fishburn and Rubinstein (1982), the parameter r can be normalized to be any arbitrary positive constant by applying a monotone transformation to the representation V. We will often set r appropriately to simplify the form of the representation. In contrast, the monotone additive statistic Φ is uniquely determined by the preference.

Over the domain of deterministic time lotteries (i.e. dated rewards), V coincides with an exponentially discounted utility representation. For general time lotteries, $\Phi(T)$ is the certainty equivalent of T, i.e. the unique deterministic time that satisfies $(x, T) \sim (x, \Phi(T))$. The monotonicity and continuity axioms ensure that such a certainty equivalent exists, and it is an implication of stochastic stationarity that $\Phi(T)$ is independent of the reward x. As we further show in the proof of Theorem 3, the monotonicity and stochastic stationarity axioms formally translate into the certainty equivalent Φ being a monotone additive statistic.

Proposition 7 in the appendix shows that the representation in Theorem 1 extends to the domain of non-negative bounded random variables. Thus every MSTP can be represented in the following form:

$$V(x, T) = u(x) \cdot \mathrm{e}^{-r \int K_a(T)\, \mathrm{d}\mu(a)}. \tag{5}$$

[6] It is worthwhile to note that we implicitly assume agents to be indifferent with respect to the timing of resolution of uncertainty. We think of the choice as being made at time 0, and we do not distinguish between situations where the realization of the random time T is revealed immediately, gradually until time T, or only at time T. Modeling preferences over the timing of resolution of uncertainty would require enlarging the choice domain beyond time lotteries.

We recover expected discounted utility when μ is a point mass concentrated on a point $-a < 0$, in which case Φ takes the form

$$\Phi(T) = K_{-a}(T) = \frac{1}{-a} \log \mathbb{E}\left[e^{-aT}\right].$$

This certainty equivalent, with the normalization $r = a$, yields the familiar representation $V(x, T) = u(x)\mathbb{E}[e^{-aT}]$. For a general measure μ, the statistic $\Phi(T) = \int K_a(T)\, d\mu(a)$ aggregates different discount rates by mixing over their corresponding certainty equivalents.

The key feature of the representation (5) is that the average is not over discount factors, but instead over certainty equivalents induced by the discount factors. The resulting representation is behaviorally distinct from expected discounted utility whenever μ is not a point mass. Indeed, as we formally prove in §F in the online appendix, this representation satisfies the independence axiom if and only if μ is a point mass.

3.3 Implications for Risk Attitudes toward Time

Theorem 3 demonstrates that there are many ways to extend discounted utility to the domain of time lotteries, while maintaining stochastic stationarity. As is well known, standard expected discounted utility preferences are *risk-seeking* over time, in the sense that a decision maker prefers receiving a reward at a random time T rather than at the deterministic expected time $t = \mathbb{E}[T]$. But other monotone additive statistics lead to stationary time preferences that are not risk-seeking. As an example, for every $a > 0$ the statistic

$$\Phi(T) = K_a(T) = \frac{1}{a} \log \mathbb{E}\left[e^{aT}\right] \tag{6}$$

leads, with the normalization $r = a$, to the representation

$$V(x, T) = \frac{u(x)}{\mathbb{E}\left[e^{aT}\right]}, \tag{7}$$

which is in fact *risk-averse* over time. Under this preference, the decision maker applies a *negative* discount rate $-a$ within the monotone additive statistic Φ, and yet is impatient. These two aspects are compatible because in the representation $u(x)e^{-r\Phi(T)}$ the statistic Φ controls the risk attitude, while the decision maker still prefers receiving prizes earlier rather than later, since Φ appears with a negative coefficient.

Another key distinctive property of monotone stationary time preferences is their flexibility in allowing for risk attitudes that are not uniform across time lotteries. To illustrate this point, consider two decision problems with a fixed common reward $x = \$1000$, where in the first problem the choice is between

(I) receiving the reward after 1 day for sure, versus

(II) receiving the reward immediately with 99% probability and after 100 days with 1% probability.

In the second decision problem the choice is between

(I') receiving the reward after 99 days for sure, versus

(II') receiving the reward immediately with 1% probability and after 100 days with 99% probability.

In both problems, the times at which the safe options I and I' deliver the prize are equal to the expected delay of the lotteries II and II', and thus a decision maker who is globally risk-averse or risk-seeking must either choose the safe options or the risky options in both problems. Nevertheless, it does not seem unreasonable for a person to choose I over II in order to avoid the risk of a long delay, but also choose II' to I', since the time lottery offers at least a chance of avoiding an otherwise very long delay.[7]

Preferences based on monotone additive statistics are not necessarily globally risk-averse or risk-seeking, and can accommodate the aforementioned behavior. For example, the statistic

$$\Phi(T) = \frac{1}{2}K_1(T) + \frac{1}{2}K_{-1}(T) = \frac{1}{2}\log \mathbb{E}\left[\mathrm{e}^{T}\right] - \frac{1}{2}\log \mathbb{E}\left[\mathrm{e}^{-T}\right]$$

leads the decision maker to choose the safe option I in the first problem and the risky option II' in the second.

Empirically, both risk-averse and risk-seeking behavior over time lotteries are observed. For example, the experiment by Ebert (2021) finds that there are risk-seeking and risk-averse subjects: "Overall, therefore, and in contrast to the evidence on wealth risk preferences, there is substantial heterogeneity in preferences toward delay risk." Moreover, DeJarnette, Dillenberger, Gottlieb, and Ortoleva (2020) find that even the same subject often exhibits both risk aversion and risk seeking depending on the choice at hand.

In §4.2 below we provide a detailed analysis of the risk attitudes of preferences represented by monotone additive statistics, including a characterization of those statistics that give rise to mixed risk attitudes, as in the above example.

3.4 Stationarity, Time Invariance and Dynamic Consistency

In the absence of risk, it was shown by Halevy (2015) that stationarity can be understood as the implication of two more basic principles: that preferences are not affected by calendar time, and that the decision maker is dynamically consistent. As we next explain, Axiom 3.3 is related to a particular notion of dynamic consistency for time lotteries.

[7] We are grateful to Weijie Zhong for suggesting this example to us.

We consider an enlarged framework where the decision maker is endowed with a profile $(\succeq_t)$ of preferences over time lotteries, with $\succeq_t$ representing the preference the decision maker expresses at time t. Formally, $\succeq_t$ is a preference over $\mathbb{R}_{++} \times L_+^\infty$, where in the context of $\succeq_t$ the pair (x, T) represents a payoff of x received at time $t + T$. Adapting the definitions from Halevy (2015) to our setting, we define *time invariance* and *dynamic consistency* below:[8]

Definition. *The collection of preferences* $(\succeq_t)$ *satisfies* time invariance *if all the preferences* $\succeq_t$ *are identical.*

Intuitively, if the agent chooses (x, T) over (y, S) at some time t then she makes the same choice at all other times.

Definition. *The collection* $(\succeq_t)$ *satisfies* deterministic dynamic consistency *if, for every pair of time lotteries* (x, T) *and* (y, S)*, and every* $d, t \in \mathbb{R}_+$ *it holds that*

$$(x, T) \succeq_{t+d} (y, S) \text{ implies } (x, T + d) \succeq_t (y, S + d).$$

That is, the decision maker does not reverse her choice between time t and time $t + d$. Time invariance together with deterministic dynamic consistency imply stationarity with respect to deterministic delays, namely $(x, T) \succeq_t (y, S)$ implies $(x, T + d) \succeq_t (y, S + d)$.

Our next definition proposes a generalization of dynamic consistency to a choice between (x, T) and (y, S) made after a random delay D. What we call *weak stochastic dynamic consistency* requires that if, at the random time $t + D$, the decision maker always prefers (x, T) over (y, S), then she would not revert her choice if asked to make the decision at time t for her future self. In general, the realization of the delay D could affect the distributions of S and T faced by the decision maker. Weak stochastic dynamic consistency considers only the case where the decision maker always faces the same choice independent of the delay, which mathematically corresponds to D being independent of both S and T.

Definition. *The collection* $(\succeq_t)$ *satisfies* weak stochastic dynamic consistency *if, for every pair of time lotteries* (x, T) *and* (y, S)*, every* $t \in \mathbb{R}_+$*, and every* $D \in L_+^\infty$ *independent of* S, T *it holds that*

$$(x, T) \succeq_{t+d} (y, S) \text{ for almost every realization } d \text{ of } D \implies (x, T + D) \succeq_t (y, S + D).$$

As we record in the next claim, our stochastic stationarity axiom is immediately implied by time invariance and weak stochastic dynamic consistency.

[8] These definitions are slightly different from his, and in particular his (deterministic) dynamic consistency axiom is slightly stronger, requiring the implication to hold in both directions.

Claim 1. *Suppose the collection* $(\succeq_t)$ *satisfies time invariance, so that* $\succeq_t = \succeq$ *for every* t*, and also satisfies weak stochastic dynamic consistency. Then the preference* $\succeq$ *satisfies stochastic stationarity.*

Indeed, by time invariance $(x,T) \succeq_t (y,S)$ implies $(x,T) \succeq_{t+d} (y,S)$ for every realization d of D. Thus by weak stochastic dynamic consistency, $(x,T) \succeq_t (y,S)$ implies $(x,T+D) \succeq_t (y,S+D)$ whenever D is independent of S,T. Conversely, if $(x,T) \succeq_{t+d} (y,S)$ for *any* realization d of D, then $(x,T) \succeq_t (y,S)$ by time invariance, and $(x,T+D) \succeq_t (y,S+D)$ would follow from stochastic stationarity. So stochastic stationarity also implies weak stochastic dynamic consistency under the assumption of time invariance.

Weak stochastic dynamic consistency considers the case where D is independent of S and T, which means that at the delayed time $t+D$ the agent always chooses between the same two time lotteries. A stronger dynamic consistency axiom would impose the same condition, but for an arbitrary delay D that need not be independent of S and T. To make this dependency more explicit we write S_d, T_d for random variables that have the conditional distributions of S,T when conditioning on $D = d$.

Definition. *The collection* $(\succeq_t)$ *satisfies* strong stochastic dynamic consistency *if, for every pair of time lotteries* (x,T) *and* (y,S)*, every* $t \in \mathbb{R}_+$*, and every* $D \in L^\infty_+$ *it holds that*

$$(x,T_d) \succeq_{t+d} (y,S_d) \text{ for almost every realization } d \text{ of } D \implies (x,T+D) \succeq_t (y,S+D).$$

Intuitively, strong stochastic dynamic consistency requires consistency at different times across different decision problems, while weak stochastic dynamic consistency only requires it over the same decision problem. For instance, imagine a traveler who must choose between a train and a flight, which involve travel times S and T respectively, and who does not know the specific day of the month D when they will need to travel. Dynamic consistency compares a traveler who must buy their ticket at the start of the month to one who can make the decision on the day of travel. Weak stochastic dynamic consistency applies when the distributions of travel times S and T are not dependent on the day of the month. Strong stochastic dynamic consistency applies further to cases where travel times S_d and T_d do depend on the day d.

The following result shows that under time invariance, strong stochastic dynamic consistency constrains the preference over time to be represented by K_a for some $a \in \overline{\mathbb{R}}$, rather than a general monotone additive statistic Φ as in Theorem 3.

Proposition 1. *Suppose* $\succeq$ *is an MSTP. Then the collection* $(\succeq_t)$ *with* $\succeq_t = \succeq$ *for every* t *satisfies strong stochastic dynamic consistency if and only if* $\succeq$ *can be represented by*

$$V(x,T) = u(x) \cdot \mathrm{e}^{-rK_a(T)}$$

for some $a \in \overline{\mathbb{R}}$*,* $r > 0$*, and* $u \colon \mathbb{R}_{++} \to \mathbb{R}_{++}$*.*

In words, the preference over time is either risk-neutral, expected discounted utility, the discounted maximum or minimum, or the negatively discounted preference described in (7). In particular, strong stochastic dynamic consistency would rule out the kind of mixed risk attitudes described in §3.3.

Proposition 1 follows from the fact that strong stochastic dynamic consistency, in combination with monotonicity and continuity, implies the classic independence axiom as we discuss in §F of the online appendix. Weak stochastic dynamic consistency does not imply the independence axiom and thus allows for a richer set of time preferences.

3.5 Aggregation of Preferences over Time Lotteries

In this section we apply monotone stationary time preferences to collective decision problems. A company making a choice among projects with different expected completion dates, a public agency choosing which research projects to fund, or a family deciding which highway to take, are all examples of social decisions where the alternatives at hand can be seen as time lotteries. In such situations, even if individuals share the same views about the desirability of the possible outcomes, there still exists a need to compromise between different degrees of patience and risk tolerance.

We model this type of problem by studying a group of individuals where each agent, denoted by i, is equipped with a preference relation $\succeq_i$ over time lotteries. These preferences may display different degrees of patience. Following the approach in social choice, we ask how individual preferences can be aggregated into a social preference relation $\succeq$ that is aligned to the individual preferences by the Pareto principle. In this context, the Pareto principle requires that if all individuals agree that one time lottery is better than another, so should the social preference:

Axiom 3.5 (Pareto)**.** *If* $(x, T) \succeq_i (y, S)$ *for every* i*, then* $(x, T) \succeq (y, S)$*.*

We first consider the case where each individual preference admits a standard expected discounted utility representation $u_i(x)\mathbb{E}[\mathrm{e}^{-a_i T}]$, where $u_i \colon \mathbb{R}_{++} \to \mathbb{R}_{++}$ is agent i's utility function and $a_i > 0$ is her discount rate.[9] The next result shows that if one insists that the social preference also conforms to expected discounted utility, then dictatorship is the only admissible aggregation procedure satisfying the Pareto axiom whenever the individual discount rates are distinct.[10]

[9] It is important to note that the parameter a_i here is uniquely pinned down by agent i's preference—when restricting to a fixed reward x, the preference is expected utility over random times T, so the discounting functions $\mathrm{e}^{-a_i t}$ are unique up to a linear transformation.

[10] When some agents have the same discount rate, Paretian aggregation need not be dictatorial. For example, if $a_1 = a_2$, then $u(x)\mathbb{E}[\mathrm{e}^{-aT}]$ with $u = \frac{u_1+u_2}{2}$ and $a = a_1 = a_2$ satisfies the Pareto axiom.

Proposition 2. *Let* $(\succeq_1, \ldots, \succeq_n, \succeq)$ *be expected discounted utility preferences over time lotteries, where each* $\succeq_i$ *is represented by* $u_i(x)\mathbb{E}[\mathrm{e}^{-a_i T}]$ *and* $\succeq$ *is represented by* $u(x)\mathbb{E}[\mathrm{e}^{-aT}]$*. Suppose* $a_1, \ldots, a_n$ *are distinct positive numbers. Then the Pareto axiom is satisfied if and only if* $\succeq = \succeq_i$ *for some agent* i*.*

As the proof shows, this impossibility result is a consequence of Harsanyi's utilitarian theorem (Harsanyi, 1955). Similar impossibility results have been obtained in the setting of preferences over consumption streams (Gollier and Zeckhauser, 2005; Zuber, 2011; Jackson and Yariv, 2014, 2015; Feng and Ke, 2018; Chambers and Echenique, 2018).

The next result offers a solution to this impossibility result. It shows that Paretian aggregation and stochastic stationarity are compatible, and do not necessarily result in a dictatorship, if we allow preferences to belong to the larger class of MSTPs.

Proposition 3. *Let* $(\succeq_1, \ldots, \succeq_n, \succeq)$ *be MSTPs, where each* $\succeq_i$ *is represented by* $u_i(x)\mathrm{e}^{-r_i \Phi_i(T)}$ *and* $\succeq$ *is represented by* $u(x)\mathrm{e}^{-r\Phi(T)}$ *for some monotone additive statistics* (Φ_i) *and* Φ*. If there exists* $\lambda_1, \ldots, \lambda_n \in \mathbb{R}_+$ *such that*

$$r = \sum_{i=1}^{n} \lambda_i r_i, \qquad r\Phi = \sum_{i=1}^{n} \lambda_i r_i \Phi_i, \tag{8}$$

and $u = \Pi_{i=1}^{n} u_i^{\lambda_i}$*, then the Pareto axiom is satisfied.*

Note that as long as the individual certainty equivalents $\Phi_1, \ldots, \Phi_n$ are not all identical, then for generic values of $\lambda_1, \ldots, \lambda_n$, the social certainty equivalent Φ constructed from (8) is distinct from each of the individual certainty equivalents. Thus the resulting social preference $\succeq$ is generically not a dictatorship.

The key insight of Proposition 3 is that a linear aggregation of certainty equivalents preserves both stochastic stationarity and the Pareto axiom; as we show below, this is in fact the only way to preserve these properties. In the special case where individuals have expected discounted utility preferences, the proposition implies that we can aggregate preferences without violating stochastic stationarity by allowing the social preference to be an MSTP. This approach complements alternative solutions that have been proposed in the literature to resolve the tension between Paretian aggregation and stationarity.[11]

The next result gives a characterization of all social preferences that admit an MSTP representation and respect the Pareto axiom, in the special case where all agents share the same utility function and it satisfies a mild richness assumption. The assumption that all

[11] For example, Feng and Ke (2018) define a different notion of Pareto efficiency that takes into account the preferences of individuals across generations. They show that a standard expected discounted social preference can satisfy this weaker Pareto axiom so long as it is more patient than all the individuals. Chambers and Echenique (2018) study a number of representations that weaken stationarity and generalize expected discounted utility.

agents (and the social planner) share the same utility function is common in the literature on aggregating discount factors, following Weitzman (2001) and Chambers and Echenique (2018).

Proposition 4. *Let* $(\succeq_1, \ldots, \succeq_n, \succeq)$ *be MSTPs, where each* $\succeq_i$ *is represented by* $u(x)\mathrm{e}^{-r_i\Phi_i(T)}$ *and* $\succeq$ *is represented by* $u(x)\mathrm{e}^{-r\Phi(T)}$ *for some monotone additive statistics* (Φ_i) *and* Φ*. Suppose that the common utility function satisfies either* $\lim_{x\to 0} u(x) = 0$ *or* $\lim_{x\to\infty} u(x) = \infty$*.*

Then, the Pareto axiom is satisfied if and only if (8) *holds for some* $\lambda_1, \ldots, \lambda_n \in \mathbb{R}_+$ *that sum to* 1*.*

It follows from our proof that even if agents had different utility functions, the Pareto axiom would still require the social certainty equivalent Φ to be a convex combination of the individual $(\Phi_i)_{i=1}^n$. However, the implications of the Pareto axiom on the social utility function seem difficult to characterize in general.[12] We leave this question for future work.

4 Preferences Over Gambles

In the theory of risk, CARA utility functions form a restrictive but useful class of expected utility preferences. Their usefulness stems from the analytical tractability of the exponential form, as well as from their invariance properties.

CARA utility functions are invariant to changes in wealth, so that a prospect X is preferred to Y if and only if $X + w$ is preferred to $Y + w$ for all wealth levels w. They are more generally invariant to the addition of background risk: if X is preferred to Y then $X + W$ is preferred to $Y + W$ for every independent random variable W.

This property makes CARA utility functions a good approximation whenever stakes are small. In addition, they are used in empirical settings in which wealth is unknown. For example, when estimating risk preferences from insurance choices, the CARA family "*has the advantage that it implies a household's prior wealth w, which frequently is unobserved, is irrelevant to the household's decisions.*" (Barseghyan, Molinari, O'Donoghue, and Teitelbaum, 2018). The stronger property of invariance to background risk is also important, since households' additional background risk—arising from, say, investments in the stock market or health conditions—may be unobservable.

[12] To illustrate the difficulty, consider two individual EDU preferences represented by $u_1(x)\mathbb{E}\left[\mathrm{e}^{-T}\right]$ and $u_2(x)\mathbb{E}\left[\mathrm{e}^{-T}\right]$, as well as a social preference represented by $u(x)\mathbb{E}\left[\mathrm{e}^{-T}\right]$, all with the same discount rate. In this case, one can show that the Pareto axiom reduces to the inequality condition $\frac{u(x)}{u(y)} \geq \min\{\frac{u_1(x)}{u_1(y)}, \frac{u_2(x)}{u_2(y)}\}$ for every pair of rewards x, y. Now suppose $u_2(x) = u_1(x)^2$ for every x, then the previous condition simplifies to $\frac{u_1(x)}{u_1(y)} \leq \frac{u(x)}{u(y)} \leq \left(\frac{u_1(x)}{u_1(y)}\right)^2$ for every $x > y$. A wide range of u functions satisfies this condition, including u_1^α for any power $\alpha \in [1, 2]$ and all convex combinations of such powers. This multiplicity of possible social utility functions makes it challenging to generalize Proposition 4.

The invariance properties of CARA utility functions are conceptually distinct from the assumption that preferences obey the axioms of expected utility. In this section, we apply monotone additive statistics to study the general class of preferences that are monotone with respect to stochastic dominance and are invariant to background risk.

4.1 Background-risk Invariant Preferences

We consider a complete and transitive preference relation $\succeq$ over L^∞, interpreted here as the space of monetary gambles. We assume that for every gamble X there exists a unique certainty equivalent $\Phi(X)$ such that $\Phi(X) \sim X$. If the preference $\succeq$ is monotone with respect to first-order stochastic dominance then so is Φ. We say that $\succeq$ is *invariant to background risk* when it has the property that $X \succeq Y$ if and only if $X + Z \succeq Y + Z$ for Z independent of X and Y.

As we now explain, a preference $\succeq$ is monotone and invariant to background risk if and only if its certainty equivalent is a monotone additive statistic. Indeed, invariance implies that $X + Y \sim \Phi(X) + Y$ for any two independent random variables X and Y. Likewise, $Y + \Phi(X) \sim \Phi(Y) + \Phi(X)$. Combining the two indifferences yields $X + Y \sim \Phi(X) + \Phi(Y)$. So, the certainty equivalent of $X + Y$ is given by the sum $\Phi(X) + \Phi(Y)$, and thus Φ is an additive. The converse is immediate to verify.

By Theorem 1, the certainty equivalent Φ of such a preference is a weighted average $\Phi(X) = \int K_a(X)\,\mathrm{d}\mu(a)$ of the certainty equivalents of multiple CARA expected utility agents, where μ is a probability measure over the coefficient of absolute risk aversion.

4.2 Risk Aversion

In this section we characterize risk-averse and risk-seeking behavior for preferences that are represented by monotone additive statistics. A preference relation $\succeq$ over gambles is risk-averse if its certainty equivalent Φ satisfies $\Phi(X) \leq \mathbb{E}[X]$ for every gamble X, and risk-seeking if the opposite inequality holds. Risk aversion translates into a property of the support of the corresponding mixing measure μ:

Proposition 5. *A monotone additive statistic satisfies $\Phi(X) \leq \mathbb{E}[X]$ for every $X \in L^\infty$ if and only if $\Phi(X) = \int_{\overline{\mathbb{R}}} K_a(X)\,\mathrm{d}\mu(a)$ for a Borel probability measure μ supported on $[-\infty, 0]$. Likewise, $\Phi(X) \geq \mathbb{E}[X]$ for every X if and only if μ is supported on $[0, \infty]$.*

In words, a preference that is invariant to background risk is additionally risk-averse (resp. risk-seeking) if and only if it ranks gambles by aggregating the certainty equivalents of risk-averse (resp. risk-seeking) CARA utility functions.[13]

[13] A corollary of Proposition 5 is that an additive statistic Φ is risk averse if and only if it is monotone with respect to *second-order* stochastic dominance. This is perhaps surprising, since the two properties are

4.3 Mixed Risk Aversion

As pointed out in the classical work of Friedman and Savage (1948), it is not uncommon to observe behavior that is neither risk-averse nor risk-seeking, such as that of a person who buys both lottery tickets and insurance. For concreteness, in analogy with our discussion in the time domain, consider a decision maker faced with the following two choices.

In the first, the choice is between (I) facing a risk of losing \$100 with probability 1%, or (II) paying \$1 and being fully insured against that risk. In the second decision problem the choice is between (I') paying \$1 dollar for a lottery ticket that yields \$100 with probability 1%, or (II') not participating in the lottery.

Preferences represented by monotone additive statistics can model a decision maker who chooses (II) over (I) but (I') over (II'), while at the same time remaining invariant to background risk. This is the case, for example, for a preference whose certainty equivalent $\Phi(X)$ takes the form $\Phi(X) = \frac{1}{2}K_{-a}(X) + \frac{1}{2}K_a(X)$, with a mixing measure that puts equal weights on two coefficients of risk aversion a and $-a$.

4.4 Comparative Risk Attitudes

We now proceed to compare the risk attitudes expressed by different monotone additive statistics. For two preference relations $\succeq_1$ and $\succeq_2$ over gambles, with corresponding certainty equivalents Φ_1 and Φ_2, the preference $\succeq_1$ is *more risk-averse* than $\succeq_2$ if $\Phi_1(X) \leq \Phi_2(X)$ for every gamble $X \in L^\infty$. That is, if the first decision maker assigns to every gamble a lower certainty equivalent. The next proposition characterizes comparative risk aversion for preferences represented by monotone additive statistics:

Proposition 6. *Let $\succeq_1$ and $\succeq_2$ be represented by monotone additive statistics with mixing measures μ_1 and μ_2, respectively. Then $\succeq_1$ is more risk-averse than $\succeq_2$ if and only if*

(i) For every $b > 0$, $\int_{[b,\infty]} \frac{a-b}{a}\,\mathrm{d}\mu_1(a) \leq \int_{[b,\infty]} \frac{a-b}{a}\,\mathrm{d}\mu_2(a)$.

(ii) For every $b < 0$, $\int_{[-\infty,b]} \frac{a-b}{a}\,\mathrm{d}\mu_1(a) \geq \int_{[-\infty,b]} \frac{a-b}{a}\,\mathrm{d}\mu_2(a)$.

Since $K_a(X)$ increases in the parameter a, a sufficient condition for $\succeq_1$ being more risk-averse than $\succeq_2$ is that μ_2 first-order stochastically dominates μ_1. First-order stochastic dominance is, however, only a sufficient condition. The reason is that the cone generated by the functions of the form $K_{(\cdot)}(X)$, as we vary X, does not contain all increasing functions, and hence defines a strictly finer stochastic order over the mixing measures.[14]

in general not equivalent for a preference over gambles.

[14]For a concrete example that the order characterized by Proposition 6 is strictly finer than first-order stochastic dominance, consider μ_1 to be a point mass at $a = 2$ and μ_2 to have $1/4$ mass at $a = 1$ and $3/4$ mass at $a = 3$. Clearly, neither one first-order dominates the other. Condition (ii) in Proposition 6 is

Proposition 6 characterizes this stochastic order by showing that the convex cone generated by the set of normalized cumulant generating functions is equal to the convex cone generated by a simple one-parameter family of test functions, of the form $g(a) = \frac{a-b}{a} 1_{a \geq b}$ or $g(a) = -\frac{a-b}{a} 1_{a \leq b}$.

4.5 Betweenness

A disadvantage of the class of preferences represented by monotone additive statistics is that it is large, with the entire measure μ as an infinite-dimensional parameter of the preference. In this section we identify a small subset of such preferences that is indexed by only two parameters, and yet retains enough flexibility to accommodate interesting risk attitudes such as mixed risk aversion.

To this end we study preferences that satisfy the *betweenness* axiom. This well-known property, first studied by Dekel (1986) and Chew (1989), requires that the decision maker's preference over probability distributions displays indifference curves that are straight lines. In comparison, the standard independence axiom (which we study in §F of the online appendix) would additionally require the indifference curves to be parallel to each other.

Given two random variables X and Y, we denote by $X_\lambda Y$ any random variable whose distribution is a convex combination that assigns weight λ to the distribution of X and weight $1-\lambda$ to the distribution of Y.

Axiom 4.1 (Betweenness)**.** *For all X, Y and all $\lambda \in (0,1)$, $X \sim Y$ if and only if $X_\lambda Y \sim Y$.*

The betweenness axiom characterizes the following class of preferences:

Theorem 4. *Suppose a preference $\succeq$ on L^∞ is represented by a monotone additive statistic $\Phi(X) = \int_{\mathbb{R}} K_a(X)\,\mathrm{d}\mu(a)$. Then $\succeq$ satisfies the betweenness axiom if and only if*

$$\Phi(X) = \beta K_{-a\beta}(X) + (1-\beta) K_{a(1-\beta)}(X)$$

for some $\beta \in [0,1]$ and $a \in [0,\infty)$.

This family of preferences is much smaller, as it is parameterized by only two numbers. It retains the properties of monotonicity, invariance to background risk, as well as the tractability of the CARA representation. Yet it is versatile enough to describe the kind of mixed risk attitude that leads to buying both insurance and lottery tickets.

The risk-attitude parameter β weights the levels of risk aversion/seeking, with $\beta = 1$ corresponding to pure CARA risk aversion and $\beta = 0$ corresponding to pure CARA risk seeking. For internal β, the preference exhibits mixed risk aversion as guaranteed by the

trivially satisfied, whereas condition (i) reduces to $\frac{1}{2}(2-b)^+ \leq \frac{1}{4}(1-b)^+ + \frac{1}{4}(3-b)^+$, which holds because the function $(a-b)^+ = \max\{a-b, 0\}$ is convex in a.

previous Proposition 5. Moreover, a simple calculation shows that for any $\beta \in (0,1)$, such a preference would buy both insurance and lottery tickets of the kind described in §4.3 whenever those gambles entail a small probability of a large loss or gain.[15]

The parameter a is a scale parameter. It can be understood as the scale at which the preference deviates from risk neutrality. For gambles whose sizes are much smaller than $1/a$, the preference is very close to being risk-neutral. While for gambles that vary by much more than $1/a$, behavior will be far from risk-neutral.

4.6 Combined Choices over Gambles

In large organizations, risky prospects are not always chosen through a deliberate, centralized process. Rather, they are combinations of independent choices, often carried out with limited coordination among the different actors.

Consider, for example, a bank that employs two workers. The first is a trader who must choose between two contracts, the Lean Hog futures X and X'. The second is an administrator who must choose between two insurance policies Y or Y' for the bank's building. Assuming the first worker chooses X and the second Y, the resulting revenue for the bank is given by the random variable $X+Y$. When the agents face choice problems that belong to independent domains, so that X and X' are stochastically independent from Y and Y', it is natural to ask to what extent coordination is necessary for the organization.

In this section we make this question precise by asking under what conditions the agents' combined choices respect first-order stochastic dominance. Our result shows this is true if and only if individual preferences are identical and represented by a monotone additive statistic. Thus, this is the only class of preferences with the property that choices over independent domains can be decentralized without obvious harm to the organization.

We study the following model. We are given two preference relations $\succeq_1$ and $\succeq_2$ over L^∞, the set of bounded gambles, that are complete and transitive (our result immediately generalizes to three or more agents). As in the example above, we think of each preference relation as describing the choices of a different agent, so that $X \succeq_i X'$ if agent i chooses X over X'. These preferences can be interpreted as being endogenous or as the result of exogenous incentives; for example, the bank trader's preferences could be driven by her contract with the employer.

Our main axiom requires that whenever the two agents face independent decision problems, their choices, when combined, do not violate stochastic dominance:

[15] It can be shown that if $\beta \neq 0.5$, then lottery tickets and insurance as described in §4.3 are preferred if and only if the probability of gain/loss (0.01 in the example) is smaller than $\min(\beta, 1-\beta)$, and the corresponding gain/loss amount (100 in the example) is sufficiently large. If $\beta = 0.5$, then the same holds for any probability of gain/loss < 0.5, and for any gain/loss amount.

Axiom 4.2 (Consistency of Combined Choices)**.** *Suppose X, X' are independent of Y, Y'. If $X \succ_1 X'$ and $Y \succ_2 Y'$, then $X' + Y'$ does not strictly dominate $X + Y$ in first-order stochastic dominance.*

If we interpret $\succeq_1$ and $\succeq_2$ as decision-making rules that are determined by the organization, then Axiom 4.2 requires such rules to never result in an outcome that is stochastically dominated. That collective choices should not violate stochastic dominance is clearly a desirable requirement for a rational organization. A similar axiom was first introduced by Rabin and Weizsäcker (2009) in the context of a model of narrow framing.

In addition to this axiom, we assume individual preference relations $\succeq_i$ satisfy basic continuity and monotonicity assumptions:

Axiom 4.3 (Continuity)**.** *If $X \succ_i Y$ then there exists $\varepsilon > 0$ such that $X \succ_i Y + \varepsilon$ and $X - \varepsilon \succ_i Y$.*

Axiom 4.4 (Responsiveness)**.** *$X + \varepsilon \succ_i X$ for every $\varepsilon > 0$.*

We next show that under these axioms, the two preference relations must be represented by monotone additive statistics. Moreover, the statistic must be the same for both agents.

Theorem 5. *Two preference $\succeq_1, \succeq_2$ on L^∞ satisfy Axioms 4.2, 4.3, and 4.4 if and only if there exists a monotone additive statistic that represents both $\succeq_1$ and $\succeq_2$.*

Thus, when individual choices are not coordinated, their combination will, in general, lead to violations of stochastic dominance, even when agents' choices concern independent decision problems. The theorem singles out preferences represented by monotone additive statistics as the only class of preferences that are robust to this lack of coordination.

Theorem 5 admits an alternative interpretation, closely related to the work of Rabin and Weizsäcker (2009) on narrow framing. In their paper, a decision maker faces multiple decisions and engages in "narrow bracketing" by choosing separately, in each problem, according to a fixed preference relation $\succeq$ over gambles. This is a special case of our model where $\succeq \, = \, \succeq_1 \, = \, \succeq_2$. They show that the decision maker's combined choices result in dominated outcomes whenever $\succeq$ is not invariant to changes in wealth (i.e. for some X, Y and $c \in \mathbb{R}$, $X \succ Y$ and $Y + c \succ X + c$), but leave open the question of characterizing the class of preferences, beyond expected utility, that satisfy Axiom 4.2. Theorem 5 provides a complete characterization of those preferences over gambles for which narrow framing does not lead to dominated choices.

5 Overview of the Proof of Theorem 1

Our approach to the proof of Theorem 1 is via a stochastic order known as the *catalytic stochastic order* (see Fritz, 2017, and references therein). Given $X, Y \in L^\infty$, we say that X

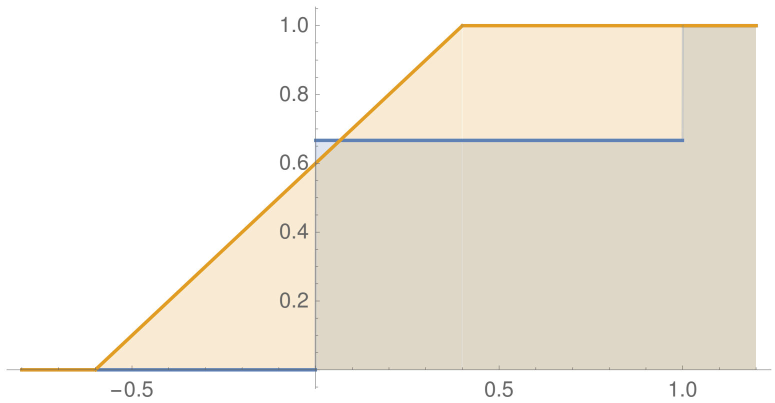

Figure 1: The c.d.f.s of X (blue) and Y (orange).

dominates Y in the catalytic stochastic order on L^∞ if there exists a $Z \in L^\infty$, independent of X and Y, such that $X + Z$ dominates $Y + Z$ in first-order stochastic dominance.

The applicability of this order to our problem is immediate. If X dominates Y in the catalytic stochastic order then

$$\Phi(X + Z) \geq \Phi(Y + Z)$$

for some Z, independent of X and Y. If Φ is also additive, then $\Phi(X + Z) = \Phi(X) + \Phi(Z)$ and $\Phi(Y + Z) = \Phi(Y) + \Phi(Z)$, and so we have that $\Phi(X) \geq \Phi(Y)$. Thus, any monotone additive Φ is monotone with respect to this order.

Clearly, if $X \geq_1 Y$ then X also dominates Y in the catalytic stochastic order, as one can take $Z = 0$. A priori, one may conjecture that this is also a necessary condition. But as Figure 1 shows, it is easy to give examples of two random variables X and Y that are not ranked with respect to first-order stochastic dominance, but are ranked with respect to the catalytic stochastic order.[16] The random variable X equals 1 with probability 1/3 and 0 with probability 2/3, while Y is uniformly distributed on $[-\frac{3}{5}, \frac{2}{5}]$. As the figure shows, their c.d.f.s are not ranked, and hence they are not ranked in terms of first-order stochastic dominance.[17] However, if we let Z assign probability half to $\pm\frac{1}{5}$, then $X + Z >_1 Y + Z$. Intuitively, since the c.d.f. of $X + Z$ is the average of the two translations (by $\pm\frac{1}{5}$) of the c.d.f. of X, and since the same holds for the c.d.f. of Y, the result of adding Z is the disappearance of the small "kink" in which the ranking of the c.d.f.s is reversed. This is depicted in Figure 2.

[16]We are indebted to the late Kim Border for helping us construct this example.

[17]Pomatto, Strack, and Tamuz (2020) give examples of random variables X and Y that are not ranked in stochastic dominance, but are ranked after adding an *unbounded* independent Z. In fact, they show that this is possible whenever $\mathbb{E}[X] > \mathbb{E}[Y]$. As we explain below, this result no longer holds when Z is required to be bounded.

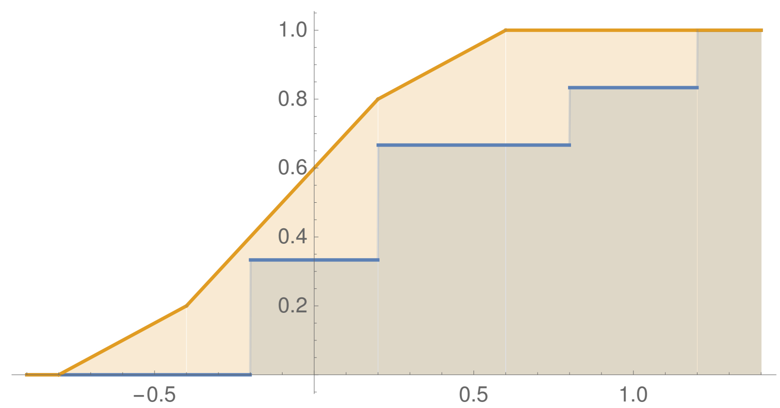

Figure 2: The c.d.f.s of $X + Z$ (blue) and $Y + Z$ (orange).

Every monotone additive statistic Φ provides an obstruction to dominance in the catalytic stochastic order. That is, if $\Phi(X) < \Phi(Y)$, then it is impossible that $X + Z \geq_1 Y + Z$ for some independent Z, since monotonicity would imply that $\Phi(X+Z) \geq \Phi(Y+Z)$, and additivity would then imply that $\Phi(X) \geq \Phi(Y)$. In particular, the existence of an a for which $K_a(X) < K_a(Y)$ forms an obstruction to the existence of such a Z. The following result shows that these are, in a sense, the only possible obstructions:[18]

Theorem 6. *Let $X, Y \in L^\infty$ satisfy $K_a(X) > K_a(Y)$ for all $a \in \overline{\mathbb{R}}$. Then there exists a c.d.f. H such that any independent $Z \in L^\infty$ with c.d.f. H satisfies $X + Z \geq_1 Y + Z$.*

To prove Theorem 6 we explicitly construct H as a truncated Gaussian c.d.f. with appropriately chosen parameters. The idea behind the proof is as follows. Denote by F and G the c.d.f.s of X and Y, respectively, and suppose that they are supported on $[-N, N]$. Let $h(x) = \frac{1}{\sqrt{2\pi V}}\mathrm{e}^{-\frac{x^2}{2V}}$ be the density of a Gaussian Z. Then the c.d.f.s of $X + Z$ and $Y + Z$ are given by the convolutions $F * h$ and $G * h$, and their difference is equal to

$$
\begin{aligned}
[G * h - F * h](y) &= \int_{-N}^{N} [G(x) - F(x)] \cdot h(y - x)\,\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi V}}\mathrm{e}^{-\frac{y^2}{2V}} \cdot \int_{-N}^{N} \underbrace{[G(x) - F(x)] \cdot \mathrm{e}^{\frac{y}{V}\cdot x}}_{(*)} \cdot \underbrace{\mathrm{e}^{-\frac{x^2}{2V}}}_{(**)}\,\mathrm{d}x
\end{aligned}
$$

If we denote $a = \frac{y}{V}$, then by integration by parts, the integral of just $(*)$ is equal to $\frac{1}{a}\left(\mathbb{E}\left[\mathrm{e}^{aX}\right] - \mathbb{E}\left[\mathrm{e}^{aY}\right]\right)$, which is positive by the assumption that $K_a(X) > K_a(Y)$ and is

[18] In fact, except for the trivial case where X and Y have the same distribution, the strict inequality $K_a(X) > K_a(Y)$ for all $a \in \mathbb{R}$ is necessary for the existence of a Z such that $X + Z \geq_1 Y + Z$. This is because $X + Z \geq_1 Y + Z$ implies the strict inequality $K_a(X + Z) > K_a(Y + Z)$ for finite a whenever $X + Z$ and $Y + Z$ have different distributions. Thus, Theorem 6 below implies that for distributions with different minima and maxima, the condition $K_a(X) > K_a(Y)$ for all $a \in \mathbb{R}$ is both necessary and sufficient for dominance in the catalytic stochastic order.

in fact bounded away from zero by the Extreme Value Theorem. The term $(**)$ can be made arbitrarily close to 1—uniformly on the integral domain $[-N, N]$—by making V large. This implies that $[G * h - F * h](y) > 0$ for all y, and we further show that the inequality still holds if we modify H by truncating its tails, ensuring that it is in L^∞.

Theorem 6 leads to the following lemma, which is a key component of the proof of Theorem 1:

Lemma 1. *Let $\Phi\colon L^\infty \to \mathbb{R}$ be a monotone additive statistic. If $K_a(X) \geq K_a(Y)$ for all $a \in \overline{\mathbb{R}}$ then $\Phi(X) \geq \Phi(Y)$.*

Proof. Suppose $K_a(X) \geq K_a(Y)$ for all $a \in \overline{\mathbb{R}}$. Given $\varepsilon > 0$, let $\hat{X}$, $\hat{Y}$ and Z in L be such that: $\hat{X}$ has the same c.d.f. as $X + \varepsilon$, $\hat{Y}$ has the same c.d.f. as Y, and Z has the c.d.f. obtained by applying Theorem 6 to $\hat{X}$ and $\hat{Y}$. We can indeed apply the theorem, since $K_a(\hat{X}) = K_a(X) + \varepsilon > K_a(Y) = K_a(\hat{Y})$ for all a. Hence, $\hat{X} + Z \geq_1 \hat{Y} + Z$. Thus, by monotonicity of Φ, $\Phi(\hat{X} + Z) \geq \Phi(\hat{Y} + Z)$, and by additivity $\Phi(\hat{X}) \geq \Phi(\hat{Y})$. This means that $\Phi(X) + \varepsilon = \Phi(\hat{X}) \geq \Phi(\hat{Y}) = \Phi(Y)$ for all $\varepsilon > 0$, and hence $\Phi(X) \geq \Phi(Y)$. □

Once we have established Lemma 1, the remainder of the proof uses functional analysis techniques (in particular the Riesz Representation Theorem) to deduce the integral representation in Theorem 1. See §A in the appendix for the complete proof.

An alternative proof of Lemma 1 can be given based on a different stochastic order. Given two random variables X and Y, let $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ be i.i.d. copies of X and Y, respectively. We say that X dominates Y in *large numbers* if

$$X_1 + \cdots + X_n \geq_1 Y_1 + \cdots + Y_n$$

for all n large enough. Using large-deviations techniques, it was shown by Aubrun and Nechita (2008) that if $K_a(X) > K_a(Y)$ for all $a \in \overline{\mathbb{R}}$, then X dominates Y in large numbers. This implies Lemma 1 since, by the additivity of Φ, $\Phi(X) \geq \Phi(Y)$ holds if and only if $n\Phi(X) = \Phi(X_1 + \cdots + X_n) \geq \Phi(Y_1 + \cdots + Y_n) = n\Phi(Y)$.

Compared to this alternative argument, our proof of Lemma 1 based on Theorem 6 is self-contained and more elementary. More importantly, (an analogue of) the catalytic stochastic order established in Theorem 6 is essential for studying monotone additive statistics defined on a domain of unbounded random variables, for which the large numbers order is difficult to characterize as far as we know.[19] This generalization of Theorem 6 is presented in Lemma 9 in the online appendix, as a key step toward the proof of Theorem 2.

[19]One particular challenge is that the large numbers order require a uniform comparison between the tail probabilities of $X_1 + \cdots + X_n$ versus those of $Y_1 + \cdots + Y_n$, for a fixed large n. For a given threshold of the tail, large-deviations theory can be used to show the desired comparison when n is large enough. But making the required n uniform across all thresholds becomes nontrivial when the random variables X and Y are unbounded.

References

C. D. Aliprantis and K. Border. *Infinite dimensional analysis: A hitchhiker's guide*. Springer, 2006.

G. Aubrun and I. Nechita. Catalytic majorization and ℓ_p norms. *Communications in Mathematical Physics*, 278(1):133–144, 2008.

L. Barseghyan, F. Molinari, T. O'Donoghue, and J. C. Teitelbaum. Estimating risk preferences in the field. *Journal of Economic Literature*, 56(2):501–564, 2018.

P. Bickel and E. Lehmann. Descriptive statistics for nonparametric models I. Introduction. *Annals of Statistics*, 3(5):1038–1044, 1975a.

P. Bickel and E. Lehmann. Descriptive statistics for nonparametric models II. Location. *Annals of Statistics*, 3(5):1045–1069, 1975b.

V. I. Bogachev. *Measure theory*, volume 1. Springer Science & Business Media, 2007.

S. Cerreia-Vioglio, D. Dillenberger, and P. Ortoleva. Cautious expected utility and the certainty effect. *Econometrica*, 83(2):693–728, 2015.

S. Cerreia-Vioglio, D. Dillenberger, and P. Ortoleva. An explicit representation for disappointment aversion and other betweenness preferences. *Theoretical Economics*, 15(4): 1509–1546, 2020.

C. P. Chambers and F. Echenique. When does aggregation reduce risk aversion? *Games and Economic Behavior*, 76(2):582–595, 2012.

C. P. Chambers and F. Echenique. On multiple discount rates. *Econometrica*, 86(4): 1325–1346, 2018.

C. P. Chambers and F. Echenique. Spherical preferences. *Journal of Economic Theory*, 189:105086, 2020.

H. W. Chesson and W. K. Viscusi. Commonalities in time and ambiguity aversion for long-term risks. *Theory and Decision*, 54(1):57–71, 2003.

S. H. Chew. Axiomatic utility theories with the betweenness property. *Annals of Operations Research*, 19(1):273–298, 1989.

J. H. Curtiss. A note on the theory of moment generating functions. *The Annals of Mathematical Statistics*, 13(4):430–433, 1942.

B. de Finetti. *Theory of probability*. Wiley, 1970.

P. DeJarnette, D. Dillenberger, D. Gottlieb, and P. Ortoleva. Time lotteries and stochastic impatience. *Econometrica*, 88(2):619–656, 2020.

E. Dekel. An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom. *Journal of Economic Theory*, 40:304–318, 1986.

S. Ebert. Decision making when things are only a matter of time. *Operations Research*, 68 (5):1564–1575, 2020.

S. Ebert. Prudent discounting: Experimental evidence on higher-order time risk preferences. *International Economic Review*, 62(4):1489–1511, 2021.

T. Feng and S. Ke. Social discounting and intergenerational Pareto. *Econometrica*, 86(5): 1537–1567, 2018.

P. C. Fishburn and A. Rubinstein. Time preference. *International Economic Review*, 23: 677–694, 1982.

H. Föllmer and T. Knispel. Entropic risk measures: Coherence vs. convexity, model ambiguity and robust large deviations. *Stochastics and Dynamics*, 11(02n03):333–351, 2011.

H. Föllmer and A. Schied. Convex measures of risk and trading constraints. *Finance and Stochastics*, 6(4):429–447, 2002.

H. Föllmer and A. Schied. *Stochastic finance: An introduction in discrete time.* Walter de Gruyter, 2011.

M. Friedman and L. J. Savage. The utility analysis of choices involving risk. *Journal of Political Economy*, 56(4):279–304, 1948.

T. Fritz. Resource convertibility and ordered commutative monoids. *Mathematical Structures in Computer Science*, 27(6):850–938, 2017.

T. Fritz, X. Mu, and O. Tamuz. Monotone homomorphisms on convolution semigroups, 2020. Working Paper.

C. Gollier and R. Zeckhauser. Aggregation of heterogeneous time preferences. *Journal of Political Economy*, 113(4):878–896, 2005.

M. J. Goovaerts, R. Kaas, R. J. Laeven, and Q. Tang. A comonotonic image of independence for additive risk measures. *Insurance: Mathematics and Economics*, 35(3):581–594, 2004.

F. Gul. A theory of disappointment aversion. *Econometrica*, 59(3):667–686, 1991.

Y. Halevy. Time consistency: Stationarity and time invariance. *Econometrica*, 83(1): 335–352, 2015.

J. C. Harsanyi. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy*, 63(4):309–321, 1955.

M. O. Jackson and L. Yariv. Present bias and collective dynamic choice in the lab. *American Economic Review*, 104(12):4184–4204, 2014.

M. O. Jackson and L. Yariv. Collective dynamic choice: The necessity of time inconsistency. *American Economic Journal: Microeconomics*, 7(4):150–78, 2015.

L. Kantorovich. On the moment problem for a finite interval. In *Dokl. Akad. Nauk SSSR*, volume 14, pages 531–537, 1937.

L. Mattner. What are cumulants? *Documenta Mathematica*, 4:601–622, 1999.

L. Mattner. Cumulants are universal homomorphisms into Hausdorff groups. *Probability Theory and Related Fields*, 130(2):151–166, 2004.

X. Mu, L. Pomatto, P. Strack, and O. Tamuz. From Blackwell dominance in large samples to Rényi divergences and back again. *Econometrica*, 89(1):475–506, 2021.

R. B. Myerson and E. Zambrano. *Probability models for economic decisions*. MIT Press, 2019.

S. Onay and A. Öncüler. Intertemporal choice under timing risk: An experimental approach. *Journal of Risk and Uncertainty*, 34(2):99–121, 2007.

L. Pomatto, P. Strack, and O. Tamuz. Stochastic dominance under independent noise. *Journal of Political Economy*, 128(5):1877–1900, 2020.

M. Rabin and G. Weizsäcker. Narrow bracketing and dominated choices. *American Economic Review*, 99(4):1508–43, 2009.

I. Ruzsa and G. J. Székely. *Algebraic probability theory*. John Wiley & Sons Inc, 1988.

M. L. Weitzman. Gamma discounting. *American Economic Review*, 91(1):260–271, 2001.

L. Zhou. Harsanyi's utilitarianism theorems: General societies. *Journal of Economic Theory*, 72(1):198–207, 1997.

S. Zuber. Can social preferences be both stationary and Paretian? *Annals of Economics and Statistics*, 101/102:347–360, 2011.