In defence of Pigou-Dalton for chances

H. Orri Stefánsson

Stockholm University,
Swedish Collegium for Advanced Study,
and Institute for Futures Studies,

Stockholm, Sweden

orri.stefansson@philosophy.su.se

Abstract

I defend a weak version of the Pigou-Dalton principle for chances. The principle says that it is better to increase the survival chance of a person who is more likely to die rather than a person who is less likely to die, assuming that the two people do not differ in any other morally relevant respect. The principle justifies plausible moral judgements that standard *ex post* views, such as prioritarianism and rank-dependent egalitarianism, cannot accommodate. However, the principle can be justified by the same reasoning that has recently been used to defend the core axiom of *ex post* prioritarianism and egalitarianism, namely, Pigou-Dalton for well-being. The arguably biggest challenge for proponents of Pigou-Dalton for chances is that it violates State Dominance for social prospects. However, I argue that we have independent reason for rejecting State Dominance for social prospects, since it prevents a social planner from properly respecting people's preferences.

KEYWORDS: *Pigou-Dalton*; *survival chances*; *priority*; *equality*; *State Dominance*.

Acknowledgements: I am grateful for suggestions I received when presenting this paper at the Institute for Futures Studies (in December 2022). Special thanks to Matt Adler, Richard Bradley, John Broome, and Ben Eggleston for extremely helpful written comments. Funding from Riksbankens Jubileumsfond is gratefully acknowledged.

1 Introduction

Suppose that through no fault of their own, Ahmed and Bogart both face some fatality risk that does not affect their well-being in any way other than potentially shortening their lives. They face this risk to somewhat different degrees, however. While Ahmed has a 50-50 chance of surviving the next twelve months, Bogart's chance is two-thirds (0.667). In all other morally relevant respects, Ahmed and Bogart's situation is identical. Their lifetime well-being up to this point is the same and their total expected lifetime well-being conditional on surviving the next twelve months is also the same (and positive).[1] Furthermore, the effect that their survival and death would have on the well-being of others is exactly symmetric.[2]

The good news is that we have the resources to improve the survival chances of Ahmed or Bogart by ten percentage points (0.1). The bad news is that we can only make the improvement for one of Ahmed and Bogart. The choice we face is represented in Table 1. What should we do?

	Ahmed	**Bogart**
Equal	0.60	0.667
Unequal	0.50	0.767

Table 1. The numbers in the table represent survival chances. Only Ahmed and Bogart are affected. Pigou-Dalton for chances implies that Equal is strictly better than Unequal.

[1]We could add that Ahmed and Bogart have led equally prudent and virtuous lives, are equally deserving, and so on. To simplify the discussion, I shall in what follows keep such features fixed, and assume that the people we consider at most differ in how well off they are, their survival chances, and their attitudes to risk.

[2]For instance, if Bogart's death reduces someone's expected well-being from 3 to 2, then Ahmed's death also reduces someone's expected well-being from 3 to 2, and similarly for all other effects. This is evidently an unrealistic assumption, but it is useful for isolating what is at stake, morally, in the trade-off I am about to describe.

To me it seems evident that we should choose Equal over Unequal. The expected number of fatalities is the same whatever we choose and so is the expected distribution of well-being. And recall that the *only* morally relevant difference between Ahmed and Bogart is that Ahmed is more likely to die, no matter what we choose. It therefore seems to me that we could offer a justification to Bogart for choosing Equal over Unequal: he is already more likely to survive, so his claim in this case is weaker than Ahmed's. But we could not offer any similar justification to Ahmed for making the opposite choice. This gives us a reason to choose Equal over Unequal.[3]

That we should choose Equal over Unequal is implied by an arguably very weak principle of ethical risk distribution, which I call *Weak Pigou-Dalton for chances*. (Here I am assuming—as I will do for simplicity through this paper—that we should choose what is better. But we can of course avoid this assumption by reformulating the previous discussion in terms of betterness.)

Weak Pigou-Dalton for chances. *For any two persons i and j, and any point in time t, and any $x, y, z \in [0, 1]$ such that $z > 0$ and $y + z \leq x$, if:*

- *i and j's expected lifetime well-being conditional on surviving to t is the same,*
- *i and j's survival chance at any time after t is the same, and*
- *the effect of i's death (survival) on the probability distribution over anonymous well-being distributions is exactly the same as the effect of j's death (survival),*

[3]Some might think that this reason depends on what type of probability the relevant survival chances are, e.g., whether they are objective or subjective (and, in the latter case, *whose* subjective probabilities). In my view, the reason for choosing Equal over Unequal is strongest if the chances in question are objective. However, in some cases (e.g., in the lottery example discussed in section 2), the chances in question are part of a deterministic system. Therefore, I shall assume that the survival chances in question are objective but still consistent with non-trivial chances (i.e., chances strictly between 0 and 1) in deterministic systems, for instance, *Humean chances* as developed by Roman Frigg and Carl Hoefer (see, e.g., 2010; 2015; see also Hoefer, 2007) and Barry Loewer (2001).

then it is better that i's chance of surviving to t is x and j's is $y + z$ than that i's chance of surviving to t is $x + z$ and j's is y (assuming that others are not directly affected).[4]

Informally, the principle implies that if two people are expected to have the same chance and well-being conditional on surviving, and if in addition their survivals (and deaths) are expected to have symmetric effects on others, then it is better to improve (by magnitude z) the survival chance of the person whose chance is lower than to improve (by the same magnitude z) the survival chance of the person whose chance is already higher (by at least z).

To see how weak Weak Pigou-Dalton for chances is, note that it says nothing about how we should reason about trade-offs where one person's well-being conditional on surviving the period in question is expected to be higher than another person's. Similarly, it says nothing about cases where we can, given a fixed budget, make greater improvements to the chances of those whose chances are, say, better compared to the improvements we can make for those whose chances are worse.[5]

Although Weak Pigou-Dalton for chances is seemingly very weak, it is inconsistent with the standard *ex post* views in distributive ethics. Perhaps most interestingly, Weak Pigou-Dalton for chances is inconsistent with *inequality*-averse *ex post* views, such as prioritarianism (see, e.g., Rabinowicz, 2002, Adler, 2012) and rank-dependent egalitarianism (see, e.g., Asheim and Zuber, 2014).

[4]There may be reason to add the restriction that $x + z < 1$. In other words, there may be reason to favour the person who is better off in terms of chances when we can ensure that they survive, but we cannot do the same for the worse off. I discuss this complication in Stefánsson (2023), but I shall not discuss it further in this paper.

[5]Weak Pigou-Dalton for chances is weaker than what Adler (2012: 502) calls the "*ex ante* Pigou-Dalton principle", which is formulated in terms of people's expected well-being (and thus implies Weak Pigou-Dalton for chances). For the purposes of this paper, I neither endorse nor reject the Pigou-Dalton principle for expected well-being. Nor do I endorse, or reject, the Pigou-Dalton principle for chances of 'goods' other than survival. My argument in this paper may carry over to, say, chances for well-being and for certain health outcomes, but I am not yet prepared to commit to that.

The reason, to put it simply, is that according to *ex post* prioritarianism and egalitarianism, as these are typically formulated (see, e.g., the references in the previous sentence), the moral value of a risky alternative (or 'prospect') is found by, first, figuring out which sums of priority- or equality-weighted well-being the alternative *might* result in, second, multiplying each possible sum of priority- or equality-weighted well-being with its probability, and, third, summing these probability weighted sums of priority- or equality-weighted well-being. In other words, the moral value of risky alternatives is determined by their *expected* sums of priority- or equality-weighted well-being. But then since Ahmed and Bogart will be exactly as well off if they survive, and since their deaths and survivals are expected to have symmetric effects on others, Equal and Unequal offer exactly the same increase in expected priority and equality weighted well-being. I briefly return to the conflict between Weak Pigou-Dalton for chances and these *ex post* views in section 3.

That Weak Pigou-Dalton for chances is inconsistent with *ex post* prioritarianism and egalitarianism may be particularly significant in light of the fact that reasoning that seems to support the characteristic axiom of these two *ex post* views—namely, the Pigou-Dalton principle for *well-being*—supports Weak Pigou-Dalton for chances too. Informally put, the Pigou-Dalton principle for well-being implies that if one person is better off than another person by at least well-being magnitude x, then it is better to improve the well-being of the worse-off person by magnitude x than to improve the well-being of the better-off person by that same magnitude x.[6]

In several influential books and articles,[7] Matthew Adler justifies the Pigou-

[6] There are of course views that satisfy both Pigou-Dalton for well-being and for chances, in particular, inequality-averse *ex ante* views.

[7] For instance, in Adler (2012, 2013, ta).

Dalton principle for well-being by arguing that well-being *levels* affect people's claims, and that we should thus not exclusively focus on effects on well-being *differences* (contrary to what utilitarianism suggests). As a result, when choosing between interventions that affect people's well-being, we should not focus exclusively on the magnitude of the difference we could make to people's well-being; we should also consider how well or badly off the people whose well-being we can affect already are, and give some priority to improving the situation of those who are worse off.

Analogous reasoning seems to support Weak Pigou-Dalton for chances.[8] When comparing Equal and Unequal, Ahmed and Bogart's survival chance levels should, I contend, have some impact on our reasoning, in favour of the person who is worse off with respect to survival chance. In fact, the reasoning I already used above to justify choosing Equal over Unequal—namely, that since Bogart is already more likely to survive, his claim is weaker than Ahmed's—parallels reasoning Adler offers in favour of Pigou-Dalton for well-being. Adler makes the plausible observation that, at least when all else is equal, the person who is worse off has a stronger claim on our assistance than the person who is better off. That, I contend, is just as plausible when it comes to survival chance as when it comes to well-being.[9] So, since the magnitude of the difference we can make to Ahmed and Bogart's survival chance is equal—and so is everything else that is of moral relevance—we should help Ahmed, whose chances are worse.

[8] Adler (2012: 505) in fact acknowledges this, although in the context of the logically stronger *ex ante* Pigou-Dalton principle (recall fn. 5).

[9] Not everyone is convinced by the above type of argument in favour of Pigou-Dalton for well-being (see, for instance, Broome, 1991b, and Greaves, 2015). For the present purposes, we can however take for granted that this is a convincing argument for Pigou-Dalton for well-being, since the aim here is just to show that analogous reasoning would support Pigou-Dalton for chances.

In the next section I further motivate Pigou-Dalton for chances by showing that it implies two attractive moral principles (*Minimal Fairness* and *Rule of Rescue*) that are known to create trouble for standard *ex post* views in distributive ethics. Section 3 confronts what is arguably the greatest challenge to Pigou-Dalton for chances, namely, that it violates State Dominance for social prospects. I argue that proponents of Pigou-Dalton for chances need not be too worried by this clash, since we have independent reason for being skeptical of State Dominance for social prospects: it prevents a social planner from properly respecting people's preferences. Some may object that my argument needs a cardinal measure of well-being which is hard to construct without assuming State Dominance, while others may contend that I use a too-narrow definition of outcomes. In section 4 I respond to these two objections. I conclude, in section 5, by briefly discussing what a full theory of distributive ethics that incorporates Pigou-Dalton for chances might look like.

2 Minimal Fairness and Rule of Rescue

Despite being very weak, Weak Pigou-Dalton for chances implies two much discussed moral principles, that seem quite plausible but are nevertheless inconsistent with standard *ex post* views in distributive ethics (including *ex post* versions of egalitarianism and prioritarianism). I call the first of these principles *Minimal Fairness* and the second *Rule of Rescue*.

As an illustration of Minimal Fairness, suppose that Annabelle and Beatrice are both in need of a kidney. Without it they will die. Sadly, the local hospital has only one kidney available for transplantation, and there is no chance of obtaining another kidney in the time Annabelle and Beatrice are expected to live without a new kidney. Further suppose that Annabelle and Beatrice's situation is identical in every morally relevant respect.

Now compare the following two alternatives: Unfair gives the kidney to Annabelle for sure, thus condemning Beatrice to death. Fair however gives the kidney to Annabelle if a fair coin comes up heads when tossed but gives the kidney to Beatrice if the coin comes up tails. So, the latter gives both Annabelle and Beatrice a 0.5 survival chance. The choice is represented by table 2.

	Annabelle	**Beatrice**
Fair	0.5	0.5
Unfair	1	0

Table 2. Numbers again represent survival chances. One and only one person is sure to live. Weak Pigou-Dalton for chances implies that Fair is strictly better than Unfair.

I take it that Fair is better than Unfair.[10] And that is indeed what is implied by what I call Minimal Fairness, which says that if two people have an equal claim to some indivisible good, and if moreover there is no moral reason to favour one person getting the good to the other person getting it, then we should—because it would be better—give both people some chance of getting the good.[11] Note that the principle is very weak: it says nothing about what we should do in cases where it would be the tiniest bit better if one person rather than the other received the good (in which case there would be *some* moral reason to favour one person getting the good rather than the other).

[10] Admittedly, not all philosophers agree with this claim. For instance, Eyal (na), Wasserman (1996), and Segall (2016) explicitly argue against it. Properly defending the claim that Fair is better than Unfair would take me too far off topic, since the aim for now is simply to show how we can derive Minimal Fairness from a Pigou-Dalton principle for chances.

[11] Since I have restricted the Pigou-Dalton principle to *surivival* chances, it won't, strictly speaking, imply Minimal Fairness formulated as a principle for the distribution of *any* good. However, a Pigou-Dalton principle for goods more generally would imply the fully general Minimal Fairness principle, by argument analogous to the one I present in this section.

Weak Pigou-Dalton for chances implies that Fair is better than Unfair. The conditions in the three bullet points in the principle are by assumption satisfied (otherwise Annabelle and Beatrice's situation would not be identical in every morally relevant respect, as assumed). Moreover, 0 and 0.5 satisfy the condition that $0 + 0.5 \leq 0.5$. Therefore, it is better, according to Weak Pigou-Dalton for chances, that Annabelle's survival chance is 0.5 and Beatrice's is also 0.5 than that Annabelle's is 1 and Beatrice's is 0.

However, since the conditions in the three bullet points in Weak Pigou-Dalton for chances are satisfied, standard *ex post* versions of prioritarianism and egalitarianism imply that Fair and Unfair are equally good. This conflict is in fact the main observation of a very influential paper by Diamond (1967).[12] I shall briefly return to this conflict in section 3.

Let's now consider the Rule of Rescue. The rule, which was coined by Jonsen (1986), is defined somewhat differently by different authors, but the general motivation behind the rule (understood normatively) is that we should sometimes prioritise saving a person identified as being at high risk of great harm even if it would be just as cost-effective to benefit people not yet at great risk (see, e.g., McKie and Richardson, 2003, Orr and Wolff, 2015).[13] The precise definition of this rule is not important for my purposes, since I shall only focus on a simple implication of it, illustrated by the following example.[14]

Suppose that a mine has collapsed and has trapped a single miner who is sure to die if we don't rescue him. Doing so is however quite expensive, and it means that we won't afford to increase safety at the mine. Further suppose that

[12] The philosophical discussion that Diamond's paper has given rise to includes Broome (1984, 1991a,b), Stone (2007, 2011), Stefánsson (2015), Stefánsson and Bradley (2015), Vong (2015, 2020), and Nissan-Rozen (2019).

[13] The rule of rescue is closely related to the 'identified lives bias'; see Cohen et al. (2015) for an overview.

[14] This example is discussed in detail by, for instance, Frick (2015) and Stefánsson (2023).

there are 100 miners whose fatality risk would be reduced if we were to spend money on increasing safety rather than saving the trapped miner. We estimate that the safety measures would, for each miner, increase the chances that they never have a fatal accident on the job by slightly less than one percentage point, from 0.98 to 0.9898. The trapped miner would also face a 0.98 survival chance if rescued. So, the total fatality risk reduction is the same whatever we do. Finally, suppose that all the 101 miners are equivalent in every morally relevant respect; so, saving each is equally morally valuable.[15] Which should we choose: to improve safety at the mine or save the trapped miner?

Given that the trade-off in expected fatalities is 1-1, and since saving each miner would be equally valuable, I contend we should—since it would be better—prioritise rescuing the trapped miner over improving safety.[16] Many people and policy makers have this intuition (but the judgements of philosophers are more mixed; see, e.g., the variety of views in Cohen, Daniels, and Eyal (eds.), 2015). And this is indeed what the Rule of Rescue implies (see, e.g., Orr and Wolff, 2015). In fact, the rule would have this implication even if improving safety would result in a somewhat higher total fatality risk reduction than rescuing the trapped miner (so, the rule is not just a tie-breaking consideration).

Now, the Pigou-Dalton principle for chances does not, by itself, imply that it would be better to prioritise the trapped miner. But it does so in the

[15]In particular, their annual well-being is exactly the same and so is their life expectancy at each point in time at which they are not trapped in the mine; similarly, each miner's death and survival has exactly the same effect on the probability distribution over anonymous distributions of well-being (in the sense previously defined).

[16]Now, by not improving safety we slightly increase the risk of multiple (up to 100) fatalities. Concentrations of fatalities can have catastrophic *indirect* effects. However, we can set such effects aside for now. The Pigou-Dalton principle for chances is not supposed to be a complete axiology. So, when risks of indirect 'catastrophe effects' arise, they have to be weighed against the good of providing a chance benefit to a person whose chance is low. But for now, let's simply stipulate that such effects are not present in the example we are considering.

presence of an assumption that in effect allows us to aggregate benefits in some circumstances. In particular, the assumption implies that we should give a large ('aggregated') benefit to one person rather than a number of small benefits to some other people, if however we divide the large benefit into smaller parts, and for each resulting part of the large benefit and for each of the small benefits, it would be morally better to give that part of the large benefit to the one person rather than the small benefit to any of the other people. The assumption is rather weak and is indeed satisfied by merely *partially* aggregative views (e.g., Voorhoeve, 2014). Since questions about whether and how to aggregate benefits are really outside the scope of my present argument, I have left the discussion and precise formulation of this assumption, and the illustration of how it allows us to derive the Rule of Rescue, to the Appendix (A.1).

By contrast, *ex post* views imply that it would be equally good to improve safety at the mine as to rescue the trapped miner. This should be evident in the case of utilitarianism, but perhaps less so in the case of prioritarianism and egalitarianism. Above, however, we already saw the explanation: *ex post* prioritarians and egalitarians only care about how our actions affect the expected sums of priority- or equality-weighted well-being; which chance distributions for individuals (i.e., which *individual* prospects) give rise to these expected sums are of no moral importance. And by the assumptions of this case, rescuing the trapped miner and improving safety give rise to the same expected sums of priority and equality weighted well-being; so, these alternatives are judged equally good by these *ex post* views. More generally, these *ex post* views violate the Rule of Rescue.

3 Pigou-Dalton for Chances vs. State Dominance

The arguably biggest challenge raised by Pigou-Dalton for chances is that it violates a principle known as *State Dominance* (for social prospects). The latter principle implies that if alternative (or prospect) A results in a better outcome than alternative B no matter how the world turns out, then A is better than B. We can state the principle more formally by letting $\mathbf{S} = \{s_1, ..., s_m\}$ denote the set of all m states of the world and $A(s_i)$ the outcome of alternative A in state s_i.

State Dominance. *For any alternatives A and B, if for any $s_i \in \mathbf{S}$, $A(s_i)$ is at least as good as $B(s_i)$, then A is at least as good as B; if, in addition, for some $s_j \in \mathbf{S}$ with positive probability, $A(s_j)$ is better than $B(s_j)$, then A better than B.*

State Dominance may seem unassailable for consequentialists (see, e.g., Adler 2012: 514). For each consequentialist view, if all $A(s_i)$ and $B(s_i)$ have been described in sufficient detail to account for anything that the view takes to be morally relevant, then it may be hard to see how it could violate State Dominance. For instance, a well-being-consequentialist can hardly reject State Dominance if all $A(s_i)$ and $B(s_i)$ completely describe the well-being distributions in these states. And, indeed, State Dominance is implied by, for instance, standard formulations of *ex post* prioritarianism and egalitarianism,[17] which are two instances of (inequality-averse) welfare-consequentialism. When evaluating an alternative, these two views first consider the alternative's risk-free moral value (respectively sums of priority and equality weighted well-being)

[17] See, for instance, Zuber and Asheim (2012), Asheim and Zuber (2014), and Buchak (2017) for canonical examples of *ex post* egalitarianism. Since these views satisfy Pigou-Dalton for well-being but do not satisfy *Separability* between persons, they are egalitarian according to the terminology of Parfit (1991); see also, e.g., Otsuka and Voorhoeve (2018). Separability ensures that we can evaluate the moral value of improving a person's well-being without considering other (unaffected) people.

See Rabinowicz (2002) and Adler (2012) for canonical examples of *ex post* prioritarianism. These views satisfy both Pigou-Dalton for well-being and Separability between persons.

in each of the possible states of the world, and then weight each risk-free moral value by its probability. In other words, the moral value in each state is determined independently of any other state. And that implies State Dominance.[18]

But State Dominance is inconsistent with Pigou-Dalton for chances. One way to see this is to consider again the choice between Fair and Unfair, in table 2. If the outcome where Annabelle receives the kidney is equally good as the outcome where Beatrice receives it, as we have been assuming, then for any state of the word, the outcome of Unfair is just as good as the outcome of Fair; so, by State Dominance, Unfair is at least as good as Fair, contra Weak Pigou-Dalton for chances.[19]

In what follows I shall argue that we have good reason to reject State Dominance for social prospects independently of the desirability of accommodating a Pigou-Dalton principle for chances. The reason is that a 'social planner' who satisfies State Dominance cannot properly respect people's preferences, and must thus be (I contend) problematically anti-liberal. In particular, such a social planner is insensitive to the fact that people may *differ* in their risk attitudes (and therefore differ in their preferences). But, I contend, being sensitive to the differences between people and their preferences—in the sense ruled out by State Dominance—is necessary for properly respecting people and their preferences.

Note that the argument I am about to make allows that State Dominance could be an attractive principle of individual rationality even if it should be

[18]This implication is of course well-known, and familiar from any axiomatisation of *expected utility theory* (in state-based frameworks), which assumes that the utility of an outcome in one state of the world can be determined independently of outcomes in other states of the world. (See, e.g., Savage, 1954 and Kreps, 1988.)

[19]It may be tempting to redescribe the outcomes of Fair, for instance by including in the description of its outcomes that they were brought about by a fair procedure, in a way that makes the preference for Fair over Unfair consistent with State Dominance (see, e.g., Broome, 1991b). In section 4 I explain why I want to resist that temptation.

rejected for moral (or social) choices. Thus we have an answer to a question Adler poses in this context: "Why should the norms for evaluating actions, in light of an ordering of outcomes, be different in the case where the ordering is a moral rather than non-moral ordering?" (2012: 480). The answer is that only the moral ordering must respect differences between people and their attitudes; the same is not true of an individual's ordering of, say, consumption bundles.

My argument against State Dominance is based on a strengthening of a result by Bradley (2022). I will illustrate the general result with the following example. Suppose that Ann is risk seeking with respect to her own well-being while Bell is risk averse. Now consider two alternatives, one that gives Ann a 50-50 chance of ending up with lifetime well-being of either 30 or 70 while ensuring that Bell ends up with lifetime well-being 50, and another that is the reverse of the first, giving Bell a 50-50 chance of ending up with lifetime well-being of either 30 or 70 while ensuring that Ann ends up with lifetime well-being 50. The two alternatives are represented in table 3.

	s_1, $p = 0.5$	s_2, $p = 0.5$
Pref	Ann: 30, Bell: 50	Ann: 70, Bell: 50
Disp	Ann: 50, Bell: 30	Ann: 50, Bell: 70

Table 3. Ann is risk seeking, Bell is risk averse. Both thus prefer Pref to Disp.

Since Ann is risk seeking with respect to her own well-being while Bell is risk averse, they both prefer Pref (for 'preferred') to Disp (for 'dispreferred'). We can further suppose that both Ann and Bell know the probabilities of the two states, and have no false beliefs that are relevant to their preferences between these alternatives. So, their unanimity is not "spurious" (Mongin, 2016).

However, in the presence of a weak assumption, a social planner who satisfies State Dominance must be indifferent between Pref and Disp; which, I shall argue in a moment, means that they do not properly respect people's preferences. The assumption in question is what I will call *Two-Person Outcome Anonymity*, which informally says that an outcome is just as good even though we switch the well-being levels of two individuals in that outcome.[20] For instance, an outcome where Ann is at well-being level 50 and Bell is at well-being level 30 is exactly as good as an outcome where Bell is at well-being level 50 and Ann is at well-being level 30. Assuming that we have an interpersonally comparable measure of well-being, for which one unit of well-being is equally valuable (i.e., represents the same amount of good) to any one person as to any other person, Two-Person Outcome Anonymity seems unassailable—unless one is partial towards specific individuals, which I shall assume a social planner should not be.[21] Interpersonal comparability is a strong assumption, but it is routinely made in distributive ethics. And, indeed, all the standard *impartial* theories in distributive ethics, from egalitarianism to utilitarianism, imply both interpersonal comparability and Two-Person Outcome Anonymity.

Now consider what Two-Person Outcome Anonymity implies when applied to Pref and Disp. Pref(s_1) is a permutation of Disp(s_1) and Pref(s_2) is a permutation of Disp(s_2). So, by Two-Person Outcome Anonymity, Pref(s_1) is equally good as Disp(s_1) and Pref(s_2) is equally good as Disp(s_2). But then State Dominance implies that Pref and Disp are equally good. More generally, State

[20] Bradley (2022) assumes the stronger Outcome Anonymity that is not limited to two-person switches. Why do I only assume the weaker condition? Because the stronger condition might be questioned in light of the type of case discussed by Parfit (2003), footnote 16. Since every finite permutation can be expressed as a series of two person switches, Two-Person Outcome Anonymity entails Outcome Anonymity in the presence of *transitivity*. If one is willing to give up transitivity, then one can however assume only Two-Person Outcome Anonymity.

[21] Adler (2012: 500) for instance says that any "minimally plausible" social welfare function, that is, any minimally plausible ranking of well-being distributions, satisfies the stronger Outcome Anonymity.

Dominance and Two-Person Outcome Anonymity together entail *Two-Person Prospect Anonymity*, which informally says that a planner should be indifferent between two alternatives where one can be obtained from the other by permuting the prospects it assigns to two individuals. For instance, the social planner should be indifferent between on the one hand Pref, which assigns to Ann a 50-50 gamble between well-being levels 30 and 70 and assigns to Bell level 50 for sure,[22] and on the other hand Disp which assigns to Ann the prospect that Pref assigns to Bell and assigns to Bell the prospect that Pref assigns to Ann.

I think most would agree that a social planner should respect their subjects' preferences, in particular when nothing is at stake other than the extent to which people's preferences are satisfied, and assuming that the preferences in question are sufficiently rational (or 'reasonable') and self-regarding (which for instance ensures that the people in question don't prefer that someone else be harmed). Most liberals would certainly agree with this. But, I contend, being indifferent between Pref and Disp means that the social planner fails to respect Ann and Bell's preferences even though nothing is at stake other than how well their preferences are satisfied. The only difference between Pref and Disp is that the former assigns the only people affected the prospects that suit their attitudes to risk while the latter does not. Disregarding this difference is inconsistent with respecting the fact that Ann and Bell have different risk attitudes. But a risk attitude is just a particular type of preference, namely, a preference about how outcomes are spread across the different possible states of the world (Stefánsson and Bradley, 2019). Further, there is no uniquely correct attitude to risk, any more than there is, say, a correct attitude to the relative desirability of vanilla vs. chocolate flavoured ice cream. So, being indifferent between Pref

[22]Note that getting well-being level 50 for sure is a 'trivial prospect', that is, one that gives Bell well-being level 50 with probability 1 and any other well-being level with probability 0.

and Disp is analogous to being indifferent between, on the one hand, Ann getting chocolate flavoured ice cream and Bell getting vanilla flavoured ice cream, and, on the other hand, Ann getting vanilla flavoured ice cream and Bell getting chocolate flavoured ice cream, even though Ann prefers the vanilla flavour to chocolate and Bell prefers chocolate to vanilla, and even though nothing is at stake other than the extent to which Ann's and Bell's preferences are satisfied.

Note in addition that reasons that have been offered for sometimes disregarding unanimity in people's preferences—for instance, concern for outcome equality (Fleurbaey and Voorhoeve, 2013) or worries that the unanimity is based on some person having false beliefs about the relevant probabilities (Mongin, 2016)[23]—do not justify disregarding Ann and Bell's unanimous preference for Pref over Disp. After all, the two prospects are perfectly symmetric as far as outcome inequality is concerned and we are assuming that the probabilities of the states are known to both Ann and Bell. More generally, it is hard to see what could justify going against unanimous preference in this case, since there would be no benefit of doing so. Still, Two-Person Outcome Anonymity and State Dominance together imply that, in this case, we should be indifferent between satisfying and not satisfying Ann and Bell's unanimous preference.

I find it hard to see how the above implication of State Dominance combined with Two-Person Outcome Anonymity could be part of a plausible (impartial) theory of distributive ethics. An impartial social planner would not disregard differences between people's attitudes in cases where there is no trade-off (everyone can get what they want) and nothing is at stake other than peo-

[23]These arguments were originally directed at the principle of *ex ante* Pareto, which implies that if every subject prefers one prospect to another, then the social planner should also prefer the one prospect to the other.

ple getting the prospect they prefer. But that means that either Two-Person Outcome Anonymity or State Dominance has to be rejected. I find it very hard to understand why an impartial social planner would violate Two-Person Outcome Anonymity when only two people and two prospects are involved. Instead, we should allow for violations of State Dominance.

Now, someone might object that my argument merely shows that we should give up State Dominance in cases where people have different risk attitudes, but State Dominance will conflict with a Pigou-Dalton principle for chances even when people's risk attitudes do not differ. In response, I would contend that once we have accepted that there are exceptions to State Dominance—i.e., when people have different risk attitudes—it is hard to see why Pigou-Dalton for chances would not also warrant such exceptions. State Dominance may be unproblematic in many cases. But a social planner cannot safely apply State Dominance when their subjects have different attitudes to risk, nor when it comes to chance trade-offs between people with different chances.

Finally, note that the argument against State Dominance that I have presented in this section is logically independent of the plausibility of a Pigou-Dalton principle for chances. The argument in this section assumed that the social planner should be impartial, as captured by Two-Person Outcome Anonymity, that people can have different attitudes to risk, and that a social planner should respect differences between people's attitudes. These assumptions are independent of (the plausibility of) Weak Pigou-Dalton for chances. Therefore, we have a reason for rejecting State Dominance that is independent of the desirability of accommodating Weak Pigou-Dalton for chances.

4 Two objections

In this section I state and respond to two objections.[24] The first of these objections may on the face of it seem to target in particular my positive argument in favour of Pigou-Dalton for chances (in sections 1 and 2), while the second objection may seem directed against my argument against State Dominance (in section 3). But as will become apparent, these two objections are not unrelated. In particular, my response to them will be similar. Generally put, my response will be that in this article I have made some modelling choices that I think are attractive as a package—in particular, given the purpose to which I put them—even though it is not, as far as I can see, possible to conclusively defend each of these choices against a rival choice (nor vice versa).

4.1 *First objection*

Consider again the choice between Fair and Unfair, in table 2. Some may find that we should individuate the outcomes more finely than I have done there. For instance, one could include in the description of Fair's outcomes that they were brought about by a fair procedure, and/or include in the description of Unfair's outcomes that they were brought about by an unfair procedure. Doing so makes the preference for Fair over Unfair consistent with State Dominance. For then it is no longer necessarily the case that the outcome where Annabelle lives is as good as the outcome where Beatrice lives; it may depend on whether the outcomes in question have the property of being fair.

On the face of it, this re-individuation of outcomes may not seem to be an option for *ex post* egalitarians and prioritarians, nor for welfare-consequentialists

[24] I am very grateful to John Broome for making me see the need to properly address both objections.

more generally, according to whom moral value is fully determined by the distribution of well-being. But it could be argued that being treated unfairly harms a person, which should then be accounted for in the measure of her well-being (see, e.g., Broome 1991b: 198). By contrast, some non-welfarists may want to include the unfairness that Beatrice experiences if Unfair is chosen without assuming that it affects her well-being. Either way, the conflict between Minimal Fairness and State Dominance would be avoided.

Now, an indiscriminate re-individuation of outcomes threatens to trivialise principles such as State Dominance and Pigou-Dalton: for any seeming violation of these principles, we can always re-individuate the outcomes in a way that avoids the violation. So, we need to find some principled way of determining what individuations are permissible. Broome (1991b: 103) suggests such a principle: "Outcomes should be distinguished as different if and only if they differ in a way that makes it rational to have a preference between them." This, he argues, makes what I am calling Minimal Fairness compatible with State Dominance (and the stronger Sure Thing Principle). He points out that unfairness is "a property with modal elements" (1991b: 114) and suggests that these elements are properties of outcomes. I agree with the first point—that unfairness is a property with modal elements—but, unlike Broome, I have elsewhere suggested that such modal elements should not be thought of as properties of outcomes, but instead as relations between outcomes. Moreover, I have defended a principled re-individuation strategy that is consistent with Pigou-Dalton for chances and Minimal Fairness but violates State Dominance (Stefánsson, 2015; Stefánsson and Bradley, 2015; Bradley and Stefánsson, 2017).

Can a conclusive argument be given in favour of my way of individuating outcomes over, say, Broome's? I think not. Nor can a conclusive argument can be given for Broome's individuation over mine. Broome himself suggests that his way of individuating outcomes is "the best way for the theoretical purpose of understanding rationality". But he admits that "[f]or many purposes, this may not be the most convenient way of individuating" (ibid: 109). The way I understand this is that we—as theoreticians—face a modelling choice. And as with most modelling choices, there is in this case no uniquely correct solution, independently of the purpose to which we want to put the model. Nevertheless, we can ask what modelling choices are most convenient, given our theoretical purpose.

I would suggest that for the purposes of this article, Broome's is not the most convenient way of individuating outcomes. Recall that my aim here is to ask questions about what attitudes social planners should take to risks to people's lives and well-being. To examine such questions, it is, I think, best to distinguish as clearly as is possible between, on the one hand, the risk or uncertainty that an alternative involves, and, on the other hand, the risk-free determinants of the goodness that is contained in the alternative's outcomes. Otherwise we risk mixing up the description of the alternatives the social planner faces with the attitudes they should take to the risks and uncertainties involved in the alternatives. But as previously indicated, I take this ultimately to be a modelling choice: neither individuation is better than the other, independently of the purpose. I shall return to this issue after discussing the second objection.

4.2 *Second objection*

The second objection is that I have been using a cardinal measure of well-being without explaining where it comes from. Moreover, the measure in question can evidently not be the one that scholars typically refer to when justifying the use of such a cardinal measure, namely the one based on expected utility theory (as e.g. developed by von Neumann and Morgenstern, 1944). One reason for this is that expected utility theory implies State Dominance (as previously mentioned). So, it might seem disingenuous to appeal to expected utility theory to justify a cardinal measure of well-being, while rejecting State Dominance as a constraint on the evaluation of social prospects.[25]

Another and perhaps more important reason for why I cannot appeal to expected utility theory to justify the use of a cardinal measure of well-being is that the axioms underlying expected utility theory imply that rational people are risk neutral with respect to the quantities (i.e., utilities) of which the theory offers a cardinalization. Let's make this more concrete by considering again the choice between Pref and Disp in table 3. If Ann's preferences satisfy the expected utility axioms, then Ann is risk neutral with respect to the values (i.e., utilities) that properly (i.e, cardinally) represent her preferences. Therefore, if she rationally strictly prefers Pref over Disp, then the quantities in table 3 don't represent the relative desirability of these outcomes, according to her, and so these quantities must acquire meaning and structure independently of Ann's preferences. And similarly for Bell.

In response, I should emphasise that I do not take the axioms of expected utility theory to be requirements of rationality. This part of my response is however not something I can conclusively defend in this article; in fact, there are very influential book-length arguments on both sides of the debate over the rationality of the expected utility axioms (see, e.g., Broome, 1991b, Buchak, 2013, and Bradley, 2017). But in short, my view on this issue is that,

[25]This would strictly speaking not be inconsistent, however, since one could insist that State Dominance should hold for individuals' preferences but not for the 'social preference'. This is in fact Diamond's (1967) view on the Sure Thing Principle. By contrast, Adler (2012: 480) suggests that this would be an unstable position.

first, the expected utility axioms—in particular, the Sure Thing Principle—are inconsistent with preferences that are, I contend, perfectly rational (see, e.g., Allais, 1953, Buchak, 2013, Bradley and Stefánsson, 2017). Secondly, and relatedly, these axioms—and, again, in particular the Sure Thing Principle[26]—prevent the expected utility theorist from distinguishing people's attitudes to risk from people's evaluations of risk-free outcomes. But these are, I contend, importantly distinct attitudes (see, e.g, Hansson, 1988, Buchak, 2013, Stefánsson and Bradley, 2019).

How can I then justify a cardinal measure of well-being if I cannot appeal to the cardinalization method offered by expected utility theory? More importantly, are there methods for generating such a measure without assuming State Dominance? Yes, there are. For instance, representation theorems for *rankings of pairs* of alternatives establish that there are such methods (see, e.g., Alt 1936, Suppes and Winet 1955, and Krantz et. al. 1971: 151).[27] In the Appendix (A.2) I state a slightly simplified version of one such representation theorem.

The most important property of representation theorems of this type is that they state conditions that are sufficient (and, in Alt's case, also necessary) for the existence of a cardinal utility function that represent a preference relation, without either entailing or assuming anything about how the person whose preferences are being represented views risk. In fact, in these theorems, all objects of preference can be "risk-free", in the sense that there is no uncertainty as to how they will turn out. Instead of assuming that an agent ranks

[26] Or the related Independence axiom of the expected utility theory of von Neumann and Morgenstern (1944).

[27] Alternatively, one can use a version of Jeffrey's (1965) decision theory, with some of Jeffrey's prospects playing the role of chance distributions, in a way that generates a cardinal measure but allows for a great deal of flexibility in how the person evaluates prospects (Bradley, 2017; Stefánsson and Bradley, 2019). This too delivers a cardinal measure that allows for non-neutral attitudes to risk and that does not assume State Dominance.

risky prospects (or 'lotteries'), from which expected utility theorists infer cardinal utilities, these theorems assume that the agent can directly rank pairs of alternatives in terms of *the difference in benefit* that they bring.

A natural interpretation of such a ranking of pairs is that the agent compares the value of *exchanges* between different pairs of alternatives. To take an example, suppose that the person of interest prefers an apple to an orange and also prefers Volvo XC90 to Mercedes EQB. We assume that the utility of each piece of fruit is independent of the car owned, and vice versa. Finally, suppose we give them an orange and a Mercedes EQB. Then one way to understand the claim that, according to this person, the difference between an orange and an apple is smaller than the difference between a Mercedes EQB and a Volvo XC90, is that if the person were allowed to exchange *either* their Mercedes EQB for a Volvo XC90 *or* their orange for an apple, then they would choose the former exchange (see Suppes and Winet 1955: 260-261).

Note that the assumption is merely that the agent can intuit which exchange would be more valuable, but not by how much. So, this is not to *assume* a cardinal measure of the relevant value. However, if the ranking of such exchanges satisfies the appropriate axioms, then a cardinal utility measure is *implied*. And this measure represents both the ranking of such exchanges and the ranking of individual alternatives (see the theorem stated in Appendix A.2).

The resulting (risk-free) quantities, that is, these derived cardinal utilities, do not necessarily correspond to the person's well-being, of course. But once one has this cardinal measure, one can use arguments analogous to those that have been used to justify identifying a (family of) utility function(s) given by expected utility theory with a measure of the person's well-being (see, e.g., Broome 2004: ch. 5), to justify identifying a (family of) utility function(s) derived from risk-free difference rankings with a measure of the person's well-being. This delivers a cardinal measure of well-being without assuming State Dominance.

Moreover, this cardinal measure can be used to make sense of the idea that a rational person prefers an outcome whose 'true' utility—i.e., whose value according to a utility function that truly represents the person's attitudes to risk-free outcomes—is X to a gamble whose expected utility is X. The reason is that the measure does not imply risk neutrality with respect to its quantities. So, we can make precise and meaningful the claim that Ann and Bell both strictly prefer Pref over Disp, even though the numbers in the table truly represent the relative desirabilities of these outcomes, according to them.

So far I have only pointed out that *there exist* ways of constructing cardinal measures that could be identified with a person's well-being, without assuming State Dominance and without the person being risk neutral with respect to the measure's quantities. But one might of course ask whether these measures are *as good* as the one provided by expected utility theory. To that question, I shall give an answer analogous to the one I gave in the last section. And again I quote Broome in support of my conciliatory approach:

> The natural, intuitive meaning we assign to quantities of wellbeing is vague. When we come to precise argument, we do not need to discover our actual precise meaning, because we do not have one. Instead we can make our vague meaning precise.
>
> We may have a choice about how to do that. Adopting Bernoulli's hypothesis [i.e., risk-neutrality with respect to well-being] is one option, and I am taking that one. [...]
>
> I think my definition is a reasonable way to make our vague quan-

titative notion of wellbeing precise. But it is important to remember I have taken this step. I have made my notion of wellbeing more precise than the intuitive one, and I have chosen one particular way of doing so. This means that all my conclusions about the aggregation of wellbeing must be interpreted through this precisification. (Broome 2004: 90-91)

So, using the expected utility framework and adopting Bernoulli's hypothesis is one way of making precise our notion of well-being. But using, say, risk-free difference rankings is another way of making this notion precise. We cannot decide between these two precisifications by trying to figure out which corresponds to "our actual precise meaning" (ibid.), because none exists: the actual notion is vague. Instead, we need to construct one. A comparison between two different constructed precise meanings will, I think, unavoidably be (at least partly) based on methodological considerations. So, again we face a modelling choice, to which there may be no uniquely correct solution.

But one can ask which precisification is most useful for the task at hand. When it comes to the task I set for myself in this article, the answer is, I think, that the precisification based on expected utility theory and Bernoulli's hypothesis is not the most useful one. The task at hand is to ask ethical questions about the attitudes that a social planner takes to risks to her subjects' well-being and to their lives. Insisting that we cannot meaningfully assume that the subjects can be anything other than risk neutral with respect to their own well-being severely limits the types of questions we can ask. For instance, then we cannot meaningfully ask how the social planner should react to subjects' conflicting attitudes to risk.

4.3 *Choosing between packages*

Finally, let me briefly explain how I think my answers to the two objections relate to each other; and, correspondingly, what I take to be the common theme of these two objections.

Suppose that a theorist is committed to expected utility theory as the correct theory of both individual and social rationality—in other words, they insist that both the social planner's preferences and her subjects' preferences should satisfy the axioms of expected utility theory. Further suppose that they are committed to Minimal Fairness. Then the only available choice when it comes to individuating outcomes is to include the bad that is done when Unfair is chosen in the description of the alternative's outcomes (and/or to include the good that is done when Fair is chosen in the description of its outcomes). To undermine the argument against State Dominance for social prospects that I presented in section 3, the theorist could then insist that well-being should be cardinalized by using both expected utility theory and Bernoulli's hypothesis. So, one coherent package of assumptions is, first, to individuate outcomes in a way that dissolves the conflict between Minimal Fairness and State Dominance, and, second, to use expected utility theory and Bernoulli's hypothesis to cardinalize well-being.

But suppose instead that a theorist does not accept expected utility theory as the correct theory of either individual rationality or social rationality. Then they can without difficulties accommodate Minimal Fairness while describing outcomes in terms of only whether Annabelle and Beatrice die or survive. Further, then they can cardinalize well-being in a way that makes the argument against State Dominance in section 3 perfectly meaningful.

Which of these packages is more attractive? Again, I think that depends on the purposes to which they are being put. I have argued in the previous two subsections that the two parts of this package that I have assumed—one part being my preferred individuation of outcomes, the other part being my preferred ways of cardinalizing well-being—are most suitable given the purposes of this article. And I have now explained how these two parts naturally hang together. But as already indicated, I do not think that one can give a conclusive argument in favour of either part of this package that is independent of the uses to which one wants to put it. This is all familiar from scientific modelling.[28] We can ask which model is most fruitful given the task at hand. But there will rarely be one correct model independently of such tasks.

5 Concluding remarks

To conclude, let's briefly consider what type of complete theory of distributive ethics is consistent with a Pigou-Dalton principle for chances. As previously noted, Weak Pigou-Dalton for chances is consistent with *ex ante* versions of prioritarianism and egalitarianism (recall fn. 6). However, Weak Pigou-Dalton for chances neither implies *ex ante* prioritarianism nor egalitarianism. For instance, Weak Pigou-Dalton for chances says nothing about the relative contribution that on the one hand people's survivals chances and on the other hand risk-free outcomes make towards the overall moral value of an alternative. So, the principle is compatible with the idea that *ensuring* that *someone* survives takes lexical priority over equality in survival chances. Similarly, the principle is compatible with well-being taking absolute priority over equality in sur-

[28]For a recent application to moral philosophy, see Roussos (2022).

vival chances, for instance in the sense that Weak Pigou-Dalton for chances is only used as a tie-breaker when alternatives give rise to the same probability distributions over anonymous well-being distributions.

Since the standard, *ex post* extensions of egalitarianism and prioritarianism to situations of risk imply State Dominance, they are inconsistent with granting Weak Pigou-Dalton for chances such a tie-breaking role (and also, of course, inconsistent with granting the principle a stronger role). That should, I contend, make us question the plausibility of these views. But it might also provide a new reason for *ex post*-leaning egalitarians and prioritarians to consider pluralism (which, of course, many of them already do). One of the virtues of the standard welfarist theories, such as *ex post* egalitarianism and prioritarianism, is admittedly their simplicity. But giving up Weak Pigou-Dalton for chances seems to me to be a price that is too high too pay for the theoretical virtue of simplicity.

Appendix: Deriving the Rule of Rescue

A.1 Deriving the Rule of Rescue

Recall from section 2 that the Weak Pigou-Dalton principle for chances does not, by itself, imply the Rule of Rescue. But it does so in the presence of a weak principle of aggregation. As an illustration of what the principle implies, consider the situation represented by the table 4, where we can either give a large benefit, of magnitude g, to Ann (alternative Large), or we can give a smaller benefit, of magnitude g^{-}, to each of Bob, Cat, and Dan (alternative Small). Suppose that we can divide the large benefit we are considering giving to Ann into three equally large parts, g_1, g_2, and g_3. And imagine that we compare giving g_1 to Ann with giving g^- to Bob and find that the former would be better; we compare giving g_2 to Ann, given that we have already given her g_1, with giving g^- to Cat and find that the former would be better; and finally we compare giving g_3 to Ann, given that we have already given her g_1 and g_2, with giving g^- to Dan and find that the former would be better. Further suppose that we would have gotten the same result had we done the comparison in a different order (say, by comparing giving g_1 to Ann with giving g^- to Cat, etc.). Then the weak principle of aggregation implies that Large is better than Small.

	Ann	Bob	Cat	Dan
Large	$g = \{g_1, g_2, g_3\}$	0	0	0
Small	0	g^-	g^-	d g^-

Table 4. Illustration of Weak Aggregation. '0' means that the alternative makes no difference to the person in question.

Now let's say that a good G is 'finely divisible' if for any integer n and for any quantity q of G, we can divide quantity q of G into n equally good parts. So, for instance, money is finely divisible, and so are chances. Here is then a more formal statement the principle of weak aggregation:

Weak Aggregation. *For any finely divisible good G, and for any quantities g and g^- of G where $g^- < g$, for any person i and for any number n of other people, if for any way of ordering the n people from j_1 to j_n and for any way of dividing g of G up into n equally good parts, g_1 to g_n, we find that:*

- *it is better that i receives g_1 than that j_1 receives g^-,*
- *and for any $k \in [2, n]$ it is better that i receives g_k, given that she has already received $g_1 \& g_2 \& ... \& g_{k-1}$, than that j_k receives g^-,*

then it is better that i gets quantity g of G than that the other n people each get quantity g^- of G.

Fully aggregative theories, such as utilitarianism, prioritarianism, and rank-dependent egalitarianism, all of course imply Weak Aggregation. But to see that Weak Aggregation does not imply full aggregation, note that it never implies that we should provide a number of people each with a small benefit rather than providing one person with a large benefit. This is because its consequent only refers to it being better to benefit the one person. Weak Aggregation will also be silent when considering goods and bads that are not finely divisible, such as the survival and death of a person.

Let's finally consider what Weak Pigou-Dalton for chances implies in the presence of Weak Aggregation. Recall that each of not-yet-trapped miner is assumed to have a total work-life survival chance of 0.98. Further, recall that this is also the survival chance that the trapped miner would face after the rescue. Finally, let's number the 100 not-yet-trapped miners from 1 to 100; we simply call them 1, 2, 3, etc. Now compare increasing the trapped miner's survival chance by 0.0098 from 0 with increasing 1's survival chance by 0.0098 from 0.98. Evidently Weak Pigou-Dalton for chances implies that the former would be better. Next compare increasing the trapped miner's survival chance by 0.0098 from 0.0098 with increasing 2's survival chance by 0.0098 from 0.98. Again, Weak Pigou-Dalton for chances implies that the former would be better. Continuing this reasoning, we eventually get to 100, and find that Weak Pigou-Dalton for chances implies that it would be better to increase the survival chance of the trapped miner from 0.9702 to 0.98 than to increase the survival chance of 100 by 0.0098 from 0.98. Since we get this result no matter how we order the 100 not-yet-trapped miners, Weak Pigou-Dalton for chances and

Weak Aggregation together imply that it would be better to rescue the trapped miner than to increase safety at the mine.

A.2 Alt's representation theorem

The theorem I state here was originally proven by Franz Alt (1936). I focus on Alt's theorem because I find its axioms to be relatively intuitive (in particular, the slightly simplified versions that I present here).

The following notation will be used from now on. A, B, C, etc., now denote risk-free alternatives, and $\mathcal{A}$ the set of these alternatives. $\precsim$ denotes a weak preference relation on $\mathcal{A}$, such that $A \precsim B$ means that the agent in question does not prefer A to B; and $\sim$ and $\prec$ denote the corresponding indifference and strict preference relations, defined from $\precsim$ in the usual way.[29] Following Alt, I will abuse notation slightly, and use these relations for comparisons of both alternatives and exchanges. I shall assume that just as a person has a preference for one alternative over another, so she has a preference for one exchange, from one alternative to another, over another exchange, from a third alternative to a fourth one. Finally, $A \to B$ denotes an exchange from A to B.[30]

The first two of Alt's axioms state that the preference relation is complete, with respect to both alternatives and exchanges.

ALT 1. *For any* $A, B \in \mathcal{A}$*:* $A \precsim B$ *or* $B \precsim A$

ALT 2. *For any* $A, B, C, D \in \mathcal{A}$*:* $(A \to B) \precsim (C \to D)$ *or* $(C \to D) \precsim (A \to B)$

[29] That is, $A \sim B$ just in case $A \precsim B$ and $B \precsim A$; and $A \prec B$ just in case $A \precsim B$ and $\neg(B \precsim A)$.

[30] It is worth emphasizing that it is assumed that agents get neither pleasure nor displeasure from exchanges. So, an exchange from A to B is evaluated only in terms of the added benefit that having B brings over having A; but the exchange itself is of neutral value. This assumption corresponds to an assumption that must be made for the soundness of the Reduction assumption in von Neumann and Morgenstern's expected utility theory, which ensures that lotteries can be simplified by calculating the total probabilities they confer on consequences: in that framework, it is assumed that agents get neither pleasure nor displeasure from gambling as such. (See von Neumann and Morgenstern 1944: 28; for a discussion, see Binmore 2009: 54.)

The next two axioms state that the preference relation is transitive, with respect to both alternatives and exchanges.

ALT 3. *For any* $A, B, C \in \mathcal{A}$*: if* $A \precsim B$ *and* $B \precsim C$ *then* $A \precsim C$

ALT 4. *For any* $A, B, C, D, E, F \in \mathcal{A}$*: if* $(A \rightarrow B) \precsim (C \rightarrow D)$ *and* $(C \rightarrow D) \precsim (E \rightarrow F)$ *then* $(A \rightarrow B) \precsim (E \rightarrow F)$

The next axiom connects the relation between exchanges to the relation between alternatives, and is necessary for the value of an exchange being measurable as the difference between the values of the exchanged alternatives.

ALT 5. *For any* $A, B, C \in \mathcal{A}$*:* $A \prec B$ *just in case* $(C \rightarrow A) \prec (C \rightarrow B)$ *and* $(B \rightarrow C) \prec (A \rightarrow C)$

The next axiom ensures that preferences over exchanges have an additive structure, and can thus be seen as a weak version of von Neumann and Morgenstern's Independence axiom:

ALT 6. *For any* $A, A', B, B', C, C' \in \mathcal{A}$*: if* $(A \rightarrow B) \prec (A' \rightarrow B')$ *and* $(B \rightarrow C) \precsim (B' \rightarrow C')$ *then* $(A \rightarrow C) \prec (A' \rightarrow C')$

The next two axioms are a continuity condition and an Archimedean condition.[31]

ALT 7. *For any* $A, B \in \mathcal{A}$*, if* $A \prec B$*, then there is a* $C \in \mathcal{A}$*, such that* $(B \rightarrow C) \sim (C \rightarrow A)$

ALT 8. *For any* $A, B \in \mathcal{A}$*, if* $A \prec B$*, then there exists a finite sequence of equally preferable exchanges from* A *to* B*.*[32]

[31] I present simplified versions of both axioms.

[32] In other words, if $A \prec B$, then there exists a sequence $X_1 \rightarrow X_2$, $X_2 \rightarrow X_3$, ..., $X_{n-1} \rightarrow X_n$, such that for any $i, j \in \{1, ..., n\}$, $X_i \rightarrow X_{i+1} \sim X_j \rightarrow X_{j+1}$, and $X_1 = A$, $X_n = B$.

Alt's representation theorem can be stated as follows:

Theorem (Alt, 1936)**.** *The following two claims are logically equivalent:*

1. *$\precsim$ satisfies all of Alt 1 to Alt 8.*

2. *There exists a function u, unique up to a positive affine transformation,[33] such that for any $A, B, C, D \in \mathcal{A}$:*

$$u(A) \leq u(B) \Leftrightarrow A \precsim B$$

$$u(B) - u(A) \leq u(D) - u(C) \Leftrightarrow (A \rightarrow B) \precsim (C \rightarrow D)$$

In other words, if (and only if) a preference relation satisfies Alt 1 to Alt 8, then there exists a cardinal utility function that represents the relation.

References

Adler, M. (2012). *Well-Being and Fair Distribution: Beyond Cost-Benefit Analysis*. Oxford University Press.

Adler, M. (2013). The Pigou-Dalton principle and the structure of distributive justice. Available at: https://ssrn.com/abstract=2263536.

Adler, M. (ta.). *Risk, Death and Well-Being: The Ethical Foundations of Fatality Risk Regulation*. Oxford University Press.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école Américaine. *Econometrica*, 21(4):503–546.

Alt, F. (1971/1936). On the measurability of utility. In Chipman, J. S., editor, *Preferences, utility, and demand: A Minnesota Symposium*. Harcourt Brace Jovanovich.

[33]That is, if another function u' also satisfies the two biconditionals in claim 2, then for any $A \in \mathcal{A}$, $u'(A) = au(A) + b$ for some constants $a > 0$ and b.

Asheim, G. B. and Zuber, S. (2014). Probability Adjusted Rank-Discounted Utilitarianism. CESifo Working Paper Series 4728, CESifo.

Binmore, K. (2009). *Rational Decisions*. Princeton University Press.

Bradley, R. (2017). *Decision Theory with a Human Face*. Cambridge University Press.

Bradley, R. (2022). Impartial evaluation under ambiguity. *Ethics*, 132(3):541–569.

Bradley, R. and Stefánsson, H. O. (2017). Counterfactual desirability. *British Journal for the Philosophy of Science*, 68(2):485–533.

Broome, J. (1984). Uncertainty and fairness. *Economic Journal*, 94(375):624–32.

Broome, J. (1991a). Fairness. *Proceedings of the Aristotelian Society*, 91:87–101.

Broome, J. (1991b). *Weighing Goods*. Basil Blackwell.

Broome, J. (2004). *Weighing Lives*. Oxford University Press.

Buchak, L. (2013). *Risk and Rationality*. Oxford University Press.

Buchak, L. (2017). Taking risks behind the veil of ignorance. *Ethics*, 127(3):610–644.

Cohen, I. G., Daniels, N., and Eyal, N. (2015). Statistical versus identified persons: An introduction. In Cohen, I. G., Daniels, N., and Eyal, N., editors, *Identified versus Statistical Lives: An Interdisciplinary Perspective*. Oxford University Press.

Cohen, Daniels, and Eyal (eds.) (2015). *Identified versus Statistical Lives: An Interdisciplinary Perspective*. Oxford University Press.

Diamond, P. (1967). Cardinal welfare, individualistic ethics, and interpersonal comparison of utility: Comment. *Journal of Political Economy*, 75(5):765–766.

Eyal, N. (na). Against fair chances.

Fleurbaey, M. and Voorhoeve, A. (2013). Decide as you would with full information! An argument against ex ante Pareto. In Norheim, O., Hurst, S., Eyal, N., and Wikler, D., editors, *Inequalities in Health: Concepts, Measures, and Ethics*. Oxford University Press.

Frick, J. (2015). Contractualism and social risk. *Philosophy and Public Affairs*, 43(3):175–223.

Frigg, R. and Hoefer, C. (2010). Determinism and chance from a Humean perspective. In Dieks, D., Gonzalez, W., Hartmann, S., Weber, M., Stadler, F., and Uebel, T., editors, *The Present Situation in the Philosophy of Science*, pages 351–271. Springer.

Frigg, R. and Hoefer, C. (2015). The best Humean system for statistical mechanics. *Erkenntnis*, 80(S3):551–574.

Greaves, H. (2015). Antiprioritarianism. *Utilitas*, 27(1):1–42.

Hansson, B. (1988). Risk aversion as a problem of conjoint measurement. In Gärdenfors, P. and Sahlin, N.-E., editors, *Decision, Probability, and Utility*. Cambridge University Press.

Hoefer, C. (2007). The third way on objective proabability: A sceptic's guide to objective chance. *Mind*, 116(463):449–496.

Jeffrey, R. (1990/1965). *The Logic of Decision*. The University of Chicago Press.

Jonsen, A. R. (1986). Bentham in a box: Technology assessment and health care allocation. *The Journal of Law, Medicine & Ethics*, 14(3-4):172–174.

Krantz, D. H., Luce, R. D., Suppes, P., and Tversky, A. (1971). *Foundations of Measurement*, volume 1. Academic Press.

Kreps, D. (1988). *Notes On The Theory Of Choice*. Avalon Publishing.

Loewer, B. (2001). Determinism and chance. *Studies in History and Philosophy of Science Part B 32*, 32(4):609–620.

McKie, J. and Richardson, J. (2003). The rule of rescue. *Social Science & Medicine*, 56(12):2407–2419.

Mongin, P. (2016). Spurious unanimity and the pareto principle. *Economics and Philosophy*, 32(3):511–532.

Nissan-Rozen, I. (2019). The value of chance and the satisfaction of claims. *Journal of Philosophy*, 116(9):469–493.

Orr, S. and Wolff, J. (2015). Reconciling cost-effectiveness with the rule of rescue: the institutional division of moral labour. *Theory and Decision*, 78(4):525–538.

Otsuka, M. and Voorhoeve, A. (2018). Equality versus priority. In Olsaretti, S., editor, *Oxford Handbook of Distributive Justice*. Oxford University Press.

Parfit, D. (2002/1991). Equality or priority? In Clayton, M. and Williams, A., editors, *The Ideal of Equality*. Palgrave.

Parfit, D. (2003). Justifiability to each person. *Ratio*, 16(4):368–390.

Rabinowicz, W. (2002). Prioritarianism for prospects. *Utilitas*, 14(1):2–21.

Roussos, J. (2022). Modelling in normative ethics. *Ethical Theory and Moral Practice*, (25):1–25.

Savage, L. (1972/1954). *The Foundations of Statistics*. Dover Publication.

Segall, S. (2016). *Why Inequality Matters: Luck Egalitarianism, its Meaning and Value*. Cambridge, United Kingdom: Cambridge University Press.

Stefánsson, H. O. (2015). Fair chance and modal consequentialism. *Economics and Philosophy*, 31(3):371–395.

Stefánsson, H. O. (2023). Identified lives 'bias' as decreasing marginal value of chances. Unpublished manuscript.

Stefánsson, H. O. and Bradley, R. (2015). How valuable are chances? *Philosophy of Science*, 82(4):602–625.

Stefánsson, H. O. and Bradley, R. (2019). What is risk aversion? *British Journal for the Philosophy of Science*, 70(1):77–102.

Stone, P. (2007). Why lotteries are just. *Journal of Political Philosophy*, 15(3):276–295.

Stone, P. (2011). *The Luck of the Draw: The Role of Lotteries in Decision Making*. Oxford University Press.

Suppes, P. and Winet, M. (1955). An axiomatization of utility based on the notion of utility differences. *Management Science*, 1(3/4):259–270.

von Neumann, J. and Morgenstern, O. (2007/1944). *Theory of Games and Economic Behavior*. Princeton University Press.

Vong, G. (2015). Fairness, benefiting by lottery and the chancy satisfaction of moral claims. *Utilitas*, 27(4):470–486.

Vong, G. (2020). Weighing up weighted lotteries: Scarcity, overlap cases, and fair inequalities of chance. *Ethics*, 130(3):320–348.

Voorhoeve, A. (2014). How should we aggregate competing claims? *Ethics*, 125(1):64–87.

Wasserman, D. (1996). Let them eat chances: Probability and distributive justice. *Economics and Philosophy*, 12(1):29–49.

Zuber, S. and Asheim, G. B. (2012). Justifying social discounting: The rank-discounted utilitarian approach. *Journal of Economic Theory*, 147(4):1572–1601.