

# Fairness and Risk attitudes

Richard Bradley<sup>1</sup> & H. Orri Stefánsson<sup>2</sup>

**Abstract.** According to a common judgement, a social planner should often use a lottery to decide which of two people should receive a good. This judgement undermines one of the best-known arguments for utilitarianism, due to John C. Harsanyi, and more generally undermines axiomatic arguments for utilitarianism and similar views. In this paper we ask which combinations of views about (a) the social planner’s attitude to risk and inequality, and (b) the subjects’ attitudes to risk are consistent with the aforementioned judgement. We find that the class of combinations of views that can plausibly accommodate this judgement is quite limited. But one theory does better than others: the theory of chance-sensitive utility.

## 1. Introduction

Imagine being in charge of matching patients with donated kidneys. Unfortunately, there is only a single kidney available, but you have two patients who are in equal need of the kidney, who would benefit equally from it, and who have waited equally long for a kidney. More generally, you take the two patients to have an equal claim to the kidney. How would you decide who gets the kidney? A common thought is that you should hold a fair (50-50) lottery to decide who gets the kidney. The idea is that although it would be equally good that one patient gets the kidney as that the other patient gets it, it is better if the person who doesn’t get the kidney had a fair chance of getting it than that they did not.

This judgement in favour of using lotteries extends to indivisible goods more generally (see, e.g., Stone 2007). The idea is that an impartial, or fair-minded, distributor of goods—an impartial “social planner”—would, whenever two people have an equal claim to some good  $G$ , prefer to hold a lottery to decide who gets  $G$ , rather than giving  $G$  to either person without holding a lottery.

Natural as this idea in favour of using lotteries seems to be, it turns out to have very important implications for theories of distributive ethics. In fact, an example capturing this idea was used by Peter Diamond (1967) to question John C. Harsanyi’s (1955) argument for utilitarianism, and has ever since been one of the greatest obstacles to deriving utilitarianism, and other theories of a similar kind, from the theory of rational decision-making under uncertainty. The problem is that any such derivation has to assume a condition known as *State Dominance*, which implies that if two alternatives lead to equally good outcomes no matter how the world turns out, then the two alternatives are equally good. But that is inconsistent with the

---

<sup>1</sup> London School of Economics and Political Science, [r.bradley@lse.ac.uk](mailto:r.bradley@lse.ac.uk).

<sup>2</sup> Stockholm University, Swedish Collegium for Advanced Study, and Institute for Futures Studies, [orri.stefansson@philosophy.su.se](mailto:orri.stefansson@philosophy.su.se).

judgement in favour of the lottery—a judgement we below refer to as “Diamond Fairness”—as illustrated by Table 1.

	State 1	State 2
Fair	Person 1 gets $G$	Person 2 gets $G$
Unfair	Person 1 gets $G$	Person 1 gets $G$

Table 1: Diamond Fairness

Let state 1 and 2 be two equiprobable “states of the world”, that is, resolutions of uncertainty, that in this case determine which person gets  $G$  when the lottery (“Fair”) is chosen. Now, if the outcome where person 1 gets  $G$  is equally good as the outcome where person 2 gets  $G$ , then by State Dominance, the two alternatives, Fair and Unfair, are equally good. But they are not, according to Diamond Fairness: Fair is better than Unfair. So, Diamond Fairness violates State Dominance. This means that it is not consistent with many theories of the social good.

A common response to Diamond’s objection is that if a lottery has benefits of some kind, then we should include them in the descriptions of the outcomes of the lottery (e.g., Broome 1991). For instance, if the lottery is preferable because it is procedurally fair, then we should include in the descriptions of the outcomes of the lottery that a person gets  $G$  by means of a fair procedure (just as we would distinguish obtaining  $G$  by honest means from obtaining it by dishonest ones). Doing so would make the preference for the lottery consistent with State Dominance. At the same time, it would make it possible to account for, say, the harms of being treated unfairly in the measure of individual wellbeing (utility) at each state of the world, thus making the preference for the lottery formally consistent with utilitarianism.

There is a core of truth to this objection: accommodation of Diamond’s insight requires careful individuation of outcomes. But this observation does not by itself resolve the question of how precisely outcomes should be re-described, and in fact whether they can be re-described in a way that accommodates the judgement in question without resulting in contradictions. We shall get back to that issue later. For now, let us assume that Diamond was right in his critique of Harsanyi’s utilitarianism; that is, assume that the value of giving people a “fair shake” (Diamond 1967: 766) means that the social planner should violate State Dominance in situations like those described above. The general question this paper seeks to answer is: What follows for theories of individual and social good under uncertainty?

So far, those who have explored similar questions have done so by merely relaxing State Dominance as a requirement on the social planner’s preference, or, to put it differently, as a requirement of social welfare or “betterness” judgements. This results in distributive theories or judgements that are averse to *ex ante* inequality, that is, averse to inequality in the distributions of people’s *expected* good (or chances for good). However, these relaxations have assumed that the betterness-rankings for individuals whose good the social planner is concerned with satisfy the axioms of expected utility theory (see, e.g., Epstein and Segal 1992 and Grant et. al. 2010).

Our aim is more radical than the aforementioned relaxations. We explore the consequences of allowing non-standard risk attitudes *both* (and simultaneously) on the part of individuals and the social planner; and, in particular, we consider which combinations of such theories can accommodate Diamond Fairness. In other words, we consider giving up standard expected utility theory for both the social planner and for her subjects, and we allow that the latter can have different attitudes to risk. In the present context, one important implication of alternatives to expected utility theory for individuals' preferences is that they make it (in principle) possible for a social planner to satisfy Diamond Fairness while aggregating people's *ex ante* good in, say, a utilitarian manner—that is, even though the planner's preferences between social prospects corresponds to the sum total of *ex ante* good that the prospects offer individuals. In other words, it makes possible a version of utilitarianism that, unlike Harsanyi's, is immune to Diamond's objection.

To see this, note that since these alternatives to expected utility theory allow that the value for individuals of a lottery may differ from its expected utility, they make room for the possibility that the value to a person of a 0.5 chance of a good  $G$  is greater than 0.5 times the ultimate value (or utility) of  $G$  to that person. For instance, getting a half chance at a kidney can be more valuable, to a person, than half the value of getting the kidney. That would mean that a social planner who needs to decide who of two patients gets a kidney, and wants to make the decision that maximises good, may strictly prefer to use a lottery to decide who gets the kidney rather than giving the kidney to either patient without holding such a lottery. More generally, such a social planner may satisfy Diamond Fairness even though their aggregation schema is (*ex ante*) utilitarian. Later we shall see that the same holds for many other families of *ex ante* aggregation schemes.

So, these rivals to expected utility theory may offer the possibility of accounting for Diamond Fairness in terms of individuals' risk attitudes. In what follows we look more closely at precisely how they do so, and ask whether the account they offer is plausible. As we shall see, the answer is mixed. We focus initially on the family of rank-dependent utility theories (see, e.g., Buchak 2013),<sup>3</sup> since these are arguably the most commonly defended *normative* (as opposed to descriptive) alternatives to expected utility theory. We show that these views typically imply that a concern for people's *ex ante* good will not result in Diamond Fairness.

Moreover, we show that the only way for these rank-dependent views to accommodate Diamond Fairness is to make some quite specific and seemingly *ad hoc* assumptions about the relationship between a social planner's attitudes and her subjects' attitudes. So, the potential justification of Diamond Fairness that rank-dependent views offer is not robust across small changes in the risk and inequality attitudes of the social planner and her subjects. Finally, the potential justification in question can only use rank-dependent utility theory for *individuals'* prospects, not for social prospects, since the theory would then imply State Dominance for social prospects. So, rank-dependent utility theorists who want to accommodate Diamond Fairness can only use their favourite theory to navigate *some* types of uncertainty.

---

<sup>3</sup> The theory is called “risk-weighted expected utility theory” by Buchak (2013), but its formal predecessor (Quiggin 1982) is typically called “rank-dependent utility theory”.

In contrast, it turns out that the theory of chance-sensitive utility developed in Stefánsson and Bradley (2015, 2019) for individual risk attitudes *can* deliver a robust justification of Diamond Fairness that only uses assumptions that have been independently justified. The distinguishing feature of chance-sensitive utility theory is that it includes “chance propositions”, that is, descriptions of chance distributions, in the domain of the preference relation. This opens the door to two natural (and in fact mutually consistent) explanations of Diamond Fairness. First, given the account of *ambiguity aversion* in Bradley (2016) and Stefánsson and Bradley (2019),<sup>4</sup> it follows that a social planner who is not inequality seeking<sup>5</sup> has preferences that align with Diamond Fairness. Second, given that chances are in this framework included as goods about which both individuals and the social planner can have non-instrumental attitudes, a social planner who is concerned with equality can have preferences that align with Diamond Fairness, independently of (any concern for) subjects’ attitudes to ambiguity. Finally, Diamond Fairness might, in this framework, result due to a combination of concern for people’s ambiguity averse preferences and aversion to inequality. So, the explanation of Diamond Fairness within this framework is both quite robust across different assumptions about the risk and inequality attitudes of a social planner and her subjects, and moreover is implied by other uses to which the theory has been put.

The rest of the paper is structured as follows. We start (section 2) by informally laying out the framework we use in this paper, as well as stating the conditions we will be discussing and some initial results. (More formal statements and results can be found in the Appendix.) Our first result is that Diamond Fairness is characterised by a simple relationship between individuals’ sensitivity to risk and the social planner’s sensitivity to inequality. This gives rise to the question: what combinations of theories of individuals’ risk attitudes and social inequality attitudes can accommodate this relationship? In section 3 we start answering this question by laying out the main theories of individuals’ risk attitudes. We next (in section 4) explain why the family of rank-dependent views has great difficulties in accommodating the aforementioned relationship, and thus cannot plausibly account for Diamond Fairness. After that (in section 5) we show that chance-sensitive utility theory does better in this regard, and finally conclude (in section 6) with some remarks on the moral implications of heterogenous risk attitudes.

## 2. Framework and conditions

Our concern in this paper is with the question of what makes distributions of outcomes to individuals in different states of the world better or worse. A distribution of outcomes to individuals in different states of the world will be called a *prospect*: an *individual* prospect when the outcomes are assigned to just one

---

<sup>4</sup> To be ambiguity averse is to prefer prospects with known probabilities over prospect with unknown ones (other things being sufficiently equal), and, more generally, to prefer a smaller rather than a greater spread in the epistemically possible probabilities of prospects’ outcomes (again, other things being sufficiently equal).

<sup>5</sup> This means that if  $X$  offers both a higher total good and a more equally distributed good than prospect  $Y$ , then  $X$  is deemed better than  $Y$ . This is of course implied by any plausible distributive view, such as utilitarianism, egalitarianism, and prioritarianism, and would thus seem justified independently of any goal to accommodate Diamond Fairness.

individual, a *social* prospect when they are assigned to more than one. A social prospect thus assigns an individual prospect to each individual and a social outcome to each state of the world, a social outcome being a distribution of outcomes to individuals. Table 1 for instance displays two social prospects. In one of these prospects (Fair) the outcome in state 1 for Person 1 is that they get the good and for Person 2 the outcome is that they do not get the good, while in state 2 the outcome for Person 1 is that they don't get the good and for Person 2 the outcome is that they do. In the other prospect in table 1 (Unfair) the outcome in both states for Person 1 is that they get the good and for Person 2 the outcome is that they do not.

In what follows, we will talk about prospects as being better or worse both for individuals and overall, but we will not put forward any substantive theory of what makes outcomes or prospects better or worse for individuals. In general, this will depend both on properties of the outcomes and the probability with which they are yielded and on characteristics of the individuals and of their circumstances; for instance, on whether their circumstances make them more or less vulnerable to various kinds of risks. Plausibly the relevant characteristics of an individual will include their well-informed, self-regarding preferences. But insofar as their actual preferences reflect, say, false beliefs (e.g., about the risks they face) or their concern for others, they need not track what is better or worse for her.<sup>6</sup>

In line with standard practice in philosophy, when we talk about prospects being better or worse (full stop) we mean *morally* or all-things-considered better or worse. This can be thought of in terms of the preference relation of the ideal social planner who always prefers a morally better prospect to a worse one. The main focus of our paper is on the ethical condition proposed by Diamond that is intended as a part of an answer to the above question of what makes distributions of outcomes to individuals in different states of the world better or worse in this sense. The condition can be informally stated as:

**Diamond Fairness:** Suppose that some good  $G$  is to be distributed between two people and that it is no better that one of them gets  $G$  than that the other does.<sup>7</sup> Then making the distribution by means of a lottery that gives each person an equal chance at getting  $G$  is better than simply giving either person  $G$  (without letting a lottery decide).

Suppose that Diamond is right about this. What kind of theory of what is good (*ex ante*) for individuals and of what is socially good (good overall) can explain it? In this section we take the first step towards answering this question by giving a characterisation of Diamond Fairness. In particular, we characterise Diamond Fairness in terms of the relationship between, on the one hand, the sensitivity of an individual's good to

---

<sup>6</sup> Whether or not interpersonal comparisons of preference strength is meaningful is a hotly contested issue. But since comparisons of how *good* prospects are for different individuals depend on objectively measurable factors (the relevant properties of the prospect and of the individual), we don't think that there is any problem in principle of doing so, however difficult it may be in practice. (We thank a referee for making us see the need to clarify this.)

<sup>7</sup> For a utilitarian this would be the case when the two people benefit equally from  $G$ . Those with prioritarian or egalitarian (*ex post*) intuitions might want to complicate this claim however. For instance, a prioritarian might say that the fair lottery should be used when the two people would receive an equal *priority weighted* benefit from  $G$ , whereas an egalitarian might want to say that Diamond Fairness should (at least) be respected as long as  $G$  doesn't affect the rank-order of the population. As these further complications are irrelevant to our argument, we shall set them aside.

the risk they face, and, on the other hand, the sensitivity of the social good to inequalities in the goods distributed to individuals.

First however we need to introduce the objects and conditions with which we will be concerned. (A more precise formulation of everything in this section can be found in the Appendix.) Let  $G$  be any good that comes in different quantities (e.g., money, years in good health, etc). Consider a set of states of the world  $S$  and a set  $I$  of individuals. Then a social prospect  $X$  is as an assignment to each individual  $i$  and state of world  $s$  a quantity  $X(i, s)$  of  $G$ . Correspondingly  $X(s)$  is the social outcome determined by  $X$  in state of world  $s$ , and  $X(i)$  the individual prospect assigned to individual  $i$  by social prospect  $X$ .

The following two, often assumed conditions on betterness between prospects will be important in what follows:

**State Dominance:** For any two prospects  $X$  and  $Y$ , if in each state of the world, the outcome  $X(s)$  is at least as good as the outcome  $Y(s)$ , then  $X$  is at least as good as  $Y$  overall.<sup>8</sup>

**Ex Ante Pareto:** For any two social prospects,  $X$  and  $Y$ , if for each person  $i$ , the individual prospect  $X(i)$  is at least as good as (strictly better than) the individual prospect  $Y(i)$ , then  $X$  is at least as good as (strictly better than)  $Y$  overall.

In line with standard terminology, let the *certainty equivalent* of a probability distribution of a quantifiable good (say money or wellbeing) across different states of the world (i.e., of a prospect) be the quantity of that good that is equally good as the probability distribution. So, for instance, if a 50-50 gamble between \$10 and \$0 is, for some individual, as good as \$4 for sure, then \$4 is the certainty equivalent to the gamble in question, for that individual. The *equality equivalent* of a distribution of some good is the amount of that good such that the distribution is equally good as one where every person gets that amount of the good. For instance, if a social planner judges that a distribution where one person has wellbeing level 100 and another person has wellbeing level 10 is equally good as one where both people are at level 50, then 50 is the equality equivalent to the first distribution, according to this social planner.

A social betterness relation is said to be *inequality neutral* with respect to some good just in case it implies that the equality equivalent to a distribution of quantities of that good to individuals is equal to the mean value of the distribution, *inequality averse* just in case the implied equality equivalent is less than the mean value of the distribution, and *inequality seeking* otherwise.<sup>9</sup> For instance, a social betterness relation is

---

<sup>8</sup> Often the following condition is added to the statement of State Dominance:

if, in addition, there is some state of the world in which the outcome of  $X$  is strictly better than the outcome of  $Y$ , then  $X$  is strictly better than  $Y$ .

The above addition is however not needed for our argument.

<sup>9</sup> What we introduce here is a rather standard definition of *general* inequality aversion. As a referee rightly notes, this definition is insensitive to the fact that someone might be averse to inequality only when, say, very badly off people are involved. Although we acknowledge that this general definition ignores subtleties that often are important when characterising betterness relations in terms of their attitude to inequality, we shall ignore these subtleties here, as they do not affect our main argument.

inequality seeking if it judges that a distribution where one person has wellbeing level 10 and another person has wellbeing level 100 is better than one where both have wellbeing level 60.

In similar fashion, a betterness relation is said to be *risk neutral* with respect to quantities of some good just in case it implies that the certainty equivalent to a probability distribution over quantities of that good is the same as the probability distribution's expectation of that good, *risk averse* just in case it implies that the certainty equivalent to a probability distribution is less than the probability distribution's expectation, and *risk seeking* otherwise. So, an example of risk aversion with respect to money would be that \$4 for sure is better than a 50-50 gamble between \$0 and \$10.

We are now in a position to address the question of what kind of theory of the social good might countenance Diamond Fairness. It should be evident that no theory that is what we might call “thoroughly<sup>10</sup> *ex post*” can do so, where a theory is thoroughly *ex post* if it evaluates social prospects by, first, determining the value of each social outcome in a way that is insensitive to the chance of that outcome, and then aggregates these values somehow (not necessarily in a utilitarian fashion). To see why this is so, recall first the simple illustration of Diamond Fairness, where what is at stake is some indivisible good  $G$  that two people get equal benefit from. Since it would, by assumption, be equally good that one person gets  $G$  as the other person getting it, a theory that first determines the value of each person receiving  $G$  without accounting for the chance that that person had of receiving  $G$ , and then does the aggregation to determine the value of a non-trivial social prospect that gives both people a chance at receiving  $G$ , must find that this non-trivial social prospect is as good as a trivial “prospect” that gives one person  $G$  for sure. Similarly, if wellbeing in a state is insensitive to the chance of that state, then an *ex post* social evaluation of prospects for wellbeing cannot accommodate Diamond Fairness.

It is not just *ex post* theories that cannot satisfy Diamond Fairness: as we saw in the introduction, no theory that respects State Dominance can do so, unless the value of (and hence the description of) an outcome in a state is sensitive to the chance of that outcome. So, let's instead work with social judgements that respect *ex ante* Pareto with respect to the underlying theory of individual good without assuming State Dominance. Assume initially that individuals have the same certainty equivalent  $c$  for the prospects they face in Fair (i.e., the lottery that gives them an equal chance at the good). And let  $e$  denote the social planner's equality equivalent to Unfair (that is, the alternative that gives one person the good for sure). Finally, let's call the view that the lottery (in Diamond Fairness) is strictly worse than giving  $G$  to either person outright

---

Similar remarks apply to the standard definition of (general) risk aversion that we use. For some purposes, it is for instance important to specify the range of outcomes with respect to which a person is risk averse, since people are often risk averse when it comes to some outcomes but not others. We can however set such complexities aside for now.

<sup>10</sup> Why “thoroughly”? Because a theory could be said to be an *ex post* theory if it evaluates social prospects by, first, determining the value of each outcome *in a way that can be sensitive to the chance of the outcome*, and then aggregates these values. Such a theory could accommodate Diamond Fairness.

“Diamond *Unfairness*”. Then, as demonstrated in the Appendix (Theorems 2 and 3), the following relationships hold between  $c$  and  $e$ :

- a) Social betterness satisfies Diamond Fairness just in case  $e < c$ , that is, just in case social betterness is more inequality averse than individual betterness is risk averse.
- b) Social betterness satisfies Diamond Unfairness just in case  $c < e$ , that is, just in case social betterness is less inequality averse than individuals are risk averse.
- c) Social betterness satisfies State Dominance just in case  $e = c$ , that is, just in case it satisfies neither Diamond Fairness nor Unfairness.

In other words, given *ex ante* Pareto and homogeneity in individual’s risk attitudes, **Diamond Fairness is equivalent to greater inequality aversion in social betterness than risk aversion in individual betterness.** (In a moment we’ll see that a similar, but slightly more complicated, relationship holds even if we do not assume homogeneity in individual’s risk attitudes.)

To illustrate, consider a case in which \$100 can be given to either Ann or Bob (Unfair), or it can be distributed by a lottery which confers both a 50% chance of getting the \$100 (Fair). Suppose Ann and Bob are both risk neutral with respect to money. Then by definition the certainty equivalent for both of (their prospects in) Fair is receiving \$50. But if social betterness is inequality averse in money then the equality equivalent of Unfair will be strictly less than \$50. So Fair must be better than the equality equivalent of Unfair on pain of violating *ex ante* Pareto, and hence than Unfair itself. In contrast, if social betterness was inequality neutral say, but individual betterness risk averse, then Unfair would be socially better than Fair. For Ann and Bob’s certainty equivalent for Fair would in this case be less than \$50 while the equality equivalent of Unfair would be exactly \$50. More generally, Fair would be better than Unfair if and only if the welfare effects of subjecting Ann and Bob to risk were offset by the benefit (from the social perspective) of equalising their individual wellbeing gains.

The same considerations apply even when individuals have different risk attitudes and hence different certainty equivalents for the Fair lottery. But the relationship between them and the equality equivalent for the Unfair lottery that characterises the conditions under which social betterness respects Diamond Fairness is then a bit more complicated, namely:

- a) If the equality equivalent of the lottery is less than the certainty equivalent of the most risk averse individual, (i.e.,  $e < \min [c_i]$ ), then social betterness satisfies Diamond Fairness.
- b) If the equality equivalent of the lottery is greater than the certainty equivalent of the most risk seeking individual, (i.e.,  $e > \max [c_i]$ ), then social betterness satisfies Diamond Unfairness.

These relationships are special cases of the more general results proven in the Appendix as Theorem 3, which hold for any lottery, and not just for the equal-weighted ones.

Whether or not Diamond Fairness can be satisfied in cases in which  $e$  lies between the different  $c_i$  will depend on the details of how individual good should be aggregated. For the most part therefore we will restrict ourselves to claims that do not depend on such details.

### 3. Theories of risk

We have seen that Diamond Fairness is characterised by a simple relationship between the sensitivity of individual goodness to risk and the sensitivity of social goodness to inequality. This gives rise to the question: what combination of theories of individual and social good can explain why this relationship does or should hold? In this section we make a start on answer this question by laying out the main alternative theories of risk. The first two of these, expected utility theory and rank-dependent utility theory, only require that we use two sets over which the preference (or “betterness”) relation ranges: a set of outcomes (or “prizes”) and a set of probability distribution over these outcomes (i.e., “lotteries”), the latter of which are the analogues to prospects when the relevant probabilities are known (or given by the decision situation).

The most widely used theory of rational choice between the risky alternatives picked out by lotteries is the expected utility theory of von Neumann and Morgenstern (1944) and it is this theory that Harsanyi assumed both for individuals and for the social planner. In this section we first state this theory and then introduce the alternatives to it that we put to work in the subsequent sections. To do so we now use a set  $\{z^1, \dots, z^m\}$  of outcomes and assume that they are ordered by the preference relations ( $\succsim$ ) we examine, meaning that for any  $i$ ,  $z^{i+1} \succsim z^i$ .<sup>11</sup> We denote canonical lotteries  $P$  and  $Q$ , where for instance  $P$  is a probability distribution that gives probability  $P^1$  to  $z^1$ ,  $P^2$  to  $z^2$ , and so on.

According to expected utility theory, a (rational) person’s preferences between lotteries correspond to the lotteries’ expectations of utility. More formally:

**Expected Utility (EU) theory:** Let  $u$  be a real valued (utility) function on the set outcomes that represents the agent’s preferences over them. Then:

$$P \succsim Q \Leftrightarrow \sum_{j=1}^m u(z^j) \cdot P^j \geq \sum_{j=1}^m u(z^j) \cdot Q^j$$

Next, we state two alternatives to EU theory. To introduce the first of these, rank-dependent utility theory, we need some additional notation. Let  $r$  be an increasing and real-valued *risk function* on probabilities satisfying the constraint that  $r(0) = 0$  and  $r(1) = 1$ .

**Rank-Dependent Utility (RDU) theory:** Let  $u$  be a real valued (utility) function on the set of outcomes that represents the agent’s preferences over them and let  $r$  be a risk function on probabilities. Then:

---

<sup>11</sup>  $A \succsim B$  can be read as either “ $A$  is better than  $B$ ” or “ $A$  is preferred to  $B$ ”, depending on the context (recall our remarks at the beginning of the last section).

$$\begin{aligned}
P \succeq Q &\Leftrightarrow u(z^1) + \sum_{j=2}^m [u(z^j) - u(z^{j-1})] \cdot r \left( \sum_{i=j}^m P^i \right) \\
&\geq u(z^1) + \sum_{j=2}^m [u(z^j) - u(z^{j-1})] \cdot r \left( \sum_{i=j}^m Q^i \right)
\end{aligned}$$

Informally, rank-dependent utility theory says that the value of a lottery equals the value of the worst outcome that the lottery might result in plus a weighted average of the improvements on that worst outcome that are made possible by the lottery, where the weighing is determined by the improvements' risk weighted probabilities.

To state our own preferred theory, chance-sensitive utility theory, we need some additional machinery. We will be working with a Jeffrey-desirability function,  $V$ , defined on a Boolean algebra  $B$  of *propositions*. Some of these propositions express the fact that an agent receives a particular outcome, some that she receives an outcome with some particular chance (the latter we call "chance propositions"). We should emphasise, however, that by "chance" we do not necessarily mean *physical* or *indeterministic* chances. By "chance" we mean whatever probability is such that once a decision-maker knows it, she ought to let it guide her decisions. In some cases these probabilities will be physical chances, but in other cases they may be *evidential* probabilities or relative frequencies.

Since  $B$  is a Boolean algebra, it is closed under conjunction and negation (and thus disjunction). Lotteries in this framework are conjunctions of chance propositions (e.g., "the chance of Bob winning the prize is 0.1 and the chance of him winning no prize is 0.9"). A point that is worth emphasising, since it will become important for our argument, is that because  $B$  includes chance propositions, this framework is very flexible when it comes to the attitudes to chances that it can accommodate. For instance, the framework is compatible with chances having both decreasing and increasing marginal value (just like any other good).

**Chance-sensitive utility theory:** Let  $V$  be a real valued (desirability) function on  $B$  that represents the agent's preferences. Then:

$$P \succeq Q \Leftrightarrow \sum_{j=1}^m V(z^j, P) \cdot P^j \geq \sum_{j=1}^m V(z^j, Q) \cdot Q^j$$

For now, the important difference to note between chance-sensitive utility theory and the previous two theories is that the desirability function,  $V$ , takes as its input pairs like outcome  $z^j$  and the lottery  $P$ , which for instance makes room for the possibility that the outcome  $z^j$  from lottery  $P$  is valued differently than the same outcome  $z^j$  from another lottery  $Q$ . In other words, chance-sensitive utility theory makes it possible that the values of final outcomes are sensitive to the chances of these outcomes.

In all three theories the value of a lottery depends on two kinds of factors: the desirability or goodness of the lottery outcomes and the weight placed on them, the latter being a function of the chances with which

they are yielded by the lottery. The explanations that they afford of patterns in an agent’s rankings of lotteries then draw on properties of the measure of the goodness of outcomes and/or properties of the weighting function and the relationship between them. But they do so in different ways.

In expected utility theory, the agent’s attitude to risky prospects is accounted for entirely in terms of the shape of the utility function over outcomes measuring how good the agent takes them to be. Risk aversion with respect to some good arises when quantities of the good have diminishing marginal value for the agent, something which corresponds to them having a concave utility function over these quantities. So, for instance, a preference against an actuarially fair monetary bet is captured by a concave utility function over the monetary amounts at stake.

Although expected utility is very widely used, it has been argued that the treatment of risk aversion that it offers conflates attitudes to risk with attitudes to risk-free quantities of a good, in addition to not being able to capture some well-documented attitudes to risk (see, e.g., Hansson 1988, Buchak 2013, Bradley 2017, Stefánsson and Bradley 2019). More importantly for the present purposes, expected utility theory cannot accommodate Diamond Fairness, unless outcomes are described in a way that makes reference to their chances (which results in additional challenges for expected utility theory; see, e.g., Stefánsson 2015). One way to see this, is that expected utility theory implies both that a lottery obtained by randomising over two equally ranked lotteries<sup>12</sup> (e.g., by tossing a coin) is no better or worse than the original two lotteries (a condition called *Betweenness*) and that if you replace an outcome in a lottery with one that is equally good, then the lottery thereby obtained is no better or worse than the original one (a condition called *Substitution*). Both conditions are lottery-analogues to State Dominance and indeed, as is demonstrated formally in the Appendix (Theorem 2), follow from State Dominance under the natural assumption that the lotteries induced by prospects are ranked in accordance with them.

To illustrate the inconsistency with Diamond Fairness very simply, consider the choice problem between Fair and Unfair, which we introduced above (as Table 1 in the introduction). Recall that the two states were assumed to be equiprobable. Since only one person gets  $G$ , “person 1 gets  $G$ ” implies that person 2 does not get  $G$  (and similarly for “person 2 gets  $G$ ”). Then the (social or overall) expected utility of the two alternatives is given by:

$$EU(\text{Fair}) = 0.5u(\text{person 1 gets } G) + 0.5u(\text{person 2 gets } G)$$

$$EU(\text{Unfair}) = u(\text{person 1 gets } G)$$

But by assumption, the outcome where person 1 receives  $G$  is equally good (results in the same utility) as the outcome where person 2 receives  $G$ . Therefore, the expected utility of the lottery, Fair, is the same as the “expected” utility of Unfair. But Diamond Fairness, of course, requires that Fair be ranked strictly above Unfair. Note that Unfair is obtained from Fair by substituting the outcome “person 2 receives  $G$ ” with the

---

<sup>12</sup> Note that the same holds for outcomes, which can be seen as trivial lotteries.

outcome “person 1 receives  $G$ ”, and that Fair can be obtained by randomising over Unfair and the equally unfair trivial “lottery” which gives  $G$  to person 2 for sure. So both Substitution and Betweenness entail that Fair and Unfair are equally good.

In rank-dependent utility theory, attitudes to risky prospects are accounted for by the risk function together with attitudes to quantities of goods as measured by the utility function. So, in particular, a preference against an actuarially fair monetary bet can be captured by either a concave utility function over money or a convex risk function (or both). Unlike expected utility theory, therefore, rank-dependent utility theory does not conflate attitudes to risk with attitudes to risk-free quantities and, by careful calibration of the risk and utility functions, can account for some patterns of risk preferences that have been observed and that contradict EU theory (see, e.g., Buchak 2013). However, the theory cannot account for some well-documented attitudes to *ambiguity*, that is, attitude to prospects where the precise probabilities are unknown (see, e.g., Stefánsson and Bradley 2019).

More importantly for the present purposes, rank-dependent utility theory has trouble accommodating Diamond Fairness. We discuss this in more detail in the next section, but simply put the problem is twofold. On the one hand, if rank-dependent utility theory is applied to social prospects, then State Dominance is satisfied, and thus Diamond Fairness is violated (as we showed before in the context of expected utility theory). On the other hand, if rank-dependent utility theory is applied only to individual betterness, then if the social planner applies the standard assumption of rank-dependent utility theory that risk functions are convex, then the lottery may be so much worse for each person than getting the good for sure is, that Diamond Fairness will require the social planner to be more inequality averse than may seem independently justified (more on this in the next section).

Finally, in chance-sensitive utility theory, risk attitudes are thought of as attitudes to the chances for goods and are accounted for in terms of the shape of the Jeffrey-desirability function over chance-propositions. A preference against an actuarially fair bet can be captured by either a concave desirability function over the risk-free quantities at stake or as a convex desirability function over chances of these quantities (and indeed by both). This can be seen as an improvement over rank-dependent utility theory, both in that it doesn’t postulate some primitive risk attitude (represented by the risk function) that is independent of the agent’s preferences and in that it accounts for the aforementioned attitudes to ambiguity (see Bradley 2016, Stefánsson and Bradley 2019).<sup>13</sup> In addition, chance-sensitive utility theory allows for a more natural account of Diamond Fairness than rank-dependent utility theory, or so we argue in section 5. The flip-side of this increased permissiveness is that the theory may be *too* permissive, unless additional constraints are imposed.

---

<sup>13</sup> However, chance-sensitive utility theory is not strictly speaking inconsistent with rank-dependent utility theory. One could include a risk function in the extended Jeffrey framework in which chance-sensitive utility theory is formulated, and for instance use that function to capture risk aversion, while also allowing that agents have non-linear attitudes to chances in a way that accommodates ambiguity aversion. (We thank a referee for encouraging us to clarify this.)

Since our aim here is to simply explore which theories can plausibly accommodate Diamond Fairness, we shall have to leave the discussion of such additional constraints for another occasion.

#### 4. Rank-dependent views and Diamond Fairness

We observed in section 2 that no theory of risk implying State Dominance is consistent with Diamond Fairness and that, if *ex ante* Pareto is assumed, then satisfaction of Diamond Fairness requires that the social betterness relation be more inequality averse than individual betterness relations are risk averse. Two lessons about the possibility of rank-dependent theories accounting for Diamond Fairness follow from this:

- a) Since rank-dependent theories assume State Dominance, they can apply to individuals' evaluations of lotteries consistently with Diamond Fairness only if they do not apply to social evaluations too.
- b) Since standard applications of rank-dependent theories explain individual risk aversion by assuming that individuals' risk functions are convex, to account for Diamond Fairness they must postulate high inequality aversion.

It is worth elaborating on these two observations, since they are crucial to understanding how non-robust an explanation of Diamond Fairness rank-dependent views can offer. Let's start with the second observation. As mentioned above, standard applications of rank-dependent utility theories explain risk aversion—including non-orthodox attitudes to risk that expected utility theory cannot account for—by postulating that the risk function is convex (see, e.g., Buchak 2013). Within the rank-dependent framework this is equivalent to assuming that individuals are *intrinsically risk or chance averse* in the sense that the value of any lottery over quantities of it is less than the expected value of its outcomes. For instance, it is worse, for an agent who is intrinsically risk averse, to have a half chance of winning \$10 than to have something “worth” half the utility of \$10.

If individuals are intrinsically risk averse then if a social planner wants to *maximise ex ante* good understood in terms of rank-dependent utility, she will avoid using lotteries to distribute goods because of the harm that the risk imposed by randomisation causes.<sup>14</sup> Consequently the social planner's preferences will imply Diamond *Un*fairness. We can demonstrate this with the kidney example with which we opened this article. Let's now call the two patients Ann and Bob. Recall that we assumed that they derived equal benefit from the kidney; that is, denoting getting the kidney  $G$ , we assume  $u_{Ann}(G) = u_{Bob}(G)$ . Now suppose that  $r(p) = p^2$  for both Ann and Bob, which means they are both intrinsically risk averse, to the same degree. (In the concluding section we discuss a variation of this example, where Ann and Bob do not have the same risk attitude.) Moreover, suppose the utility function for each person is zero-normalised around the

---

<sup>14</sup> For this to be meaningful, we need to assume that *ex ante* good is interpersonally comparable. This is far from being a trivial assumption, in particular when *ex ante* good is understood in terms of rank-dependent utility (as a referee for this journal points out). We however need to make this assumption for the sake of the argument, since it is impossible to explore the *implications* for social preferences of maximising and equalising *ex ante* good understood in terms of rank-dependent utility without it.

outcome where the person does not get the kidney. Then the aggregated rank-weighted *ex ante* value of a lottery that gives Ann and Bob an equal chance at receiving the kidney is:  $(0.25)u_{Ann}(G) + (0.25)u_{Bob}(G) = (0.5)u_{Ann}(G)$ , that is, only half the value of giving Ann (or equivalently Bob) the kidney for sure. In other words, a half chance of a kidney, given to each of two intrinsically risk averse patients who would get equal benefit from the kidney, is worth less, in terms of aggregated rank-dependent *ex ante* individual good, than giving the kidney to either patient without giving the other patient a chance. More generally, any lottery has a lower sum of rank-weighted utility than either of the unfair options, for any convex  $r$ .

Now the social planner may not simply wish to maximise *ex ante* good, but could seek also to take account of inequalities in the distribution of *ex ante* good. If such a social planner were sufficiently inequality averse she might prefer to distribute outcomes by means of a lottery despite the harm that it imposes on risk averse individuals. So, the assumption that individuals are risk averse rank dependent utility maximisers (or, for that matter, that they are expected utility maximisers) is not inconsistent with satisfaction of Diamond Fairness by the social betterness relation. The point is simply that a social planner, who understands people's *ex ante* good in terms of rank-dependent utility, will only display Diamond Fairness if she is sufficiently inequality averse for it to counterbalance her concern for maximising *ex ante* good.

Now, some might think that Buchak's (2017) "veil of ignorance" argument for egalitarianism,<sup>15</sup> which she bases on her rank-dependent theory of individuals' risk preferences, might help account for Diamond Fairness within a rank-dependent framework (as long as the rank-dependent view is not applied to social prospects; more on this below). She suggests that, when choosing between social arrangements under a veil of ignorance, one should adopt the most risk averse attitude "within reason". The upshot of this is that the degree of the social planner's *ex post* inequality aversion will correspond to the most risk averse individual attitude within reason, which is surely more risk averse than the average risk aversion within the population. And, from there, a natural next step might be to argue that the social planner's *ex ante* inequality aversion will also be greater than the average risk aversion. But unfortunately, this will not deliver Diamond Fairness in cases where the social planner doesn't know her subjects' attitudes to risk. For in such cases the social planner should, on this view, use the most risk averse attitude within reason when estimating subjects' *ex ante* good and will moreover aggregate *ex ante* good by using the *same* degree of inequality aversion. So, her preferences will then not satisfy Diamond Fairness. By contrast, it is possible but far from guaranteed that the social planner satisfies Diamond Fairness in the case where she does know her subjects' attitudes to risk. In that case, Buchak suggests, the social planner should defer to the subjects' risk attitudes. So, then the social planner will only display Diamond Fairness if she happens to be more averse to inequality than the most risk averse subject is risk averse. But it is hard to see what that would necessarily be the case.

---

<sup>15</sup> Here we follow Parfit's (1991) terminology, according to which an inequality averse theory that does not satisfy separability between person is called "egalitarianism". Buchak herself however calls her view "relative prioritarianism".

While rank-dependent utility theory is often presented with a risk function that is convex throughout (as we, following e.g. Buchak 2013, have assumed so far), some theorists assume that the risk function is “snake shaped”, i.e. is concave over small probabilities and convex elsewhere, with an inflection point somewhere below the 0.5 mark (see Abdellaoui 2000, Tversky and Kahnemann 1992, Tversky and Wakker 1995). Now, this is not generally presented as a risk function that is rational; rather, the observation is simply that a risk function with this shape predicts commonly observed choices. But for our purposes, the important point about a snake-shaped risk function is that an  $n$  chance of  $q$  is worth more than  $n$  times the value of  $q$  if  $n$  is low (say less than 0.5), otherwise an  $n$  chance of  $q$  is worth less than  $n$  times the value of  $q$ . This understanding of people’s risk preferences, combined with a desire to maximise *ex ante* good, would result in a social planner displaying Diamond Fairness only when *many* people’s interests are at stake. For instance, such a social planner might use a lottery when three patients need the same kidney, but she would be willing to pay a price to avoid having to use a lottery when only two patients need the same kidney.

Finally, as per the first observation above, it is impossible to apply rank-dependent utility theory on social prospects and satisfy Diamond Fairness—irrespective of what degree of risk and inequality aversion is assumed—for the simple reason that Diamond Fairness requires violating State Dominance at the social level (as we have already seen) while rank-dependent utility theory implies State Dominance (Buchak 2013: 94). So, while it is logically possible—although somewhat awkward—for a social planner who is concerned with *ex ante* welfare understood in a rank-dependent way to satisfy Diamond Fairness, it is not logically possible for a social planner to herself evaluate social prospects in a rank-dependent way while also satisfying Diamond Fairness.

Together, we take it that the above tells against the plausibility of using rank-dependent utility theory to account for Diamond Fairness. First, one would at the very least have to give up a rank-dependent view for the evaluation of social prospects (just as Epstein and Segal 1992 give up expected utility theory for social prospects while retaining it for individual prospects). Second, even if the rank-dependent view is reserved for individual prospects, one would have to assume that the social planner is more inequality averse in individual goodness than the subjects are intrinsically risk averse (assuming that the social planner satisfies *ex ante* Pareto). So, a rank-dependent account of Diamond Fairness is not particularly robust to different choices about the degree of risk and inequality aversion, and requires seemingly *ad hoc*—or at least not independently motivated—assumptions about the relationship between these degrees.

## 5. Chance-sensitive utility and Diamond Fairness

While rank-dependent views have great difficulty in accounting for Diamond Fairness, chance-sensitive utility theory delivers a justification of Diamond Fairness that is both more robust with respect to assumptions about risk and inequality aversion and only uses assumptions that have been independently motivated. The perhaps most important difference between on the one hand chance-sensitive utility theory

and on the other hand rank-dependent utility theory and expected utility theory, is that chance-sensitive utility theory includes propositions about chances in the domain of the preference relation, thus allowing that agents take a much more varied set of attitudes to chances for goods than either expected utility theory or rank-dependent utility theory does.<sup>16</sup> In particular, chance-sensitive utility theory allows that the chances of obtaining a good can have decreasing marginal value (or utility) within some ranges and increasing marginal value within other ranges, just like any other goods. (When they are increasing (decreasing) for some range, we will say that the individual is chance seeking (averse) in that range.) This makes it possible to account for ambiguity aversion in terms of the relationship between attitudes to quantities of goods and attitudes to the chances of obtaining them, thus disarming the so-called *Ellsberg paradox* (Bradley 2016). At the same time, the theory has the resources to account for the type of risk attitude that gives rise to the *Allais paradox* (Stefánsson and Bradley 2019).<sup>17</sup>

For the present purposes, the importance of the above is that the way in which ambiguity aversion is accounted for in chance-sensitive utility theory, namely by assuming that individuals are chance seeking over most of the zero-one range, delivers Diamond Fairness for free, as it were, assuming that the social planner is not inequality seeking in individual good.<sup>18</sup> We demonstrate this informally below, by assuming that the social planner is not inequality seeking in *ex ante* good; but strictly speaking, we only need this to hold for *ex post* good, as we demonstrate in Theorem 4 of the Appendix. (Note however that it does not suffice for our purposes that the social planner is inequality averse with respect to the resources (in this case  $G$ ) that constitute the individuals' good; we must assume that the social planner is not inequality seeking with respect to individuals' good itself.) Recall that not being (*ex ante*) inequality seeking means that if prospect  $X$  offers both a higher total good and a more equally distributed good than another prospect  $Y$ , then  $X$  is deemed better than  $Y$ . This is implied of course by utilitarianism, prioritarianism, and egalitarianism; moreover, this is an assumption that surely is intuitively plausible. For what could be the value of to sacrificing good in order to *increase* inequality?

For a simple demonstration that these assumptions—that is, the assumption that the individuals are chance seeking and that the social planner is not inequality seeking—suffice to deliver Diamond Fairness, consider the re-formulation below of the decision between using a fair procedure and an unfair procedure to decide who gets good  $G$ , where the columns now represent persons (rather than representing states of the world, as they did in Table 1).

---

<sup>16</sup> Expected utility theory only allows that an agent's attitude to a chance  $n$  of a good  $G$  corresponds to the utility of  $G$  multiplied by  $n$ . Rank-dependent utility theory only allows that an agent's attitude to a chance  $n$  of a good  $G$  corresponds to the utility of  $G$  multiplied by  $r(n)$ . In contrast, chance-sensitive utility theory allows that a non-trivial chance ( $0 < n < 1$ ) of good  $G$  is not a function of the pair consisting of the utility of  $G$  and the numerical value of  $n$ .

<sup>17</sup> See Cohen et. al (2022) for an empirical study of the distinguishing feature of the theory of chance-sensitive utility.

<sup>18</sup> See also Nissan-Rozen (2019) for a discussion of the connection between chance-sensitive utility and the fairness of lotteries.

	Person 1	Person 2
Fair	Gets $G$ with chance 0.5	Gets $G$ with chance 0.5
Unfair	Gets $G$ with chance 1	Gets $G$ with chance 0

Table 2: Diamond Fairness reformulated

By the assumption that chances (on average) have decreasing marginal value to each person, together with the assumption that the two people get an equal benefit from the  $G$ , it follows that Fair offers a higher *ex ante* good than Unfair. (Note that this does not require that the value of chances decreases *to the same extent* according to both people.) Moreover, Fair distributes *ex ante* good more equally than Unfair does. Therefore, chance-sensitive utility theory delivers Diamond Fairness in this case with the help of independently motivated assumptions. We demonstrate this fact more generally and formally in the Appendix (as Theorem 4).

In sum, the assumption about individual's risk attitudes that has previously been used to explain ambiguity aversion delivers Diamond Fairness by only assuming that the social planner is not inequality seeking. So, both assumptions needed to derive Diamond Fairness from chance-sensitive utility theory are independently motivated. Moreover, the explanation is robust, in the sense that it holds for *any* planner who is not inequality seeking given any population of subjects who display any degree of ambiguity aversion.

Let us get clear on what precisely we are claiming about the relationship between the fairness of lotteries and attitudes to chances. Our claim is that in order to explain a social planner's concern with fairness, in the types of examples we have been discussing, we should postulate that the chance of obtaining a good is itself something of (intrinsic) value. But that means that chances are appropriate objects of concern on the part of both individuals and the planner. The assumption that chances of a good are themselves goods has two implications that are relevant for explaining Diamond Fairness:

- On the one hand, a social planner who seeks to maximise total good must give appropriate weight to chances (whether or not she cares about equality in the distribution of goods).
- On the other hand, a social planner concerned with equality in the distribution of good should care about equality in the distribution of chances (whether or not she cares about maximising the good).

These two observations can, individually or jointly, serve as explanations, or justifications, for Diamond Fairness. The first observation suffices for a social planner who wants to maximise *ex ante* good to have preferences in line with Diamond Fairness, assuming that chances provide decreasing marginal good *for individuals*. (So, we get a "utilitarian" justification of an equal distribution of chances that is analogous to the utilitarian justification of an equal distribution of resources under the assumption that resources have decreasing marginal utility.) The latter observation however suffices for a social planner who is concerned with equality in the distribution of goods to have preferences in line with Diamond Fairness, *even if* chances *don't* provide decreasing marginal good for individuals. Finally, the most "robust" justification of Diamond Fairness follows when the social planner is both sensitive to *ex ante* good—assuming that chances are goods

of decreasing marginal value to individuals—and in addition cares about equality in the distribution of good. As previously mentioned, for the final justification it suffices that the social planner is not inequality seeking, but it does not require them to be inequality averse.

Now, one might of course question the assumption that chances provide decreasing marginal good *for individuals*. What if individuals simply don't care about chances? Indeed, what if they don't care about the chances in the way that we postulate, that is, what if they take chances to have increasing or constant marginal value? Two things should be said in response. First, as previously alluded to, there is good evidence that people actually do take chances to have decreasing marginal value; in particular, some of the (ambiguity averse) behaviour that has troubled decision theorists for decades—in particular, the previously mentioned *Ellsberg paradox*—can be naturally explained by assuming that people take chances to have decreasing marginal value (Bradley 2016, Stefánsson and Bradley 2019).

Secondly, and maybe more importantly, we should emphasise that what we have been talking about in previous sections is *goodness* which might be different from utility, i.e., the quantity that represents what people in fact prefer. In other words, we claim that the decreasing marginal good of chances *justifies* the standard “Ellsberg preferences”. Now, this second point is not completely unrelated to the first. Whatever one thinks of the exact relationship between what a person prefers and what is good for her, most would agree that the two often at least partly coincide. In particular, when a person after careful reasoning and consideration of all the relevant facts displays a self-regarding preference that is based on them taking some chances to have decreasing marginal utility, then (typically, at least) these chances have decreasing marginal goodness for that person. But evidently people don't always prefer that which is best (or even good) for them: most obviously since they may be mistaken about some relevant facts or because their preferences reflect their concern for others.

Relatedly, we contend that having a chance is good for a person whether or not she prefers it. This has two implications, corresponding to the above two observations. On the one hand, a utilitarian planner who maximises people's good—as, for instance, Broome (1991) argues the social planner should—will include chances amongst the goods in their calculation. On the other hand, since many egalitarians are concerned with equalising what is good for people rather than equalising that which people prefer getting (see, e.g., Rawls 1971), they should want to equalise chances for goods, even if people themselves don't take chances to be of intrinsic value.

## 6. Concluding remarks

Let us conclude by summarising the main findings of our paper, revisiting the main upshot for social ethics under risk, and discussing some final implications. Our main question in this paper has been: what combinations of, on the one hand, theories about individuals' risk attitudes and, on the other hand, theories of social aggregation, are consistent with Diamond Fairness; that is, consistent with favouring lotteries to

distribute indivisible goods amongst those who derive equal benefit from them. As we have seen, the set of such pairs of theories is quite limited. For instance, the theory of social aggregation cannot be what we called thoroughly *ex post*. In other words, the theory cannot, in situations of uncertainty, model the social good as a weighted average of individual's (chance-insensitive) good in different states of the world. Nor can the theory of social aggregation satisfy State Dominance. But we have also seen how hard it is to find a plausible combination of a theory of social aggregation with rank-dependent theories of individuals' risk attitudes, that satisfies Diamond Fairness. In particular, accommodating the social preference for using lotteries while using these rank-dependent views to represent individuals' risk attitudes requires some seemingly *ad hoc* assumptions about how the social planner's inequality attitude relates to the individuals' attitudes; moreover, rank-dependent views cannot be used to evaluate social prospects. In contrast, we have seen that the theory of chance-sensitive utility can accommodate Diamond Fairness in a way that is internally coherent and that uses only assumptions that are independently motivated. That, we think, should be seen as a great advantage of chance-sensitive utility theory as a theory of rational risk attitudes.

Before concluding, we would like to draw attention to the fact that although the theory of chance-sensitive utility vindicates Diamond Fairness, in the sense of justifying distributing goods via lotteries, the theory does not necessarily imply that a *perfectly equal* lottery is always best. Recall that one of the two justifications of Diamond Fairness that this theory offers is based on sensitivity to people's *ex ante* good, coupled with the assumption that chances provide decreasing marginal goodness for people. But it is possible that people may differ in this regard; the marginal goodness of chances may decrease faster for one person than it does for another. Now, we should emphasise that this is not something that we need to assume; in fact, as we shall see in a moment, this possibility creates a challenge for us. However, we think the assumption is (perhaps unfortunately) quite plausible, under the assumption that how and why chances are good for a person depends on characteristics of hers that she may not perfectly share with everyone else. For instance, a chance may be better for someone who is venturesome than for someone who is not; conversely, risk may be worse for someone who is anxious than it is for someone who is bold.

We will not attempt here a full assessment of the possibility that the marginal goodness of chances may decrease faster for one person than it does for another. Instead we shall simply make two observations about its implications, if true. First, this possibility does not affect the main message of this paper: it is still true that chance-sensitive utility theory delivers a robust justification of Diamond Fairness and one that is much more so than rank-dependent utility theory. Second, this possibility does nevertheless have the implication that equal-weighted lotteries are not always ranked higher than biased lotteries, even when the people involved would benefit equally from the good that is at stake.

As an illustration of the second observation, suppose that Ann and Bob get equal benefit from good  $G$ . However, the marginal goodness of chances decreases faster for Ann than it does for Bob; or, put another way, Ann's goodness relation with respect to quantities of  $G$  is less risk averse than Bob's. Then some lottery that gives Ann a higher chance at  $G$  than Bob produces more good than a lottery that gives both a

0.5 chance at  $G$ . The “pure” egalitarian planner would of course still prefer the equal-weighted lottery—and a pluralist egalitarian planner who is sensitive to both equality and total good *might* still prefer (at least something close to) an equal-weighted lottery. But the same is not true of a utilitarian planner whose concern is only to maximise total good.

Is this a problematic implication of our view? We think not. If one accepts that risks and chances are themselves bads and goods, as we have been arguing, then it is perfectly natural that if we are concerned with people’s good, then we may sometimes prefer biased lotteries over equal-weighted lotteries. After all, if people differ in the extent to which risks and chances are bad and good for them, then it is natural that we do not (even if we could) give perfectly equal risks and chances to all, just like the fact that people differ in the extent to which they derive benefits from resources implies that we should not provide all people with the same set of resources. This will admittedly have seemingly anti-egalitarian implications that may seem counterintuitive; in particular, that a chance-sensitive utilitarian social planner should sometimes give greater chances to those who are more risk seeking.

There might nevertheless be two distinct reasons for why even such a utilitarian social planner should generally choose *as if* people’s betterness relations were equally risk averse. First, in cases where a social planner does not have the type of information that would allow them to determine differential aversion to (or love for) risk, the equal-weighted lottery seems like a reasonable default position. Second, by taking differential aversion to (or love for) risk into account, the social planner incentivises people to develop less cautious personalities. So, the theoretical possibility that a biased lottery maximises *ex ante* good might have limited practical implications, first, due to social planners’ informational limitations, and, second, given that social planners are concerned with what incentives they give and what character traits people develop in the long run.

## Appendix: Formal results

### A) PROSPECTS

Let  $\mathcal{J}$  be a set of  $n$  individuals with canonical member  $i$  and  $\mathcal{S}$  be a set of  $m \geq n$  possible states of the world, where  $s^j$  denotes a canonical state. Let  $\mathbb{G} = \{g\}$  be a set of quantities of some good  $G$ , assumed to be convex and containing 0 and 1. An elementary outcome is simply a quantity of  $G$ , while a social outcome is an  $n$ -tuple  $(g_1, \dots, g_n)$  representing the distribution of quantities of the good to individuals. A *prospect* is a distribution of outcomes to states of the world. A prospect for *individual*  $i$ ,  $X(i) = \{g^1, \dots, g^m\}$ , is a distribution of a quantity  $g^j$  of  $G$  to individual  $i$  in each state of the world  $s^j$ . A *social* prospect  $X = \{g_1^1; \dots; g_i^j; \dots; g_n^m\}$  is a distribution of a quantity  $g_i^j$  of  $G$  to each individual  $i$  in each state of the world  $s^j$  and thereby a distribution  $\{X(s^1); \dots; X(s^m)\}$  of social outcomes to states, with  $X(s^j)$  being the social outcome of  $X$  in state  $s^j$ . No distinction will be made between a prospect that yields an outcome with certainty and the outcome itself.

Let  $\succeq$  be the weak social betterness or “at least as good as” binary relation on social prospects. Let  $\succ$  and  $\sim$  be the corresponding “strictly better than” and “as good as” relations on social prospects defined by

$X \succ Y$  iff  $X \succeq Y$  but not  $Y \succeq X$  and  $X \sim Y$  iff  $X \succeq Y$  but not  $Y \succeq X$ . For any individual  $i$ , let  $\succeq_i$  be  $i$ 's weak betterness relation on the set of individual prospects, with  $\succ_i$  and  $\sim_i$  defined in the same way.

We assume that  $\succeq$  and each  $\succeq_i$  are complete, transitive and continuous and that they are monotonic in quantities of  $G$ , i.e., that no distribution is better than another unless it confers more of the good to at least one individual in at least one state of the world. (Monotonicity admittedly excludes some egalitarian views, namely those that endorse “levelling down”.) It follows from this and the convexity of  $\mathbb{G}$  that the certainty equivalent of any social or individual prospect exists and so does the equality equivalent of any social outcome. The following two conditions on these relations will be important:

**State Dominance:** Let  $X$  and  $Y$  be any two prospects and suppose that for all states  $s^j$ ,  $X(s^j) \succeq Y(s^j)$ . Then  $X \succeq Y$ .

**Ex Ante Pareto:** Let  $X$  and  $Y$  be any two social prospects and suppose that for all individuals  $i$ ,  $X(i) \succeq_i Y(i)$ . Then  $X \succeq Y$ .

## B) LOTTERIES

A lottery  $P = \{(p^1, z^1); \dots; (p^m, z^m)\}$  is a probability distribution over a set of outcomes  $\{z^j\}$  with the probability of outcome  $z^j$ ,  $P(z^j)$ , denoted by  $p^j$ . By convention the lottery  $\{(1, z)\}$  that yields outcome  $z$  with probability one will be denoted by  $\bar{z}$ . No distinction will be made between  $\bar{z}$  and  $z$ .

Let  $Z^1, \dots, Z^k$  be any  $k$  lotteries and  $\gamma = \{\gamma_1, \dots, \gamma_k\}$  be a set of real numbers such that  $(\sum_{j=1}^k \gamma_j) = 1$ . Then the  $\gamma$ -mixture of these lotteries, denoted by  $\{(\gamma_j, Z^j)\}$  is defined by, for all outcomes  $z$ :

$$\{(\gamma_j, Z^j)\}(z) = \gamma_1(Z^1(z)) + \dots + \gamma_k(Z^k(z))$$

Note that any lottery over a set of outcomes  $\{z^j\}$  is equivalent to some mixture of the corresponding constant lotteries  $\bar{z}^j$ .

When the probabilities of states are given, prospects induce lotteries over the set of outcomes achieved in the different states; individual outcomes in the case of individual lotteries, social outcomes in the case of social lotteries. For any probability function  $P$  on the states of the world and for any prospect  $X$ , let  $X^P = \{(P(s^1), X(s^1)); \dots; (P(s^m), X(s^m))\}$  be the lottery over the outcomes that they induce. We assume in the usual manner that when the probability  $P$  on states is given then the social betterness relation is extended to the set of lotteries by:

$$X^P \succeq Y^P \Leftrightarrow X \succeq Y$$

Let's call the above assumption “prospect-lottery coherence”. The following conditions will be compelling to those who have what we called “thoroughly” *ex post* intuitions:

**Substitution:** Let  $X^P = \{(p^1, z^1); \dots; (p^j, x^j); \dots; (p^m, z^m)\}$  and  $Y^P = \{(p^1, z^1); \dots; (p^j, y^j); \dots; (p^m, z^m)\}$  be two lotteries that differ only with respect to the  $j$ -th outcome, i.e.,  $Y^P$  can be obtained from  $X^P$  by replacing outcome  $x^j$  with outcome  $y^j$ . Then  $X^P \sim Y^P$  iff  $x^j \sim y^j$ .

**Betweenness:** Let  $\bar{Z}^1, \dots, \bar{Z}^m$  be any  $m$  constant social lotteries (i.e. social outcomes), one for each state of the world, such that  $\bar{Z}^1 \sim \dots \sim \bar{Z}^m$ . Then for any  $\gamma$ -mixture,  $\{(\gamma_j, Z^j)\}$ , of these constant social lotteries,  $\bar{Z}^1 \sim \{(\gamma_j, Z^j)\}$ .

As the reader will recall, we informally explained in section 3 why both Substitution and Betweenness are inconsistent with Diamond Fairness.

**Theorem 1:** Given prospect-lottery coherence, State Dominance implies Substitution.

**Proof:** By construction, for all  $i \in \{1, \dots, m\} - \{j\}$ ,  $X(s^i) = Y(s^i)$ . Hence by State Dominance  $X \sim Y$  iff  $X(s^j) \sim Y(s^j)$ , i.e., iff  $x^j \sim y^j$ . In virtue of prospect-lottery coherence,  $X^P \sim Y^P$  iff  $X \sim Y$ . So  $X^P \sim Y^P$  iff  $x^j \sim y^j$ . ■

**Theorem 2:** Given prospect-lottery coherence, State Dominance implies Betweenness.

**Proof:** Let prospect  $X$  be such that for any  $s^j$ ,  $X(s^j) = \bar{Z}^j$  and let probability distribution  $P$  be such that  $P(s^j) = \gamma_j$  for any  $s^j$ . Then it follows from State Dominance, the transitivity of betterness, and the fact that  $\bar{Z}^1 \sim \dots \sim \bar{Z}^m$ , that  $X \sim \bar{Z}^1$ . Hence by prospect-lottery coherence and the transitivity of betterness,  $X^P = \{(\gamma_j, Z^j)\} \sim \bar{Z}^1$ . ■

### C) RANDOMISATION

Of central concern is the comparison between the unfair lotteries corresponding to an assignment of some quantity  $g > 0$  of the good  $G$  to a single individual and none to the others and the fairer ones in which each individual gets a chance of receiving  $g$ . The unfair ones, one for each individual  $i$ , are the constant social lotteries, which we will denote  $U_1 := \{(g; 0, \dots, 0)\}$ ,  $U_2 := \{(0; g; \dots; 0)\}$ , ...,  $U_n := \{(0; \dots; 0; g)\}$ , where  $U_i$  distributes the quantity  $g$  to individual  $i$  and nothing to the others. The fairer lotteries are the  $\gamma$ -mixtures of the  $U_i$ , i.e., the lotteries  $F^\gamma = \{(\gamma_i, U_i)\}$  defined by a set of chance weights  $\gamma = \gamma_1, \dots, \gamma_k$  such that  $(\sum_{j=1}^k \gamma_j) = 1$ .

Now let  $e_i$  be the equality equivalent of  $U_i$  and, for each individual  $i$ , let  $c_i^\gamma$  be the certainty equivalent of her prospect under  $F^\gamma$ , i.e., such that  $F^\gamma(i) = \{(\gamma_i, U_i)\} \sim \{c_i^\gamma\}$ .

**Theorem 3:** Assume *ex ante* Pareto. Then:

1.  $F^\gamma \sim U_i$  iff  $(c_1^\gamma, \dots, c_n^\gamma) \sim (e_i, \dots, e_i)$
2. If the  $c_i^\gamma = c$ , then  $F^\gamma \sim U_i$  iff  $c = e_i$ .
3. If  $e_i \leq \min [c_i^\gamma]$  then  $F^\gamma \succeq U_i$
4. If  $e_i \geq \max [c_i^\gamma]$  then  $U_i \succeq F^\gamma$

**Proof:** By *ex ante* Pareto,  $F^\gamma \sim (c_1^\gamma, \dots, c_n^\gamma)$ . Then by the definition of the equality equivalent, for all  $i$ ,  $U_i \sim (e_i, \dots, e_i)$ . Hence (1)  $F \sim U_i$  iff  $(c_1^\gamma, \dots, c_n^\gamma) \sim (e_i, \dots, e_i)$ . (2) Suppose now that the  $c_i^\gamma$  are all equal to some constant  $c$ . Then it follows from (1) that:  $F^\gamma \sim U_i$  iff  $(c, \dots, c) \sim (e_i, \dots, e_i)$ . Hence by monotonicity,  $F^\gamma \sim U_i$  iff  $c = e_i$ . (3) Suppose that  $e_i \leq \min [c_i^\gamma]$ . Then by the monotonicity of  $\succeq$ ,  $(c_1^\gamma, \dots, c_n^\gamma) \succeq (e_i, \dots, e_i)$ . Hence  $F^\gamma \succeq U_i$ . (4) Same as (3). ■

### D) EXPLAINING DIAMOND FAIRNESS

As before let  $g$  be some quantity of the good  $G$ , the  $U_i = \{0; \dots; g; \dots; 0\}$  be the unfair lotteries giving all of  $g$  to individual  $i$ , and  $F = \{(\frac{1}{n}, U_i)\}$  be the fair lottery giving each individual an equal chance of obtaining  $g$ . Then:

**Diamond Fairness:** Suppose quantities of  $G$  are equally good for all individuals and that  $U_1 \sim \dots \sim U_n$ . Then for any  $i$ ,  $F \succ U_i$ .

We now show that a betterness relation will satisfy Diamond Fairness whenever individual betterness is chance loving in quantities of  $G$  and social betterness is not inequality seeking in individual goodness. To define these, let the  $V_i$  be cardinal representations of the individual betterness relations over individual lotteries (how they are constructed will depend on the underlying theory of individual risk attitudes). Then the social betterness relation  $\succeq$  is *inequality seeking (averse)* in individual goodness iff the goodness of the equality equivalent of any social outcome is greater (less) than the mean goodness of the individual outcomes constituting it. And an individual  $i$ 's betterness relation is *chance seeking (averse)* over a set of outcomes iff the goodness of the certainty equivalent of any lottery over these outcomes is greater (less) than its expected goodness.

**Theorem 4:** Assume that the social betterness relation is not inequality seeking in goodness and that the individual betterness relations are chance seeking over the set  $\{0, g\}$  of quantities of  $G$ . Then social betterness will satisfy Diamond Fairness.

**Proof:** Suppose that the quantities of  $G$  are equally good for all individuals and let  $w: \mathbb{G} \rightarrow \mathcal{R}$  be a common measure of their goodness. For convenience let  $w(0) = 0$ . Let the  $V_i$  be numerical representations of the individual betterness relations over individual lotteries co-scaled so that the  $V_i(g) = w(g)$ . (This co-scaling assumes interpersonal comparability of *ex ante* betterness, which is far from trivial but commonly made in the literature on distributive ethics under risk.) Let  $V$  be any numerical representation of the social betterness relation  $\succeq$  on social lotteries.

Recall that the  $e_i$  are the equality equivalents of the  $U_i$  for the social betterness relation and the  $c_i$  are the certainty equivalents of  $F$  for the individual betterness relations. Then  $V_i(F(i)) = V_i(c_i) = w(c_i)$  and  $V_i((e_i, \dots, e_i)(i)) = V_i(e_i) = w(e_i)$  (this follows from the definition of certainty/equality equivalents plus the co-scaling assumption). Assume that  $\succeq$  is not inequality seeking in individual goodness. It follows that  $w(e_i) \leq \frac{1}{n}w(g)$  and hence that for all individuals  $i$ :

$$V_i((e_i, \dots, e_i)(i)) = w(e_i) \leq \frac{1}{n}w(g) = \frac{1}{n}V_i(g)$$

But if individuals are chance seeking over the set  $\{0, g\}$  then for all individuals  $i$ :

$$V_i(F(i)) = V_i(c_i) > \frac{1}{n}V_i(g)$$

Hence for all  $i$ ,  $F(i) \succ_i (e_i, \dots, e_i)(i)$ . So by *ex ante* Pareto  $F \succ (e_i, \dots, e_i)$ , But by definition  $U_i \sim (e_i, \dots, e_i)$ . Hence, by transitivity  $F \succ U_i$  in accordance with Diamond Fairness. ■

**Acknowledgements** We thank two reviewers for this journal for very useful comments. Stefánsson gratefully acknowledges funding from Riksbankens Jubileumsfond (Pro Futura Scientia XIII).

## Bibliography

- Abdellaoui, M. (2000). Parameter-Free Elicitation of Utility and Probability Weighting Functions. *Management Science* 46(11): 1497-1512.
- Allais, Maurice (1953). Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine. *Econometrica* 21(4): 503-546.
- Bradley, Richard (2016). Ellsberg's Paradox and the value of chances. *Economics and Philosophy* 32(2): 231-248.
- Bradley, Richard (2017). *Decision Theory with a Human Face*. Cambridge University Press.
- Broome, John (1991). *Weighing Goods: Equality, Uncertainty and Time*. Wiley-Blackwell.
- Buchak, Lara (2013). *Risk and Rationality*. Oxford University Press.
- Buchak, Lara (2017). Taking Risks Behind the Veil of Ignorance. *Ethics* 127 (3): 610-644.
- Cohen, Haim, Maril, Anat, Bleicher, Sun & Nissan-Rozen, Ittay (2022). Attitudes toward risk are complicated: experimental evidence for the re-individuation approach to risk-attitudes. *Philosophical Studies* 179 (8):2553-2577
- Diamond, Peter (1967). Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility: Comment. *Journal of Political Economy* 75(5): 765-766.
- Ellsberg, Daniel (1961). Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics* 75(4): 643-669.
- Epstein, Larry G, & Segal, Uzi (1992). Quadratic Social Welfare Functions. *Journal of Political Economy* 100(4):691-712.
- Grant, Simon, Kajii, Atsushi, Polak, Ben & Safra, Zvi (2010). Generalized Utilitarianism and Harsanyi's Impartial Observer Theorem. *Econometrica* 78(6): 1939-1971.
- Hansson, Bengt (1988). Risk Aversion as a Problem of Conjoint Measurement', in P. Gärdenfors and N.-E. Sahlin (eds), *Decision, Probability, and Utility*. Cambridge University Press, pp. 136-58.
- Harsanyi, John C. (1955). Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility. *Journal of Political Economy* 63(4): 309-321.
- Jeffrey, Richard C. (1965). *The Logic of Decision*. University of Chicago Press.
- Nissan-Rozen, Ittay (2019). The Value of Chance and the Satisfaction of Claims. *Journal of Philosophy* 116 (9):469-493.
- Parfit, Derek (1991). Equality or Priority? *The Lindley Lecture*. University of Kansas.
- Quiggin, John (1982). A Theory of Anticipated Utility. *Journal of Economic Behaviour & Organization* 3(4): 323-343.
- Rawls, John (1971). *A Theory of Justice: Original Edition*. Belknap Press.
- Savage, Leonard J. (1954). *The Foundations of Statistics*. Wiley Publications in Statistics.
- Stefánsson, H. Orri (2015). Fair Chance and Modal Consequentialism. *Economics and Philosophy* 31(3): 371-395.
- Stefánsson, H. Orri & Bradley, Richard (2015). How Valuable Are Chances? *Philosophy of Science* 82(4): 602-625.
- Stefánsson, H. Orri & Bradley, Richard (2019). What Is Risk Aversion? *British Journal for the Philosophy of Science* 70 (1): 77-102.
- Stone, Peter (2007). Why Lotteries are Just. *Journal of Political Philosophy* (15)3: 276-295.
- Tversky, Amos & Kahneman, Daniel (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5(4): 297-323

Tversky, Amons & Wakker, Peter (1995). Risk attitudes and decision weights. *Econometrica* 63(6): 1255-1280.

Von Neumann, John & Morgenstern, Oskar (1944). *Theory of Games and Economic Behavior*. Princeton, NJ, USA: Princeton University Press.