Theory Dec.
DOI 10.1007/s11238-015-9504-1

Choice under aggregate uncertainty

Nabil I. Al-Najjar[1] · **Luciano Pomatto**[1]

Abstract We provide a simple model to measure the impact of aggregate risks. We consider agents whose rankings of lotteries over vectors of outcomes satisfy expected utility and separability. Such rankings are characterized in terms of *aggregative utilities* that measure sensitivity to aggregate uncertainty in a straightforward way. We consider applications to models of product variety, portfolio choice, and public attitudes towards catastrophic risks. The framework lends support to precautionary measures that penalize policies for exposure to correlation. The model rationalizes a number of behavioral and policy patterns as attempts to hedge against aggregate uncertainty.

Keywords Aggregate risks · Risk and uncertainty

1 Introduction

An important question facing individuals and economic institutions is how to react to correlation. A recent example is the debate over the role of systemic risk in financial markets. Although the financial positions of individual institutions may appear sound, recent crises make a compelling case for heightened regulatory scrutiny of banks'

✉ Nabil I. Al-Najjar
al-najjar@northwestern.edu
http://www.kellogg.northwestern.edu/faculty/alnajjar/htm/index.htm

Luciano Pomatto
http://www.kellogg.northwestern.edu/faculty/pomatto/index.htm

[1] Department of Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, Evanston, IL 60208, USA

Published online: 24 June 2015

exposure to the correlation in these positions.[1] The role of correlation appears, of course, in other even more basic contexts. In consumer theory, for example, a consumer will likely treat uncorrelated changes in relative prices differently from aggregate shifts in consumption levels.[2]

A common modeling practice is to assume an additively separable utility:

$$V(s) = \frac{1}{n} \sum_{i=1}^{n} v_i(s_i). \tag{1}$$

Here, the agent faces n coordinates, $s = (s_1, \ldots, s_n)$ is a generic profile, v_i is the utility derived from the ith coordinate, and $\frac{1}{n}$ is an innocuous normalization. A lottery P over profiles is evaluated based on its expected utility $E_P V$. Although widely used for their tractability and appealing foundations, additively separable utilities suffer from a major limitation, namely their insensitivity to correlation.[3] As noted above, concern about aggregate uncertainty is natural in many settings. It is entirely reasonable for a public authority to treat correlated pandemic risk (such as the reaction to the recent Ebola scare) differently from uncorrelated health incidents. The fact that additively separable utilities cannot distinguish the two is potentially an important limitation.

There is, of course, an easy way out: simply replace V by a general von Neumann–Morgenstern utility function $U(s_1, \ldots, s_n)$ that is not separable in its n coordinates. While this approach can succeed in introducing sensitivity to correlation, it is intractable without further structure on U. For example, one would like to answer questions like: What is the impact of increasing n? Is randomization beneficial? or What are good quantitative measures of the attitude towards correlation? These questions are central to understanding the impact of aggregate risks, yet they can be difficult, if not impossible, to answer with a general utility function.

This paper provides a very simply way to identify and measure the sensitivity of economic decisions to aggregate uncertainty. Our starting point is an agent who ranks lotteries based on their expected utilities with respect to a von Neumann–Morgenstern utility U. A tractable model is obtained by requiring, further, that the ranking of *deterministic* profiles satisfies the conditions in Debreu's (1960) classic characterization of separable preferences. His theorem then implies that this ranking has a cardinal representation V, in the sense of (1) above.[4] We are therefore given two cardinal utilities, U and V, with identical ordinal ranking of profiles. It follows that there must be a (cardinally unique) strictly increasing function u such that U has the *aggregative utility* form $U(s) = u(V(s)) = u(n^{-1} \sum_i v_i(s_i))$.

---

[1] See, for instance, Acharya et al. (2010) and Adrian and Brunnermeier (2011) for discussions of the policy issues surrounding, and measures to correct for, correlation.

[2] See Sect. 3.4's discussion of the seminal Dixit–Stiglitz model of consumer choice with product varieties.

[3] Since the operations of summing over the population and integrating over states commute, the utility of a lottery depends only on its marginals on the coordinates $s_1, \ldots, s_n$, forcing indifference to correlation. See equation (4).

[4] It is crucial that the utility function (1) is unique up to a *common* positive affine transformation. Multiplication by $\frac{1}{n}$ is innocuous and simplifies comparisons across different n's. The formal structure of Debreu's theorem is reviewed below.

We interpret u as reflecting the attitude towards aggregate uncertainty. To justify this claim, note that the function U captures two conceptually distinct aspects of the problem: (1) uncertainty about which profile will obtain; and (2) the aggregation of n coordinates into a utility of a profile. These two components are intertwined in a general U. Combining the von Neumann–Morgenstern and Debreu's cardinal theories is a natural way to disentangle the two. This approach is very simple, making it all the more surprising that, to our knowledge, it has not been taken before.

Indifference to correlation obtains if and only if u is affine, while aversion to aggregate uncertainty corresponds to the strict concavity of u. More generally, with obvious caveats to be discussed below, the standard machinery of utility theory can be used to quantify the attitude towards correlation. We provide a detailed example in the case of the Dixit–Stiglitz CES aggregator.

Aggregate uncertainty plays a central role in evaluating public policies, especially as they relate to catastrophic risks. Catastrophes represent, almost by definition, risks that are correlated across individuals (firms, assets, or investments). The increasing interconnectedness of modern economies sharpened the impact of old sources of correlation, such as systemic risk in financial markets, medical treatment uncertainty, and product recalls. Political and technological changes also created new sources of correlation, including climate change risk and global terrorism. Many authors noted the fact that American public opinion reacts differently to these risks compared to more familiar ones like car accidents and house fires.[5] A large debate on the proper public policy attitude to catastrophic risk centers around the status of the Precautionary Principle, a policy position that has been widely adopted in many laws, international treaties, and government regulations.[6] This and other related principles that focus on worst-case scenarios have been criticized by Sunstein (2005) and others as incoherent.

Aggregative utility can also provide a foundation for randomized decision rules. We consider three instances where randomization arises in the literature. First, Manski (2004, 2011) makes a normative case for such rules when the effectiveness of a treatment is uncertain. He derives the optimality of randomized rules by assuming a utilitarian social planner who uses a non-Bayesian minimax regret criterion. We show, consistently with Manski and Tetenov (2007), that a Bayesian planner with an aggregative utility may strictly prefer to randomize as a way to hedge against aggregate uncertainty. Second, sensitivity to aggregate uncertainty relates to an insight of Schmeidler (1989) and Gilboa and Schmeidler (1989) that individuals may randomize to hedge against unknown probabilities. Our work is related to Halevy and Feltkamp's (2005) finding that strict preference for randomization arises when utility depends on the outcomes of multiple correlated urns. We note that the decision maker in Halevy

[5] See, for example, Sunstein (2005). Robson (1996) provides an evolutionary justification for why Nature might have designed individuals with differing attitudes towards aggregate vs. idiosyncratic risks.

[6] For example, the United Nations Framework Convention on Climate Change, the 1992 Rio Declaration, the Treaty Establishing the European Community, the U.S. National Environmental Policy Act, and the U.S. Clean Water Act. One of the best known statements of this policy principle is "Wingspread Declaration" which states that "When an activity raises threats of harm to human health or the environment, precautionary measures should be taken even if some cause-and-effect relationships are not fully established scientifically." Ashford et al. (1998). See Sunstein (2005) for a review and critique of this principle and Al-Najjar (2015) for a different perspective on the issue.

and Feltkamp (2005) has aggregative utility in our sense. Finally, a surprising context where randomization appears is evolutionary dynamics. In Bergstrom (1997), for example, Nature may introduce heterogeneity in preferences when there is aggregate uncertainty. Section 5.4 discusses how randomization by introducing heterogeneity in a population may be interpreted within our analysis.

The plan of the paper is as follows: Sect. 2 introduces notation and a motivating example, while Sect. 3 states the main theorems. Section 4 considers asymptotics as n increases. Finally, Sect. 5 considers public choice questions, fractional allocations, and connects our model to some of the findings in the literatures.

2 Notation and motivating example

2.1 Notation and mathematical structure

An agent's utility depends on the realization of a profile of n coordinates, indexed by $i \in \{1, \ldots, n\}$. To each coordinate i corresponds a set of *(individual) outcomes* X_i. The product $S = X_1 \times \cdots \times X_n$ is the set of *profiles*, with typical element $s = (s_1, \ldots, s_n)$. For an individual outcome $x \in \cap_{i=1}^{n} X_i$, the constant profile where all coordinates are equal to x is denoted by $\bar{x} = (x, \ldots, x)$.

We assume that each X_i is a complete separable metric space. For any such space Z, we let $\Delta(Z)$ denote the set of probability distributions (lotteries) on Z.[7] We shall refer to elements of $\Delta(S)$ as *lotteries* and elements of $\Delta(X_i)$ as *individual lotteries*.

Given a lottery P, we use p_i to denote its marginal on coordinate i. Define $P^\circ \in \Delta(S)$ to be the product of the marginals $p_1 \times \cdots \times p_n$. The lottery P° preserves the marginal distributions of P but removes any correlation that might exist between the coordinates. A lottery P is *independent* if, under P, the random elements $s_1, \ldots, s_n$ are independent. Note that a lottery P is independent if and only if $P = P^\circ$.

2.2 Example: the "life and death" lottery

Consider the problem of choosing among medical treatments with uncertain effectiveness. A common approach is to assume that each individual has a utility function v_i over X_i and to evaluate policies based on the additively separable criterion $V(s) = \frac{1}{n} \sum_{i=1}^{n} v_i(s_i)$ introduced in (1). A treatment that gives rise to a lottery P has expected utility:

$$E_P V = \int_S \left(\frac{1}{n} \sum_{i=1}^{n} v_i(s_i) \right) \mathrm{d}P(s). \tag{2}$$

This criterion, often referred to as "utilitarian," is extensively used (e.g., in optimal taxation, treatment problems, among other areas). We caution, however, that it repre-

[7] To avoid unnecessary repetition, probability measures on complete separable metric spaces are defined on their Borel σ-algebras. All product spaces, such as S, are given their product topologies and corresponding σ-algebras. All subsets are assumed to be measurable unless otherwise noted.

sents a very narrow form of utilitarianism that requires individuals to be indifferent to others' outcomes. Section 5 discusses this problem; for now, this simple criterion is convenient for the purpose of the present example.

An important concern with (2) is its insensitivity to aggregate uncertainty. A stark example illustrates this point:

Example 1 Consider medical treatments whose individual outcomes are either full recovery, x, or death, y! Assume that $v_i = v$ for all i, and normalize utilities so that $v(x) = 1$ and $v(y) = 0$. Fix $0 < \alpha < 1$ and consider the following treatments:

- *Treatment 1* yields a lottery where:
  - $\bar{x} = (x, \ldots, x)$ obtains with probability α, and
  - $\bar{y} = (y, \ldots, y)$ obtains with probability $1 - \alpha$.
- *Treatment 2* yields a lottery that is uniform over the set $S_\alpha \subset S$ of *fractional profiles* s_α where $\alpha\%$ of the population recovers, and the rest dies.[8]

Under the additively separable criterion (2), Treatment 1 generates an expected utility $\alpha V(\bar{x}) + (1 - \alpha)V(\bar{y}) = \alpha$ through a *probability mixture* of the two constant profiles, $\bar{x}$ and $\bar{y}$. Under Treatment 2, a utility α is achieved via a *population mixture*: a profile $s_\alpha \in S_\alpha$ is drawn where a *known* population fraction α gets x, and a fraction $(1 - \alpha)\%$ gets y. Given s_α, no probabilities are involved.

Since an additively separable utility treats the perfectly correlated probability mixtures and (essentially) idiosyncratic population mixtures identically, it is indifferent to a treatment where half the population lives and the other half dies, and a treatment where either everyone lives or everyone dies depending on the outcome of a coin toss! While it is not irrational for a planner to be indifferent to the two, it is also not irrational to consider the two treatments to be very different. The notion of aggregative utility we now introduce provides a way to break the indifference to aggregate uncertainty.

3 Aggregative utility: characterization and properties

3.1 Expected utility

We assume an agent with an expected utility preference over the set of lotteries $\Delta(S)$. First, some structural assumptions:

Assumption 1 (*Structural Assumptions*) The following structural assumptions will be maintained throughout the paper:

(i) $n \geq 3$;
(ii) each X_i is a complete separable metric space;
(iiii) each X_i is connected.

8 This formulation simplifies the exposition by forcing the conclusion of the law of large numbers to hold exactly for any n. Since s_α is chosen uniformly, individual outcomes are identically distributed, but not independent. They are, however, "nearly" independent, in the sense that correlation between any two outcomes decreases to zero as n increases. Finally, note that we implicitly assume α to be a multiple of $\frac{1}{n}$ to ensure that S_α is well-defined.

This assumption will be maintained throughout the paper. Parts (i) and (iii) are required to apply the characterization theorem of Debreu (1960). Part (ii) simplifies the introduction of lotteries.

Assumption 2 (*Expected Utility*) The agent ranks lotteries $P \in \Delta(S)$ according to their expected utility:

$$\mathcal{U}(P) = \int_S U(s)\, dP(s),$$

where $U : S \to \mathbb{R}$ is a bounded and continuous function that is unique up to a positive affine transformation (or *cardinally unique* for short).

In the discussion below, we will refer to a specific function U for convenience. It is to be understood, however, that this function is well-defined only up to positive affine transformation.

Since any profile s can be identified with the degenerate lottery δ_s that puts unit mass on s, we may view the set of deterministic profiles S as a subset of $\Delta(S)$. The aggregation properties of U concern the restriction of utility to S. For every non-empty proper subset of indices $I \subset \{1, \ldots, n\}$, let S_I denote the set of profiles defined for members in I only. Given such an I, define $S = S_I \times S_{I^c}$, where I^c is the complement of I. Using this notation, write any profile as $s = (s_I, s_{I^c})$, where $s_I \in S_I$ and $s_{I^c} \in S_{I^c}$.

Assumption 3 (*Non-triviality*) For every non-empty proper subset of indices I, there exists s_I, y_I, z_{I^c} such that

$$U(s_I, z_{I^c}) > U(y_I, z_{I^c}).$$

3.2 Main theorem

The next assumption is Debreu's separability condition over deterministic profiles. Formally, given a profile t, let $S_{t_I} = \{s : s_i = t_i, i \in I\}$ be the set of profiles that coincide with t on I, and let U_{t_I} be the restriction of U to S_{t_I}. Thus, U_{t_I} is the utility over profiles whose extensions to the subpopulation I are made to coincide with t.

Assumption 4 (*Separability*) For every non-empty proper subset of indices I and profiles t and z, we have

$$U_{t_I} = U_{z_I}.$$

This condition says that the ranking of profiles that have a common value on a subset I does not depend on that common value. It is clear that this condition is necessary for an additive representation over S. Debreu (1960) shows that it is also sufficient.[9] Our main representation result is

[9] Assuming non-triviality, continuity, and a version of the structural conditions above, Gorman (1968) extends Debreu's theorem to partial separability.

Theorem 1 *The ranking over the set of lotteries $\Delta(S)$ is non-trivial, has an expected utility representation, and is separable if and only if it can be represented by*

$$\mathcal{U}(P) = \int_S u\left(\frac{1}{n}\sum_{i=1}^{n} v_i(s_i)\right) \mathrm{d}P(s) \tag{3}$$

for some continuous and non-constant functions $v_i : X_i \to \mathbb{R}$, $i = 1, \ldots, n$, and an increasing, continuous, bounded function u : $\mathrm{range}\left(\frac{1}{n}\sum_{i=1}^{n} v_i\right) \to \mathbb{R}$. In addition, the v_i's are unique up to a common positive affine transformation, and, for fixed $v_1, \ldots, v_n$, u is cardinally unique.

The theorem says that the agent evaluates a distribution on profiles in three steps: (1) an individual utility $v(s_i)$ is calculated for each i; (2) the utilities are additively aggregated to $\frac{1}{n}\sum_{i=1}^{n} v_i(s_i)$; and (3) the agent's final utility is a function u of this aggregate. The factor $\frac{1}{n}$ is a convenient normalization to simplify comparisons when n changes.

Definition 1 (i) The utility $\mathcal{U}$ (underlying preference) satisfying the properties in Theorem 1 is called an *aggregative utility* (preference).
(ii) An aggregative utility is *additively separable* if u is affine.
(iii) A utility is *additively separable* if it has an additively separable representation.

Debreu's theorem is usually used to justify additively separable consumer utility over commodity bundles in a deterministic setting. In our case, we start with an expected utility preference over lotteries represented by a utility function U. An expected utility preference is necessarily separable across states. We further assume that the restriction of this preference to deterministic profiles is separable across coordinates. The interplay between these two forms of separability is what underlies Theorem 1.

Our companion paper develops a similar theory for infinite populations (Al-Najjar and Pomatto (2014), available on our websites). An important advantage is that an exact law of large numbers holds in the infinite setting. The two are complementary, but independent, theories. Debreu's theorem, which underlies the present model, requires the number of coordinates to be finite, so our arguments do not extend to the infinite population setting. We instead use Savage's (1954) theory of subjective probability to provide a cardinal population aggregator. An important limitation of Savage's theory is that the v_i's must all be equal. Since applying his theory requires the population to be infinite, the (more complex) machinery in our companion paper cannot be used to deal with the finite case. Finite models are important in many applications. Regulation of an industry consisting of a handful of firms (such as pharmaceutical companies or investment banks) is one example. Another example is modeling attitudes to different sources of revenue, as discussed below.

3.3 Characterizing the sensitivity to aggregate uncertainty

In this section, we formalize the intuition that u captures the agent's sensitivity to aggregate uncertainty.

Definition 2 A utility U is *indifferent to aggregate uncertainty* if $\mathcal{U}(P) = \mathcal{U}(Q)$ for any two lotteries $P, Q \in \Delta(S)$ that have the same marginals (*i.e.,* $p_i = q_i$ for all i).

In particular, such a planner is indifferent to a lottery P and the product of its marginals P°. The following theorem says that additive separability is both necessary and sufficient for indifference to aggregate uncertainty:

Theorem 2 *A utility U is indifferent to aggregate uncertainty if and only if it is additively separable.*

To further appreciate the role of aggregate uncertainty, assume $v_i = v$ for all i and that u is affine. Then, after an appropriate normalization, we can express aggregative utility as

$$\int_S \frac{1}{n} \sum_{i=1}^{n} v(s_i) \mathrm{d}P(s) = \frac{1}{n} \sum_{i=1}^{n} E_P v(s_i). \tag{4}$$

If we interpret $\frac{1}{n}$ in the RHS as a uniform distribution on the finite set $\{1, \ldots, n\}$, then the above expression can be interpreted as the expected utility obtained from interacting with a single randomly drawn individual.[10] In other words, indifference to aggregate uncertainty eliminates the distinction between two intuitively different problems: one where decisions are based on the entire profile, and another where they are based on a randomly drawn individual.

To illustrate the theorems, we revisit the stylized life-and-death lottery introduced in Sect. 2.2. In that example, we had

$$V(\bar{x}) = 1 > V(s_\alpha) = \alpha > V(\bar{y}) = 0.$$

Assume, further, that the preference *over lotteries* has an expected utility representation $\int_S U(s) \mathrm{d}P(s)$, normalized so that:

$$U(\bar{x}) = 1 > U(s_\alpha) = u(\alpha) > U(\bar{y}) = 0.$$

Since V and U agree on the *ordinal* ranking of profiles, U can be decomposed as $U(s) = u(V(s)$, where V represents how profiles are aggregated, and u reflects attitude towards aggregate uncertainty.

In Fig. 1, population mixtures are represented on the horizontal axis, and probability mixtures on the vertical axis. Given a fractional profile s_α, the additive separability of V implies

$$V(s_\alpha) = \alpha V(\bar{x}) + (1 - \alpha) V(\bar{y}).$$

10 The assumption that $v_i = v$ for all i is not completely innocuous. Suppose that, for some vector $(q_1, \ldots, q_n)$ of probabilities on $\{1, \ldots, n\}$, the function $\hat{V}(s) = \sum_i q_i \hat{v}_i(s_i)$ is another cardinal representation of $V(s) = n^{-1} \sum_i v_i(s_i)$. Since utility is unique up to a *common* positive affine transformation, requiring $v_i = v$ for all i in V forces $q_i = \frac{1}{n}$ for all i in $\hat{V}$. Without this requirement, the "probabilities" with which various individuals are drawn are indeterminate. While this indeterminacy does affect the point being made here, it also adds a degree of freedom that may confound the issue.

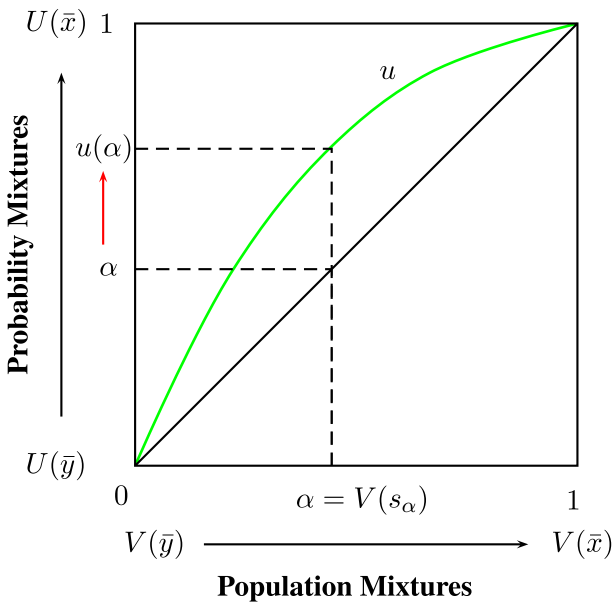

Fig. 1 Population versus probability mixtures

On the other hand, the number $u(\alpha)$ must satisfy:

$$U(s_\alpha) = u(\alpha)V(\bar{x}) + (1 - u(\alpha))V(\bar{y}).$$

That is, the utility $U(s_\alpha)$ of the α-fractional profile is an average of the utility of the constant profiles but with the weights $\alpha\%$, $(1-\alpha)\%$ distorted by u. Theorem 2 states that the 45° line, defined by $u(\alpha) = \alpha$, characterizes indifference to aggregate uncertainty. In this case, α-population mixtures and α-probability mixtures are treated as equivalent. On the other hand, to convince a planner averse to aggregate uncertainty to give up s_α, a higher probability of $\bar{x}$ is required. This means that $u(\alpha) > \alpha$ and u must be concave, as shown in the figure. For a concave u, the difference $u(\alpha) - \alpha$ measures the compensation needed to offset the impact of aggregate uncertainty.

The construction of Theorem 1 works only because Assumptions 2 and 4 provide *two independent cardinal measurements* of utility. To illustrate this, suppose that U is the cardinal von Neumann–Morgenstern utility as before, but that $\hat{V}$ is just an ordinal utility whose ranking of profiles is consistent with U.[11] While it is still true that $U = \hat{u}(\hat{V})$ for some increasing function $\hat{u}$, the shape of this function no longer captures anything useful. When $\hat{V}$ is only ordinal, we can modify it with increasing transformations to make $\hat{u}$ take just about any shape. Debreu (1960)'s theorem is key because it fixes a cardinal ranking over profiles.

As further illustration of the results, we briefly discuss one of the most familiar contexts where aggregative utility appears, namely portfolio allocation among n assets.

[11] In the sense that $\hat{V}(s) \geq \hat{V}(s')$ if and only if $U(s) \geq U(s')$ for any s, s'.

Let $s_i \in \mathbb{R}$ denote the random payoff to investing \$1 in assets $i = 1, \ldots, n$, and $\alpha = (\alpha_1, \ldots, \alpha_n)$ a portfolio where α_i is the amount of money invested in asset i. The standard assumption is that the investor evaluates portfolios based on the expected utility:

$$\int u\left(\frac{1}{n}\sum_{i=1}^{n} \alpha_i s_i\right) \mathrm{d}P(s).$$

This, of course, is an aggregative utility where $v_i(\alpha_i s_i) = \alpha_i s_i$ (and the normalization $\frac{1}{n}$ is suppressed).

The assumption that the v_i's are the identity builds-in perfect fungibility of different sources of income. Experimental studies (and introspection) suggest that individuals often treat different sources of returns as different "accounts." When the v_i's are possibly different, we obtain

$$\int u\left(\frac{1}{n}\sum_{i=1}^{n} v_i(\alpha_i s_i)\right) \mathrm{d}P(s).$$

Barberis and Huang (2001) use a functional form with similar features to model investors who focus narrowly on sources of gains and losses. The reader is referred to their paper for further motivation and references.[12] Violating fungibility in our model is not "irrational," in the sense that it is not inconsistent with expected utility (Assumption 2) and can even be consistent with separability (Assumption 4).

3.4 Sensitivity to correlation with CES aggregators

In application, it is often important to quantify the sensitivity to correlation in terms of a preference parameter. For concreteness, consider Dixit and Stiglitz (1977)'s seminal model of consumer choice with product varieties. A representative consumer has a CES utility:

$$\sum s_i^{\rho},$$

where $s_i \in (0, \infty)$ is the number of units of variety $i = 1, \ldots, n$, and $\rho \in (0, 1)$ measures the representative consumer's taste for variety.[13] This is often normalized using the transformation $r \mapsto \left(\frac{r}{n}\right)^{\frac{1}{\rho}}$ to obtain

12 In Barberis and Huang (2001), a narrow framing may reflect "utility unrelated to consumption. [...] An investor may interpret a big loss on a stock as a sign that he is a second-rate investor, thus dealing his ego a painful blow, and he may feel humiliation in front of friends and family when word about the failed investment leaks out."

13 This is one of many aggregators considered in their paper. We focus on the finite-variety version of the Dixit–Stiglitz aggregator; the framework of our companion paper, Al-Najjar and Pomatto (2014), can be used to model the infinite variety case.

$$\left(\frac{1}{n}\sum s_i^{\rho}\right)^{\frac{1}{\rho}}. \tag{5}$$

The exponent $\frac{1}{\rho}$ in (5) has, in principle, no cardinal meaning. In models with no uncertainty, the interpretation of $\frac{1}{\rho}$ is a moot issue. But when consumption levels are random, say due to price fluctuations, a cardinal measure is needed. To obtain such a measure, apply Theorem 1 and require, further, that $v_i = v$ takes the form $v(r) = r^{\rho}$. This yields a cardinal utility of the form:

$$U(s) = u\left(\frac{1}{n}\sum s_i^{\rho}\right). \tag{6}$$

Note that Debreu's theorem does not restrict the shape of the v_i's, so the CES form must either be assumed directly or derived from additional assumptions.

To place more structure on (6), assume that (the cardinally unique) u has the form $u(r) = r^{\frac{\kappa}{\rho}}$ for some $\kappa > 0$ so (6) can be rewritten as

$$\left(\frac{1}{n}\sum s_i^{\rho}\right)^{\frac{\kappa}{\rho}}. \tag{7}$$

The utility function (7) distinguishes between the representative consumer's taste for variety and his attitude towards uncertainty about consumption levels.

To make this precise, recall that the elasticity of substitution between any pair of varieties is $\frac{1}{1-\rho}$. To incorporate uncertainty about consumption levels, fix a consumption profile s and consider proportional changes in consumption of the form αs for $\alpha > 0$. We can quantify the sensitivity of utility to such changes by the coefficient of relative risk aversion with respect to changes in α:

$$R(\alpha) \equiv \frac{-\alpha\, U''(\alpha s)}{U'(\alpha s)},$$

where derivatives are with respect to α. It is easy to see that $R(\alpha) = 1 - \kappa$.[14]

For $\kappa = 1$, we obtain the usual normalization (5), except that the exponent $\frac{1}{\rho}$ has a cardinal meaning corresponding to risk neutrality to lotteries over consumption levels. For $\kappa = \rho$, aversion to changes in consumption level is the inverse of the elasticity of substitution between varieties. Aversion to lotteries in income and lotteries in relative prices are represented by the same parameter ρ. Finally, when $\kappa < \rho$, the representative consumer is more averse to changes in consumption level than to substitution between varieties.

[14] Note that $U(\alpha s) = \left(\frac{1}{n}\sum_i (\alpha\, s_i)^{\rho}\right)^{\frac{\kappa}{\rho}} = \alpha^{\kappa}\, U(s)$ and differentiate with respect to α.

4 Aggregating idiosyncratic risks

Does idiosyncratic risk become irrelevant as n increases? This conclusion does not directly follow from the law of large numbers, which concerns the statistical properties of profiles $(s_1, \ldots, s_n)$, rather than the utility they generate. Even if the s_i's are i.i.d., the fact that the v_i's are different means that a small number of individuals may continue to have a large utility impact regardless of n.

Whether idiosyncratic risk is *irrelevant to utility* is not automatic, as it requires conditions to ensure that utility is not too sensitive to individual coordinates. The theorems in this section introduce such conditions and use concentration-of-measures results to derive bounds on the impact of idiosyncratic risk for any given n (rather than just asymptotically).

4.1 Idiosyncratic risk in large finite populations

Given the aggregative utility

$$\mathcal{U}(P) = \int_S u\left(\frac{1}{n}\sum_{i=1} v_i(s_i)\right) \mathrm{d}P(s), \tag{8}$$

it will be useful to define

$$\bar{\mathcal{U}}(P) = u\left(\frac{1}{n}\sum_{i=1}^{n} E_{p_i} v_i\right) \tag{9}$$

as the utility obtained by replacing random outcomes by their expectations $E_{p_i} v_i$ with respect to the marginal p_i.

Idiosyncratic risk disappears when $\mathcal{U}(P)$ is approximately equal to $\bar{\mathcal{U}}(P)$. Instead of an asymptotic argument that takes limits as n increases, we use concentration inequalities to establish a bound for any n. First we need the following:

Assumption 5 The aggregative utility $\mathcal{U}$ satisfies

(i) The range of v_i is contained in [0, 1] for every i;
(ii) u is Lipschitz continuous, with Lipschitz constant K;
(iii) The range of u is contained in [0, 1].

It is important that the bounds on the functions $v_i, i = 1, \ldots, n,$ and u are independent of n (the specific values, 0 and 1, are for convenience). The first two parts of the assumption ensure that the population average utility is not dominated by the utility of a single individual. By restricting the ranges of the $v_i's$, part (i) rules out the possibility that the utility impact of a single individual is so large that idiosyncratic risk remains significant even as n increases without bound. Part (ii) rules out the case where a small change in one individual's outcome can have large impact on the planner's utility. Part (iii) is an innocuous normalization.

Theorem 3 *For every aggregative utility $\mathcal{U}$ satisfying Assumption 5, for every $\epsilon > 0$, and every independent distribution P*

$$P\left\{s : \left|u\left(\frac{1}{n}\sum_{i=1}^{n} v_i(s_i)\right) - \bar{\mathcal{U}}(P)\right| < \epsilon\right\} > 1 - 2\mathrm{e}^{-2n\left(\frac{\epsilon}{K}\right)^2}.$$

In particular,

$$\left|\mathcal{U}(P) - \bar{\mathcal{U}}(P)\right| < \epsilon + 2\mathrm{e}^{-2n\left(\frac{\epsilon}{K}\right)^2}.$$

Theorem 3 formalizes the intuition that independent risks disappear as n increases *without assuming* that the lotteries are i.i.d. or that the v_i's are equal (much less linear). The key feature of the theorem is that the bounds hold simultaneously over all n's, u's, v_i's, and independent lotteries P. No assumptions are made relating, for instance, the profile of individual utility functions $(v_1, \ldots, v_n)$ for different values of n.

4.2 The conditionally i.i.d. case

A natural way to model correlation is to assume that outcomes are i.i.d. with an unknown parameter. This class of conditionally i.i.d. distributions is widely used in practice and has the advantage of being easy to parameterize.

Definition 3 (*Conditionally i.i.d. Distributions*) Fix n, assume that $X = X_i$ for all i, and let $\Theta = \Delta(X)$ denote the set of probability distributions on X.

(1) For $\theta \in \Theta$, a distribution P^θ is i.i.d.(-θ) if the random objects $s_1, \ldots, s_n$ are independent with a common marginal θ.
(2) For $\mu \in \Delta(\Theta)$, a distribution P^μ is conditionally i.i.d.(-μ) if for every event $E \subset S$, we have $P^\mu(E) = \int_\Theta P^\theta(E)\, d\mu(\theta)$.[15]
(3) The marginal distribution of a conditionally i.i.d. P^μ on coordinate i is

$$p_i^\mu(E) = \int_\Theta \theta(E)\mathrm{d}\mu(\theta), \;\; E \subset X_i. \tag{10}$$

Note that p_i^μ in (10) does not depend on i, and that when μ puts unit mass on a single θ, we have $p_i^\mu = p_i^\theta = \theta$. Nevertheless, we continue writing p_i^μ throughout for notational consistency.

With this notation, aggregative utility is

$$\mathcal{U}_n(P^\mu) = \int_\Theta \int_S u\left(\frac{1}{n}\sum_{i=1}^{n} v_i(s_i)\right)\mathrm{d}P^\theta(s)\mathrm{d}\mu(\theta). \tag{11}$$

[15] If n were infinite, then de Finetti's theorem characterizes conditionally i.i.d. distributions as those that are invariant to permutations (or exchangeable). This equivalence does not hold for finite n.

The usual interpretation is that the outcomes are i.i.d. with an unknown parameter θ, with μ representing the planner's belief about this parameter.

From Theorem 3, we know that independent (in particular, i.i.d.) variations wash out when n is large. Thus, for any parameter θ, the realized random utility concentrates around the utility of the expectation, and the rate at which this concentration occurs is independent of θ. The next theorem uses Theorem 3 to conclude that, when n is large, idiosyncratic risk disappears, but aggregate uncertainty remains.

Theorem 4 *Assume $X = X_i$ for all i, j, and that Assumption 5 holds. Then, for every $\mu \in \Delta(\Theta)$,*

$$\left| \mathcal{U}(P^{\mu}) - \int_{\Theta} u\left(\frac{1}{n}\sum_{i=1}^{n} E_{p_i^{\theta}} v_i\right) \mathrm{d}\mu(\theta) \right| < \epsilon + 2\mathrm{e}^{-2n\left(\frac{\epsilon}{K}\right)^2}.$$

Idiosyncratic risk disappears in the sense that variability conditional on the parameter moves inside u, but aggregate uncertainty, represented by μ, does not.

5 Planning under aggregate uncertainty

A large literature evolved to explore whether special public policy measures are warranted when dealing with aggregate uncertainty.[16] Compare, for example, the introduction of a fire code with the implementation of a new medical procedure. Assume that both policies give rise to conditionally i.i.d. lotteries with identical marginals. In the case of fire code, abundant past data support the belief that the causes of fires and the effectiveness of safety measures are well-understood, while for a new medical procedure, it is likely that a significant residual uncertainty about its effectiveness remains.

We explore some of these issues from the perspective of aggregative utility in a large population. Unless indicated otherwise, we assume that policies give rise to conditionally i.i.d. distributions and that the policy maker has an aggregative utility that takes the form:

$$\mathcal{U}(P) = \int_{\Theta}\int_{S} u\left(\frac{1}{n}\sum_{i=1}^{n} v(s_i)\right) \mathrm{d}P^{\theta}(s)\mathrm{d}\mu(\theta). \tag{12}$$

This makes a number of implicit assumptions: (1) individuals have symmetric preferences that do not depend on the size of the population; (2) the dependence of the planner's utility on the average individual utilities also does not vary with n.

[16] An example is the Precautionary Principle, according to which "When an activity raises threats of harm to human health or the environment, precautionary measures should be taken even if some cause-and-effect relationships are not fully established scientifically." *"Wingspread Statement on the Precautionary Principle,"* Ashford et al. (1998). See Sunstein (2005) for a review and critique of this principle and Al-Najjar (2015) for a different perspective on precautionary policies.

5.1 Social versus private attitudes to risk

Consider the problem of designing a fire code. An individual outcome in this case contains all relevant information about the costs of safety measures, property damages, ... etc. The function v summarizes the various individual costs and benefits trade-offs. A fire code then gives rise to a social lottery P^{μ} characterized by a distribution μ on Θ.

It may be reasonable to think that there is enough data on the causes of fires and the effectiveness of safety measures that the distribution on the outcomes θ is known. Contrast this with the problem of approving a new drug. Here, there is usually significant residual uncertainty about outcomes, such as side effects, costs, and effectiveness, even after lengthy medical trials. This uncertainty can be captured by a diffuse μ, reflecting the correlation in the drug's performance across users. Call a policy P^{μ} *risky* if μ puts unit mass on a single θ (the fire code example), and *uncertain* otherwise.

Can a planner with an aggregative utility rely on self-interested individuals to carry out his policy choices? To answer this question, we need to specify agents' preferences. An individual with utility v over individual outcomes cares only about s_i and will thus rank social lotteries P^{μ} based on their marginal distributions p_i^{μ}. Intuitively, p_i^{μ} is the "average" lottery, where the average is taken with respect to the distribution μ. The distribution p_i^{μ} completely ignores correlation.

The following definition formalizes the relationship between public and private attitudes towards risk:

Definition 4 The choice between two social lotteries P^{μ} and P^{ν} can be *decentralized in large populations* if there is an N such that for all $n \geq N$,

$$\mathcal{U}^n(P^{\mu}) \geq \mathcal{U}^n(P^{\nu}) \iff E_{p_i^{\mu}} v \geq E_{p_i^{\nu}} v.$$

Implicit in the definition is the assumption that all individuals and the planner agree on μ. Under this assumption, the planner is not more (or less) informed than the individuals, ruling out tensions in public policy due to differing subjective beliefs. While we do not believe that this assumption is always reasonable, maintaining it makes it possible to focus on the role of aggregate uncertainty.

Theorem 5 *Fix an aggregative utility $\mathcal{U}$, as in* (12)*, satisfying Assumption* 5.

(1) *The choice between any two risky social lotteries θ_0, θ_1 can be decentralized in large populations.*
(2) *Unless $\mathcal{U}$ is additively separable, there exist social lotteries P^{μ} and P^{ν} such that the choice between them cannot be decentralized in large populations.*

Part 1 says that individual and social ranking of *risky* lotteries coincide in large populations. This, of course, is trivial for any population size if u were affine. The point is that this holds even when u is not affine, provided that the population is large. Consider the case where u is strictly concave. In this case, the planner is more risk averse than any individual agent; indeed, for small n, the planner would prefer more cautious alternatives than the individual.[17] However, when n is large, Theorem 3 can be used to show that the impact of i's utility on the social planner diminishes, bringing the planner's ranking closer to the agent's.

Part 2 says that the planner's sensitivity to aggregate uncertainty creates a tension between the planner's and the individual's preferences. For concreteness, consider a planner with a strictly concave u. An individual will be indifferent to a social lottery P^{μ} and an i.i.d. lottery with marginal p^{μ}. However, the planner is willing to settle for a worse social lottery P^{μ} in exchange for removing aggregate uncertainty, while the individual, who is indifferent to aggregate uncertainty, would consider such move as inferior.

5.2 Hedging and fractional allocations

In many instances the set of options is discrete: vaccines are either taken or not, airbags are installed or not, ... etc. Motivated by Manski's work,[18], we explore the use of randomization and fractional allocations in treatment problems when the planner has aggregative utility.

For concreteness, we focus on a very simple setting with just two policies: a status quo policy that guarantees a constant profile $\bar{x} = (x, \ldots, x)$ and an uncertain policy P^{μ} with an unknown distribution of consequences. Consider the following variation on Example 1:

Example 2 **(Treatment with Status Quo)** A new medical treatment is proposed to treat an illness.

- With probability β the treatment succeeds and everyone receiving this treatment recover completely, but with probability $1 - \beta$ it fails and everyone receiving this treatment suffers significant complications and diminished state of health. Normalize the payoffs of these two outcomes to 1 and 0 per patient, respectively.
- The status quo is a constant profile $\bar{x}$ with utility $v = v(x) \in (0, 1)$ per patient.[19]

For $\beta > v$, every individual will strictly prefer to use the new treatment, as would a planner with separable utility.

In the example, assigning all individuals to the new treatment exposes the planner to considerable aggregate uncertainty, yet this uncertainty is irrelevant under additive separability. If the planner is averse to aggregate risk (u strictly concave), he will prefer to use a mixture of the two treatments. One way to mix the two treatments is to assign only a subset of individuals to the new treatment. We define this more generally:

Definition 5 Consider a social lottery P^{μ} and a status quo $\bar{x} = (x, \ldots, x)$.

(a) A *treatment assignment* is a vector $a = (a_1, \ldots, a_n) \in \{0, 1\}^n$;

[17] When the population is small (e.g., $n = 2$) and u is non-linear, there will typically be "aggregate uncertainty" even if P is independent.

[18] Especially, Manski (2004) and Manski and Tetenov (2007). For a survey, see Manski (2011).

[19] An important restriction, discussed below, is that the status quo is deterministic.

(b) The *fractional treatment* (relative to an assignment a) is the social lottery

$$\bar{x} \oplus_a P^{\mu}$$

defined by:

- $s_i = x$ with probability whenever $a_i = 0$; and
- $\bar{x} \oplus_a P^{\mu}$ and P^{μ} have the same marginal distribution on $\prod_{\{i:a_i=1\}} X_i$.

We define fractional assignment only with respect to a deterministic status quo. Fractional assignment of two general social lotteries P^{μ_1} and P^{μ_2} is more involved since one must take into account the correlation between the two treatments. Note also that the assignment $a = (a_1, \ldots, a_n)$ is deterministic; Sect. 5.3 below considers random assignments.

A planner with a strictly concave u will strictly prefer to hedge against the aggregate uncertainty by assigning some individuals to an ex ante inferior status quo. The next example illustrates the point:

Example 3 **(Fractional Assignments)** Consider the setting of Example 2 with $\beta \geq v$. Assume that the planner has a differentiable and strictly concave u, with $\lim_{r \downarrow 0} u'(r) = \infty$ and $\lim_{r \uparrow 1} u'(r) = 0$.[20]

Let δ denote the fraction of the population the planner assigns to the new treatment. Then his utility is

- $u\,(\delta \times 1 + (1-\delta)v)$ with probability β;
- $u\,(\delta \times 0 + (1-\delta)v)$ with probability $1-\beta$.

Differentiating with respect to δ and setting the derivative equal to 0, the optimal fraction $\bar{\delta}$ must satisfy

$$\beta\,(1-v)u'\left(\bar{\delta} + (1-\bar{\delta})v\right) = v(1-\beta)\,u'\left((1-\bar{\delta})v\right).$$

It is clear that this cannot be satisfied for δ close to 1.

In the example, assigning a positive fraction of the population to the ex ante inferior status quo can strictly increase utility because a population mixture can be valuable for hedging purposes. Manski (2004; 2011) calls such rules *fractional treatment rules*. He points out that fractional treatments are inconsistent with a Bayesian decision criterion and an additively separable social welfare function, but shows that they can be derived from a minimax regret criterion. The above example points out that fractional treatments can arise in a Bayesian setting with aggregative utility.

It is useful to record the general point illustrated by the example for future reference:

Corollary 1 *There exists a social lottery P^{μ} and a deterministic status quo $\bar{x}$ such that*

$$\mathcal{U}(\bar{x} \oplus_a P^{\mu}) > \mathcal{U}(\bar{x}) \geq \mathcal{U}(P^{\mu}) \tag{13}$$

for some fractional assignment a.

[20] This assumption simplifies the example, but may be inconsistent with the new treatment having a well-defined utility. Note that we do not need to calculate the expected utility of the new treatment in this example.

The conditions under which mixing with a strictly inferior treatment might help are difficult to investigate in the finite population setting because the assessment of utility can confound aggregate and idiosyncratic risks.[21] We investigate these issues more thoroughly in our companion paper, Al-Najjar and Pomatto (2014), where the population is infinite and an exact law of large numbers holds.

Manski and Tetenov (2007) consider a social planner who evaluates treatments based on the expectation of a concave increasing transformation f of the treatments success rate. They interpret the concavity of f as the planners risk aversion. Using our results, it can be shown that this functional form corresponds to an aggregative utility (3) when the set of outcomes is binary, treatments give rise to conditionally i.i.d. social lotteries, the function u is concave, and the size of the population n increases to infinity. In addition to providing foundations for the social welfare functional in Manski and Tetenov (2007), we also address the finite n case. As noted earlier, a finite population is more appropriate in applications such as financial regulations of investment banks, drug approval policy for pharmaceutical companies, among others.

5.3 Preference for randomization

Consider a planner who finds it optimal to assign a positive fraction of individuals to an inferior treatment for hedging purposes, in the sense of Corollary 1. Since a planner with aggregative utility is only concerned about the average of the individuals' utilities, in a symmetric environment, such a planner is indifferent to all assignment profiles that keep the fractions allocated to each treatment fixed.

Individuals assigned to the ex ante inferior treatment may feel justified in viewing such an assignment as "unfair." A completely satisfactory solution to this problem seems difficult: hedging against aggregate uncertainty requires differential treatment of ex ante identical individuals, and some will have to be assigned to an inferior treatment.

While fairness considerations do not receive a formal support in our setting, it is still natural to impose them indirectly in the form of a constraint set on what the planner can do. For example, suppose we impose the requirement that the planner can only choose policies that induce symmetric social lotteries. Formally, we require the social lottery to have the property that $p_i = p_j$ for all i, j. In a symmetric environment, this ensures ex ante equity, in the sense that each individual gets the same expected payoff as any other.

Ex ante equity can be implemented by assigning treatments at random. To formalize this observation, fix n and consider any number $\delta = \frac{j}{n}$, $j \in \{0, \ldots, n\}$. Let A_δ denote the uniform distribution on the (finite) set of assignments where a fraction δ receives treatment 1 and a fraction $1 - \delta$ receives treatment 0.[22] For a constant profile $\bar{x}$, define

$$\bar{x} \oplus_{A_\delta} P^\mu$$

as the social lottery $\bar{x} \oplus_a P^\mu$ where the assignment profile is drawn at random uniformly.

[21] This issue was artificially eliminated in Examples 2 and 3 by assuming away idiosyncratic risk.

[22] That is, A_δ is the uniform distribution on the finite set $\left\{a : \sum_i a_i = n\delta\right\}$.

Corollary 2 *There exist a social lottery P^μ and a deterministic status quo $\bar{x}$ such that*

$$\mathcal{U}(\bar{x} \oplus_{A_\delta} P^\mu) > \mathcal{U}(\bar{x}) \geq \mathcal{U}(P^\mu). \tag{14}$$

Randomized rules play an important role in many contexts. In addition to Manski's work discussed earlier, randomization also underlies the fundamental insight of Schmeidler (1989) and Gilboa and Schmeidler (1989) that individuals may strictly prefer to randomize as a way to hedge against unknown probabilities. Our approach is closely related to Halevy and Feltkamp (2005) who show that randomization can be strictly preferred when the decision maker's payoff is the sum of two draws from two conditionally i.i.d. urns. The large population context makes it clear that randomization yields higher utility not because it is intrinsically desirable, but because (a) deterministic δ-fractional rules hedge against aggregate uncertainty; and (b) uniform randomization using A_δ always picks one such fractional rule.

It is clear from (14) that aggregative utility fails the Pareto criterion. This is because a planner with an aggregative utility is concerned about the joint distribution induced by a policy, while individuals only care about the marginals. From Theorem 5, the difference is immaterial if the risk associated with the policy is idiosyncratic, but can be significant when policy risk is correlated. For example, consider a central bank regulating a financial sector. Private banks are concerned about their own profits and financial positions, not about systemic risk that impacts the entire financial sector. In this case, weakening the Pareto criterion seems reasonable. A better understanding of the sensitivity of welfare criteria to aggregate uncertainty is an important challenge for future work.

5.4 Hedging via population heterogeneity

The strict preference for randomization discussed earlier has an intriguing connection to bet-hedging strategies in evolutionary biology. Take the point of view of Nature as a principal who designs individuals' utilities.[23] If Nature is anthropomorphized as a planner with an aggregative utility, then the same logic that led to a strict preference for mixed treatments implies that evolution would favor heterogeneity in behavior to hedge against environmental uncertainty. One way for Nature to induce heterogenous behavior is to introduce heterogeneity in preferences. The injection of heterogeneity in preferences as a bet-hedging strategy is well-known in the evolutionary biology literature; an early paper is Cooper and Kaplan (1982). Robson and Samuelson (2010)'s survey provides further discussion and recent references.

Bergstrom (1997) provides a nice example: a population (of squirrels) infrequently faces winters of varying intensity. Most winters are mild and require storing only the normal reserves of food. Harsh winters are infrequent but can wipe out all members except for those who accumulated large reserves. By assumption, accumulating large reserves is an inferior strategy most of the time, except in those infrequent episodes of harsh winters. Similar to the randomized treatment discussed above, where a frac-

[23] See Robson and Samuelson (2010)'s survey for discussion of the principal-agent approach.

tion of the population is assigned to the ex ante inferior treatment, in Bergstrom's model, Nature designs random utilities where a subpopulation has a preference for accumulating large reserves. This subpopulation is Nature's insurance policy against the unpredictability of winters.

6 Concluding remarks

Aggregate risks appear in greater frequency and significance in contexts ranging from epidemics, to product recalls and systemic financial risk. A fundamental weakness of standard additive decision criteria is their inability to capture sensitivity to correlation. This paper provided a theoretical framework to isolate and measure the impact of these risks. We use two classical cardinal measurements of utility, one reflecting aggregation and the other risk, to provide a simple theory that can be adapted to a wide range of problems. We view the simplicity of the theory as an advantage, making it all the more surprising that this approach has not been taken before.

7 Proofs

Proof of Theorem 1 Sufficiency: Given Assumptions 3, 4, and the continuity of U in Assumption 2, Fishburn (1970, Theorem 5.5, p. 71) implies that the utility U restricted to S can be represented by the function:

$$V(s) \equiv \sum_{i=1}^{n} v_i(s_i)$$

for some continuous, and non-constant functions $v_i : X_i \to \mathbb{R}$, $i = 1, \ldots, n$, and that these functions are unique up to a common positive affine transformation.

The functions V and U represent the same preference on S, so $V(s) \geq V(t)$ if and only if $U(s) \geq U(t)$. Therefore, there exists a strictly increasing function $u : V(S) \to \mathbb{R}$ such that $U(s) = u\left(\sum_{i=1}^{n} v_i(s_i)\right)$ for every $s \in S$.

The continuity of u follows from the continuity of the functions $\sum_{i=1}^{n} v_i$ and U and the connectedness of S. See Wakker (1991, Lemma 2.1, p. 1). Finally, because the function U is cardinally unique, u is also cardinally unique given $v_1, \ldots, v_n$.

Necessity: the preference obviously satisfies expected utility with $U(s) = u\left(\sum_{i=1}^{n} v_i(x_i)\right)$. Since the v_i's and u are continuous, so is U. Non-triviality follows from the fact that the v_i's are not constant. To prove separability, we first note that $\mathcal{U}$ restricted to profiles is ordinally equivalent to the function V above. Given two profiles s and t,

$$\begin{aligned}\mathcal{U}(s) \geq \mathcal{U}(t) &\iff u\left(\sum_{i=1}^{n} v_i(s_i)\right) \geq u\left(\sum_{i=1}^{n} v_i(t_i)\right) \\ &\iff \sum_{i=1}^{n} v_i(s_i) \geq \sum_{i=1}^{n} v_i(t_i),\end{aligned}$$

where the last equality follows from the fact that u is strictly increasing. The preference this function represents is obviously separable. □

Proof of Theorem 2 Suppose that the agent's preference has an additively separable representation $\mathcal{U}$ and that P, Q are two lotteries with identical marginals. Then

$$\begin{aligned}\mathcal{U}(P) &= \int \left(\sum_{i=1}^{n} v_i(s_i)\right) \mathrm{d}P(s)\\ &= \sum_{i=1}^{n}\left(\int_S v_i(s_i)\mathrm{d}P(s)\right)\\ &= \sum_{i=1}^{n} E_{p_i} v_i\\ &= \sum_{i=1}^{n} E_{q_i} v_i .\\ &= \mathcal{U}(Q). \end{aligned} \tag{15}$$

The key step is the equality (15) where the sum over the population and the expectation over the uncertainty are interchanged. This is possible when u is affine, but is generally not true otherwise.

In the other direction, assume that $P \sim Q$ whenever P and Q have identical marginals. Fishburn (1970, Theorem 11.1, p. 149) shows that $\mathcal{U}$ must have an additive representation.[24] □

Proof of Theorem 3 Apply McDiarmid's concentration inequality (McDiarmid (1998, Theorem 3.1)) to the function: $V(s) = \frac{1}{n}\sum_{i=1}^{n} v_i(s_i)$ to obtain:

$$P\left\{s : |V(s) - E_P V| < \frac{\epsilon}{K}\right\} > 1 - 2e^{-2n\left(\frac{\epsilon}{K}\right)^2}$$

for every independent distribution P.

The first claim of the theorem is now proved by applying Lipschitz continuity and noting that because $\bar{\mathcal{U}}_n(P) = u(E_P V)$, we have that

$$|V(s) - E_P V| < \frac{\epsilon}{K} \implies \left| u\left(\frac{1}{n}\sum_{i=1}^{n} v_i(s_i)\right) - \bar{\mathcal{U}}(P)\right| < \epsilon.$$

The second conclusion follows from the fact that the distance between $u\left(\frac{1}{n}\sum_{i=1}^{n} v_i\right)$ and $\bar{\mathcal{U}}_n(P)$ is bounded by 1 and is less than ϵ with probability at least $1 - 2\mathrm{e}^{-2n\left(\frac{\epsilon}{K}\right)^2}$. □

24 Fishburn's theorem requires only that indifference obtains for lotteries that are either degenerate or assigns equal weights to the two deterministic profiles.

Proof of Theorem 4 For every θ, Theorem 3 implies

$$\left|\mathcal{U}(P^\theta) - \bar{\mathcal{U}}(P^\theta)\right| \leq \epsilon + 2\mathrm{e}^{-2n\left(\frac{\epsilon}{K}\right)^2}$$

By definition, $\mathcal{U}(P^\mu) = \int \mathcal{U}(P^\theta)\mathrm{d}\mu(\theta)$. Hence

$$\begin{aligned}\left|\mathcal{U}(P^\mu) - \int \bar{\mathcal{U}}(P^\theta)\mathrm{d}\mu(\theta)\right| &= \left|\int \left(\mathcal{U}(P^\theta) - \bar{\mathcal{U}}(P^\theta)\right)\mathrm{d}\mu(\theta)\right| \\ &\leq \int \left|\mathcal{U}(P^\theta) - \bar{\mathcal{U}}(P^\theta)\right|\mathrm{d}\mu(\theta) \\ &\leq \epsilon + 2\mathrm{e}^{-2n\left(\frac{\epsilon}{K}\right)^2}.\end{aligned}$$

□

Proof of Theorem 5 Since u and v are bounded, it is without loss of generality to assume that their range is included in [0, 1]. For each lottery P^μ, Theorem 4 implies that $\mathcal{U}_n(P^\mu)$ converges to $\int u(E_p^\theta v)\mathrm{d}\mu(\theta)$.

Therefore, given any two lotteries P^μ and P^ν, we have that $\mathcal{U}_n(P^\mu) \geq \mathcal{U}_n(P^\nu)$ for each n large enough if and only if

$$\int u(E_{p^\theta} v)\mathrm{d}\mu(\theta) \geq \int u(E_{p^\theta} v)\mathrm{d}\nu(\theta).$$

Because u is strictly increasing, this means that when μ and ν are degenerate (*i.e.,* the lotteries are i.i.d.), the choice between the two lotteries can be decentralized.

We now prove that if the choice between any two social lotteries can be decentralized, then u must be affine. Let ξ_0 and ξ_1 be two points in $v(X)$, and let θ_0 and θ_1 satisfy $E_{p^{\theta_0}} v = \xi_0$ and $E_{p^{\theta_1}} v = \xi_1$. Now fix $\alpha \in [0, 1]$ and let μ satisfy $\mu(\theta_0) = \alpha$ and $\mu(\theta_1) = 1 - \alpha$. Finally, define θ_α to be the i.i.d. social lottery with marginal equal to p^μ. By construction, we have

$$E_{p^{\theta_\alpha}} v = E_{p^\mu} v$$

If the choice between P^{θ_α} and P^μ can be decentralized, then it must be that

$$u(E_{p^{\theta_\alpha}} v) = \alpha u(E_{p^{\theta_0}} v) + (1 - \alpha)u(E_{p^{\theta_1}} v).$$

That is,

$$u(\alpha\xi_0 + (1 - \alpha)\xi_1) = \alpha u(\xi_0) + (1 - \alpha)u(\xi_1)$$

which proves that u is affine. □

Acknowledgments We thank Robert Gary-Bobo, Yoram Halevy, Chuck Manski, Mallesh Pai, and Phil Reny for their comments. We also thank Hasat Cakkalkurt for her research assistance.

References

Acharya, V. V., Pedersen, L. H., Philippon, T., & Richardson, M. P. (2010). *Measuring systemic risk*. New York: NYU Stern School of Business.

Adrian, T., & Brunnermeier, M. K. (2011). *CoVaR*. Cambridge: National Bureau of Economic Research.

Al-Najjar, N. I. (2015). A Bayesian framework for the precautionary principle. *Journal of Legal Studies*, Forthcoming.

Al-Najjar, N. I., & Pomatto, L. (2014). *Aggregative utility in large populations*. Evanston: Northwestern University.

Ashford, N. et al. (1998). Wingspread statement on the precautionary principle. Retrieved from http://www.gdrc.org/u-gov/precaution-3.html.

Barberis, N., & Huang, M. (2001). Mental accounting, loss aversion, and individual stock returns. *The Journal of Finance*, *56*(4), 1247–1292.

Bergstrom, T. C. (1997). *Storage for good times and bad: Of squirrels and men*. Santa Barbara: University of California, Santa Barbara.

Cooper, W. S., & Kaplan, R. H. (1982). Adaptive "coin-flipping": A decision-theoretic examination of natural selection for random individual variation. *Journal of Theoretical Biology*, *94*(1), 135–151.

Debreu, G. (1960). Topological methods in cardinal utility theory. In K. Arrow, G. Karlin, & P. Suppes (Eds.), *Mathematical methods in the social sciences* (pp. 16–26). Stanford: Stanford University Press.

Dixit, A. K., & Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity. *The American Economic Review*, *67*(3), 297–308.

Fishburn, P. (1970). *Utility theory for decision making*. New York: Wiley.

Gilboa, I., & Schmeidler, D. (1989). Maxmin expected utility with nonunique prior. *Journal of Mathematical Economics*, *18*(2), 141–153.

Gorman, W. M. (1968). The structure of utility functions. *The Review of Economic Studies*, *35*, 367–390.

Halevy, Y., & Feltkamp, V. (2005). A Bayesian approach to uncertainty aversion. *Review of Economic Studies*, *72*(2), 449–466.

Manski, C. F. (2004). Statistical treatment rules for heterogeneous populations. *Econometrica*, *72*(4), 1221–1246.

Manski, C. F. (2011). Choosing treatment policies under ambiguity. *Annual Review of Economics*, *3*(1), 25–49.

Manski, C. F., & Tetenov, A. (2007). Admissible treatment rules for a risk-averse planner with experimental data on an innovation. *Journal of Statistical Planning and Inference*, *137*, 1998–2010.

McDiarmid, C. (1998). *Concentration*. Oxford: University of Oxford.

Robson, A. (1996). A biological basis for expected and non-expected utility. *Journal of Economic Theory*, *68*, 397–424.

Robson, A., & Samuelson, L. (2010). The evolutionary foundations of preferences. In J. Benhabib, A. Bisin, & M. Jackson (Eds.), *The social economics handbook*. Amsterdam: Elsevier Press.

Savage, L. J. (1954). *The foundations of statistics*. New York: Wiley.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, *57*(3), 571–587.

Sunstein, C. (2005). *Laws of fear: Beyond the precautionary principle*. Cambridge: Cambridge University Press.

Wakker, P. (1991). Continuity of transformations. *Journal of Mathematical Analysis and Applications*, *162*(1), 1–6.