Marc Fleurbaey

Assessing risky social situations

Working paper

Assessing risky social situations*

Marc Fleurbaey†

August 2009

Abstract

This paper re-examines the welfare economics of risk. It singles out a class of criteria, the "expected equally-distributed equivalent", as the unique class which avoids serious drawbacks of existing approaches. Such criteria behave like ex-post criteria when the final statistical distribution of well-being is known ex ante, and like ex-ante criteria when risk generates no inequality. The paper also provides a new result on the tension between inequality aversion and respect of individual ex ante preferences, in the vein of Harsanyi's aggregation theorem.

Keywords: risk, social welfare, ex ante, ex post, fairness, Harsanyi theorem.

JEL Classification: D63, D71, D81.

1 Introduction

Welfare economics has not yet reached a consensus about how to assess social situations involving risk. There are three competing approaches. The utilitarian approach maximizes the sum of individual expected utilities, which is also, identically, the expected value of the sum of utilities. Such identity is singled out in Harsanyi's (1955) aggregation theorem, which shows that any criterion that is based on expected social utility (corresponding to rationality of the evaluation) and on individual expected utilities

*This paper has benefited from comments on a preliminary draft by G. Arrhenius, R. Bradley, J. Broome, A. Casella, J. Drèze, T. Gajdos, N. Gravel, P. Mongin, A. Trannoy, G. Valletta, J. Weymark, and from discussions with P. Hammond, A. Harel, G. Laroque, J. Leach, F. Maniquet, S. Orr, M. Otsuka, J. Roemer, A. Voorhoeve, A. Williams, as well as from reactions from audiences at the LSE, Carlos III (Madrid), GREQAM (Marseilles), PSE (Paris), Osnabrück, NHH (Bergen) and Hitotsubashi (Tokyo).

†CNRS and University Paris Descartes, CORE (University Louvain-la-Neuve) and IDEP. Email: marc.fleurbaey@parisdescartes.fr.

(i.e., obeying the Pareto principle) must be affine in individual utilities.[1] The most disturbing feature of utilitarianism is its indifference to the distribution of utilities, and many analysts consider this to be a fatal flaw. But if one wants to introduce preference for equality in the criterion, Harsanyi's theorem forces one to make a choice between rationality and Pareto—which is why there are two main contenders.

A first option is "ex-ante egalitarianism" (advocated by Diamond 1967, Epstein and Segal 1992, among others), which consists in retaining the Pareto principle and applying inequality aversion to the distribution of individual expected utilities, while dropping the assumption of expected social utility. This has implications in terms of rationality that appear more problematic than a mere violation of expected utility. Consider a situation in which an impending climate change will alter the distribution of well-being on Earth. Suppose that only two scenarios are considered possible. In one scenario, the extreme latitudes gain and the low latitudes suffer, whereas the reverse occurs in the other scenario. Suppose that in either scenario the distribution of well-being is ultimately much worse than in absence of climate change. Therefore one is sure that such climate change is harmful. However, if individual expected utilities are not diminished ex ante, because everyone may gain or lose depending on which scenario is realized, ex-ante egalitarianism considers that climate change is harmless. This is strange as the *same* criterion considers that the change will ultimately be catastrophic.[2]

The other option is "ex-post egalitarianism" (advocated by Adler and Sanchirico 2006),[3] which uses an inequality averse social welfare function and evaluates prospects in terms of the expected value of social welfare. Obviously, in view of Harsanyi's theorem this approach must relax the Pareto principle, and this is generally defended on the ground that individual expected utilities fail to record the ex-post inequalities that risk-taking may entail. However, if one takes a standard social welfare function $\sum_{i=1}^{n} \varphi(u_i)$, where φ is a concave function, the violation of Pareto is extreme and hard to justify, because individual ex-ante preferences will not be respected even when no inequality arises in any of the possible consequences.

[1] Hammond (1983) and Broome (1991) have offered additional arguments in favor of this criterion, and made a careful scrutiny of the scope of this result. See also Weymark (1991, 2005) for a broad examination of Harsanyi's arguments in favor of utilitarianism.

[2] Ex-ante egalitarianism has also been criticized for implying time inconsistency: wouldn't it be better to always toss a new coin before giving a prize, as this would restore equality of expected utilities (Myerson 1981, Broome 1984, 1991, Hammond 1988, 1989)? However, Epstein and Segal (1992) argue that ex-ante egalitarianism can be made dynamically consistent by letting decisions depend on the past and compute expected utilities at a relevant time only. Once a coin has been tossed once, there is then no need to do it again.

[3] It is also favorably considered and applied in Meyer and Mookherjee (1987), Harel et al. (2005), Otsuka and Voorhoeve (2007), Roemer (2007).

Indeed, in absence of inequalities this approach maximizes the expected value of $\varphi(u_i)$ instead of u_i, implying a stronger risk aversion than warranted by individual preferences. Can a concern for ex-post inequalities justify this kind of paternalism *in absence of* inequalities?

A variant of this option must be mentioned. Proposed by Broome (1991) and Hammond (1983), it consists in adopting utilitarianism as a formal criterion but with a measure of individual utility which incorporates the evaluator's concern for the distribution. This approach is formally elegant, attractive, and promises to combine inequality aversion and rationality. But how exactly to incorporate inequality aversion remains to be worked out and some simple ways of doing it make the approach boil down to ex-post egalitarianism. In this paper, we will retain the standard assumption that individual utilities only refer to personal situations and that the Pareto principle refers only to personal interests. If this modified utilitarian approach were the solution, it would emerge from our analysis, together with a specific way of incorporating distributional concerns. As it turns out, this is not the case.[4]

The purpose of this paper is to explore the possibility of a coherent approach that would avoid the three severe drawbacks which have just been pointed out for each of the three main contenders: 1) the utilitarian indifference to the distribution of utility; 2) the irrationality displayed by ex-ante egalitarianism when an option is declared good whereas each of its possible realizations is declared bad; 3) the violation of individual preferences in absence of inequalities as observed for ex-post egalitarianism. A particular class of criteria will be singled out, which behave like ex-ante criteria when risk does not generate inequalities and like ex-post criteria otherwise. Such criteria do compute the expected value of a social welfare function, like ex-post criteria, but the function that is used must take the form of the "equally distributed equivalent" (Kolm 1968, Atkinson 1970), a classical example being the function $\varphi^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\varphi(u_i)\right)$.

Having identified the general form of a criterion that jointly satisfies the three requirements, we then examine how much inequality aversion and how much respect of individual expected utilities is possible in this approach. We will in particular highlight a special family of criteria which satisfies a relatively strong version of the Pareto principle and encompasses all possible degrees of inequality aversion. But we will also uncover that the new approach has a practical drawback, compared to the other approaches, because it drastically reduces the possibility to make separate evaluations for subpopulations. A critical

[4] Yet another approach, proposed by Ben Porath et al. (1997) and Gajdos and Maurin (2004), consists in taking a convex combination of an ex-ante criterion and an ex-post criterion. The problem with this approach is that it adds up the drawbacks of the two combined criteria.

example of this issue concerns the present and future generations, whose choices will not affect the past generations and for which, as pointed out in Blackorby et al. (2005), it would be cumbersome to be forced to take account of the "utility of the dead" in their decisions. Exploring this problem, a strong form of Harsanyi's utilitarian theorem will be obtained. Finally, at the end of the paper the various approaches are put to work in economic problems of allocation under risk. The new criteria will be shown to behave reasonably in problems of insurance and value of life.

In summary, the paper is structured as follows. In Section 2 the model and the main assumptions are introduced. In Section 3 the key requirements are formally expressed and the main result is derived. Section 4 focuses on a relatively strong version of the Pareto principle and identifies the subfamily which obeys it. Section 5 addresses the challenge of the separability of subpopulations in the evaluation. Section 6 examines three short applications to the welfare economics of risk. Section 7 concludes. Appendix A contains some proofs and checks the tightness of the results. Appendix B explains how the rationality assumptions used in this paper connect to decision theory.

2 The framework

The framework is as simple as possible for this analysis. The population is finite and fixed, $N = \{1, ..., n\}$. The set of states of the world is finite, $S = \{1, ..., m\}$, and the evaluator has a fixed probability vector $\pi = (\pi^s)_{s\in S}$, with $\sum_{s\in S} \pi^s = 1$. This probability vector corresponds to the evaluator's best estimate of the likelihood of the various states of the world.[5] Since what happens in zero-probability states can be disregarded, we simply assume that $\pi^s > 0$ for all $s \in S$. Vector inequalities are denoted $\geq, >, \gg$.

The evaluator's problem is to rank prospects $U = (U_i^s)_{i\in N, s\in S} \in \mathbb{R}^{nm}$, where U_i^s describes the utility attained by individual i in state s. Let $\mathcal{L} \subseteq \mathbb{R}^{nm}$ denote the relevant set of such prospects over which the evaluation must be made. The social ordering (i.e., a complete, transitive binary relation) over the set $\mathcal{L}$ is denoted R (with strict preference P and indifference I).

Let U_i denote $(U_i^s)_{s\in S}$ and U^s denote $(U_i^s)_{i\in N}$. Let $[U^s]$ denote the riskless prospect in which vector U^s occurs in all states of the world. Two subsets of $\mathcal{L}$ are worth distinguishing. $\mathcal{L}^c$ will denote the subset of riskless prospects (i.e., $U^s = U^t$ for all $s, t \in S$); $\mathcal{L}^e$ will denote the subset of egalitarian prospects (i.e.,

[5]There is no problem of aggregation of beliefs here, since the evaluator uses his own probability vector in order to compute expected values. The evaluator's beliefs may have been influenced by the population beliefs, but we work here with his fully updated probability vector.

$U_i = U_j$ for all $i, j \in N$).

The utility numbers U_i^s are assumed to be fully measurable and interpersonally comparable, and they may measure any subjective or objective notion of advantage that the evaluator considers relevant for social evaluation. In particular, it is not necessary for the analysis that they derive from individual Bernoulli utility functions. However, it is assumed that, for one-person evaluations, the evaluator considers that the expected value $EU_i = \sum_{s\in S} \pi^s U_i^s$ correctly measures agent i's ex-ante interests.

A key assumption here is that for every $s \in S$, the vector U^s fully describes the relevant features of the final situation occurring in state s. This assumption is of paramount importance because it implies two things. First, the vector U^s containing all the relevant information, the social preferences over final situations need not and must not be state dependent. If U^s is deemed better than U'^s, this must be true independently of the state that is realized, otherwise the state itself would be a piece of information missing from utilities. This means that U^s is deemed better than U'^s in state s if and only if $[U^s]\ R\ [U'^s]$. In other words, there is no need to introduce preferences over final consequences in the model as they are equivalent to the social ordering R restricted to riskless prospects.

The second implication of the assumption is that if part of the relevant information includes how the final situation came about or what would have occurred in counterfactual states of the world, such information has been incorporated in U^s. This means that Diamond's (1967) criticisms against certain aspects of Harsanyi's utilitarianism do not apply. Consider his classical example of a two-individual, two-state problem in which one must decide to give a prize to one individual or to flip a coin. If the utilities are as in the following tables (with individuals named Ann and Bob and equiprobable states named Heads and Tails), Diamond argues that U' is better than U even though the utilitarian criterion is indifferent. Note that the utilitarian criterion is not the only one that is indifferent here, as the distribution of utility that U' yields is (4,0) in every state, just as for U. Any ex-post egalitarian criterion will also be indifferent.

U	Heads	Tails
Ann	4	4
Bob	0	0

U'	Heads	Tails
Ann	4	0
Bob	0	4

If it is preferable to give fair chances to the individuals, however, this can only mean that (4,0) *after a fair lottery* is preferable to (4,0) *in absence of a lottery.* The utility figures do not then contain all the relevant information. If they contained all the relevant information, then Diamond's criticism would not apply. Broome (1984) defends the view that it is possible to incorporate all fairness considerations into utilities.[6] The readers who do not accept this view may nonetheless take interest in this paper if they consider that all prospects in the analysis are equally good in terms of fairness. For instance, if U depicts a prospect after Ann has won the lottery (and Heads and Tails refer to flipping the coin again with no change in the prize assignment) while U' depicts a prospect before the lottery is drawn, it makes a lot of sense to be indifferent between them.

3 A general solution

We now introduce the three main requirements that R should satisfy. Utilitarianism displays no aversion to inequality, and we will be interested in the possibility to introduce inequality aversion in the social ordering. But there is no need to formalize a requirement to that effect in this section.

In contrast, it is important to formalize the two other requirements mentioned in the introduction. First, unlike ex-ante egalitarianism, the criterion must not at the same time say that a prospect is at least as good as another and that it produces a worst situation for sure, or that a prospect is better than another and that it never produces a better situation. This can be written in the form of two dominance conditions.

Axiom 1 (Weak Dominance) *For all* $U, U' \in \mathcal{L}$*, one has* $U\,R\,U'$ *if for all* $s \in S$, $[U^s]\ R\ [U'^s]$.

Axiom 2 (Strict Dominance) *For all* $U, U' \in \mathcal{L}$*, one has* $U\,P\,U'$ *if for all* $s \in S$, $[U^s]\ P\ [U'^s]$.

Such dominance conditions are quite familiar in decision theory.[7] To see how compelling they are, imagine an omniscient being who has the same preferences over final consequences as the decision-maker and knows everything about the state of the world. If the decision-maker could be told which act the omniscient being would like him to choose, he should obviously follow suit.[8] Now, in situations of

[6] In the introduction we rejected the idea that individual utility should incorporate distributional concerns because an explicit solution of the problem of incorporating inequality aversion in the criteria is preferable. The idea of incorporating lottery fairness into utility is different because it refers to "having had the chance of winning", a purely personal data. (As suggested in Deschamps and Gevers (1978), expected utility is also a part of final utility when uncertainty is resolved late in life.) Certainly, one can argue that an explicit solution to the problem of incorporating lottery fairness in the analysis would also be preferable. This is a project for future research, as mentioned in the conclusion.

[7] More on how dominance relates to decision theory, and in particular to Savage's sure-thing principle, can be found in Appendix B.

[8] A very basic principle in decision theory is that, if the set of feasible actions is not affected by the acquisition of information, it is always good to acquire more information. See Savage (1972), Marschak (1954). Drèze (1987) provides

dominance between two acts, the decision-maker can actually guess the omniscient being's preference about these two acts (because it is the same, whatever the true state of the world). Following these preferences is equivalent to respecting dominance.

Note that it would be much less compelling to require $U\,P\,U'$ when for all $s \in S$, $[U^s]\ R\ [U'^s]$ and $[U^s]\ P\ [U'^s]$ for some s only. Indeed, in such a case it may be that for the state s that is eventually realized one has $[U^s]\ I\ [U'^s]$. The dominance conditions adopted here only deal with cases in which one is absolutely sure of the final preference.

Second, unlike ex-post egalitarianism based on a concave social welfare function such as $\sum_{i=1}^{n} \varphi(u_i)$, the criterion must respect individual expected utility in cases in which no inequalities arise. If whatever the state of the world, perfect equality is obtained, the population is similar to just one person taking a risk, and it appears compelling to follow the unanimous expected utility of the population in this special case.

Axiom 3 (Weak Pareto for Equal Risk) *For all $U, U' \in \mathcal{L}^e$, one has $U\,P\,U'$ if for all $i \in N$, $EU_i > EU'_i$.*

This principle, however, sometimes meets skepticism on the ground that when the whole population takes a risk, bad luck translates into a widespread catastrophe, which should be avoided at all cost. This idea of "catastrophe avoidance", however, may appear attractive only when one's intuition is driven by the worry that a catastrophe for a particular generation puts the existence of the next generations at risk.[9] No such externality can occur in the framework of this paper, because the population under consideration is supposed to include all the relevant individuals, including future generations if needed.

The Pareto principle is also compelling for non-risky situations. This is satisfied by all approaches under consideration in this paper. It is useful to write this requirement formally.

Axiom 4 (Weak Pareto for No Risk) *For all $U, U' \in \mathcal{L}^c$, one has $U\,P\,U'$ if for all $i \in N$, $U_i > U'_i$.*

A valuable implication of Pareto for No Risk is the existence of an "equally-distributed quasi-equivalent" (EDQE) function e, as defined in the following lemma.

many examples in which the acquisition of information reduces the set of feasible actions, inducing a negative value of information. Such examples do not apply to our "evaluator", and more generally do not question the dominance principle for the ranking of options.

[9] Keeney (1980) also emphasizes that catastrophe avoidance implies concentrating risk on some individuals and thereby goes against the most basic idea of equity in the allocation of risk. See also Fishburn (1984). The approach developed here concurs with Keeney and opposes the concentration of risk, as shown in Section 6.

Lemma 1 *Let $\mathcal{L} = \mathbb{R}^{nm}$ and let R satisfy Weak Pareto for No Risk. For all $U \in \mathcal{L}$, all $s \in S$, there is a unique $e(U^s) \in \mathbb{R}$ such that for all $\varepsilon \in \mathbb{R}_{++}$,*

$$[U^s] \ P \ [(e(U^s) - \varepsilon, ..., e(U^s) - \varepsilon)],$$

$$[(e(U^s) + \varepsilon, ..., e(U^s) + \varepsilon)] \ \ P \ \ [U^s].$$

For all $x \in \mathbb{R}$, $e(x, ..., x) = x$.

Proof. Take $U^s \in \mathbb{R}^n$. By Weak Pareto for No Risk, the sets $\{x \in \mathbb{R} \mid [U^s] \ P \ [(x, ..., x)]\}$ and $\{x \in \mathbb{R} \mid [(x, ..., x)] \ P \ [U^s]\}$ are not empty. Moreover, necessarily

$$\sup \{x \in \mathbb{R} \mid [U^s] \ P \ [(x, ..., x)]\} \leq \inf \{x \in \mathbb{R} \mid [(x, ..., x)] \ P \ [U^s]\}.$$

Suppose that

$$\sup \{x \in \mathbb{R} \mid [U^s] \ P \ [(x, ..., x)]\} < \inf \{x \in \mathbb{R} \mid [(x, ..., x)] \ P \ [U^s]\}.$$

Then let $z, z' \in \mathbb{R}$ be such that

$$\sup \{x \in \mathbb{R} \mid [U^s] \ P \ [(x, ..., x)]\} < z < z' < \inf \{x \in \mathbb{R} \mid [(x, ..., x)] \ P \ [U^s]\}.$$

One cannot have $[U^s] \ P \ [(z, ..., z)]$ or $[(z, ..., z)] \ P \ [U^s]$, so that $[U^s] \ I \ [(z, ..., z)]$. Similarly, $[U^s] \ I \ [(z', ..., z')]$. By transitivity, $[(z, ..., z)] \ I \ [(z', ..., z')]$, which contradicts Weak Pareto for No Risk. As a consequence, one must have

$$\sup \{x \in \mathbb{R} \mid [U^s] \ P \ [(x, ..., x)]\} = \inf \{x \in \mathbb{R} \mid [(x, ..., x)] \ P \ [U^s]\}.$$

This defines a function $e(U^s)$, and by definition of sup and inf (and Weak Pareto for No Risk) one necessarily has, for all $\varepsilon \in \mathbb{R}_{++}$,

$$e(U^s) - \varepsilon \in \{x \in \mathbb{R} \mid [U^s] \ P \ [(x, ..., x)]\},$$

$$e(U^s) + \varepsilon \in \{x \in \mathbb{R} \mid [(x, ..., x)] \ P \ [U^s]\}.$$

It is immediate that for all $x \in \mathbb{R}$, $e(x, ..., x) = x$. ■

When the social ordering R is continuous, the EDQE becomes the "equally-distributed equivalent", a continuous function satisfying, for all $U^s \in \mathbb{R}^n$,

$$[U^s] \ I \ [(e(U^s), ..., e(U^s))].$$

We can now state a first result, which partly describes how R must be computed in order to satisfy these axioms.

Theorem 1 *Let $\mathcal{L} = \mathbb{R}^{nm}$ and R satisfy Weak or Strict Dominance, Weak Pareto for Equal Risk and Weak Pareto for No Risk. For all $U, U' \in \mathcal{L}$, one has $U\,P\,U'$ if $\sum_{s\in S} \pi^s e(U^s) > \sum_{s\in S} \pi^s e(U'^s)$, where e is the EDQE function defined in Lemma 1.*

Proof. Let $U, U' \in \mathcal{L}$ be such that $\sum_{s\in S} \pi^s e(U^s) > \sum_{s\in S} \pi^s e(U'^s)$. Let $\varepsilon \in \mathbb{R}_{++}$ be such that

$$\sum_{s\in S} \pi^s e(U^s) - \varepsilon > \sum_{s\in S} \pi^s e(U'^s) + \varepsilon.$$

For all $s \in S$, one has

$$[U^s]\ P\ [(e(U^s) - \varepsilon, ..., e(U^s) - \varepsilon)],$$
$$[(e(U'^s) + \varepsilon, ..., e(U'^s) + \varepsilon)]\ P\ [U'^s].$$

Let $e\langle U\rangle$ denote the prospect in which every U_i^s is replaced by $e(U^s)$. Let $\mathbf{1} \in \mathbb{R}^{nm}$ be such that for all $i \in N$, all $s \in S$, $\mathbf{1}_i^s = 1$. By Weak or Strict Dominance, one has $U\ R\ e\langle U\rangle - \varepsilon\mathbf{1}$ and $e\langle U'\rangle + \varepsilon\mathbf{1}\ R\ U'$. One computes, for all $i \in N$

$$E(e\langle U\rangle - \varepsilon\mathbf{1})_i = \sum_{s\in S} \pi^s e(U^s) - \varepsilon \text{ and } E(e\langle U'\rangle + \varepsilon\mathbf{1})_i = \sum_{s\in S} \pi^s e(U'^s) + \varepsilon.$$

As $e\langle U\rangle - \varepsilon\mathbf{1}$,$e\langle U'\rangle + \varepsilon\mathbf{1} \in \mathcal{L}^e$ and $E(e\langle U\rangle - \varepsilon\mathbf{1}) > E(e\langle U'\rangle + \varepsilon\mathbf{1})$, by Weak Pareto for Equal Risk one has $e\langle U\rangle - \varepsilon\mathbf{1}\ P\ e\langle U'\rangle + \varepsilon\mathbf{1}$. By transitivity, $U\,P\,U'$. ■

The statement of the theorem becomes an "iff" when the social ordering R is required to be continuous.

A noteworthy feature of this result is that the social criterion must, at least partly, rely on the expected value of some social welfare function even though the only rationality condition at the social level is a dominance condition. The expected value comes here from Weak Pareto for Equal Risk, which is much weaker than standard Pareto requirements but is sufficient here to force this result. For further reference, it is convenient to give a name to the class of criteria covered by this result. Although this is a little abusive, let us call them "expected equally distributed equivalent" (EEDE).

The classical utilitarian criterion satisfies all the axioms of Theorem 1, and corresponds to the situation in which $e(U^s)$ equals the average utility in U^s. But this result is also compatible with incorporating inequality aversion into R, as developed in the next section. Ex ante egalitarian criteria violate the dominance requirements, while ex-post egalitarian criteria based on concave social welfare functions such as $\sum_{i=1}^n \varphi(u_i)$ violate Weak Pareto for Equal Risk. In contrast, an ex-post egalitarian criterion based on the expected value of $\varphi^{-1}\left(\frac{1}{n}\sum_{i=1}^n \varphi(u_i)\right)$ does satisfy all the requirements.

An EEDE criterion behaves like an ex-ante criterion in absence of ex-post inequalities and like an ex-post criterion in absence of risk on the final distribution. Let us illustrate it with the $\varphi^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\varphi(u_i)\right)$ formula. The expected value of $\varphi^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\varphi(u_i)\right)$ behaves like an ex-ante criterion in absence of ex-post inequalities because $\varphi^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\varphi(u_i)\right)$ is then equal to u_i, so that its expected value coincides with individual expected utility. It behaves like an ex-post criterion when the statistical distribution of final utilities is known, because $\varphi^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\varphi(u_i)\right)$, which is neutral to permutations, then has the same value in all states of the world, and maximizing it is equivalent to maximizing $\sum_{i=1}^{n}\varphi(u_i)$.

At this point it is perhaps important to address the potential objection that abandoning the full Pareto principle is too high a price for the sake of combining dominance and inequality aversion. The response to this objection is that the full Pareto principle is not compelling when it is applied to expected utilities as opposed to final utilities. Consider the following prospects.

U'	Heads	Tails
Ann	4	0
Bob	0	4

U''	Heads	Tails
Ann	2	2
Bob	2	2

For both individuals considered in isolation, U' and U'' are equally good ex ante. The Pareto principle applied to expected utilities would then declare U' and U'' to be equivalent. But as only one state of the world can be true ex post, a fully informed observer would actually say that U' is better for one individual and worse for another. If the individuals themselves knew the state of the world, they would not be indifferent between U' and U'', and by applying the Pareto principle to expected utilities the evaluator only respects their *ignorant* preferences and does not particularly respect their *informed* preferences. As individual informed preferences normally take precedence in welfare economics over preferences based on ignorance,[10] the Pareto principle applied to expected utilities is much less compelling than the Pareto principle applied to final utilities.[11] An ex-post egalitarian criterion that prefers U'' to U' can be justified on the ground that it gives priority to the informed preferences of the worse-off (these preferences indeed

[10] Someone who believes in a truly indeterministic world would oppose describing the informational situation before the resolution of uncertainty as a situation of "ignorance" in general. It is fair to say, however, that in most practical cases risk corresponds to ignorance of the relevant facts. Moreover, even in a truly indeterministic situation knowing the final state of the world by definition involves more information about the final situation, by comparison to which the ex-ante situation is a situation of ignorance.

[11] It is already widely recognized that, when individuals have different beliefs, their unanimity over prospects may be spurious and carries little normative weight. See e.g. Broome (1999, p. 95), or Mongin (2005). But the point here is that even if all the individuals have the same beliefs about the probabilities of the various possible states of the world, their ex-ante preferences over prospects do not call for the same respect as their ex-post preferences. Hammond (1981) almost

rank U'' above U') over the informed preferences of the better-off (which rank U' first) as well as over the uninformed preferences of both (which are indifferent). The full Pareto principle is therefore not a serious challenge to the EEDE approach. We will see in Section 5 that there is a weaker version of the Pareto principle which raises a more serious challenge.

4 A generalized Gini family

This section highlights a particular family of EEDE criteria. It is special because it satisfies a stronger version of the Pareto principle than Weak Pareto for Equal Risk. This stronger version leaves quite some room for respecting individual expected utilities in the presence of ex post inequalities. Interestingly, it is shown here that this is nevertheless compatible with an arbitrarily high degree of inequality aversion.

This section is motivated by the observation that the EEDE approach based on a social welfare function $\varphi^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\varphi\left(u_i\right)\right)$, for a strictly concave function φ, respects individual expected utility only in cases of perfect equality in every state of the world. This may appear rather restrictive. In this section we show that the EEDE approach is compatible with a greater scope for the respect of individual expected utility.

The idea is to extend the application of the Pareto principle from situations of perfect equality to cases in which the individuals may be arbitrarily unequal but retain the same relative positions in all states of the world. Let us say that U involves no reranking if for all $s,t \in S$ and all $i,j \in N$, $\left(U_i^s - U_j^s\right)\left(U_i^t - U_j^t\right) \geq 0$. Let $\mathcal{L}^{nr}$ be the subset of $\mathcal{L}$ that contains the prospects involving no reranking. As $\mathcal{L}^c \cup \mathcal{L}^e \subseteq \mathcal{L}^{nr}$, the following axiom encompasses the two cases of riskless prospects and equal risks that were distinguished in the previous section. Moreover, it introduces the Strong Pareto principle, which implies social indifference when all individuals are indifferent and strict social preference when some individuals are indifferent and some have a strict preference (in terms of expected utility).

Axiom 5 (Strong Pareto for Non Reranking Risk) *For all $U,U' \in \mathcal{L}^{nr}$, one has $U\,R\,U'$ if for all $i \in N$, $EU_i \geq EU_i'$; in addition one has $U\,P\,U'$ if for some $i \in N$, $EU_i > EU_i'$.*

Note that individuals may occupy different relative positions in the two prospects U and U' mentioned in this axiom. The no reranking condition applies only across states of the world for each prospect

makes this point when warning that the context of risk makes the Pareto principle more demanding. Among other things he notes that misperceptions of probabilities by individual agents undermine the Pareto principle. If we consider that we live in a deterministic world, this problem is very serious—in such a world all probabilistic beliefs are mistaken!

separately.

We first show that this is compatible with a strong aversion to inequality. Let us introduce a requirement expressing inequality aversion over individual *prospects*. Such a condition can just as well be satisfied by ex-ante and ex-post criteria. When one agent has greater utility than another in all states of the world, it is not controversial who is the better off. The axiom defined below is adapted from Hammond's (1976) Equity condition. The original axiom is about inequality reduction between two agents with unequal utilities. We apply the same idea when one agent has greater utility in all states of the world, and inequality reduction is performed in every state of the world. This axiom requires an infinite inequality aversion in changes affecting only two individuals.

Axiom 6 (Hammond Equity) *For all $U, U' \in \mathcal{L}$, one has $U\, R\, U'$ if for some $i, j \in N$,*

$$U_i' \gg U_i \geq U_j \gg U_j',$$

while $U_k = U_k'$ for all $k \neq i, j$.

We also need a requirement of impartiality across individuals, saying that permutations of individual prospects do not affect the value of social prospects.

Axiom 7 (Anonymity) *For all $U, U' \in \mathcal{L}$, one has $U\, I\, U'$ if U' differs from U only by permuting the vectors U_i.*

We can now define the ex-post leximin social ordering, which satisfies all of the requirements introduced so far. Let $U_{(i)}^s$ denote the utility of ith rank (by increasing order) in vector U^s, and $U_{(i)}$ denote $\left(U_{(i)}^s\right)_{s \in S}$ The symbol $\geq_{lex}$ denotes the ordinary leximin criterion, which compares two vectors by comparing the smallest component, and if they are equal it compares the second smallest component, and so on. The ex-post leximin criterion weakly prefers U to U' iff

$$\left(EU_{(i)}\right)_{i \in N} \geq \left(EU_{(i)}'\right)_{i \in N}.$$

In other words, this criterion computes the expected value of the utility of ith rank in U, for all $i = 1, ..., n$, and applies the standard leximin criterion to such vectors. Note that the utility of ith rank may befall different individuals in different states of the world. This criterion is characterized by some of the requirements introduced in this section and the previous one.

Theorem 2 *Let $\mathcal{L} = \mathbb{R}^{nm}$. The ordering R satisfies Weak Dominance, Strong Pareto for Non Reranking Risk, Hammond Equity and Anonymity if and only if it is the ex-post leximin criterion. It then also satisfies Strict Dominance.*

Proof. By Strong Pareto for Non Reranking Risks, there is an ordering $\tilde{R}$ over $\mathbb{R}^n$ such that for all $U, U' \in \mathcal{L}^{nr}$, $U \, R \, U'$ iff

$$(EU_i)_{i\in N} \; \tilde{R} \; (EU_i')_{i\in N} \, .$$

By Hammond Equity, Anonymity and Strong Pareto for Non Reranking Risk, this ordering satisfies the standard Hammond Equity, Anonymity and Strong Pareto properties over n-vectors (these properties are immediate adaptations to n-vectors of the axioms defined here for prospects). By Hammond (1979, Th. 4-5), $\tilde{R}$ is then the leximin ordering.

Let $U, U' \in \mathcal{L}$. Let $V, V' \in \mathcal{L}^{nr}$ be defined by $V_i^s = U_{(i)}^s$ and $V_i'^s = U_{(i)}'^s$ for all $s \in S$ and all $i \in N$. By Anonymity, for all $s \in S$, $[U^s] \; I \; [V^s]$ and $[U'^s] \; I \; [V'^s]$. Therefore, by Weak Dominance, $U \, I \, V$ and $U' \, I \, V'$. Therefore, $U \; R \; U'$ iff $V \; R \; V'$. By Strong Pareto for Non Reranking Risks, $V \; R \; V'$ iff $(EV_i)_{i\in N} \; \tilde{R} \; (EV_i')_{i\in N}$, which, by the previous paragraph, holds iff $(EV_i)_{i\in N} \geq_{\text{lex}} (EV_i')_{i\in N}$.

Conversely, all the axioms of the theorem, plus Strict Dominance, are satisfied by the ex-post leximin criterion. ■

The utilitarian criterion satisfies all the axioms of this theorem except Hammond Equity. The ex-ante leximin criterion (i.e., applying the leximin criterion to expected utilities) satisfies all the axioms except the dominance requirements. It is shown in the appendix that replacing Strong Pareto for Non Reranking Risk by a weaker requirement taking the form of combining two Strong Pareto axioms, one restricted to equal risks and one to riskless prospects, would fail to characterize the ex-post leximin criterion.

Let us now turn to the case of finite inequality aversion. An ex-post generalized Gini criterion weakly prefers U to U' iff

$$\sum_{i\in N} \alpha_i EU_{(i)} \geq \sum_{i\in N} \alpha_i EU_{(i)}',$$

for some fixed parameters $(\alpha_i)_{i\in N} \in \mathbb{R}_{++}^n$. Inequality aversion is obtained when $\alpha_1 > ... > \alpha_n$. The utilitarian criterion and the ex-post leximin criterion are both on the boundary of the subfamily of inequality-averse ex-post generalized Gini criteria.

The next result involves a continuity axiom. Continuity is a requirement which, combined with a Strong Pareto axiom, excludes infinite inequality aversion.

Axiom 8 (Continuity) *Let $U, U' \in \mathcal{L}$ and $(U(t))_{t\in\mathbb{N}} \in \mathcal{L}^{\mathbb{N}}$ be such that $U(t) \to U$. If $U(t)\, R\, U'$ for all $t \in \mathbb{N}$, then $U\, R\, U'$. If $U'\, R\, U(t)$ for all $t \in \mathbb{N}$, then $U'\, R\, U$.*

One then obtains the following characterization.[12]

Theorem 3 *Let $\mathcal{L} = \mathbb{R}^{nm}$. The ordering R satisfies Weak or Strict Dominance, Strong Pareto for Non Reranking Risk, Anonymity and Continuity if and only if it is an ex-post generalized Gini criterion.*

Proof. Strict Dominance, Continuity and Weak Pareto for No Risk (which is implied by Strong Pareto for Non Reranking Risk) imply Weak Dominance. It is therefore enough to prove the necessity part with Weak Dominance.

Let e denote the EDQE function associated to R (its existence is proved in Lemma 1). It is immediate from Th. 1 that, under Continuity, for all $U, U' \in \mathcal{L}$, $U\ R\ U'$ iff $\sum_{s\in S} \pi^s e(U^s) \geq \sum_{s\in S} \pi^s e(U'^s)$. Moreover e must then be continuous.

Let $C = \{u \in \mathbb{R}^n \mid u_1 \leq ... \leq u_n\}$, and $\mathcal{L}_0 = C^m$. Let $U, V \in \mathcal{L}_0$ be such that for all $i \in N$, all $s \in S$, $V_i^s = EU_i$. By Strong Pareto for Non Reranking Risk, $U\ I\ V$. This is equivalent to $\sum_{s\in S} \pi^s e(U^s) = \sum_{s\in S} \pi^s e(V^s)$. Now, one has $\sum_{s\in S} \pi^s e(V^s) = e\left(\sum_{s\in S} \pi^s U^s\right)$. Therefore, for all $U \in \mathcal{L}_0$, $\sum_{s\in S} \pi^s e(U^s) = e\left(\sum_{s\in S} \pi^s U^s\right)$. As the set C is convex, this implies, by Lemma 3 stated and proved in Appendix A, that e is affine on C. Therefore there is $(\alpha_i)_{i\in N} \in \mathbb{R}^n$ such that for all $U, U' \in \mathcal{L}_0$, $U\ R\ U'$ iff $\sum_{i\in N} \alpha_i EU_i \geq \sum_{i\in N} \alpha_i EU_i'$. By Strong Pareto for Non Reranking Risk, $\alpha_i > 0$ for all $i \in N$.

By Weak Dominance and Anonymity, for all $U \in \mathcal{L}$, $U\ I\ \left(U_{(i)}\right)_{i\in N}$. As $\left(U_{(i)}\right)_{i\in N} \in \mathcal{L}_0$, one then obtains by transitivity of R that for all $U, U' \in \mathcal{L}$, $U\ R\ U'$ iff $\sum_{i\in N} \alpha_i EU_{(i)} \geq \sum_{i\in N} \alpha_i EU'_{(i)}$.

Conversely, all the axioms of the theorem are satisfied by any ex-post generalized Gini criterion. ■

These results show that there is substantial flexibility, as far as inequality aversion is concerned, in the class of EEDE criteria, even under Strong Pareto for Non Reranking Risk. The next section, however, shows that there are more serious trade-offs if one wants some separability of the evaluation for subpopulations in the presence of rerankings.

[12] In this result inequality aversion is not introduced, but it would be easy to obtain it with a suitable axiom based on the Pigou-Dalton principle of transfer.

5 The separability challenge

We have seen that the full Pareto principle is not compelling for risky prospects because it caters to individual uninformed preferences even when it is possible to know what the distribution of informed preferences will ultimately be, as with the prospects U' and U'' introduced at the end of Section 3. A particular feature of that example is worth emphasizing. Even though the evaluator has no information about states of the world, the informational situation is equivalent to perfectly knowing the state of the world. Indeed, knowledge of the state of the world would add nothing relevant to the evaluation, because it would only specify the names of the winner and loser, an information that is irrelevant for an impartial evaluation. Looking at U' and U'', the evaluator knows all that is needed for the evaluation, namely, the distribution of final utilities: $(4,0)$ with U', $(2,2)$ with U''.

Obviously, such knowledge depends on the correlation between individuals across states of the world. Consider the following prospects, two of which reproduce U' and U'' under different names.

V	Heads	Tails
Ann	2	2
Bob	2	2

V'	Heads	Tails
Ann	4	0
Bob	2	2

V''	Heads	Tails
Ann	4	0
Bob	0	4

V'''	Heads	Tails
Ann	4	0
Bob	4	0

In V'' there is perfect negative correlation, as opposed to V''' in which there is perfect positive correlation. The latter prospect provides no information about the final distribution of utilities. This is precisely why it makes sense to apply the Pareto principle to equal risks and declare V''' to be as good as V.

These considerations are useful when one turns to another set of issues. An important difference between the three main approaches discussed in the introduction and the EEDE approach is that with the former one can assess the prospects of subpopulations separately before making a global assessment, whereas the latter imposes a global examination of the population even when only a fraction of the population considers a change. For instance, utilitarianism, ex-ante egalitarianism based on a social welfare function $\sum_{i=1}^{n} \varphi(EU_i)$ and ex-post egalitarianism based on the expected value of a social welfare function $\sum_{i=1}^{n} \varphi(u_i)$ will all be indifferent between V'' and V''', because the difference between the two only concerns Bob, for whom the two prospects are equivalent. In contrast, a typical EEDE criterion will prefer V''' to V'', because the presence of Ann with her specific prospect makes the prospect for Bob in V'' look worse, in terms of EDQE, than in V'''. Here again, this lack of separability makes sense because the informational situation of the evaluator does depend crucially on the correlation between individuals.

Such reasoning makes it sensible to prefer V''', but also V', to V''. However, there is a problem.

A typical EEDE criterion (with strict inequality aversion) will also prefer V to V', even though one can hardly say that the evaluator gains any informational advantage from the (null) correlation between individuals. In V', as compared to V, Ann takes a risk alone while Bob keeps the same sure prospect. Why should the presence of Bob affect the evaluation? Moreover, the EEDE evaluation may depend on Bob's level of utility. For instance, if his utility were lower than 0 or greater than 4, the ex-post leximin criterion would accept the evaluation based on Ann's sole prospect. Formally, it is easy to understand why we have such a strong non-separability here. The presence of Bob affects the value of the EDQE in every state of the world. The problem is that the informational value of the presence of Bob in V', in terms of knowledge of the final distribution of utilities, seems close to zero. The informational arguments that justify the lack of separability in the comparison of V''' or V' with V'' seem to fail in the case of V and V'.

The problem takes an interesting and somewhat dramatic turn if one thinks of applying the EEDE approach to real-life cases. In real life, the only part of the population that faces absolutely no risk, like Bob in V and V', is the subgroup of individuals who are already dead at the moment of the evaluation. It appears problematic if the utility level of the past generations must play a role in the evaluation of future risks and this has motivated the axiom of "Independence of the Utilities of the Dead" in Blackorby et al. (2005).[13] The EEDE approach applied to the whole population (past, present and future) typically violates this axiom. An easy way out is to restrict attention to present and future generations in the evaluation, but it comes at the cost of intertemporal consistency. The evaluation then depends on the time at which it is made, and this seems a deep rationality failure.

In order to analyze this problem formally, let us formulate an axiom that strengthens Pareto for Equal Risk by considering cases in which the equal risk is taken by a subgroup of the population while the rest is unconcerned and faces no risk (like the past generations). For any $M \subseteq N$, let $U_M = (U_i)_{i \in M}$ and $U_{-M} = (U_i)_{i \notin M}$.

Axiom 9 (Weak Pareto for Subgroup Equal Risk) $\forall U, U' \in \mathcal{L}^e$, $\forall V \in \mathcal{L}^c$, $\forall M \subseteq N$, $(U_M, V_{-M})\ P\ (U'_M, V_{-M})$ *if* $\forall i \in M$, $EU_i > EU'_i$.

This axiom is logically weaker than jointly imposing Weak Pareto for Equal Risk and a separability axiom saying that unconcerned individuals facing no risk can be ignored in the evaluation, i.e., can be removed from the population without altering the ranking of two prospects (such an axiom requires a

[13] See also Bommier and Zuber (2008).

broader framework in which the social ordering is defined for different possible populations).

Weak Pareto for Subgroup Equal Risk is violated by EEDE criteria displaying strict aversion to inequality. Actually, one can see that, under Weak or Strict Dominance, this axiom is incompatible with the most basic form of inequality aversion, namely, inequality aversion over riskless prospects. Consider the following prospects, where $\varepsilon > 0$.

V	Heads	Tails
Ann	2	2
Bob	2	2

V'	Heads	Tails
Ann	$4+\varepsilon$	$0+\varepsilon$
Bob	2	2

Z	Heads	Tails
Ann	$3-\varepsilon$	$1-\varepsilon$
Bob	$3-\varepsilon$	$1-\varepsilon$

By Weak Pareto for Subgroup Equal Risk, V' is preferred to V and V is preferred to Z. By transitivity, therefore, V' is preferred to Z. However, in the presence of inequality aversion over riskless prospects, for ε small enough, $[Z^s]$ is preferred to $[V'^s]$ for s =Heads,Tails. Therefore, by Weak or Strict Dominance, Z is at least as good as V', a contradiction.

More specifically, if in Theorem 1 one replaces Weak Pareto for Equal Risk by Weak Pareto for Subgroup Equal Risk, one obtains a partial characterization of weighted utilitarianism, which is, in essence, much stronger than Harsanyi's theorem as it involves only a basic rationality condition (dominance) and a very restricted version of the Pareto principle.[14] Before stating the result, the following Lemma must be noted.

Lemma 2 *Let $\mathcal{L} = \mathbb{R}^{nm}$. If R satisfies Weak Pareto for Subgroup Equal Risk, then it satisfies Weak Pareto for No Risk.*

Proof. Let $U, V \in \mathcal{L}^c$ be such that $U \ll V$. Let

$$\begin{aligned} T(1) &= (V_1, U_2, ..., U_n) \\ T(2) &= (V_1, V_2, U_3, ..., U_n) \\ &\vdots \\ T(n-1) &= (V_1, V_2, ..., V_{n-1}, U_n). \end{aligned}$$

By Weak Pareto for Subgroup Equal Risk, applied to $M = \{1\}, \{2\}, ..., \{n\}$ successively, one has $T(1)\ P\ U$, $T(2)\ P\ T(1)$, and so on, until $V\ P\ T(n-1)$. By transitivity, $V\ P\ U$. ■

[14] Variants of Harsanyi's result without the expected utility hypothesis can also be found in Blackorby et al. (2004), Chambers and Hayashi (2006), and Fleurbaey (2009). Harsanyi's theorem was formulated for lotteries. Its adaptation to state-contingent alternatives has been studied by Blackorby et al. (1999).

Theorem 4 *Let $\mathcal{L} = \mathbb{R}^{nm}$ and R satisfy Weak or Strict Dominance, and Weak Pareto for Subgroup Equal Risk. There is $\alpha \in \mathbb{R}^n_+ \setminus \{0\}$ such that for all $U, U' \in \mathcal{L}$, one has $U\, P\, U'$ if $\sum_{i \in N} \alpha_i EU_i > \sum_{i \in N} \alpha_i EU'_i$.*

Proof. By Lemma 2, R also satisfies Weak Pareto for No Risk. Weak Pareto for Subgroup Equal Risk implies Weak Pareto for Equal Risk. By Th. 1, there is an EDQE function e such that $\sum_{s \in S} \pi^s e\left(U^s\right) > \sum_{s \in S} \pi^s e\left(U'^s\right)$ implies $U\ P\ U'$.

Take any $u \in \mathbb{R}^n$, any $i \in N$ and any $x, y \in \mathbb{R}$. Define the prospect V by $V_j^s = u_j$ for all $j \neq i$, all $s \in S$, $V_i^1 = x$ and $V_i^s = y$ for all $s > 1$. For some $\varepsilon > 0$, let the prospect $V^{\oplus}$ be such that $V_i^{\oplus s} = EV_i + \varepsilon$ for all $s \in S$ and $V_j^{\oplus} = V_j$ for all $j \neq i$; and the prospect $V^{\ominus}$ be such that $V_i^{\ominus s} = EV_i - \varepsilon$ for all $s \in S$ and $V_j^{\ominus} = V_j$ for all $j \neq i$. By Weak Pareto for Subgroup Equal Risk, $V^{\oplus}\ P\ V\ P\ V^{\ominus}$. This implies that

$$\sum_{s \in S} \pi^s e\left(V^{\oplus s}\right) \geq \sum_{s \in S} \pi^s e\left(V^s\right) \geq \sum_{s \in S} \pi^s e\left(V^{\ominus s}\right).$$

Let $f(z) = e(u_1, ..., u_{i-1}, z, u_{i+1}, ..., u_n)$. The inequalities read

$$f\left(\pi^1 x + \left(1 - \pi^1\right) y + \varepsilon\right) \geq \pi^1 f(x) + \left(1 - \pi^1\right) f(y) \geq f\left(\pi^1 x + \left(1 - \pi^1\right) y - \varepsilon\right). \tag{1}$$

The fact that (1) holds for all $x, y \in \mathbb{R}$, $\varepsilon > 0$ implies, by Lemma 4 stated and proved in Appendix A, that the function f is affine.

In summary, the function $e(u)$ is affine in u_i for all i and all $(u_1, ..., u_{i-1}, u_{i+1}, ..., u_n)$. This means that there exist coefficients α_K, for $K \subseteq N$, such that

$$e(u) = \sum_{K \subseteq N} \alpha_K \left(\prod_{i \in K} u_i\right),$$

with the convention that $\prod_{i \in \varnothing} u_i = 1$. Now, the fact that $e(x, ..., x) \equiv x$ implies that $\alpha_K = 0$ for K such that $|K| \neq 1$, and that $\sum_{i \in N} \alpha_{\{i\}} = 1$. The fact that $\sum_{s \in S} \pi^s e\left(V^{\oplus s}\right) \geq \sum_{s \in S} \pi^s e\left(V^s\right)$, which is then equivalent to $\alpha_{\{i\}}\left(EV_i + \varepsilon\right) \geq \alpha_{\{i\}} EV_i$, requires $\alpha_{\{i\}} \geq 0$. ■

The weakest part of Harsanyi's utilitarian theorem is the Pareto axiom applied to expected utilities because, as we have seen, prospects that look good to every individual separately may be ultimately bad for some of them for sure. It is harder to reject Weak Pareto for Subgroup Equal Risk because it is immune to this objection. The risk taken by the subgroup depicted in the axiom will be good or bad for all of them and no certain information about the final distribution is available ex ante. If present and future generations want to take a risk that may create inequalities only with respect to the (unconcerned) past generations, who is entitled to declare it a bad move? Theorem 4 is therefore a quite serious challenge to the EEDE approach and to inequality aversion in general.

One might think of avoiding the conclusion of this theorem by considering a restricted domain in which the unconcerned individuals in the application of Weak Pareto for Subgroup Equal Risk can only be the first generations up to some generation, for a fixed order of generations. Unfortunately, this restricted domain still carries the result that inequality aversion across generations is excluded.[15]

There is probably no easy answer to this challenge. But four considerations may attenuate it somewhat. First, it is actually not counterintuitive to worry about inequalities involving past generations. For an EEDE criterion the problem, really, is not that future generations may turn out to be lucky and be much better off than past generations. For the family of criteria analyzed in Section 4, for instance, provided the future generations are better off than their ancestors in all states of the world they can take all the risk they want. A problem appears only if in case of bad luck they may end up worse off than past generations, a perspective that, indeed, looks intuitively ominous. The application of an EEDE criterion would then typically imply taking precautions so that the risks taken by present and future generations do not put them at risk at falling below the utility of their ancestors. Far from being counterintuitive, this is actually strikingly similar to sustainability principles that have become fashionable recently.

Second, even if the kind of risk depicted in Weak Pareto for Subgroup Equal Risk does not make it possible for an observer to know what the final distribution will be, the presence of the unconcerned individuals does bring some information. Consider for instance the ex-post leximin criterion. In a two-agent society the presence of a moderately well-off unconcerned individual makes it problematic for the other individual to take a risk that could bring him either above or below the first one. This is because this risk may lower the utility of the worst-off and it is not compensated by possible gains because the gains accrue to the best-off and are therefore of lesser social value. The evaluator is then sure that the utility level of the worst-off cannot be above the utility level of the unconcerned individual, and there

[15] To see this, consider a simple domain with three generations and only one individual per generation, with $i = 1, 2, 3$ corresponding to the historical order of succession. A similar argument as in the proof implies that $e(u_1, u_2, u_3)$ must be affine in u_3, in u_2 when $u_2 = u_3$, and in u_1 when $u_1 = u_2 = u_3$. One must then have

$$e(u_1, u_2, u_3) = \alpha(u_1, u_2)\, u_3 + \beta(u_1, u_2).$$

As this expression must be affine in u_2 when $u_2 = u_3$, $\alpha(u_1, u_2)$ must be constant in u_2. One then has:

$$e(u_1, u_2, u_3) = \alpha(u_1)\, u_3 + \beta(u_1)\, u_2 + \gamma(u_1).$$

By a similar argument, $\alpha(u_1)$ and $\beta(u_1)$ must be constant, and one gets

$$e(u_1, u_2, u_3) = \alpha u_3 + \beta u_2 + \gamma u_1 + \delta.$$

is a positive probability that it may fall below. This information about the distribution is ignored by an isolated individual who only looks at his own personal gains and losses. In conclusion, Weak Pareto for Subgroup Equal Risk may be criticized on similar grounds as the stronger Pareto axioms because it requires ignoring some potentially relevant information about the distribution.

Third, consider the alternative strategy consisting in applying an EEDE criterion on present and future generations only, ignoring past generations. Such a strategy is unlikely to generate much time inconsistency in practice. A serious inconsistency would involve precise contingent plans that would involve choices affecting only future generations. There are very few such situations, as in most contexts of decision the alternatives under consideration are not neutral for the present generation. In sum, whether one decides to include or not to include past generations in the computation of an EEDE criterion, the consequences are unlikely to be shocking.

Finally, it must be emphasized that we are dealing here with evaluations, not with decisions made in specific feasible sets. It may be that risk taking creates regrettable inequalities but that it would be even worse, given realistic feasibility constraints, to prevent individuals from doing what they want. If the choice is not between happy risk taking and happy equality but between happy risk taking and miserable equality, an inequality averse EEDE criterion may well prefer the former.

6 Welfare economics of risk

In this section we illustrate the various approaches in three examples of welfare analysis in the presence of risk.

Limits of the second welfare theorem. The second welfare theorem says that under certain conditions, any Pareto efficient allocation can be obtained as a competitive equilibrium with suitable initial endowments. It is often remembered as meaning that the socially optimal allocation can be implemented by competitive markets, under ideal conditions. This particular formulation is no longer valid if the socially optimal allocation is not Pareto efficient. This does occur in particular when ex post social optimality and ex ante Pareto efficiency fall apart, a common situation in the presence of risk. Here is an example.

Consider a continuum of individuals whose common utility function is $U(c,h)$, where $c \geq 0$ is consumption and $h \in [0,1]$ is health. The distribution of h is known ex ante, but each individual faces a lottery on h. The socially optimal allocation of c, given the distribution of h, is computed as the solution

$c(h)$ of the problem

$$\max \int_0^1 \varphi\left(U(c(h),h)\right) f(h)dh \text{ s.t. } \int_0^1 c(h)f(h)dh = C.$$

This is what the standard ex post approach, but also the EEDE approach, would recommend in connection with an additively separable social welfare function. In contrast, an individual facing an equal chance of having the health of anyone in the population, and facing an actuarially fair insurance, would seek to maximize his expected utility:

$$\max \int_0^1 U(c(h),h)f(h)dh \text{ s.t. } \int_0^1 c(h)f(h)dh = C.$$

The two solutions coincide only if φ is affine, i.e., if the social welfare function displays no aversion to inequality. If φ is strictly concave, the socially optimal allocation cannot be implemented by insurance markets. For instance, if $U(c,h) = \ln hc$, individuals do not want to insure against health shocks because the marginal utility of consumption does not depend on health. But ex post those who have bad health wish they had taken an insurance, and the socially optimal allocation would provide them with some compensation.[16] If $U(c,h) = hv(c)$, a function similar to that used in Murphy and Topel (2006), individuals want to consume more when they are healthy because their marginal utility of consumption is greater then. In the market they would then take contracts that would reduce their consumption when their health is bad, which may go in the opposite direction from the socially optimal redistribution if the social welfare function is sufficiently averse to inequality.[17]

Adverse selection and inequalities. Consider the classical adverse selection problem with high and low risk (Rothschild and Stiglitz 1976, Wilson 1977, Dahlby 1981). The literature has focused on the fact that a small amount of compulsory insurance may be Pareto-improving. What happens is that some compulsory insurance at the pooling indemnity-premium ratio relaxes the temptation for high-risk agents to mimic the low-risk agents, so that the latter can obtain a better coverage while subsidizing the former through the compulsory part of the insurance. From the ex-post viewpoint, however, this is *not* a Pareto-improvement. Fortunately, the only agents who suffer from this policy are the lucky low-risk agents, who are better-off than the unlucky low-risk agents and also better-off than the high-risk agents if their initial wealth is identical. This is illustrated on Figure 1, where compulsory insurance makes the high-risk agents move from HR_1 to HR_2 and the low-risk agents move from LR_1 to LR_2 (the point NI

[16] A study of the optimal mix of public and private health insurance when $U(c,h) = h + v(c)$ is made in Leach (2008), following an ex-post approach compatible with EEDE.

[17] It is known since Cook and Graham (1977) that state-dependent utilities may imply puzzling consequences. Contrary to the standard case, insurance may widen ex-post inequalities, the unlucky having acted against their ex-post interests.

corresponds to consumption in absence of insurance). The arrow on the figure shows the decrease of consumption of lucky low-risk agents.

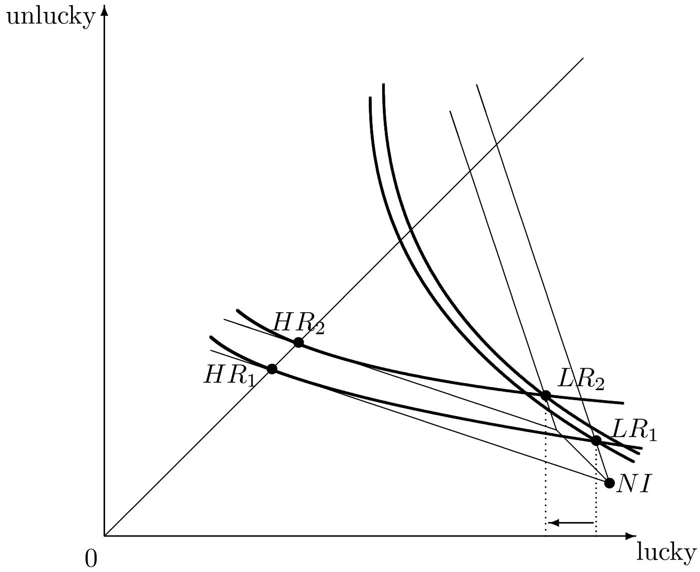

Figure 1: Introducing compulsory partial insurance

Consider now the case in which the low-risk agents are much poorer than the high-risk agents. Assume that wealth is private information and cannot be used to discriminate in the insurance system. If the proportion of high-risk agents is great enough, a complete compulsory insurance at the pooling indemnity-premium ratio would decrease the expected utility of the low-risk agents and raise the utility of the high-risk agents, in comparison with the market equilibrium. See Figure 2. An ex-ante egalitarian criterion would then reject this policy because the low-risk agents are ex ante the worst-off. From the ex post viewpoint, things look different. This full insurance policy increases the final utility of all high-risk agents but also of unlucky low-risk agents, and the only losers are, once again, the lucky low-risk agents. As the unlucky low-risk agents are the worst-off in the final distribution, a sufficiently inequality averse ex post criterion (EEDE or standard) would approve this policy.

Value of life. The two previous examples share the feature that the final distribution is known ex ante, which makes the standard ex-post approach and the EEDE coincide. Here is an example showing their difference. Suppose that individual i's lifetime utility is $U(w_i(l_i), l_i)$, where l_i is longevity and $w_i(l_i)$ is wealth (which is influenced by longevity). Let a_i denote i's current age. There is a finite number of individuals. We are interested in a particular hazard, which will strike with probability p. If it strikes, the joint distribution of longevity $l = (l_i)_{i \in N}$ is described by a probability density function $f(l)$. If it does

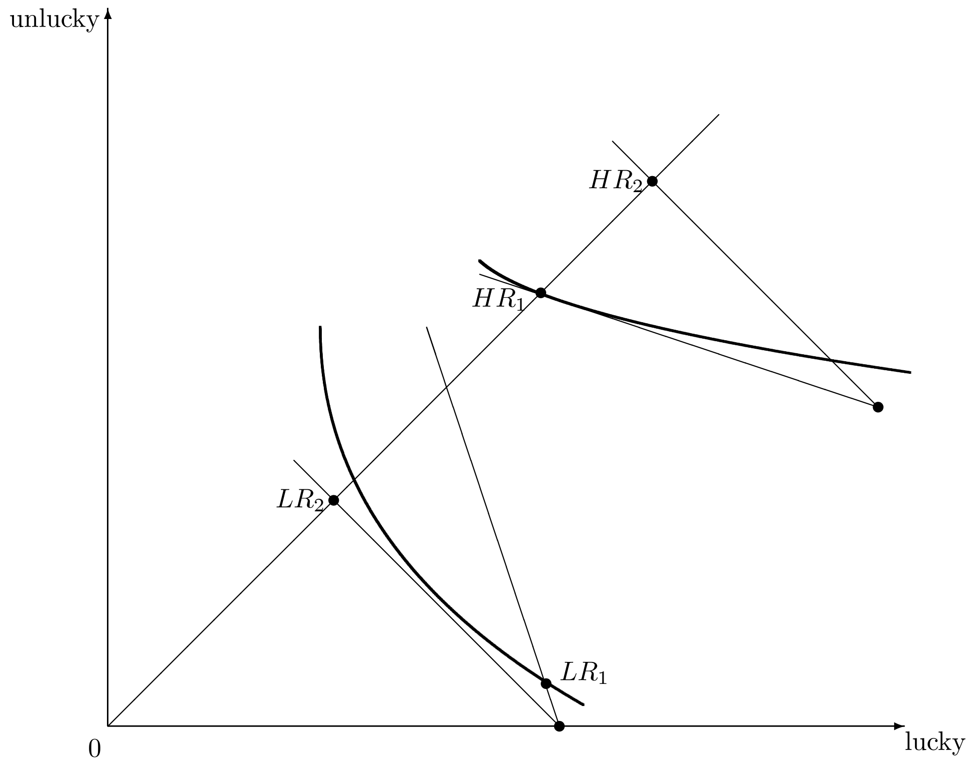

Figure 2: Introducing compulsory full insurance when low-risk agents are poorer

not strike, the distribution is described by a more favorable function $g(l)$. Let $f_i(l_i)$ and $g_i(l_i)$ denote the probability density functions of the marginal distribution of l_i. The problem is to compute the value of reducing p. To make the comparison easier, let all the criteria be based on the additively separable social welfare function $\sum_{i=1}^n \varphi(u_i)$.

An ex-ante criterion[18] computes the derivative of social welfare with respect to $-p$ as:

$$\sum_{i=1}^{n} \varphi' (EU_i) \int_{a_i}^{+\infty} U(w_i(l_i), l_i) [g_i(l_i) - f_i(l_i)] \, dl_i. \tag{2}$$

A standard ex-post criterion computes it as:

$$\sum_{i=1}^{n} \int_{a_i}^{+\infty} \varphi (U(w_i(l_i), l_i)) [g_i(l_i) - f_i(l_i)] \, dl_i. \tag{3}$$

The EEDE criterion computes it as:

$$\int_{a_1}^{+\infty} ... \int_{a_n}^{+\infty} \varphi^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} \varphi (U(w_i(l_i), l_i)) \right) [g(l) - f(l)] \, dl_n ... dl_1. \tag{4}$$

[18] The ex-ante approach, often without inequality aversion (i.e., utilitarianism), is widely used in this context. See, e.g., Viscusi (1992), Pratt and Zeckhauser (1996).

Consider a simple case. If it hits, the hazard may kill only two individuals j and k, each with probability 1/2. If it does not hit they live until a fixed L. Formula (2) becomes

$$\frac{1}{2}\varphi'\left(EU_j\right)\left[U\left(w_j(L),L\right)-U\left(w_j(a_j),a_j\right)\right]+\frac{1}{2}\varphi'\left(EU_k\right)\left[U\left(w_k(L),L\right)-U\left(w_k(a_k),a_k\right)\right],$$

where EU_i denotes the expected value of $U\left(w_i(l_i),l_i\right)$, while formula (3) becomes

$$\frac{1}{2}\left[\varphi\left(U\left(w_j(L),L\right)\right)-\varphi\left(U\left(w_j(a_j),a_j\right)\right)\right]+\frac{1}{2}\left[\varphi\left(U\left(w_k(L),L\right)\right)-\varphi\left(U\left(w_k(a_k),a_k\right)\right)\right].$$

In contrast, the EEDE criterion needs more information than the marginal distributions. Let us distinguish three interesting cases: 1) if the hazard hits, one of the two individuals is killed for sure, but only one; 2) they face independent risks of dying with probability 1/2; 3) either both dies, or both survive, with probability 1/2. The EEDE criterion also needs to know what the others' situation is, but let us assume for simplicity that their longevity is L for sure and that all individuals have the same wealth function $w(l)$. Let us moreover assume that $a_j = a_k$. Formula (4) becomes, in the three cases:

$$\begin{array}{ll}
1) & U\left(w(L),L\right)-\varphi^{-1}\left(\frac{1}{n}\left[(n-1)\varphi\left(U\left(w(L),L\right)\right)+\varphi\left(U\left(w(a),a\right)\right)\right]\right); \\
2) & \frac{1}{2}\left[U\left(w(L),L\right)-\varphi^{-1}\left(\frac{1}{n}\left[(n-1)\varphi\left(U\left(w(L),L\right)\right)+\varphi\left(U\left(w(a),a\right)\right)\right]\right)\right] \\
 & +\frac{1}{4}\left[U\left(w(L),L\right)-\varphi^{-1}\left(\frac{1}{n}\left[(n-2)\varphi\left(U\left(w(L),L\right)\right)+2\varphi\left(U\left(w(a),a\right)\right)\right]\right)\right]; \\
3) & \frac{1}{2}\left[U\left(w(L),L\right)-\varphi^{-1}\left(\frac{1}{n}\left[(n-2)\varphi\left(U\left(w(L),L\right)\right)+2\varphi\left(U\left(w(a),a\right)\right)\right]\right)\right].
\end{array}$$

If φ is strictly concave, the first expression is greater than the second, which is greater than the third. It is then more important to reduce a risk that condemns one (yet unknown) individual for sure than an independent risk, and more important to reduce an independent risk than a risk that hits everyone equally, when the expected number of casualties is the same in all three cases. It is especially intuitive that hazards that kill for sure are less acceptable than hazards with a good probability to avoid casualties,[19] and it is problematic that the ex-ante criteria and the standard ex-post criteria are insensitive to this.

7 Conclusion

This paper has introduced the EEDE class of criteria and shown that it uniquely avoids serious shortcomings of existing approaches, namely, the utilitarian indifference to inequalities in utility, the ex-ante

[19] It is less intuitive that hazards that kill everybody if they kill at all are more acceptable than independent risks, but only because of the illusion of "catastrophe avoidance" that was discussed in Section 3. Observe that such hazards minimize the probability of having casualties.

egalitarian violations of dominance, and the total disrespect, with standard ex-post egalitarian criteria, of individual expected utility. The ex-post generalized Gini family, extended to include the ex-post leximin criterion, has been highlighted in virtue of the fact that it respects individual expected utility in a rather broad set of cases. But the EEDE approach has also been shown to require dropping an advantageous feature of the three existing approaches, namely, the possibility to evaluate the prospects of subpopulations separately. Although the previous section has argued that this aspect of the EEDE approach may not be as problematic as it seems, it has presented a variant of the Harsanyi utilitarian theorem which signals a serious tension between inequality aversion, rationality of the evaluation, and the neglect of subpopulations that are not concerned and face no risk.

This paper has not proposed a full theory of social evaluation under risk. Let us briefly mention three topics which would deserve in-depth explorations. First, there is Diamond's objection to the ex-post approach, already discussed in Section 2. It has been put aside in this paper by assuming that final consequences contained a full description of the relevant information, including about the fairness of lotteries. But a full theory should obviously make it possible to formulate fairness requirements explicitly and to derive the restrictions that they impose on the evaluation. This is left for future research.

A second important topic is personal responsibility for risk-taking. Hammond (1983) and Kolm (1998) suggest that one might associate the principle of respecting individual preferences over risk with the idea that individuals have the right to take risks and should be able to assume the consequences. The ex-ante Pareto principle would then be based on an attribution of responsibility to individual agents for their risk-taking choices. This appears intuitive but, as a matter of fact, assigning responsibility to individuals generally means that the variables of responsibility are disregarded in the evaluation. For instance, if the consumer is responsible for what he buys, his situation will be evaluated in terms of budget set, not in terms of consumption vector. In the context of risk, an individual's ex-post utility presumably depends on his expected utility and his luck. Now, the individual's decisions or attitude to risk, for which he may be held responsible, can determine his expected utility, but not his luck, which by definition is neither under his control nor a part of his attitude to risk. Therefore, if anything should be disregarded in the evaluation in virtue of responsibility, it is the individual's expected utility, not his luck. This does not go in the direction of the ex-ante approach which focuses exclusively on expected utility and disregards individual luck. At any rate, certainly, incorporating notions of personal responsibility for risk-taking should be an important chapter of a full theory.[20]

[20] More on this topic can be found in Fleurbaey (2008, ch. 6). Whether, for the EEDE approach, redistribution should

A third topic has to do with the application of the EEDE approach to economic models. The framework adopted in this paper is simple but it involves an assumption that may not be satisfactory in more concrete analyses, namely, the assumption that interpersonal comparisons and expected utilities are computed with the same utility measure. It is common, for instance, in the economic theory of fair allocation to compare individual situations in terms of resources rather than utilities. It would therefore be interesting to extend the EEDE approach to cases in which interpersonal comparisons are based on a different measure. This is again left for future research.

As most of the welfare economics of risk and insurance adopts the ex-ante viewpoint, this analysis suggests that it would be worthwhile reexamining policy issues in this field in the light of EEDE criteria. The examples provided in the previous section have only shown that new criteria may indeed make a difference. In a time in which global (e.g., climate change, pandemics) and local (e.g., unhealthy lifestyles) risks combining various degrees of correlation across subpopulations are the focus of much attention, this seems an important alley of research.

References

Adler M.D., C.W. Sanchirico 2006, "Inequality and uncertainty: Theory and legal applications", *University of Pennsylvania Law Review* 155: 279-377.

Anscombe F.J., R.J. Aumann 1963, "A definition of subjective probability", *Annals of Mathematical Statistics* 34: 199-205.

Atkinson A.B. 1970, "On the measurement of inequality", *Journal of Economic Theory* 2: 244-263.

Ben-Porath E., I. Gilboa, D. Schmeidler 1997, "On the measurement of inequality under uncertainty", *Journal of Economic Theory* 75: 194-204.

Blackorby C., D. Donaldson, P. Mongin 2004, "Social aggregation without the expected utility hypothesis", *Cahier* n° 2004-020 *du Laboratoire d'Econométrie, Ecole Polytechnique.*

Blackorby C., D. Donaldson, J.A. Weymark 1999, "Harsanyi's social aggregation theorem for state-contingent alternatives", *Journal of Mathematical Economics* 32: 365–387.

Bommier A., S. Zuber 2008, "Can preferences for catastrophe avoidance reconcile social discounting with

take place between the lucky and unlucky risk-takers depends on the precise responsibility-sensitive criterion that is adopted. If one evaluates the situation of risk-takers by what would have happened to them if they had been prudent, an EEDE criterion may advocate no redistribution in favor of the unlucky risk takers.

intergenerational equity?", *Social Choice and Welfare* 31: 415-434.

Broome J. 1984, "Uncertainty and fairness", *Economic Journal* 94: 624-632.

Broome J. 1991, *Weighing Goods. Equality, Uncertainty and Time*, Oxford: Blackwell.

Broome J. 1999, *Ethics Out of Economics*, Cambridge: Cambridge University Press.

Chambers C.P., T. Hayashi 2006, "Preference aggregation under uncertainty: Savage vs. Pareto", *Games and Economic Behavior* 54: 430–440.

Cook P.J., Graham D.A. 1977, "The demand for insurance and protection: The case of irreplaceable commodities", *Quarterly Journal of Economics* 91: 143–156.

Coulhon T., P. Mongin 1989, "Social choice theory in the case of Von Neumann–Morgenstern utilities", *Social Choice and Welfare* 6: 175-187.

Dahlby B. 1981, "Adverse selection and pareto improvements through compulsory insurance", *Public Choice* 37: 548–558.

Deschamps R., L. Gevers 1979, "Separability, risk-bearing and social welfare judgments", in J.J. Laffont ed., *Aggregation and Revelation of Preferences*, Amsterdam: North-Holland.

Diamond P.A. 1967, "Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility: Comment", *Journal of Political Economy* 75: 765-766.

Drèze J.H. 1987, *Essays on Economic Decisions under Uncertainty*, Cambridge: Cambridge University Press.

Epstein L.G., U. Segal 1992, "Quadratic social welfare functions", *Journal of Political Economy* 100: 691-712.

Fishburn P.C. 1978, "On Handa's "New theory of cardinal utility" and the maximization of expected return", *Journal of Political Economy*, 86: 321-324.

Fishburn P.C. 1984, "Equity axioms for public risks", *Operations Research* 32: 901-908.

Fleurbaey M. 2008, *Fairness, Responsibility, and Welfare*, Oxford: Oxford University Press.

Fleurbaey M. 2009, "Two variants of Harsanyi's aggregation theorem", mimeo.

Gajdos T., E. Maurin 2004, "Unequal uncertainties and uncertain inequalities: An axiomatic approach", *Journal of Economic Theory* 116: 93-118.

Grant S. 1995, "Subjective probability without monotonicity: Or how Machina's Mom may also be probabilistically sophisticated", *Econometrica* 63: 159-189.

Hammond P.J. 1976, "Equity, Arrow's conditions, and Rawls' difference principle", *Econometrica* 44: 793-804.

Hammond P.J. 1979, "Equity in two-person situations", *Econometrica* 47: 1127-1136.

Hammond P.J. 1981, "Ex-ante and ex-post welfare optimality under uncertainty", *Economica* 48: 235-250.

Hammond P.J. 1983, "Ex post optimality as a dynamically consistent objective for collective choice under uncertainty", in P.K. Pattanaik, M. Salles eds., *Social Choice and Welfare*, Amsterdam: North-Holland.

Hammond P.J. 1988, "Consequentialist foundations for expected utility", *Theory and Decision* 25: 25-78.

Hammond P.J. 1989, "Consistent plans, consequentialism, and expected utility", *Econometrica* 57: 1445-1449.

Handa 1977, "Risk, probabilities, and a new theory of cardinal utility", *Journal of Political Economy* 85: 97-122.

Harel A., Z. Safra, U. Segal 2005, "Ex-post egalitarianism and legal justice", *Journal of Law, Economics and Organization* 21: 57-75.

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

Kahneman D., A. Tversky 1979, "Prospect theory: An analysis of decision under risk", *Econometrica* 47: 263-292.

Keeney R.L. 1980, "Equity and public risk", *Operations Research* 28: 527-534.

Kolm S.C. 1968, "The optimal production of social justice", in H. Guitton, J. Margolis (eds.), *Economie Publique*, Paris: Ed. du CNRS.

Kolm S.C. 1998, "Chance and justice: Social policies and the Harsanyi-Vickrey-Rawls problem", *European Economic Review* 42: 1393-1416.

Leach J. 2008, "Ex post welfare under alternative health care systems", mimeo.

Machina M.J. 1982, "Expected utility analysis without the independence axiom", *Econometrica* 50: 277-323.

Machina M.J. 1989, "Dynamic consistency and non-expected utility models of choice under uncertainty", *Journal of Economic Literature* 27: 1622-1668.

Marschak J. 1954, "Towards an economic theory of organization and information", in R.M. Thrall, C.H. Coombs, R.L. Davis (eds.), *Decision Processes*, New York: Wiley.

Meyer M.A., D. Mookherjee 1987, "Incentives, compensation, and social welfare", *Review of Economic Studies* 54: 209-226.

Mongin P. 1995, "Consistent Bayesian aggregation", *Journal of Economic Theory* 66: 313-351.

Mongin P. 2005, "Spurious unanimity and the Pareto principle", *LSE-CPNSS Working Paper*, vol. 1, no. 5.

Myerson R.B. 1981, "Utilitarianism, egalitarianism, and the timing effect in social choice problems", *Econometrica* 49: 883-897.

Otsuka M., A. Voorhoeve 2007, "Why it matters that some are worse off than others: An argument against Parfit's Priority View", *Philosophy and Public Affairs*, forthcoming.

Pratt J.W., R.J. Zeckhauser 1996, "Willingness to pay and the distribution of risk and wealth", *Journal of Political Economy* 104: 747-763.

Quiggin J. 1982, "A theory of anticipated utility", *Journal of Economic Behavior and Organization* 3: 323-343.

Quiggin J. 1989, "Stochastic dominance in regret theory", *Review of Economic Studies*, 57: 503-511.

Roemer J.E. 2006, "The global welfare economics of immigration", *Social Choice and Welfare* 27: 311-325.

Rothschild R.S., J.E. Stiglitz 1976, "Equilibrium in competitive insurance markets: An essay on the economics of imperfect information", *Quarterly Journal of Economics* 11: 629-649.

Savage L.J. 1972, *The Foundations of Statistics*, 2nd revised ed., New York: Dover Publications.

Tversky A., D. Kahneman 1992, "Advances in prospect theory: cumulative representation of uncertainty", *Journal of Risk and Uncertainty* 5: 297-323.

Viscusi W.K. 1992, *Fatal Tradeoffs: Private and Public Responsibilities for Risk*, Oxford: Oxford University Press.

Wakker P.P. 1993, "Savage's axioms usually imply violation of strict stochastic dominance", *Review of Economic Studies*, 60: 487-493.

Weymark J.A. 1991, "A reconsideration of the Harsanyi-Sen debate on utilitarianism", in J. Elster, J.E. Roemer eds., *Interpersonal Comparisons of Well-Being*, Cambridge: Cambridge University Press.

Weymark J.A. 2005, "Measurement theory and the foundations of utilitarianism", *Social Choice and Welfare* 27: 527-555.

Wilson C.A. 1977, "A model of insurance markets with incomplete information", *Journal of Economic Theory* 16: 167-207.

Appendix A: Additional proofs

Let $\hat{\pi}$ be an arbitrary positive probability vector that differs from π.

Theorem 1: We check that no axiom is redundant by exhibiting an example that satisfies all the axioms but one.

- Drop Weak and Strict Dominance. Let R_1 be defined by $U\ R_1\ U'$ iff $\min_{i\in N} EU_i \geq \min_{i\in N} EU_i'$.
- Drop Weak Pareto for Equal Risk. Let R_2 be defined by $U\ R_2\ U'$ iff $\sum_{i\in N}\sum_{s\in S}\hat{\pi}^s U_i^s \geq \sum_{i\in N}\sum_{s\in S}\hat{\pi}^s U_i'^s$ for some arbitrary $\hat{\pi}$ that differs from π.
- Drop Weak Pareto for No Risk. Let R_3 be defined by $U\ R_3\ U'$ iff either $U' \notin \mathcal{L}^e$ or $U, U' \in \mathcal{L}^e$ and $EU_1 \geq EU_1'$. Note that for R_3 no unequal U^s has an EDQE.

Theorem 2: We check that no axiom is redundant and also test some possible weakenings.

- Drop Weak Dominance. Let R_4 be defined by $U\ R_4\ U'$ iff $(EU_i)_{i\in N} \geq_{\text{lex}} (EU_i')_{i\in N}$.
- Replace Weak Dominance by Strict Dominance: Let R^{lex} denote the ex-post leximin criterion, and R_5 be defined by $U\ R_5\ U'$ iff either $U\ P^{\text{lex}}\ U'$ or $U\ I^{\text{lex}}\ U'$ and $U' \in \mathcal{L}^{nr}$.
- Drop Strong Pareto for Non Reranking Risk. Let R_6 be the ex-post leximin criterion based on some arbitrary $\hat{\pi}$ that differs from π.
- Replace Strong Pareto for Non Reranking Risk by Strong Pareto for Equal Risks and Strong Pareto for No Risk, which are defined as follows:

  Axiom (Strong Pareto for Equal Risk) *For all $U, U' \in L^e$, one has $U\,R\,U'$ if for all $i \in N$, $EU_i \geq EU_i'$; one has $U\,P\,U'$ if in addition, for some $i \in N$, $EU_i > EU_i'$.*

  Axiom (Strong Pareto for No Risk) *For all $U, U' \in L^c$, one has $U\,R\,U'$ if for all $i \in N$, $U_i \geq U_i'$; one has $U\,P\,U'$ if in addition, for some $i \in N$, $U_i > U_i'$.*

  Let $(z_i(U))_{i\in N}$ be defined by

  $$\begin{aligned} z_1(U) &= \sum_{s\in S} \pi^s U^s_{(1)}, \\ z_{i+1}(U) &= z_i(U) + \sum_{s\in S} \hat{\pi}^s \left(U^s_{(i+1)} - U^s_{(i)}\right) \text{ for } i = 1, ..., n-1. \end{aligned}$$

  Let R_7 be defined by $U\ R_7\ U'$ iff $(z_i(U))_{i\in N} \geq_{\text{lex}} (z_i(U'))_{i\in N}$.
- Drop Hammond Equity. Classical utilitarianism is a counterexample.

- Drop Anonymity. Let $\geq_{lexico}$ denote the lexicographic criterion that compares the first component of vectors, and then the second component, and so on. Let R_8 defined by $U\, R_8\, U'$ iff

$$\left(\sum_{s\in S} \pi^s U^s_{(i)}\right)_{i\in N} >_{lex} \left(\sum_{s\in S} \pi^s U'^s_{(i)}\right)_{i\in N} \quad \text{or}$$
$$\left(\sum_{s\in S} \pi^s U^s_{(i)}\right)_{i\in N} =_{lex} \left(\sum_{s\in S} \pi^s U'^s_{(i)}\right)_{i\in N} \quad \text{and } (EU_i)_{i\in N} \geq_{lexico} (EU'_i)_{i\in N}.$$

Lemma 3 *Let C be a convex subset of $\mathbb{R}^n$ and $f : C \to \mathbb{R}$ satisfy: for all $(x^1, ..., x^m) \in C^m$,*

$$f\left(\sum_{s\in S} \pi^s x^s\right) = \sum_{s\in S} \pi^s f(x^s).$$

Then f is affine.

Proof. Let us restrict attention to vectors $(x^1, ..., x^m) \in C^m$ such that $x^2 = ... = x^m$. One then has, for all $x, y \in C$,

$$f(\pi^1 x + (1 - \pi^1)\, y) = \pi^1 f(x) + (1 - \pi^1)\, f(y).$$

Let $\lambda \in (0, 1)$ and $z = \lambda x + (1 - \lambda)\, y$. One constructs a sequence in C that converges to z as follows: $s_1 = x$, $s_2 = y$, $s_3 = \pi^1 x + (1 - \pi^1)\, y$, and for $t \geq 4$, $s_t = \pi^1 s^*_t + (1 - \pi^1)\, s^{**}_t$, where s^*_t and s^{**}_t are the two adjacent members of $\{s_1, ..., s_{t-1}\}$ such that z belongs to the line segment joining them. As f is continuous, $f(s_t)$ tends to $f(z)$.

Assume that up to $t - 1$, writing $s_k = \alpha_k x + (1 - \alpha_k)\, y$ one has $f(s_k) = \alpha_k f(x) + (1 - \alpha_k)\, f(y)$. This is true for $t = 4$. Writing $s^*_t = \alpha^*_t x + (1 - \alpha^*_t)\, y$, $s^{**}_t = \alpha^{**}_t x + (1 - \alpha^{**}_t)\, y$, one then has $f(s^*_t) = \alpha^*_t f(x) + (1 - \alpha^*_t)\, f(y)$, $f(s^{**}_t) = \alpha^{**}_t f(x) + (1 - \alpha^{**}_t)\, f(y)$ and

$$\begin{aligned} f(s_t) &= \pi^1 f(s^*_t) + (1 - \pi^1)\, f(s^{**}_t) \\ &= \pi^1 \left[\alpha^*_t f(x) + (1 - \alpha^*_t)\, f(y)\right] + (1 - \pi^1) \left[\alpha^{**}_t f(x) + (1 - \alpha^{**}_t)\, f(y)\right] \\ &= \left[\pi^1 \alpha^*_t + (1 - \pi^1)\, \alpha^{**}_t\right] f(x) + \left[\pi^1 (1 - \alpha^*_t) + (1 - \pi^1)(1 - \alpha^{**}_t)\right] f(y). \end{aligned}$$

As one has $s_t = \alpha_t x + (1 - \alpha_t)\, y$ for $\alpha_t = \pi^1 \alpha^*_t + (1 - \pi^1)\, \alpha^{**}_t$, this shows that $f(s_t) = \alpha_t f(x) + (1 - \alpha_t)\, f(y)$.

The fact that s_t tends to z implies that α_t tends to λ, and one obtains that $f(z) = \lambda f(x) + (1 - \lambda)\, f(y)$. By Coulhon and Mongin (1989, Lemma, p. 183), this implies that f is affine. ■

Theorem 3: We check that no axiom is redundant and also test a possible weakening.

- Drop Weak or Strict Dominance. Let R_9 be defined by $U \; R_9 \; U'$ iff $\sum_{i \in N} \varphi\left(EU_i\right) \geq \sum_{i \in N} \varphi\left(EU'_i\right)$, where φ is an increasing, continuous, strictly concave function.

- Drop Strong Pareto for Non Reranking Risk. Let R_{11} be an ex-post generalized Gini criterion based on some arbitrary $\hat{\pi}$ that differs from π.

- Replace Strong Pareto for Non Reranking Risk by Strong Pareto for Equal Risks and Strong Pareto for No Risk. Let R_{12} be defined by $U \; R_{12} \; U'$ iff $\sum_{i \in N}(n-i) z_i(U) \geq \sum_{i \in N}(n-i) z_i(U')$, where $z_i(U)$ has been introduced in the definition of R_7.

- Drop Anonymity. Let R_{13} defined by $U \, R_{13} \, U'$ iff $\sum_{i \in N}(n-i) EU_i > \sum_{i \in N}(n-i) EU'_i$.

- Drop Continuity. Take R^{lex}.

Lemma 4 *Let $f : \mathbb{R} \to \mathbb{R}$ satisfy: there is $\alpha \in (0,1)$ such that for all $x, y \in \mathbb{R}$, all $\varepsilon > 0$,*

$$f\left(\alpha x + (1-\alpha) y + \varepsilon\right) \geq \alpha f(x) + (1-\alpha) f(y) \geq f\left(\alpha x + (1-\alpha) y - \varepsilon\right).$$

Then f is affine.

Proof. Let $g = f - f(0)$. Letting $y = 0$, one obtains

$$\begin{aligned} f\left(\alpha x + \varepsilon\right) &\geq \alpha f(x) + (1-\alpha) f(0) \geq f\left(\alpha x - \varepsilon\right), \\ f\left(\alpha x + \varepsilon\right) - f(0) &\geq \alpha \left(f(x) - f(0)\right) \geq f\left(\alpha x - \varepsilon\right) - f(0), \\ g\left(\alpha x + \varepsilon\right) &\geq \alpha g(x) \geq g\left(\alpha x - \varepsilon\right), \end{aligned}$$

and similarly, one has

$$g\left((1-\alpha) x + \varepsilon\right) \geq (1-\alpha) g(x) \geq g\left((1-\alpha) x - \varepsilon\right).$$

Letting $u = \alpha x$, $v = (1-\alpha) y$, one writes

$$\begin{aligned} f\left(u + v + \varepsilon\right) - f(0) &\geq \alpha \left(f(x) - f(0)\right) + (1-\alpha)\left(f(y) - f(0)\right), \\ g\left(u + v + \varepsilon\right) &\geq \alpha g(x) + (1-\alpha) g(y), \end{aligned}$$

and, as for all $\varepsilon > 0$ one has $\alpha g(x) \geq g\left(\alpha x - \varepsilon\right)$, $(1-\alpha) g(x) \geq g\left((1-\alpha) x - \varepsilon\right)$, one therefore has for all $\eta_1, \eta_2 > 0$

$$\alpha g(x) + (1-\alpha) g(y) \geq g\left(u - \eta_1\right) + g\left(v - \eta_2\right),$$

implying

$$g\left(u+v+\varepsilon\right) \geq g\left(u-\eta_1\right)+g\left(v-\eta_2\right).$$

Similarly,

$$g\left(u+v-\varepsilon\right) \leq g\left(u+\eta_1\right)+g\left(v+\eta_2\right).$$

Letting $a=u+\varepsilon/2$, $b=v+\varepsilon/2$, $\delta_i=\eta_i+\varepsilon/2$, one obtains that for all $a,b\in\mathbb{R}$ and $\delta_1,\delta_2>0$,

$$g\left(a+b\right) \geq g\left(a-\delta_1\right)+g\left(b-\delta_2\right),$$

and letting $a=u-\varepsilon/2$, $b=v-\varepsilon/2$, $\delta_i=\eta_i+\varepsilon/2$, one obtains for all $a,b\in\mathbb{R}$ and $\delta_1,\delta_2>0$,

$$g\left(a+b\right) \leq g\left(a+\delta_1\right)+g\left(b+\delta_2\right).$$

Letting $a-\delta_1=b-\delta_2=z$, one sees that $g(z)\leq g(z+\varepsilon)$ for all $z\in\mathbb{R}$ and $\varepsilon>0$. That is, g is nondecreasing. Therefore it has at most a countable number of discontinuity points. At all points $x,y\in\mathbb{R}$ where it is continuous, one has

$$g(x+y)=g(x)+g(y). \tag{5}$$

Suppose that g is not continuous and let u be a point of discontinuity, i.e., such that $g(u^-)<g(u^+)$. As g is nondecreasing, this means that there is no x such that $g(x)\in\left(g(u^-),g(u^+)\right)$. For all a,b close enough to u and such that $a<u<b$, one has $\left(g(a)+g(b)\right)/2\in\left(g(u^-),g(u^+)\right)$. Therefore, there must be $p<p'<u<q<q'$ such that g is continuous at p,q,p',q' and $\left(g(p)+g(q)\right)/2,\left(g(p')+g(q')\right)/2\in\left(g(u^-),g(u^+)\right)$. Let $t\in\left(\left(p+q\right)/2,\left(p'+q'\right)/2\right)$ be a point where g is continuous. By (5) one has $g(2t)=2g(t)$ and by (5) and the fact that g is nondecreasing, one has

$$g(p)+g(q)=g(p+q)\leq g(2t)\leq g(p'+q')=g(p')+g(q').$$

Dividing by 2, one obtains

$$\frac{g(p)+g(q)}{2}\leq g(t)\leq\frac{g(p')+g(q')}{2},$$

implying that $g(t)\in\left(g(u^-),g(u^+)\right)$, a contradiction. Therefore g is continuous.

This implies that g satisfies Cauchy's equation (5) for all $x,y\in\mathbb{R}$ and, being continuous, is therefore linear, implying that f is affine. ■

Theorem 4: We check that no axiom is redundant.

- Drop Weak and Strict Dominance. Take R_1.
- Drop Weak Pareto for Subgroup Equal Risk. Take R_2.

Appendix B: Dominance and the sure-thing principle

In decision theory, Weak Dominance is often called "monotonicity" (Anscombe and Aumann 1963) or "statewise dominance" (Quiggin 1989). It is worth clarifying the relation between statewise dominance and the sure-thing principle because the latter is often confused with a separability condition. In Savage's theory (Savage 1972), an act **f** is a function that associates every state of the world $s \in S$ with a consequence $f(s)$. Savage defines the sure-thing principle as follows. Take any arbitrary event E (i.e., a subset of the set S of states of the world). If the decision-maker weakly prefers act **f** to act **g** conditionally on event E being realized, and also conditionally on event E not being realized, then he must weakly prefer **f** to **g**. If a businessman considers it worthy to buy a piece of property on the assumption that the Democrats win the next election, and also considers it worthy on the assumption that the Republicans win, then he should buy it. Savage writes that he knows of "no other extralogical principle governing decisions that finds such ready acceptance" (p. 21).

Statewise dominance is weaker than the sure-thing principle because it applies only in the case when, for all $s \in S$, the decision-maker weakly prefers **f** to **g** conditionally on the singleton event $E = \{s\}$ being realized. Here is an example of a rule that is not absurd and satisfies statewise dominance but not the sure-thing principle. When the probability of death – or, more rigorously, of a lower level of utility than a certain critical level – exceeds a threshold, the decision-maker seeks to minimize this probability. Otherwise (and in case of ties for the over-threshold probability of death), he maximizes expected utility. This rule satisfies statewise dominance because when an act weakly dominates another for all states of the world, it has either a lower probability of death or an equal probability of death combined with a weakly greater expected utility. The rule violates the sure-thing principle because knowing that an event E will occur may raise the (conditional) probability of death above the threshold and reverse the preference between two acts that coincide on $S \setminus E$. That is, this rule may prefer the higher expected-utility act **f** to the lower probability-of-death act **g** even though: *i*) it prefers **g** conditionally on E being realized, and *ii*) it is indifferent between **f** and **g** conditionally on E not being realized.

Another difference between statewise dominance and Savage's theory is that, instead of introducing conditional preferences directly, Savage chooses to formalize the sure-thing principle with his P2 axiom which incorporates an additional separability requirement in order to simplify the expression of conditional preferences.

P2: If **f** is weakly preferred to **g**, $f(s) = f'(s)$ for all $s \in E$, $g(s) = g'(s)$ for all $s \in E$, $f(s) = g(s)$ for all

$s \notin E$, $f'(s) = g'(s)$ for all $s \notin E$, then $\mathbf{f}'$ is weakly preferred to $\mathbf{g}'$.

This is a separability axiom implying that preference conditional on E being realized does not depend on what happens if E is not realized. It makes it possible to derive conditional preference on E unambiguously from preference over acts that coincide on $S \setminus E$.

P2 implies the sure-thing principle by an iteration argument.[21] The sure-thing principle, however, does not require separability in itself. The maximin criterion is an example of a (not absurd) rule that satisfies the sure-thing principle but not P2.[22]

Statewise dominance is also weaker than first-order stochastic dominance, since the latter additionally incorporates the possibility to compare acts over different states of the world (e.g., assuming that $S = \{s, s'\}$ and s and s' are equiprobable, if $f(s)$ is better than $g(s')$ while $f(s')$ is better than $g(s)$, then first-order stochastic dominance applies while statewise dominance may be silent).[23]

The non-expected utility theories failing to satisfy statewise dominance or first-order stochastic dominance have been criticized (see, e.g., Handa 1977 and Fishburn 1978, or Kahneman and Tversky

---

[21] If $\mathbf{f}$ is weakly preferred to $\mathbf{g}$ conditionally on E and conditionally on not E, under P2 this means that

$$\mathbf{f} = \left\{ \begin{array}{c} f(s) \text{ for } s \in E \\ f(s) \text{ for } s \notin E \end{array} \right\} \text{ is weakly preferred to } \left\{ \begin{array}{c} g(s) \text{ for } s \in E \\ f(s) \text{ for } s \notin E \end{array} \right\}$$

and

$$\left\{ \begin{array}{c} g(s) \text{ for } s \in E \\ f(s) \text{ for } s \notin E \end{array} \right\} \text{ is weakly preferred to } \left\{ \begin{array}{c} g(s) \text{ for } s \in E \\ g(s) \text{ for } s \notin E \end{array} \right\} = \mathbf{g},$$

implying, by transitivity, that $\mathbf{f}$ is weakly preferred to $\mathbf{g}$.

[22] The next axiom in Savage's list relies on the notion of conditional preferences and requires consistency between global and conditional preferences over the subset of constant acts:

P3: If $f(s) = a$ for all $s \in S$, $g(s) = b$ for all $s \in S$, and if preference conditional on E is not universal indifference, then $\mathbf{f}$ is weakly preferred to $\mathbf{g}$ if and only if $\mathbf{f}$ is weakly preferred to $\mathbf{g}$ conditionally on E being realized.

An alternative formulation of P3, that does not incorporate P2, can be found in the literature. For instance, in Grant (1995, p. 164) one finds:

P3* If $f(s) = a$ for all $s \in E$, $g(s) = b$ for all $s \in E$, $f(s) = g(s)$ for all $s \notin E$, $f'(s) = a$ for all $s \in S$, $g'(s) = b$ for all $s \in S$, and if preference on acts that coincide on $S \setminus E$ is not universal indifference, then $\mathbf{f}$ is weakly preferred to $\mathbf{g}$ if and only if $\mathbf{f}'$ is weakly preferred to $\mathbf{g}'$.

This axiom, in its "only if" part, contains a weaker separability requirement than P2, as analyzed by Grant in terms of independence over lotteries. If one restricts P3* to its "if" part, it implies statewise dominance but not the sure-thing principle.

[23] For instance, as explained in Quiggin (1989), regret theory violates first-order stochastic dominance but satisfies statewise dominance.

1979, p. 284), and the more recent theories (e.g., Quiggin 1982, Machina 1982, Tversky and Kahneman 1992) do respect such dominance principles.[24] Interestingly, Grant (1995) notes the violation of statewise dominance in Diamond's critique for social evaluation and considers that it can be used to question statewise dominance in individual decision theory, invoking Machina's (1989) example of a mother who must give an indivisible "treat" to her daughter Abigail or her son Benjamin and seeks to be fair by the use of a lottery. This example is indeed a direct adaptation of Diamond's argument. We have assumed in this paper that fairness considerations can justify the use of lotteries but not a violation of dominance. There is an obvious difference between "giving the treat to Abigail, being unfair to Benjie" and "giving it to Abigail after a fair lottery" which can be described in terms of ex-post consequences and does not require such a strong violation of rationality.

[24] Note that, unlike the weak version, the strict version of statewise dominance (i.e., an act is strictly better if it strictly dominates in all states of nature) is not even satisfied by Savage's expected utility theory in some special contexts (finitely additive probabilities—see Wakker 1993, p. 491).