Welfare and Autonomy under Risk

Pietro Cibinel

Forthcoming in *Philosophy and Phenomenological Research*

Abstract

This paper studies the relationship between promoting people's welfare and respecting their autonomy of choice under risk. I highlight a conflict between these two aims. Given compelling assumptions, welfarists end up disregarding people's unanimous preference, even when everyone involved is entirely rational and only concerned with maximizing their own welfare. Non-welfarist theories of social choice are then considered. They are shown to face difficulties, too: either they fail to respect the value of welfare in at least one important sense, or they end up prioritizing different people's welfare differently in non-risky choices, on the basis of their attitudes to risk, which are intuitively irrelevant in this context.

1 Introduction

Everyone agrees that the well-being of individuals matters. Welfarists believe it is all that matters. Welfarism is commonly assumed in social choice theory, and has many proponents in ethics.[1] The aim of this paper is to study the connection between the value of welfare and that of respecting people's autonomy of choice under risk. I present a framework for carrying out this task, and then offer two impossibility theorems about the connection between welfare and autonomy. First, I will show that, given compelling assumptions, any welfarist theory is committed to disrespecting the unanimous preference of individuals for a given policy over another. This will pave the way for an exploration of non-welfarist theories of social choice that respect people's autonomy in conditions of risk. For illustration

[1]For the significance of welfarism in social choice theory, see, e.g., Mongin and d'Aspremont (1998, p. 394) and Adler (2019, p. 27); for discussions of welfarism in ethics, see, e.g., Moore and Crisp (1996) and Keller (2009).

purposes, I will consider two classes of non-welfarist theories, and discuss distinct problems that they each face; I will then show that there is no non-welfarist theory that avoids both of those types of problem.

Let me introduce in more detail my contribution by situating it within two broad traditions. First, there is a variety of well-established objections to welfarism based on appeals to the value of autonomy.[2] Different authors are interested in different facets of this value, but *non-interference* is usually a common focus. Famously, Sen (1970), (1979) highlights a tension between welfarism and a certain liberal value of individual discretion over personal choices. Sen shows that a unique concern for welfare in social choice rules out the possibility of a sphere of private concern over which people have freedom to act as they wish. Nozick (1974) pursues a distinct but related line of thought when he argues against "patterned" principles of distributive justice—which include all welfarist theories—on the basis that no such principle "can be continuously realized without continuous interference with people's lives" (Nozick, 1974, p. 163), since "liberty upsets patterns" (Nozick, 1974, p. 160).

Here, I will be interested in an aspect of the value of autonomy that is only loosely connected to non-interference, and refers instead to a form of deference to people's own attitudes towards risk. This notion comes from the second broad tradition that I shall address, which is the literature on social choice under risk.[3] The main question in this area is as follows: suppose we need to choose one of various possible policies, each of which will affect the welfare of a number of individuals. We are uncertain about the outcome that each policy will bring about, but can assign probabilities to the different possible states of the world. What principles should guide our decision in such a situation?

One of the most widely discussed principles encapsulates the ideal of autonomy

[2]There are then authors who deny welfarism for other reasons; I shall mention two strands of literature. One rejects welfarism—at least if understood as a unique concern with "final well-being" as a measure of "how well [a person's] life goes" (Voorhoeve & Fleurbaey, 2016, p. 936)—because of concerns with the *fairness* of how a certain distribution of welfare is attained—see, e.g., Voorhoeve and Fleurbaey (2016) and Voorhoeve (2021). The other rejects welfarism to take account of equality of opportunity, where social evaluation is based on information relating to an individual's *effort* as well as on how well off they are—see Ferreira and Peragine (2016) for an overview.

[3]For overviews of this literature, see Mongin and Pivato (2016), Fleurbaey (2018), and Adler (2019).

that will be our focus: it roughly says that, if everyone involved prefers a certain policy over another, then we shouldn't choose the latter. Versions of this thought are usually referred to as the *ex ante Pareto* principle. I shall later qualify this principle, and make it precise. What is important to note now is that this principle, together with *dominance* reasoning, gives rise to a range of striking impossibility results. Dominance reasoning connects the evaluation of acts to that of outcomes—the principle of statewise dominance, in particular, says that if an act results in at least as good an outcome as another act under every possible state of the world, then the former act is at least as good as the latter. The impossibility results of interest here all trace back to Harsanyi 's (1955) aggregation theorem, and build upon it in various ways.[4] A strengthened version of Harsanyi's result shows the following (Fleurbaey, 2009, 2010). Assume that individual preferences over policies satisfy the axioms of the orthodox view in decision theory, expected utility theory. Then, under weak assumptions, ex ante Pareto and statewise dominance entail that policies must be ranked at the social level by a sum of individual decision-theoretic "utilities"—these are the numbers representing the quantity whose expectation individuals maximize.[5] When this quantity is equated with welfare, Harsanyi's aggregation theorem becomes an argument for utilitarianism, and this is indeed how Harsanyi himself interpreted his theorem (Harsanyi, 1977, 1982). The equation of decision-theoretic utility and welfare means that individuals are expected welfare maximizers, an assumption that has come to be known as Bernoulli's Hypothesis, or *Bernoulli* for short (Broome, 1991b; Adler, 2019).

Some recent work shows that if we reject Bernoulli by relaxing expected utility theory and allow individuals to have heterogeneous preferences for policies, the conflict between ex ante Pareto and statewise dominance is not solved, but rather exacerbated: those who evaluate outcomes in a utilitarian fashion face it too (Blessenohl, 2020; Nissan-Rozen, 2020; Bradley, 2022).[6] It is tempting to interpret these recent

[4]A list that is not exhaustive but representative for our purposes includes Harsanyi (1977), (1982), Broome (1991b), Weymark (1991), Mongin (1994), Fleurbaey (2009), (2010), Mongin and Pivato (2015), Danan et al. (2015), Blessenohl (2020), Nebel (2020), Nissan-Rozen (2020), Bradley (2022), Fleurbaey and Zuber (2022), Kowalczyk (2022), Doody (manuscript).

[5]To get the mathematical expectation of an uncertain quantity, we multiply the value the quantity takes in each possible state of the world by the probability of that state, and then sum all such probability-weighted values.

[6]There are two other ways of relaxing Bernoulli in the context of Harsanyi-style results: one is to relax expected utility but assume that there is a homogenous ranking of policies among individuals (McCarthy et al., 2020), the other involves relaxing expected utility theory as an assumption but

impossibilities as arguments in favor of expected utility theory. However, restoring expected utility theory is not enough. For we can relax Bernoulli while retaining expected utility at the individual level. This was in fact part of Sen (1976), (1977) and Weymark's (1991), (2005) critique to Harsanyi's interpretation of his theorem (see also Roemer, 2008, pp. 141-45). On the Sen-Weymark line of thought, utilitarianism relies on a quantitative notion of welfare that is independent of an individual's preferences in conditions of risk (Sen, 1986, p. 1123). Given such independently given quantities of welfare, requiring Bernoulli amounts to an unmotivated substantive assumption of risk neutrality.

This paper will build upon the Sen-Weymark observation, and offer an alternative interpretation of the impossibilities stemming from Harsanyi's aggregation theorem, which is focused on questioning welfarism.[7] In Section 2, I shall relax Bernoulli but leave open that individuals may be expected utility maximizers. I then show that, given some plausible principles involving dominance reasoning and impartiality at the level of social evaluation, ex ante Pareto entails that welfarism must be false. I will lay out a framework for social choice under risk in Section 3, and use it to state the conflict between autonomy and welfarism precisely in terms of a first impossibility theorem. An intuitive illustration of this result is given in Section 4. My interest in presenting this argument is partly to utilize it as a springboard for considering non-welfarist views that respect the value of autonomy. Section 5 begins an exploration of what such views might look like; I describe two different types of non-welfarist views: one ends up disregarding the value of welfare altogether, while the other gives priority to people's welfare on the basis of their risk attitude. Section 6 generalizes the point via a second impossibility theorem, showing that any autonomy-based non-welfarist theory entails some version of one of these two implications. Section 7 concludes by considering different reactions one might have to the results of the paper.

---

deriving it back as a conclusion (Mongin & Pivato, 2015; McCarthy et al., 2020).

[7] Mogensen (2024) interprets the results by Blessenohl (2020), Nebel (2020), and Bradley (2022) as showing that the motivation for respecting people's preferences *ex ante* (in the context of Buchak's (2013) non-orthodox decision theory) is incompatible with consequentialism—cf. Frick (2013, p. 132). In this paper, I retain instead a broadly consequentialist framework and pin the blame on the claim that only welfare matters to the social evaluation of outcomes. Blessenohl (2020, pp. 505-506) briefly mentions a non-welfarist response to his impossibility result, a version of which I took up and defended in Cibinel (2022). My aim here is to explore this suggestion at a much greater level of generality.

Before we delve into the details, it will be useful to have a rough version of the conflict between welfarism and autonomy on the table; I illustrate it now with an example. There are only two individuals affected, and we need to choose between one of two policies. Policy 1 gives 10 units of welfare to the first individual, and 0 to the second, if event E occurs.[8] It gives 0 units of welfare to the first individual, and 10 to the second, if event $\neg E$ occurs. Each of E and $\neg E$ have probability 0.5 of occurring. Policy 2 gives 5 units of welfare to both individuals, no matter whether E or $\neg E$ obtains. Perhaps, both individuals are suffering from a moderate illness (welfare level 0), and there is a medicine that will fully cure the first individual (welfare level 10) and leave the second ill if E obtains, and fully cure the second individual and leave the first ill if $\neg E$ obtains. However, a second medicine will guarantee partial recovery for both (welfare level 5). We can represent the case in matrix form, as follows.[9]

	E	$\neg E$
policy 1	(10, 0)	(0, 10)
policy 2	(5, 5)	(5, 5)

Each cell in the matrix represents a certain final outcome that might obtain; a pair of numbers (r_1, r_2) stands in for the first individual getting r_1 units of welfare, and the second individual r_2 units of welfare, in the given outcome. Welfarists think that well-being is all that matters to evaluating outcomes. They say: no matter what other circumstances obtain, there is a fixed fact of the matter as to whether, for example, the outcome of policy 1 under E is better than the outcome of policy 2 under E. The matter is fixed, according to welfarists, because we know how well off each individual would be in those outcomes. Since nothing else matters, this is enough to evaluate them.

[8] The example makes certain assumptions about welfare measurement that are carefully considered only in Section 2.

[9] This kind of case is very often used to illustrate the puzzles raised by Harsanyi's aggregation theorem; see, e.g., Fleurbaey and Voorhoeve (2013, p. 116), Frick (2015, p. 191), Nissan-Rozen (2017), and Doody (manuscript). The connection it exemplifies between risk aversion and equality on the one hand, and risk-seekingness and inequality on the other, predates even Harsanyi's aggregation theorem—see, e.g., Friedman (1953, p. 278). See also the literature on "option luck", concerning whether individuals should be held responsible for the results of "deliberate and calculated gambles" (Dworkin, 1981, p. 293). Footnote 28 below compares my result based on this case, Theorem 1, and additional related work—especially Hammond (1981) and Fleurbaey and Zuber (2021b), (2022).

However, on pain of violating people's autonomy, we can show there *are* additional circumstances that should affect our evaluation of outcomes. One way to fill in the details of the case might be this. Suppose both individuals are risk-seeking and prefer policy 1 to policy 2. By the ex ante Pareto principle, policy 1 is indeed better than policy 2 on this version of the case. But, plausibly, policy 1 can't be better unless there is some state in which its outcome is better than that of policy 2. Moreover, clearly the outcome of policy 1 under E is exactly as good as the outcome of policy 1 under $\neg E$. So, on this version of the case, both of the possible outcomes of policy 1 are better than the single outcome that policy 2 is sure to result in.

Here is a different way to fill in the details of the case. Suppose both individuals are risk-averse and prefer policy 2 to policy 1. By the ex ante Pareto principle, policy 2 is indeed socially better than policy 1 on this version of the case. But, plausibly, policy 2 can't be better unless there is some state in which its outcome is better than that of policy 1. Moreover, since the two possible outcomes of policy 1 are equally good, the single outcome that policy 2 is certain to result in must be better than both, on this version of the case.

Under one set of circumstances, an outcome with distribution of welfare $(10, 0)$—or $(0, 10)$—is better than one with distribution of welfare $(5, 5)$. Under another set of circumstances, the outcome with welfare distribution $(5, 5)$ is better than that with welfare distribution $(10, 0)$—or $(0, 10)$. The facts about welfare remain the same, but our evaluation changes. Therefore, welfarism is false. Very informally, this is the argument that motivates the study that follows.

2 Welfare and Risk

This section clarifies the assumptions about welfare and risk attitudes that the argument just presented relies upon. We begin with a qualification on ex ante Pareto. As I said, this principle gives voice to a certain ideal of autonomy, to do with respecting people's attitudes to risk when their welfare is at stake. My earlier, first pass formulation was this:

> if every individual involved prefers a certain act or policy over another,

then we shouldn't choose the latter.

Without qualification, this principle is dubious. Suppose the individuals involved prefer a certain act over another, but we know that this is because they have false beliefs, either about their environment or about what their good consists in. Even if we don't know that their beliefs are false, we might know that they are based on much less evidence than we possess, and that their preferences would change if they knew more. Or perhaps we all share the same beliefs and evidence, but these individuals' preferences are in some sense irrational. A staunch opponent of paternalism might claim that even so, we should defer to what these people want. But some would concede that in such circumstances people's preferences may legitimately be overridden.[10] I will leave this issue to one side. I am concerned with a special kind of case, in which, I think, failing to respect people's preferences is much more problematic.

Suppose every individual affected shares the same evidence, and they all share the same evidence with us, who are engaged in the exercise of ranking policies. All of our beliefs are modeled by a single, given probability distribution over the possible states of the world. Moreover, the individuals affected know exactly what would be ultimately good for them, always and only preferring outcomes in which they get more welfare (we needn't make the rather implausible assumption that everyone is self-interested, but only the more innocuous one that self-interested preferences are the legitimate input to social choice). Individuals' preferences over policies with uncertain consequences are rational, in the most stringent sense of the term. First, they are coherent in the sense of satisfying the axioms of the correct decision theory. Second, they are reasonable in the sense of not being in any other way criticizable from the perspective of someone who is uniquely concerned with their interest.[11] In this context at least, the ex ante Pareto principle is very hard to resist. It is this restricted version of the principle that I shall use throughout the paper.

[10] See Broome (1987, p. 408) for an early statement of the problems that arise for ex ante Pareto when probabilities differ; cf. Mongin (2016), who suggests that ex ante Pareto should be rejected in all cases of "spurious unanimity", in which (roughly) people's preferences all agree but are based on different reasons.

[11] For this second sense of rationality, I draw on Parfit (2011, p. 33). Buchak (2017, p. 620) and Hájek (2021, p. 190) specifically give a version of this distinction in relation to attitudes to risk. Those skeptical of there being anything more to rationality than coherence can simply disregard this point: my assumption, for them, is just that all individuals are rational in the sense of having preferences that are coherent.

I shall assume that we can compare the welfare levels and improvements of different people on a single scale, before any apparatus for dealing with risk is introduced—this assumption corresponds to taking seriously the Sen-Weymark challenge to Harsanyi mentioned in Section 1. In the technical jargon, welfare is cardinally measurable and fully comparable. What this means is that it makes sense to say things like "person 1's welfare in some outcome is greater than person 2's welfare in another" and "person 1 would benefit more if we acted in this way than person 2 would if we acted in that way". These assumptions make it possible to represent welfare levels with numbers.[12] When, below, I say that a certain individual gets welfare level x in a certain outcome, with x a real number, what I mean is that this individual gets x units of welfare as measured by a scale appropriately chosen to reflect the assumptions about welfare that I just described (i.e., what is called an *interval scale*, common among all individuals).

These assumptions also entail that it makes sense to speak of risk neutrality, risk aversion, and risk-seekingness in relation to welfare. I will call someone *risk-neutral* if and only if they are always indifferent between policies that yield the same expectation of welfare for them, *risk-averse* if and only if they always prefer a policy that gives them x units of welfare for sure to a policy with uncertain consequences that gives them x units of welfare in expectation, *risk-seeking* if and only if they always prefer a policy with uncertain consequences that gives them x units of welfare in expectation to a policy that gives them x units of welfare for sure.[13] Orthodox decision theory and its rivals alike make room for risk neutrality, risk aversion, and risk-seekingness for welfare (see Broome, 1991a; Buchak, 2013; Bottomley & Williamson, 2024). So there is no doubt that both risk-averse and risk-seeking preferences, as well as risk-neutral ones, can be rational in the sense of satisfying the axioms of the correct decision theory. All I will assume about a person's preference relation over policies is that it is an ordering (namely, it is reflexive, complete, and transitive); this makes what I have to say compatible with

[12] See Krantz et al. (1971, ch. 4), for the relevant representation theorems; Adler (2019, p. 268) and Nebel (2024) specifically apply Krantz et al.'s results to the issue of welfare measurement.

[13] This is a standard characterization of risk preferences with respect to a given quantity—see, e.g., Buchak (2013, p. 21-22) for a discussion of it and of related characterizations. Note also that my definitions of risk aversion and risk-seekingness are "global"; it is worth noting that my results below actually only rely on risk aversion and risk-seekingness being displayed "locally" with respect to some interval of welfare levels.

orthodox decision theory, but also with any of its main rivals.

A further question one might ask is which patterns of preferences for risk are reasonable, in the sense of being immune from any kind of criticism from the perspective of someone concerned with the person's interest. We have already assumed that people always prefer more welfare to less. The question is whether there is any prudential advantage that one particular attitude—for example, risk neutrality—possesses over the others. I shall assume that the answer is "no". Perhaps extreme levels of proneness to risk, as well as extreme levels of risk aversion, can be ruled out on this basis. But there is certainly a range of entirely reasonable preferences for risk involving one's welfare.[14] I will call this thesis *welfare-risk permissivism*.

The thesis of welfare-risk permissivism gives a positive spin to the negative observation of Sen and Weymark, according to which Harsanyi is unwarranted in assuming Bernoulli. Before moving on, we should briefly consider two related attempts to salvage Bernoulli. The first strategy claims that we do not in fact possess any independently given quantitative notion of welfare, relative to which individuals can display differing risk attitudes. The suggestion would then be that equating decision-theoretic utility with welfare allows us to give meaning to quantitative statements concerning welfare, which would otherwise make no sense or display semantic indeterminacy (Broome, 1991b, 2004; Risse, 2002; Greaves, 2017). The second strategy claims that, while we do have some quantitative notion of individual "benefit" prior to choice under risk, an individual's *welfare* is nonetheless affected not only by what benefit they receive, but also by the probability that was associated with receiving that benefit.[15] An individual may be risk-averse with respect to what benefits they get, but should come out as risk-neutral with respect to welfare itself.

Both of these strategies, however, deliver counterintuitive verdicts, owing to the fact that risk attitudes seem irrelevant to determining how well off an individual is

[14] Versions of this thought are articulated by Buchak (2017), Hájek (2021), and (on one reading) Broome (1991a, p. 6). Broome (1991b), however, ends up accepting a version of Bernoulli, which requires reasonable people to be expected welfare maximizers. I will discuss this issue presently.

[15] This strategy could be developed by appealing to Stefánsson and Bradley's (2015) denial of what they call "chance neutrality"; see also Goldschmidt and Nissan-Rozen (2020) and Bradley and Stefánsson (2023). Cohen et al. (2022, p. 2562) specifically show how such a strategy can help respond to Nissan-Rozen's (2020) impossibility, which I mentioned above. I thank a referee for inviting me to consider this view.

(see Fleurbaey & Zuber, 2022). To illustrate, suppose you are uniquely concerned with a particular individual's welfare. You don't know whether this individual is risk-averse or risk-seeking with respect to experiencing pain but you do know that they ended up in some state of pain. Would you rather learn that they are risk-seeking and in more pain, or risk-averse and in less pain? On both of the strategies for salvaging Bernoulli just considered, there will be attitudes to risk and states of pain such that you should be indifferent between which of the two pieces of news you'd rather learn.[16] But if all you genuinely care about is this person's welfare, you should be relieved to know they are in less pain, and indifferent about what risk attitude they have. So, you should want the individual risk-averse and in less pain. To the extent that we might be inclined to prefer them risk-seeking and in "more pain", it seems to me that this is because it comes natural to associate risk-seekingness with a greater ability to endure misfortune. But this is not how we are supposed to think about this example, for surely such an ability would affect how much pain one feels. Notice, moreover, that while the example is specifically about pain, the point it makes does not rely on assuming hedonism about welfare. We could instead suppose that the individual fares better, if risk-averse, along any other objective good (such as friendship, knowledge, and so on) or with respect to desire satisfaction and strength of preference (provided that the latter is measured independently of risk, which can be done along the lines indicated by Bell and Raiffa (1988)).

This argument no doubt falls short of a conclusive refutation of the strategies considered for salvaging Bernoulli.[17] But it does at least suggest that following the Sen-Weymark critique and severing our notion of quantities of welfare from that

[16]On the first strategy, since the shape of the individual's expected utility function is different depending on whether they are risk-averse or risk-seeking with respect to pain, equal increases in pain can lead to different welfare losses. See Adler (2019, pp. 55-64) and Fleurbaey and Zuber (2021a) for more on this topic. Below we will consider a view where expected utility is not a measure of welfare differences, but rather of the moral significance of welfare differences—see Section 5.1 and especially footnote 29 (compare Nebel, 2022, p. 24). On the second strategy, if we assume that the pain resulted from a risky option, the welfare loss that the risk-seeking individual gets from the risk (i.e. the antecedent probability of that pain) will be lower than for the risk-averse (and may even be a gain, see Stefánsson & Bradley, 2015, p. 605). This will make up for the added pain the risk-seeking individual experiences.

[17]I aim to explore in much greater detail how quantities of welfare can be obtained independently of risk in another paper; see Dietrich (manuscript) for a recent proposal on how to disentangle welfare from decision-theoretic utility.

of risk attitudes has intuitive appeal. The study of the paper can then be seen as exploring the consequences of this approach for those who want to both promote welfare and respect autonomy under risk.

3 Framework and First Result

I now lay out a framework that will allow us to state the conflict between autonomy and welfarism more precisely. The framework I use combines a simplified version of Sen's (2017) approach of *social welfare functionals* with a minimal apparatus for choice under risk.[18] To reiterate, the key component of our approach will be that cardinal facts about welfare are given; people's preferences between policies are then formed on the basis of what quantities of welfare they would receive from each policy, with which probabilities.

We assume that there is a set N of n individuals whose welfare our choices might affect, with $n \geq 2$. There is a set X of outcomes, each of which includes all information relevant for purposes of ethical evaluation: any outcome will certainly tell us about every individual's welfare, but we leave open that it may contain other information, too, since we don't want to prejudge the question of welfarism. Information specifically about welfare is captured by a fixed welfare function W, which associates to each outcome a list of welfare levels: the first entry in the list is a number representing the welfare level the first individual gets in that outcome, the second entry is a number representing the welfare level of the second individual, and so on.[19] So, for example, with two individuals, there may be some outcome x such that $W(x) = (10, 0)$. To single out the i^{th} entry in this outcome's list of welfare levels, we write $W_i(x)$. Individual preferences over outcomes in X are encoded in

[18]Two early frameworks that make use of Sen's approach in the context of choice under risk are given by Mongin (1994) and Mongin and d'Aspremont (1998). For two recent frameworks, more similar to mine, see Chambers and Echenique (2012) and Fleurbaey and Zuber (2021a)—some differences are highlighted in footnotes below. Another connection is with frameworks that explicitly make room for a non-welfare component, if only to better characterize welfarism (see Blackorby et al., 2005, ch. 3, 2006). The non-welfare component, in my framework, will be given by the apparatus for choice under risk.

[19]Our assumption of a fixed welfare assignment should be acceptable for those who think that there couldn't be a welfare difference for an individual without a difference in the outcome that individual gets. Those who reject this claim should see the choice of a fixed welfare profile as a simplification to limit the number of variables in my framework; all my results could be restated in a more general setting with multiple welfare profiles (following Sen (2017) more closely).

the welfare function W, given our assumption that all individuals ultimately care about is improving their welfare—or, at least, that such self-interested concern is the legitimate input to social choice.

To model uncertainty, let S be a finite set of k states of nature, and let $\pi(s)$ be the probability of state s. This fixed probability function reflects the assumption I made that everyone shares the same evidence and beliefs. I will assume that S can be partitioned into two equiprobable events, E and $\neg E$; these could be, for instance, the two possible results of a fair coin flip. There is a set A of policies or acts, which we model as functions from states of nature to outcomes. Crucially, we do *not* make the following common assumption (following Broome's (1991b, p. 80) nomenclature).

> RECTANGULAR FIELD: A is the set of all functions from states to outcomes.

In other words, we leave open that only some of the functions from states to outcomes stand in for genuine policies we may implement. The reason for this will emerge in due course, but is roughly the following: it is important for some of the theories I will consider below that outcomes include information about how they come about. If an outcome x results from a genuinely risky act, and such information is included in the outcome itself, it would indeed make little sense for there to be a function in A that returned x in all states.

Each individual i associates to each act a a prospect a_i, which tells us how much welfare i gets in each of the states, if act a is chosen.[20] Individuals have preferences over prospects, encoding their risk attitudes: for each individual i, there is a reflexive, complete, and transitive relation R_i—that is, a "weak preference" ordering—over the prospects A offers to i. Strict preference and indifference are understood in the usual way and denoted by P_i and I_i respectively.[21] I assume that individual preferences are defined over the same set of prospects, so that it makes sense to say that two individuals share the same preference relation or risk attitude.[22] Notice that, when a prospect results in a given outcome with certainty, the

[20] Formally, for all states $s \in S$, $a_i(s) = W_i(a(s))$.

[21] For all acts $a, b \in A$, if $a_i R_i b_i$ and not $b_i R_i a_i$, we write $a_i P_i b_i$. If both $a_i R_i b_i$ and $b_i R_i a_i$, we write $a_i I_i b_i$.

[22] This imposes some constraints on the sets X of outcomes and A of acts. In particular, if $i, j \in N$ and $a \in A$ are fixed, there must be $b \in A$ such that $a_i = b_j$. In turn, this has implications for what

individual's ranking of that prospect is determined by how much welfare it gives to that individual: the more, the better.

These orderings over prospects R_i are grouped together by an n-tuple $R = (R_1, ..., R_n)$, which we call a risk profile, or simply a profile. The set of all rational risk profiles is denoted by $\mathcal{R}$. A risk profile is rational if it is made up entirely of rational orderings, both in the sense of rationality as coherence and in the sense of rationality as reasonableness. We can now state formally the claim that there are multiple rationally permissible attitudes to risk involving one's welfare.

> WELFARE-RISK PERMISSIVISM: $\mathcal{R}$ includes profiles with risk-neutral, risk-averse, and risk-seeking orderings of prospects.

Notice that all results below actually only need there to be two distinct orderings—for example, one risk-averse and one risk-seeking. With this setup in place, we can formulate the problem of social choice under risk as follows: given any profile R, we want to be able to rank acts and outcomes. Our primary focus will be on the following kind of functional relations:

- A *social quasi-ordering function* is a function that maps each profile $R \in \mathcal{R}$ to a pair of reflexive and transitive social preference relations, i.e. quasi-orderings, $(\succeq_R^A, \succeq_R^X)$, with $\succeq_R^A$ over A and $\succeq_R^X$ over X.

Strict preference and indifference over acts are denoted by $\succ_R^A$ and $\sim_R^A$, each again defined in the standard way; the analogous convention holds for preference over outcomes.[23]

A note about the need for two social preference relations, one over acts and another over outcomes. Had I assumed Rectangular Field above, I could have defined the social preference relation over outcomes in terms of how $\succeq_R^A$ ranks *sure acts*, that is, acts that return the same outcome in every state of the world. But I did not make this assumption, so there is no guarantee that, given some outcome x, we will find an act a in A that returns x in every state. If risk has a certain kind of normative relevance, there might be good reasons to think that such a search should systematically fail, as briefly hinted at above and explained more carefully below (see also

outcomes exist. However, this assumption is uncontroversial if we think of acts as *possible* policies a decision-maker might face, and of outcomes as merely possible final states of affairs.

[23]For all $a, b \in A$, and $R \in \mathcal{R}$, if $a \succeq_R^A b$ and not $b \succeq_R^A a$, we write $a \succ_R^A b$; if both $a \succeq_R^A b$ and $b \succeq_R^A a$, we write $a \sim_R^A b$. The analogous convention holds for $\succeq_R^X$.

Broome, 1991b, pp. 115-117). For readability, however, I will omit the superscripts when no confusion arises.

We can finally address the problem of social choice under risk: how should we rank acts and outcomes? We impose three axioms on social quasi-ordering functions. For each axiom, I will first describe it informally, and then give a formal statement. The first axiom is a package of dominance principles, which guarantees that our evaluation of acts coheres with our evaluation of outcomes. This axiom has two parts. Part (i) says: if an act is better than another socially speaking, then the outcome of the former is better than the outcome of the latter in at least some possible state of the world. And part (ii) says: if the outcome of an act is socially better than the outcome of another act in every state, then the former act is socially better than the latter.

> DOMINANCE PACKAGE: For all profiles $R \in \mathcal{R}$, and for all acts $a, b \in A$, (i) if $a \succ_R b$ then there exists $s \in S$ such that $a(s) \succ_R b(s)$; (ii) if $a(s) \succ_R b(s)$ for all $s \in S$, then $a \succ_R b$.[24]

Next, we have the ex ante Pareto principle discussed in Section 2. This principle ensures that if everyone affected prefers their prospect with some act when compared to their prospect with another act, then the former act is indeed socially better than the latter. Formally, we have

> EX ANTE PARETO: For all profiles $R \in \mathcal{R}$ and acts $a, b \in A$ such that $a_i \neq b_i$ for some $i \in N$, if $a_i P_i b_i$ for all i such that $a_i \neq b_i$, then $a \succ_R b$.

The third principle is required for impartiality, or equal concern for all individuals. Say that two acts a and b are permutations of one another if they distribute the very same prospects, though possibly to different people.[25] In words, the principle states that when everyone shares the same preferences for risk, and two acts are

[24] Some (e.g., Hare, 2010; Bader, 2018) will be suspicious of part (i) of Dominance Package because, given incompleteness, it conflicts with stochastic dominance in cases of "opaque sweetening", for which see Hare (2010). While I do not assume completeness for social preference, this needn't worry us. Variants of my results could be proved using the following weaker principle, which Hare's case gives us no reason to doubt: if a is better than b then *either* a stochastically dominates b *or* there exists a state in which the outcome of a is better than that of b. For my first result, alternatively, Lederman's (forthcoming) "negative dominance" could be employed instead. I use Dominance Package (i) to avoid complicating my discussion and proofs.

[25] Formally, there is a bijection $\sigma : N \rightarrow N$ such that for all $i \in N$, and for all $s \in S$, $W_i(a(s)) = W_{\sigma(i)}(b(s))$.

permutations of one another, then it can't be that one of the two acts is better than the other.

> HOMOGENOUS ANONYMITY: For all profiles $R \in \mathcal{R}$ and acts $a, b \in A$, if a and b are permutations of one another, and $R_i = R_j$ for all $i, j \in N$, then neither $a \succ_R b$ nor $b \succ_R a$.

Our three axioms constrain social quasi-ordering functions in different ways. Dominance Package constrains social quasi-ordering functions with respect to how they relate social preference over acts in A to social preference over outcomes in X, in any given profile. Ex Ante Pareto and Homogenous Anonymity constrain social quasi-ordering functions with respect to how they rank acts in any given profile. When combined with Dominance Package, these two axioms are capable of entailing verdicts about social preferences over outcomes. Notice, however, that we do not assume any axiom directly about social preference over outcomes.

Lastly, we need to characterize welfarism. In social choice theory, the welfarist stance is regimented in terms of the thesis that "[...] the ethical ranking of outcomes is determined by individual well-being. [...] It takes a well-being difference to make an ethical difference in the outcomes" (Adler, 2019, p. 27). In the current framework, this thesis can be formulated as follows: there is a single reflexive and transitive relation on lists of welfare levels that fully determines the social ranking of outcomes. Formally, call $\mathcal{W}$ the subset of $\mathbb{R}^n$ (which is the set of all n-tuples of real numbers) such that $(r_1, ..., r_n) \in \mathcal{W}$ just in case there exists an outcome x in X with $W(x) = (r_1, ..., r_n)$. Then we have

> WELFARISM: There exists a unique quasi-ordering $\succeq$ on $\mathcal{W}$ such that for all outcomes $x, y \in X$ and for all profiles $R \in \mathcal{R}$, $x \succeq_R y$ if and only if $W(x) \succeq W(y)$.[26]

We can better understand the commitments of Welfarism by noticing that it entails two distinct principles. The first, Pareto Indifference, says that—keeping the risk profile fixed—two outcomes associated with the same list of welfare levels must be ranked as equally good. The second, Independence of Irrelevant Alternatives, says that changing non-welfare aspects of a situation should not affect the ranking

[26]Cf. Bossert & Weymark, 2004, p. 1107. Standard formulations often assume that the "welfarist" relation $\succeq$ is complete; since every ordering is a quasi-ordering, but not the other way around, the formulation I use here is more general.

of outcomes. In our framework, this means that tweaking what risk attitudes individuals have should not make a difference to how outcomes are assessed: a social quasi-ordering function satisfying Independence of Irrelevant Alternatives assigns the *same* quasi-ordering over outcomes to every profile in $\mathcal{R}$. Formally, we have (cf. Bossert & Weymark, 2004, pp. 1105-1106):

> PARETO INDIFFERENCE: For all outcomes $x, y \in X$ and for all profiles $R \in \mathcal{R}$, if $W(x) = W(y)$, then $x \sim_R y$.
>
> INDEPENDENCE OF IRRELEVANT ALTERNATIVES: For all outcomes $x, y \in X$ and for all profiles $R, R' \in \mathcal{R}$, $x \succeq_R y$ if and only if $x \succeq_{R'} y$.[27]

Since these two principles are entailed by Welfarism, if either of them fails Welfarism does, too.

We are now in a position to present the first main result of the paper formally: in our framework, Welfarism is incompatible with the conjunction of Dominance Package, Ex Ante Pareto, and Homogenous Anonymity—given a technical assumption about the sets of outcomes and policies which I will call Domain Condition.

> **Theorem 1.** *Given Welfare-Risk Permissivism and Domain Condition, no social quasi-ordering function satisfies Dominance Package, Ex Ante Pareto, Homogenous Anonymity, Pareto Indifference, and Independence of Irrelevant Alternatives.*[28]

[27] Chambers and Echenique (2012, p. 587) and Fleurbaey and Zuber (2017, p. 680) consider a principle structurally identical to Independence of Irrelevant Alternatives, which they call "Invariance to risk attitudes for constant acts". Substantially, however, the two principles are different: both Chambers and Echenique and Fleurbaey and Zuber only assume, as a matter of framework, ordinal non-comparability for individual welfare. As a result, their principle is as much a restriction on well-being comparisons and measurement as it is on social evaluation. Importantly for our purposes, a denial of their principle is compatible with Welfarism. For example, they understand accepting Bernoulli as a rejection of their principle (ibid.). In my framework, where cardinal measurability and full comparability of welfare are built-in, accepting Bernoulli should be seen as claiming that $\mathcal{R}$ only includes the risk-neutral ordering of prospects. Bernoulli thus entails Independence of Irrelevant Alternatives.

[28] This theorem can be related to two strands of extant work, both connected to Harsanyi's aggregation theorem. First, the conflict it points to is reminiscent of the debate spawned by Harsanyi on whether the "social welfare function" should be applied *ex ante* or *ex post*; see especially Hammond (1981), (1982). The main difference here is that my theorem is *multiprofile*: in this context, welfarism corresponds to the idea that the social welfare function *ex post* should not depend on what profile we're in. Denying welfarism in this sense is compatible with a range of approaches to the *ex ante-ex post* debate, including a rapprochement of the two perspectives, with the ex post social

The proof of Theorem 1 is given in the Appendix, together with a statement of the Domain Condition, which is of little philosophical import. Essentially, by combining Ex Ante Pareto and Homogenous Anonymity (which concern social preference over acts) with Dominance Package (which connects social preference over acts with social preference over outcomes) we get verdicts about social preference over outcomes that contradict Welfarism (which concerns social preference over outcomes). The next section explains the impossibility that Theorem 1 gives rise to, giving a more detailed exposition of the argument sketched in Section 1.

4 A More Detailed Illustration

Suppose there are only two individuals, and that they share the same ordering of prospects. For ease of exposition, I assume completeness (i.e., that $\succeq_R^A$ and $\succeq_R^X$ are orderings); Theorem 1 shows it is not actually needed. Consider first a choice between the following two sure acts, a and b, where s^* is the disjunction of all states.

	s^*
a	(10, 0)
b	(0, 10)

Act a gives 10 units of welfare to individual 1 and 0 to individual 2 for sure, and

welfare function differing in each profile (Secion 5.1), and a denial of the ex post approach, which however retains dominance reasoning (Section 5.2). The second strand of extant work is more recent and shares with the present paper a focus on multiprofile results, together with a rejection of Bernoulli; see Chambers and Echenique (2012) and Fleurbaey and Zuber (2017), (2021b), (2022). The closest results are by Fleurbaey and Zuber (2021b), (2022): they show that social evaluation cannot both comply with statewise dominance and a weak version of ex ante Pareto while not paying attention to risk attitudes in the absence of risk (see their Theorem 1 in Fleurbaey & Zuber, 2021b, p. 14, and their Theorem in Fleurbaey & Zuber, 2022, p. 459). Here is a comparison with my Theorem 1. Formally, their results rely on social evaluation *across* profiles, e.g. "outcome x with profile R is better than outcome y with profile R'". My result, by contrast, only relies on intraprofile social evaluation. The proof strategy is nonetheless similar; each proof proceeds by considering two profiles with homogenous but different risk attitudes. At the philosophical level, they do not interpret their results as questioning welfarism. This is because, in their framework, risk attitudes are *not* with respect to independently given and interpersonally comparable quantities of welfare (as they are in mine). So, they take a denial of their principles forbidding risk attitudes to count in the absence of risk to mean that risk attitudes must play a role in the measurement of individual welfare (Fleurbaey & Zuber, 2021b, p. 25, 2022, p. 463). They do not consider the alternative route, which my Theorem 1 is precisely supposed to highlight, of claiming that while risk attitudes do not affect individual welfare measurement, they do affect social evaluation in a way that is inconsistent with welfarism.

b gives 0 units of welfare to individual 1 and 10 to individual 2 for sure. a and b are thus permutations of one another. By Homogenous Anonymity, neither is better than the other. By the contrapositive of Dominance Package (ii) and completeness, a welfarist must conclude that:

> $(*)$ any outcome with welfare distribution $(10, 0)$ is exactly as good as any outcome with welfare distribution $(0, 10)$.

Consider next a second choice, a copy of the one introduced in Section 1.

	E	$\neg E$
c	(10, 0)	(0, 10)
d	(5, 5)	(5, 5)

E and $\neg E$ partition S, the set of possible states, into two equiprobable events. By Welfare-Risk Permissivism, we may assume there is some profile $R \in \mathcal{R}$ such that both individuals prefer c to d, and some profile $R' \in \mathcal{R}$ such that both individuals prefer d to c. Consider first R. Because both individuals prefer c to d in R, by Ex Ante Pareto $c \succ_R d$. But then by Dominance Package (i) and $(*)$, a welfarist must accept that:

> $(\dagger)$ any outcome with welfare distribution $(10, 0)$ is better than any outcome with welfare distribution $(5, 5)$, and any outcome with welfare distribution $(0, 10)$ is better than any outcome with welfare distribution $(5, 5)$.

However, consider next profile R'. Because both individuals prefer d to c in R', by Ex Ante Pareto we have $d \succ_{R'} c$. Then, by Dominance Package (i) and $(*)$, a welfarist must conclude that any outcome with welfare distribution $(5, 5)$ is better than any outcome with welfare distribution $(10, 0)$ and also better than any outcome with welfare distribution $(0, 10)$. But this contradicts $(\dagger)$.

Since all of Dominance Package, Ex Ante Pareto, and Homogenous Anonymity are very plausible, the above argument may be taken as an objection to Welfarism. The particular aspect of Welfarism to be blamed is Independence of Irrelevant Alternatives. For we can avoid contradiction by accepting that in some cases (under profile R, where both individuals prefer c to d) the outcome with welfare distribution $(10, 0)$ or $(0, 10)$ is better than that with welfare distribution $(5, 5)$, but in other cases (under profile R', where both individuals prefer d to c), the outcome with welfare distribution $(5, 5)$ is better than that with welfare distribution $(10, 0)$ or $(0, 10)$. But this means that changing some non-welfare aspects of a situation—in particular, what risk attitudes people have—can make a difference to the ranking of outcomes, which is what Independence of Irrelevant Alternatives denies.

5 Non-Welfarist Possibilities

What does a non-welfarist theory that accommodates all of Dominance Package, Ex Ante Pareto, and Homogenous Anonymity look like? This section considers two classes of non-welfarist theories that satisfy these principles. One of these lets individual risk attitudes influence how much weight each person's welfare receives in the social ranking of outcomes. The other violates Pareto Indifference, which I see as the more innocuous implication of Welfarism. In Section 6, I show that this is no coincidence: in our framework, any theory that satisfies Dominance Package and Ex Ante Pareto entails one of these two *prima facie* unpalatable conclusions.

5.1 The Risk-Priority View

The first class of theories prioritizes individuals differently depending on their attitude to risk; I call theories in this class Risk-Prioritarian, and the general class of theories the Risk-Priority View. Risk-Prioritarian theories sometimes give priority to the welfare of more risk-averse individuals, and sometimes instead attach greater importance to the welfare of more risk-inclined people. This differential prioritization obtains because outcomes are ranked in terms of a transformed sum of individual welfare levels, where the transformation of welfare is potentially different for each person: it is a function that *also* represents this person's risk attitude.[29]

Unlike the class of theories considered in the next subsection, the Risk-Priority View gives us no reason to reject Rectangular Field, as defined above (see Broome,

[29]The Risk-Priority View is the theory Harsanyi would end up with if he accommodated the Sen-Weymark critique and granted that decision-theoretic utility needn't be linear in welfare, while at the same time retaining ex ante Pareto. For precedents in the literature, see especially Broome (1991b, pp. 211-213), but also Blackorby et al. (1980, especially pp. 19-31) and Dietrich (manuscript, section 5). As hinted at in footnote 16, there is also an analogy here with the literature that accepts Bernoulli and addresses the question of when risk aversion in various goods should be seen as an advantage/burden with respect to one's welfare, that is, how utility/welfare should be normalized (e.g., Adler, 2019, pp. 55-64; Fleurbaey & Zuber, 2021a). I consider these issues in more detail in Cibinel (manuscript).

1991b, p. 115). To streamline the presentation, we can thus help ourselves to this assumption—though we need to remember that it does not hold in general, but rather conditional on the Risk-Priority View. So, for each outcome there is now a "sure" act resulting in that outcome in all states.

To develop the proposal, we need more structure on the individuals' preferences over prospects. This will allow us to better measure their degree of proneness to risk. In particular, we assume that each individual i's preference relation over prospects admits of an expected utility representation U_i: a function from prospects to real numbers representing the individual's preferences, with the expected utility property—that is, the utility of an act is its expected utility. Utility functions with expectational form encode an individual's risk attitude towards their welfare (in the standard sense of Pratt (1964) and Arrow (1971)) and are unique up to positive affine transformation.[30]

Here is how the view works, in detail. Let R_* be any rational ordering of prospects an individual could have. We pick out a *privileged* expected utility representation for this ordering, and call it $\mathbf{U}_{R_*}$. Next, we do this for all other possible rational orderings of prospects. Define finally, for each profile $R \in \mathcal{R}$ and act $a \in A$,

$$U_R(a) := \mathbf{U}_{R_1}(a_1) + ... + \mathbf{U}_{R_n}(a_n),$$

where for all individuals $i \in N$, $\mathbf{U}_{R_i}$ is the privileged expected utility representation of R_i, the ordering that i has under profile R. The Risk-Priority View then considers an act as socially better than another if and only if U_R assigns it a greater value. Notice: first, privileged expected utility representations are assigned to orderings; then, orderings are assigned to individuals given any profile $R \in \mathcal{R}$. As a result, any two individuals with the same ordering get the same privileged expected utility representation (whether within or across profiles). The ranking of outcomes immediately follows from the ranking of sure acts. If an act a results in some single outcome x in all states, we abuse notation slightly and write:

$$U_R(x) := \mathbf{U}_{R_1}(W_1(x)) + ... + \mathbf{U}_{R_n}(W_n(x))$$

[30]So, in this sub-section I am assuming, for illustration purposes, orthodox decision theory's account of risk aversion, as presented by, e.g., Broome (1991a) and Wilkinson (2022).

(compare Broome, 1991b, p. 211). Then, outcomes with greater U_R value are better. This last equation, moreover, makes vivid the distinctive feature of the Risk-Priority View: outcomes are ranked in terms of a transformed sum of individual welfare levels, where the transformations are particular utility functions that also represent the individuals' risk attitudes. This class of theories satisfies Ex Ante Pareto, Dominance Package, Homogenous Anonymity, and Pareto Indifference (**Claim 1** in the Appendix). I say this is a class of theories, rather than a single theory, because different choices of privileged expected utility representations are possible.

An illustration may help us get a better grip on the Risk-Prioritarian approach, and also show why some will find it unpalatable.[31] Suppose there are two individuals, and fix some profile in which the first individual is risk-neutral and the second risk-averse. Consider the following three pairs of choices in matrix form, where s^* is the disjunction of all states.

	s^*
a	(10, 0)
b	(0, 10)

	s^*
c	(5, 0)
d	(0, 5)

	s^*
e	(2, 0)
f	(0, 2)

It's easy to check that some versions of the Risk-Priority View deliver the following verdicts: relative to this risk profile, a is better than b, c and d are equally good, and f is better than e.[32] These Risk-Prioritarian theories give more weight to the first individual in the first choice and to the second individual in the third choice. But it's natural to think that people's risk attitudes shouldn't affect our judgments about the relative merits of a and b, and of e and f: intuitively, a and b are equally good, and so are e and f, since they distribute the same amounts of welfare, only to different people. Nor is this a peculiar feature of some Risk-Prioritarian theories; all versions of the Risk-Priority View will entail similar verdicts.[33]

[31]This example is inspired by a case in Blessenohl (2020, p. 509). Blessenohl uses that case as an objection to a different theory of social choice, which incorporates Buchak's (2013)'s non-orthodox decision theory but shares with the Risk-Priority View a focus on giving weight to risk attitudes in situations where one might have thought them irrelevant.

[32]Suppose, for example, that $\mathbf{U}_{R_1}(10) = 10$, $\mathbf{U}_{R_1}(5) = 5$, $\mathbf{U}_{R_1}(2) = 2$, $\mathbf{U}_{R_1}(0) = 0$, and $\mathbf{U}_{R_2}(10) = 8$, $\mathbf{U}_{R_2}(5) = 5$, $\mathbf{U}_{R_2}(2) = 3$, $\mathbf{U}_{R_2}(0) = 0$.

[33]Incidentally, notice that the Risk-Priority View is sensitive to stakes: this means that, if we were working with multiple welfare profiles (as entertained in footnote 19 above), it would violate the relevant invariance condition, i.e., invariance with respect to a common positive affine transformation. I take this to be a feature of the view worth emphasizing, but not necessarily an objection—see Nebel's (2024) critical investigation of invariance principles, to which I owe this point.

Those tempted to reject this theory in light of these kinds of examples will want to embrace a different non-welfarist approach. In particular, we can consider a cousin of Homogenous Anonymity: the idea that, at least when there is no risk involved, equal quantities of welfare matter in the same way—independently of what risk attitudes people have.

> SURE-ACT ANONYMITY: For all sure acts $a, b \in A$, and for all profiles $R \in \mathcal{R}$, if a and b are permutations of one another, then neither $a \succ_R b$ nor $b \succ_R a$.

Requiring Sure-Act Anonymity to hold, then, will serve as an antidote to the arguably objectionable verdicts of the Risk-Priority View considered above.[34]

5.2 The Unanimity Rule

Let us turn to examine a non-welfarist theory that respects Sure-Act Anonymity, as well as Ex Ante Pareto, Dominance Package, and Homogenous Anonymity (see **Claim 3** in the Appendix). This theory can be understood as a generalization of the idea that motivates Ex Ante Pareto. While Ex Ante Pareto says that autonomy matters, this theory says that it is all that matters. According to what we can call the Unanimity Rule:

- An act is at least as good as another just in case everyone weakly prefers it (adapted from Sen, 2017, p. 75);
- An outcome is at least as good as another just in case the act that brings it about is weakly preferred by everyone.

This theory entertains much *incomparability* between acts and outcomes. If some individuals prefer a first act to a second, and others prefer the second to the first, then the two acts will be incomparable (in the sense that the social preference relation fails to hold either way). Similarly, two outcomes will be incomparable whenever there is no unanimity of preference over the respective acts that bring them about.

[34]Thus, Sure-Act Anonymity is perhaps closer to the spirit of Chambers and Echenique (2012) and Fleurbaey and Zuber's (2017) principle than Independence of Irrelevant Alternatives—see footnote 27 above. As we will see in the next section, there are views that satisfy Sure-Act Anonymity while violating Independence of Irrelevant Alternatives and also Chambers and Echenique (2012) and Fleurbaey and Zuber's (2017) principle. These views do not intuitively take notice of risk attitudes in riskless contexts, suggesting that Sure-Act Anonymity is best suited among these principles to capture a concern with not letting risk attitudes make a difference when there is no risk.

Moreover, the Unanimity Rule is not well-defined unless we assume a tight restriction on which outcomes can result from which acts (**Claim 2** in the Appendix): roughly, what we need is that each outcome is the output of exactly one act—I call this the "Act-Outcome Restriction" in the Appendix. Recall that, in laying out my framework, I explicitly avoided assuming Rectangular Field, the claim that A includes all functions from states of nature to outcomes. We can now better see the reason for this choice.

A non-welfarist who cares about the value of autonomy may think that the history of how a certain social arrangement came about is ethically relevant. If so, outcomes in X will include such historical information, since outcomes include all normatively relevant information. But then we should want to impose strict requirements on which outcomes can be combined together, on pain of allowing in our model impossible acts: an example being an act a that delivers an outcome x in only some of the states, where part of the description of x is that it results from a context of zero risk.[35] Imposing these restrictions is part and parcel of the idea that motivates certain non-welfarist, autonomy-based approaches to social choice. The Unanimity Rule exemplifies this class of views: all that matters in evaluating an outcome, according to it, is what kind of act it resulted from. A rejection of Rectangular Field and the imposition of stringent requirements on what kinds of functions from states to outcomes constitute genuine acts is easily justified given this stance.[36]

This line of thought is bolstered by the following fact, proved in the Appendix: given Welfare-Risk Permissivism and Domain Condition*, a slight modification of the technical domain assumption required for Theorem 1, if Rectangular Field holds, no social quasi-ordering function satisfies Dominance Package, Ex Ante Pareto and Sure-Act Anonymity (**Claim 4** in the Appendix, which is essentially a variant of the result in Blessenohl 2020, though I draw a different lesson from it). Since these three latter principles can be motivated by a viable ethical outlook involving concern for autonomy, a certain kind of impartiality, and coherence in

[35] That a theory allows for such impossible entities is often seen in the literature as a strong objection against it; see, e.g., Broome (1991b, pp. 115-117) and Joyce (1999, p. 108).

[36] The Risk-Priority View, by contrast, is compatible with Rectangular Field because it does not need to include in an outcome's description any information about how it came about. Different non-welfarist theories accommodate Ex Ante Pareto in different ways—including historical information in an outcome is one such way, but it is not necessary.

evaluating acts and outcomes, it seems entirely legitimate for those attracted to this outlook to take the fact just stated as an argument against Rectangular Field.

The real problem with the Unanimity Rule is that, while it fully accommodates autonomy, it gives no ethical weight at all to how well off people are. Suppose act a is preferred unanimously over b, and that everyone is equally well off in some possible outcome of a, x, as they are in some possible outcome of b, y. Then the Unanimity Rule ranks x as better than y, violating Pareto Indifference. Worse still, suppose a involves some risk that everyone will end up badly off, in outcome z. The Unanimity Rule ranks z as better than any of the outcomes that b could result in, even if in some such outcomes everyone fares very well indeed. This is implausible; as I said at the beginning of the paper, everyone agrees that the well-being of individuals matters.

6 The Value of Welfare

How then can we accommodate the thought that welfare matters within a non-welfarist framework, so as to avoid the bad implication of the Unanimity Rule? The lesson seems to be that our theory should evaluate outcomes so that, if everyone is strictly worse off in x than in y, then x isn't ranked as socially better than y; in fact, we should want y to be ranked as better than x. Formally, the requirement our social quasi-ordering function would need to satisfy is

> EX POST PARETO. For all outcomes $x, y \in X$ and for all profiles $R \in \mathcal{R}$, if $W_i(y) > W_i(x)$ for all $i \in N$, then $y \succ_R x$.

Kaplow and Shavell (2001) show that given a widely accepted "continuity" assumption, any theory that violates the principle of Pareto Indifference entails violations of Ex Post Pareto, too. We can therefore understand Pareto Indifference as a condition that must be satisfied by theories wishing to avoid the sort of insensitivity to the value of welfare exemplified by the Unanimity Rule.

Here, however, we encounter a new problem, which serves as a counterpoint to Theorem 1. Let us step back: the Unanimity Rule stands as an alternative to the Risk-Priority View, in that it satisfies Sure-Act Anonymity. But, it turns out, it violates Pareto Indifference. The Risk-Priority View, by contrast, satisfies Pareto

Indifference but violates Sure-Act Anonymity. The following theorem confirms that this is no coincidence.

> **Theorem 2.** *Given Welfare-Risk Permissivism and Domain Condition*, no social quasi-ordering function satisfies Dominance Package, Ex Ante Pareto, Sure-Act Anonymity, and Pareto Indifference.*

The proof can be found in the Appendix, and is based on a case due to Broome (1991b, p. 190) and Blessenohl (2020, p. 486); like Claim 4 above, Theorem 2 is closely related to Blessenohl's own result.[37] Domain Condition* is stated in the Appendix, too. The first main result of the paper was an impossibility involving welfarist theories: they are incompatible with Ex Ante Pareto, Dominance Package, and Homogenous Anonymity in my framework. This second result can be seen as an impossibility for non-welfarist theories. Even if we deny Independence of Irrelevant Alternatives, and so Welfarism, we need to choose between Pareto Indifference, which encodes a kind of concern for people's welfare, and Sure-Act Anonymity, which amounts to the thought that risk attitudes shouldn't affect what priority different people's welfare receives in non-risky choices.[38] More constructively, how-

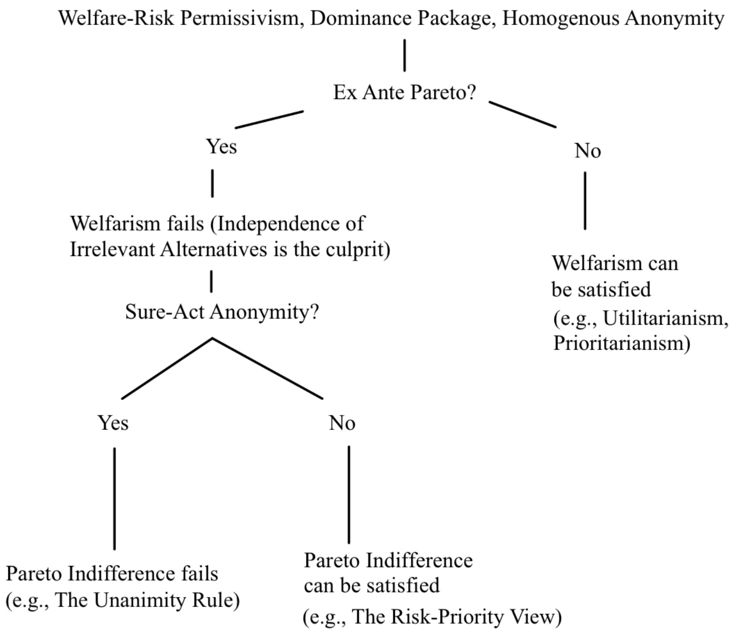
[37]The comparison is non-obvious because of differences in our respective frameworks. For one, Blessenohl's (2020) is a "single-profile" framework (individuals' risk attitudes are fixed) while mine is "multiprofile" (they vary with $R \in \mathcal{R}$). Further, Blessenohl bakes Pareto Indifference into that framework, and defines social preference over outcomes in terms of social preference over sure-acts. For reasons given above, if neither Pareto Indifference nor Rectangular Field are assumed, this definitional move is not available. Blessenohl shows that an analogue of Welfare-Risk Permissivism is inconsistent with the conjunction of statewise dominance and an analogue of Ex Ante Pareto, given the principle of "Constant Anonymity" (Blessenohl, 2020, p. 495). Restated in my framework, Constant Anonymity says that for any profile $R \in \mathcal{R}$, and any two outcomes $x, y \in X$, if $W(x)$ and $W(y)$ are permutations of one another, then $x \sim_R^X y$. Note that Pareto Indifference follows from Constant Anonymity by taking the identity permutation. Given Rectangular Field, completeness of social preference, and Dominance Package, Sure-Act Anonymity and Constant Anonymity are equivalent. Therefore, under these assumptions Sure-Act Anonymity entails Pareto Indifference, and Theorem 2 adds nothing to Claim 4, which I called a variant of Blessenohl's result. Moreover, assume that for any list of welfare levels $(r_1, ..., r_n)$ such that $W(x) = (r_1, ..., r_n)$ for some $x \in X$, there exists a sure-act $a \in A$ whose unique outcome y is such that $W(y) = W(x) = (r_1, ..., r_n)$. Then given completeness of social preference and Dominance Package, the conjunction of Pareto Indifference and Sure-Act Anonymity is equivalent to Constant Anonymity, so again we see the connection. In my proof of Theorem 2, however, I assume neither Rectangular Field nor Pareto Indifference: instead, I use Dominance Package, Ex Ante Pareto, and Sure-Act Anonymity to derive a violation of Pareto Indifference. This re-framing of the significance of Blessenohl's case reflects my interest in seeing Theorem 2 as an exploration of (im)possibilities for non-welfarism.

[38]One might wonder whether Theorem 2 opens up an alternative way out of Theorem 1: if we deny Pareto Indifference to block this second impossibility, can we retain Independence of Irrelevant Alternatives after all? I haven't formally ruled this option out, but it is unpromising. The proof of Theorem 1 uses Pareto Indifference to derive a conclusion that is anyway surely correct for

ever, Theorem 2 helps us get a better view of the space of non-welfarist theories that satisfy Ex Ante Pareto, Dominance Package, and Homogenous Anonymity.

7 Conclusion

As we conclude, it will be helpful to have a clear picture of the theoretical options that the work of this paper left open. Here it is.



It is now clear that both welfarists and non-welfarists face difficult choices. To wrap up, let's begin by summarizing the welfarist's predicament. All of Welfare-Risk Permissivism, Dominance Package, and Homogenous Anonymity seem to me extremely plausible. But we should return in closing to what is perhaps the simplest way out for the welfarist: namely, to deny Welfare-Risk Permissivism. There are two ways to do this. One is to retain our assumption that there are independently given quantities of welfare, and insist that some particular attitude to risk involving one's welfare—for example, risk neutrality—is uniquely rational.[39] The other in-

those who accept Sure-Act Anonymity; in the choice between c and d of Section 4, it is the claim that the outcome of c under E is neither better nor worse than the outcome of c under $\neg E$. This claim alone, coupled with Dominance Package and Ex Ante Pareto, allows us to derive violations of Independence of Irrelevant Alternatives.

[39]See Zhao (2021) for a recent defence of risk neutrality, though note that Zhao's argument vindicates risk neutrality in a certain range of cases only (Zhao, 2021, p. 157; Wilkinson, 2022).

volves denying that quantities of welfare exist independently of choice under risk, in one of the two ways suggested in Section 2. Welfarists who find these claims hard to believe need to reject Ex Ante Pareto, and the value of autonomy it expresses. For those antecedently committed to either of these positions, the results of the paper will look like additional evidence in their support.

But I'm inclined towards a different interpretation of the results: if both Welfare-Risk Permissivism and Ex Ante Pareto are true, somehow we need to thread the needle between the two possibilities that Theorem 2 leaves us with. Non-welfarist theories of social choice have been underexplored compared to their welfarist rivals, and much work remains to be done in figuring out what kind of non-welfarist one should be.[40] For those who reject Sure-Act Anonymity: can a plausible normative story be given for letting risk attitudes have normative relevance in non-risky choices? For those who reject Pareto Indifference: what alternative ways are there to accommodate the obvious fact that welfare matters? Without clear answers to these questions, our ability to simultaneously promote the values of welfare and autonomy remains at risk.[41]

Acknowledgments

I would especially like to thank my advisors Lara Buchak and Jake Nebel for extremely valuable guidance throughout the project, which took the form of many conversations as well as multiple rounds of written comments on different drafts of the paper. I am also grateful to Harvey Lederman, Fanhao Meng, Aidan Penn, and two referees of this journal for written comments on drafts of the paper, and to Adam Elga and Elizabeth Harman for reading a draft and providing oral feedback. The Princeton Philosophy dissertation seminar participants read a draft and provided helpful suggestions. My thanks also go to Chris Bottomley, Franz Diet-

[40]However, see Chang (2001), Fleurbaey et al. (2003), Fleurbaey and Maniquet (2008), and the literatures on fairness and equality of opportunity mentioned above for discussions of extant non-welfarist approaches in social choice. Also relevant is my proposal in Cibinel (2022) responding to Blessenohl (2020), again cited above, which however needs to be supplemented to generate a full theory of social choice. In our taxonomy, any theory incorporating that proposal will fall in the bottom-left corner of the diagram sketched, together with the Unanimity Rule.

[41]As I see it, the lesson of the paper is in line with the program propounded by Sen (2017, pp. 339-44), of critically inspecting and, as appropriate, expanding the "informational basis" of social choice.

rich, Ryan Doody, Brendan Kolb, Kacper Kowalczyk, Richard Pettigrew, Marcus Pivato, Richard Roth, Weng Kin San, Margaret Shea, Katie Steele, Teruji Thomas, John Weymark, Timothy L. Williamson, and Qichen Yan for conversations or correspondence on the topic of the paper. Finally, I would like to thank participants of the UCHV Fellowship dissertation seminar for questions on a talk based on this paper. This work was supported by a UCHV Graduate Prize Fellowship.

Appendix

This appendix states some domain assumptions left out of the main text and gives statements and proofs of the Theorems and Claims mentioned therein. As a piece of convention applying throughout, for all $p, q \in \mathbb{R}$, let (p, q) denote:

- the n-tuple assigning p to the first $\frac{n}{2}$ entries and q to the remaining ones *if* n *is even*,
- the n-tuple assigning p to the first $\frac{n-1}{2}$ entries, q to the next $\frac{n-1}{2}$ entries, and some fixed arbitrary number $r \in R$ to the last entry *if* n *is odd*.

Recall that E and $\neg E$ partition S into two equiprobable events.

> Domain Condition: For some $l, m, t \in \mathbb{R}$, with $l > m > t$ and $l - m = m - t$, there exist outcomes $v, w, x, y, z \in X$ and acts $a, b, c, d \in A : a(s) = v, \forall s \in S$, with $W(v) = (l, t)$; $b(s) = w$, $\forall s \in S$, with $W(w) = (t, l)$; $c(s) = x$ for $s \in E$ and $c(s) = y$ for $s \in \neg E$, with $W(x) = (l, t)$ and $W(y) = (t, l)$; $d(s) = z, \forall s \in S$, with $W(z) = (m, m)$.

To illustrate in matrix form, Domain Condition simply guarantees the existence of the following acts, where, e.g., "$v : (l, t)$" indicates that the outcome is v and the list of welfare levels associated to v is (l, t).

	E	$\neg E$
a	v: (l, t)	v: (l, t)
b	w: (t, l)	w: (t, l)
c	$x : (l, t)$	$y : (t, l)$
d	$z : (m, m)$	$z : (m, m)$

For Theorem 2, we need the following variation on Domain Condition.

DOMAIN CONDITION*: For some $l, m, t \in \mathbb{R}$, with $l > m > t$ and $l - m = m - t$, there exist outcomes $q, r, u, v, w, x, y, z \in X$, and acts $a, b, c, d, e, f \in A$: $a(s) = x$ for $s \in E$ and $a(s) = y$ for $s \in \neg E$, with $W(x) = (l, m)$ and $W(y) = (t, m)$; $b(s) = v$ for $s \in E$ and $b(s) = w$ for $s \in \neg E$, with $W(v) = (m, l)$ and $W(w) = (m, t)$; $c(s) = z, \forall s \in S$, with $W(z) = (l, m)$; $d(s) = u, \forall s \in S$, with $W(u) = (m, l)$; $e(s) = q, \forall s \in S$, with $W(q) = (t, m)$; and $f(s) = r, \forall s \in S$, with $W(r) = (m, t)$.

In matrix form, this condition guarantees the existence of the following acts.

	E	$\neg E$
a	x: (l, m)	y: (t, m)
b	v: (m, l)	w: (m, t)
c	$z : (l, m)$	$z : (l, m)$
d	$u : (m, l)$	$u : (m, l)$
e	$q : (t, m)$	$q : (t, m)$
f	$r : (m, t)$	$r : (m, t)$

Theorem 1. *Given Welfare-Risk Permissivism and Domain Condition, no social quasi-ordering function satisfies Dominance Package, Ex Ante Pareto, Homogenous Anonymity, Pareto Indifference, and Independence of Irrelevant Alternatives.*

Proof. I show that if the first four principles are satisfied, Independence of Irrelevant Alternatives fails. Let $l, m, t \in \mathbb{R}$, $v, w, x, y, z \in X$, and $a, b, c, d \in A$ satisfy Domain Condition, so that $l > m > t$ and $l - m = m - t$; $a(s) = v$ for all $s \in S$, with $W(v) = (l, t)$; $b(s) = w$, with $W(w) = (t, l)$; $c(s) = x$ for $s \in E$ and $c(s) = y$ for $s \in \neg E$, with $W(x) = (l, t)$ and $W(y) = (t, l)$; and $d(s) = z$ for all $s \in S$, with $W(z) = (m, m)$.

Note that for all $i \in N$ with $c_i \neq d_i$, $\sum_{j=1}^{k} c_i(s_j)\pi(s_j) = \sum_{j=1}^{k} d_i(s_j)\pi(s_j)$ but c_i has uncertain consequences while d_i does not. By Welfare-Risk Permissivism and the fact that $\mathcal{R}$ is the set of all rational risk profiles, we can select $R \in \mathcal{R}$ such that $R_i = R_j$ for all $i, j \in N$, and $c_i P_i d_i$ for all $i \in N$ with $c_i \neq d_i$. And we can select $R' \in \mathcal{R}$ such that $R'_i = R'_j$ for all $i, j \in N$, and $d_i P'_i c_i$ for all $i \in N$ with $c_i \neq d_i$.

By Homogenous Anonymity, neither $a \succ_R b$ nor $b \succ_R a$. By Dominance Package (ii), neither $v \succ_R w$ nor $w \succ_R v$. By Pareto Indifference, $v \sim_R x$, $w \sim_R y$. By transitivity, neither $x \succ_R y$ nor $y \succ_R x$.

By Ex Ante Pareto, $c \succ_R d$ and $d \succ_{R'} c$. By Dominance Package (i) either $x \succ_R z$ or $y \succ_R z$, and either $z \succ_{R'} x$ or $z \succ_{R'} y$. There are four possibilities. If $x \succ_R z$ and $z \succ_{R'} x$, or $y \succ_R z$ and $z \succ_{R'} y$, we have a violation of Independence of Irrelevant Alternatives. Suppose next $x \succ_R z$ and $z \succ_{R'} y$. Since $x \succ_R z$ and $\neg(x \succ_R y)$, it follows by transitivity that $\neg(z \succ_R y)$. But then Independence of Irrelevant Alternatives is violated, given $z \succ_{R'} y$. Suppose finally $y \succ_R z$ and $z \succ_{R'} x$. Since $y \succ_R z$ and $\neg(y \succ_R x)$, it follows by transitivity that $\neg(z \succ_R x)$, which violates Independence of Irrelevant Alternatives given $z \succ_{R'} x$. □

Theorem 2. *Given Welfare-Risk Permissivism and Domain Condition*, no social quasi-ordering function satisfies Dominance Package, Ex Ante Pareto, Sure-Act Anonymity, and Pareto Indifference.*

Proof. I show that if the first three principles are satisfied, Pareto Indifference fails. Let $l, m, t \in \mathbb{R}$, $u, v, w, x, y, z \in X$, and $a, b, c, d \in A$ satisfy Domain Condition*, so that $l > m > t$ and $l - m = m - t$; $a(s) = x$ for $s \in E$ and $a(s) = y$ for $s \in \neg E$, with $W(x) = (l, m)$ and $W(y) = (t, m)$; $b(s) = v$ for $s \in E$ and $b(s) = w$ for $s \in \neg E$, with $W(v) = (m, l)$ and $W(w) = (m, t)$; $c(s) = z, \forall s \in S$, with $W(z) = (l, m)$; $d(s) = u, \forall s \in S$, with $W(u) = (m, l)$.

Fix $R \in \mathcal{R}$ such that $a_i P_i b_i$ for all $i \in N$ such that $a_i \neq b_i$; by Welfare-Risk Permissivism and the fact that $\mathcal{R}$ is the set of all rational risk profiles, we can make this assumption. By Ex Ante Pareto, $a \succ_R b$. By Dominance Package (i), either $x \succ_R v$ or $y \succ_R w$.

Suppose that $x \succ_R v$. By Sure-Act Anonymity, neither $c \succ_R d$ nor $d \succ_R c$. By Dominance Package (ii), neither $z \succ_R u$ nor $u \succ_R z$. If $v \succ_R u$, or $u \succ_R v$, or neither $v \succeq_R u$ nor $v \succeq_R u$, Pareto Indifference is violated. So, suppose $v \sim_R u$. By transitivity, $x \succ_R u$. But given $\neg(z \succ_R u)$, it follows that $\neg(x \sim_R z)$, which violates Pareto Indifference. The case of $y \succ_R w$ is analogous, and requires a consideration of $e, f \in A$ as specified by Domain Condition*. □

Claim 1. *Risk-Prioritarian theories are social quasi-ordering functions that satisfy Ex Ante Pareto, Dominance Package, Homogenous Anonymity, and Pareto Indifference.*

Proof. It's easy to see that $\succeq_R^A$ and $\succeq_R^X$, as defined by the Risk-Priority View, are quasi-orderings. Now let $R \in \mathcal{R}$ be given. Choose $a, b \in A$ so that $a_i \neq b_i$ for some $i \in N$ and $a_i P_i b_i$ for all $i \in N$ such that $a_i \neq b_i$. Then $U_i(a_i) > U_i(b_i)$

for all i with $a_i \neq b_i$ and $U_i(a_i) = U_i(b_i)$ for all i with $a_i = b_i$. It follows that $U_R(a) > U_R(b)$. So Ex Ante Pareto is satisfied. For Dominance Package (i), suppose $a \succ_R b$. Then $U_R(a) > U_R(b)$, which cannot be true unless there exists $s \in S$ such that $U_R(a(s)) > U_R(b(s))$. For Dominance Package (ii), if $a(s) \succ_R b(s)$ for all $s \in S$, then $U_R(a(s)) > U_R(b(s))$ for all $s \in S$, and so of course $U_R(a) > U_R(b)$. For Pareto Indifference, if $W(x) = W(y)$ then $U_R(x) = U_R(y)$ and so $x \sim_R y$. Note that we are here using the assumption that $W_i(x) \geq W_i(y)$ just in case $U_i(W_i(x)) \geq U_i(W_i(y))$, for all $i \in N$. Finally, if $U_i = U_j$ for all $i, j \in N$ and a and b are permutations of one another, then $U_R(a) = U_R(b)$. So, Homogenous Anonymity follows. $\square$

We now give a formal statement of the Unanimity Rule. For all acts $a, b \in A$ and for all profiles $R \in \mathcal{R}$, define $a\overline{R}b$ to mean $a_i R_i b_i$ for all $i \in N$.
According to the Unanimity Rule:

- For all profiles $R \in \mathcal{R}$ and acts $a, b \in A$, $a \succeq_R^A b$ if and only if $a\overline{R}b$ (adapted from Sen, 2017, p. 75);
- For all profiles $R \in \mathcal{R}$ and outcomes $x, y \in X$, $x \succeq_R^X y$ if and only if there exist $a, b \in A$, with $x = a(s)$ for some state $s \in S$ and $y = b(s)$ for some state $s \in S$, such that $a\overline{R}b$.

The Unanimity Rule requires the

> ACT-OUTCOME RESTRICTION. For all outcomes $x \in X$ and acts $a, b \in A$, if $x = a(s)$ for some state $s \in S$ and $x = b(s)$ for some state $s \in S$, then $a_i I_i b_i$ for all $i \in N$ and $R \in \mathcal{R}$; for all $x \in X$, there is some $a \in A$ such that $x = a(s)$ for some $s \in S$.

Claim 2. *If Act-Outcome Restriction fails, the Unanimity Rule contradicts the reflexivity of weak social preference.*

Proof. Consider the first conjunct of Act-Outcome Restriction, and suppose there exists $R \in \mathcal{R}$ such that $\neg(a\overline{R}b)$, with $x = a(s)$ for some $s \in S$, but also $x = b(s)$ for some $s \in S$. Then the Unanimity Rule entails $\neg(x \succeq_R x)$, contradicting reflexivity (moreover, if in addition $b\overline{R}a$, we can also deduce $x \succeq_R x$, with a contradiction following even apart from reflexivity). For the second conjunct, suppose there exists some $x \in X$ for which there is no $a \in A$ such that $a(s) = x$ for some $s \in S$. Then on the Unanimity Rule for all outcomes $y \in X$, $\neg(x \succeq_R y)$. So, in particular, $\neg(x \succeq_R x)$, again in violation of reflexivity. $\square$

Claim 3. *Given Act-Outcome Restriction, the Unanimity Rule is a social quasi-ordering function that satisfies Ex Ante Pareto, Dominance Package, Homogenous Anonymity, and Sure-Act Anonymity.*

Proof. That the Unanimity Rule is a social quasi-ordering function given Act-Outcome Restriction follows from the fact that the individual preference relations over prospects $R_1, ..., R_n$ are all orderings. Ex Ante Pareto follows from the clause for $\succeq_R^A$, and Dominance Package from the clause for $\succeq_R^X$. For Homogenous Anonymity, note that if a and b are permutations, and $R_i = R_j$ for all $i, j \in N$, either every individual is indifferent, or some prefer a and some prefer b. Either way, neither $a \succ_R b$ nor $b \succ_R a$. A similar argument establishes Sure-Act Anonymity. $\square$

Claim 4. *Given Welfare-Risk Permissivism and Domain Condition*, if Rectangular Field holds, no social quasi-ordering function Satisfies Dominance Package, Ex Ante Pareto and Sure-Act Anonymity.*

Proof. We repeat the first paragraph of the proof of Theorem 2 (and so use Welfare-Risk Permissivism, Domain Condition*, Ex Ante Pareto, and Dominance Package (i)) to get $x \succ_R v$, with x, v and R defined as in that proof (note that we don't need the full strength of Domain Condition*). By Rectangular Field, we can find $c, d \in A : c(s) = x, \forall s \in S$ and $d(s) = v, \forall s \in S$. By Dominance Package (ii), since $x \succ_R v$, $c \succ_R d$, violating Sure-Act Anonymity. $\square$

References

Adler, M. D. (2019). *Measuring Social Welfare: An Introduction*. Oxford University Press. https://doi.org/10.1093/oso/9780190643027.001.0001.

Arrow, K. J. (1971). *Essays in the Theory of Risk-bearing*. Chicago: Markham Publishing Co.

Bader, R. M. (2018). Stochastic dominance and opaque sweetening. *Australasian Journal of Philosophy*, *96*, 498–507. https://doi.org/10.1080/00048402.2017.1362566.

Bell, D. E., & Raiffa, H. (1988). Marginal value and intrinsic risk aversion. In D. E. Bell, H. Raiffa, & A. Tversky (Eds.), *Decision Making: Descriptive, Normative, and Prescriptive Interactions* (pp. 384–397). Cambridge University Press.

Blackorby, C., Bossert, W., & Donaldson, D. (2005). *Population Issues in Social Choice Theory, Welfare Economics, and Ethics*. Cambridge University Press. https://doi.org/10.1017/CCOL0521825512.

Blackorby, C., Bossert, W., & Donaldson, D. (2006). Anonymous single-profile welfarism. *Social Choice and Welfare*, *27*, 279–287. https://doi.org/10.1007/s00355-006-0131-1.

Blackorby, C., Donaldson, D., & Weymark, J. A. (1980). On John Harsanyi's defences of utilitarianism. *Discussion Paper of the Center for Operations Research and Econometrics, Université Catholique de Louvain*, *1013*, 1–40. Available at: https://johnweymark.github.io/files/BDW80CORE.pdf.

Blessenohl, S. (2020). Risk attitudes and social choice. *Ethics*, *130*, 485–513. https://doi.org/10.1086/708011.

Bossert, W., & Weymark, J. A. (2004). Utility in social choice. In S. Barberà, P. J. Hammond, & C. Seidl (Eds.), *Handbook of utility theory: Volume 2* (pp. 1099–1177). Springer. https://doi.org/10.1007/978-1-4020-7964-1_7.

Bottomley, C., & Williamson, T. L. (2024). Rational risk-aversion: Good things come to those who weight. *Philosophy and Phenomenological Research*, *108*, 697–725. https://doi.org/10.1111/phpr.13006.

Bradley, R. (2022). Impartial evaluation under ambiguity. *Ethics*, *132*, 541–569. https://doi.org/10.1086/718081.

Bradley, R., & Stefánsson, O. H. (2023). Fairness and risk attitudes. *Philosophical Studies*, *180*, 3179–3204. https://doi.org/10.1007/s11098-023-02025-2.

Broome, J. (1987). Utilitarianism and expected utility. *The Journal of Philosophy*, *84*, 405–422. https://doi.org/10.2307/2026999.

Broome, J. (1991a). "Utility". *Economics and Philosophy*, *7*, 1–12. https://doi.org/10.1017/S0266267100000882.

Broome, J. (1991b). *Weighing Goods: Equality, Uncertainty and Time*. Wiley-Blackwell. https://doi.org/10.1002/9781119451266.

Broome, J. (2004). *Weighing Lives*. Oxford University Press. https://doi.org/10.1093/019924376X.001.0001.

Buchak, L. (2013). *Risk and Rationality*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199672165.001.0001.

Buchak, L. (2017). Taking risks behind the veil of ignorance. *Ethics*, *127*, 610–644. https://doi.org/10.1086/690070.

Chambers, C. P., & Echenique, F. (2012). When does aggregation reduce risk aversion? *Games and Economic Behavior*, *76*, 582–595. https://doi.org/10.1016/j.geb.2012.07.015.

Chang, H. F. (2001). Non-welfarist paretian methods of policy assessment. Available at: http://dx.doi.org/10.2139/ssrn.288085.

Cibinel, P. (2022). Risk attitudes and justifiability to each. *Ethics*, *133*, 106–121. https://doi.org/10.1086/720777.

Cibinel, P. (manuscript). The Risk-Priority View. Available upon request.

Cohen, H., Maril, A., Bleicher, S., & Nissan-Rozen, I. (2022). Attitudes toward risk are complicated: Experimental evidence for the re-individuation approach to risk-attitudes. *Philosophical Studies*, *179*, 2553–2577. https://doi.org/10.1007/s11098-022-01778-6.

Danan, E., Gajdos, T., & Tallon, J.-M. (2015). Harsanyi's aggregation theorem with incomplete preferences. *American Economic Journal: Microeconomics*, *7*, 61-69.

Dietrich, F. (manuscript). Welfare vs. utility. Available at: http://www.franzdietrich.net/papers/dietrich-welfarevsutility.pdf.

Doody, R. (manuscript). Consider the ostrich: Non-utilitarians, ex ante interests, and burying your head in the sand. Available at: http://rdoody.com/considerostrich.pdf.

Dworkin, R. (1981). What is equality? Part 2: Equality of resources. *Philosophy and Public Affairs*, *10*, 283–345.

Ferreira, F. H. G., & Peragine, V. (2016). Individual Responsibility and Equality of Opportunity. In M. D. Adler & M. Fleurbaey (Eds.), *The Oxford Handbook of Well-Being and Public Policy.* Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199325818.013.24.

Fleurbaey, M. (2009). Two variants of Harsanyi's aggregation theorem. *Economics Letters*, *105*, 300-302. https://doi.org/10.1016/j.econlet.2009.08.023.

Fleurbaey, M. (2010). Assessing risky social situations. *Journal of Political Economy*, *118*, 649–680. https://doi.org/10.1086/656513.

Fleurbaey, M. (2018). Welfare economics, risk and uncertainty. *Canadian Journal of Economics/Revue canadienne d'économique*, *51*, 5–40. https://doi.org/10.1111/caje.12314.

Fleurbaey, M., & Maniquet, F. (2008). Utilitarianism versus fairness in welfare economics. In M. Fleurbaey, M. Salles, & J. A. Weymark (Eds.), *Justice, Political Liberalism, and Utilitarianism: Themes from Harsanyi and Rawls* (pp. 263–280). Cambridge University Press. https://doi.org/10.1017/CBO9780511619595.012.

Fleurbaey, M., Tungodden, B., & Chang, H. F. (2003). Any Non-welfarist Method of Policy Assessment Violates the Pareto Principle: A Comment. *Journal of Political Economy*, *111*, 1382–1386. https://doi.org/10.1086/378533.

Fleurbaey, M., & Voorhoeve, A. (2013). Decide as you would with full information! An argument against ex ante Pareto. In O. Norheim, S. Hurst, N. Eyal, & D. Wikler (Eds.), *Inequalities in Health: Concepts, Measures, and Ethics* (pp. 113–128). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199931392.003.0009.

Fleurbaey, M., & Zuber, S. (2017). Fair management of social risk. *Journal of Economic Theory*, *169*, 666–706. https://doi.org/10.1016/j.jet.2017.03.009.

Fleurbaey, M., & Zuber, S. (2021a). Fair utilitarianism. *American Economic Journal: Microeconomics*, *13*, 370-401.

Fleurbaey, M., & Zuber, S. (2021b). Universal social welfare orderings and risk. Available at: https://dx.doi.org/10.2139/ssrn.3876726.

Fleurbaey, M., & Zuber, S. (2022). A new puzzle in the social evaluation of risk. *Economics and Philosophy*, *38*, 450–465. https://doi.org/10.1017/S0266267121000225.

Frick, J. (2013). Uncertainty and justifiability to each person: Response to Fleurbaey and Voorhoeve. In O. Norheim, S. Hurst, N. Eyal, & D. Wikler (Eds.), *Inequalities in Health: Concepts, Measures, and Ethics* (pp. 129–146). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199931392.003.0010.

Frick, J. (2015). Contractualism and social risk. *Philosophy and Public Affairs*, *43*, 175–223. https://doi.org/10.1111/papa.12058. doi: 10.1111/papa.12058

Friedman, M. (1953). Choice, chance, and the personal distribution of income. *Journal of Political Economy*, *61*, 277-290. https://doi.org/10.1086/257390.

Goldschmidt, Z., & Nissan-Rozen, I. (2020). The intrinsic value of risky prospects. *Synthese*, *198*, 7553–7575. https://doi.org/10.1007/s11229-020-02532-3.

Greaves, H. (2017). A reconsideration of the Harsanyi-Sen-Weymark debate on utilitarianism. *Utilitas*, *29*, 175–213. https://doi.org/10.1017/S0953820816000169.

Hájek, A. (2021). Risky business. *Philosophical Perspectives*, *35*, 189–205. https://doi.org/10.1111/phpe.12147.

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

Hammond, P. J. (1982). Utilitarianism, uncertainty and information. In A. K. Sen & B. A. O. Williams (Eds.), *Utilitarianism and beyond* (pp. 85–102). Cambridge University Press. https://doi.org/10.1017/CBO9780511611964.006.

Hare, C. (2010). Take the sugar. *Analysis*, *70*, 237–247. doi: https://doi.org/10.1093/analys/anp174

Harsanyi, J. C. (1955). Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy*, *63*, 309–321. https://doi.org/10.1086/257678.

Harsanyi, J. C. (1977). *Rational Behaviour and Bargaining Equilibrium in Games and Social Situations*. Cambridge University Press. https://doi.org/10.1017/CBO9780511571756.

Harsanyi, J. C. (1982). Morality and the theory of rational behaviour. In A. Sen & B. Williams (Eds.), *Utilitarianism and Beyond* (pp. 39–62). Cambridge University Press. https://doi.org/10.1017/CBO9780511611964.004.

Joyce, J. M. (1999). *The Foundations of Causal Decision Theory*. Cambridge University Press. https://doi.org/10.1017/CBO9780511498497.

Kaplow, L., & Shavell, S. (2001). Any non-welfarist method of policy assessment violates the pareto principle. *Journal of Political Economy*, *109*, 281–286. https://doi.org/10.1086/319553.

Keller, S. (2009). Welfarism. *Philosophy Compass*, *4*, 82–95. https://doi.org/10.1111/j.1747-9991.2008.00196.x.

Kowalczyk, K. (2022). People in suitcases. *Journal of Moral Philosophy*, *20*, 3–30. https://doi.org/10.1163/17455243-19030009.

Krantz, D., Luce, D., Suppes, P., & Tversky, A. (1971). *Foundations of Measurement, Vol. I*. Academic Press.

Lederman, H. (in press). Of marbles and matchsticks. *Oxford Studies in Epistemology.* Available at: https://philpapers.org/archive/LEDOMA-2.pdf.

McCarthy, D., Mikkola, K., & Thomas, J. T. (2020). Utilitarianism with and without expected utility. *Journal of Mathematical Economics*, *87*, 77–113. https://doi.org/10.1016/j.jmateco.2020.01.001.

Mogensen, A. L. (2024). Respect for others' risk attitudes and the long-run future. *Noûs.* Advanced online publication. https://doi.org/10.1111/nous.12488.

Mongin, P. (1994). Harsanyi's aggregation theorem: multi-profile version and unsettled questions. *Social Choice and Welfare*, *11*, 331–354. https://doi.org/10.1007/BF00183302.

Mongin, P. (2016). Spurious unanimity and the pareto principle. *Economics and Philosophy*, *32*, 511–532. https://doi.org/10.1017/S0266267115000371.

Mongin, P., & d'Aspremont, C. (1998). Utility theory and ethics. In S. Barberà, P. J. Hammond, & C. Seidl (Eds.), *Handbook of Utility Theory: Volume 1* (pp. 371–481). Kluwer Academic Publishers.

Mongin, P., & Pivato, M. (2015). Ranking multidimensional alternatives and uncertain prospects. *Journal of Economic Theory*, *157*, 146–171. https://doi.org/10.1016/j.jet.2014.12.013.

Mongin, P., & Pivato, M. (2016). Social evaluation under risk and uncertainty. In M. Adler & M. Fleurbaey (Eds.), *The Oxford Handbook of Well-Being and Public Policy.* Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199325818.013.23.

Moore, A., & Crisp, R. (1996). Welfarism in moral theory. *Australasian Journal of Philosophy*, *74*, 598–613. https://doi.org/10.1080/00048409612347551.

Nebel, J. M. (2020). Rank-weighted utilitarianism and the veil of ignorance. *Ethics*, *131*, 87–106. https://doi.org/10.1086/709140.

Nebel, J. M. (2022). Aggregation without interpersonal comparisons of well-being. *Philosophy and Phenomenological Research*, *105*, 18–41. https://doi.org/10.1111/phpr.12783.

Nebel, J. M. (2024). Ethics without numbers. *Philosophy and Phenomenological Research*, *108*, 289–319. https://doi.org/10.1111/phpr.12966.

Nissan-Rozen, I. (2017). How to be an ex-post egalitarian and an ex-ante Paretian. *Analysis*, *77*, 550–558. https://doi.org/10.1093/analys/anx111.

Nissan-Rozen, I. (2020). Weighing and aggregating reasons under uncertainty: A trilemma. *Philosophical Studies*, *178*, 2853–2871. https://doi.org/10.1007/s11098-020-01587-9.

Nozick, R. (1974). *Anarchy, State, and Utopia*. Basic Books.

Parfit, D. (2011). *On What Matters: Volume 1*. Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199572809.001.0001.

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, *32*, 122–136.

Risse, M. (2002). Harsanyi's 'utilitarian theorem' and utilitarianism. *Noûs*, *36*, 550–577. https://doi.org/10.1111/1468-0068.t01-1-00401.

Roemer, J. E. (2008). Harsanyi's impartial observer is not a utilitarian. In M. Fleurbaey, M. Salles, & J. A. Weymark (Eds.), *Justice, Political Liberalism, and Utilitarianism: Themes from Harsanyi and Rawls* (p. 129–135). Cambridge University Press. https://doi.org/10.1017/CBO9780511619595.006.

Sen, A. (1970). The impossibility of a Paretian liberal. *Journal of Political Economy*, *78*, 152–157. https://doi.org/10.1086/259614.

Sen, A. (1976). Welfare inequalities and Rawlsian axiomatics. *Theory and Decision*, *7*, 243–262. https://doi.org/10.1007/BF00135080.

Sen, A. (1977). Non-linear social welfare functions. In R. Butts & J. Hintikka (Eds.), *Foundational Problems in the Special Sciences. the University of Western Ontario Series in Philosophy of Science, Vol. 10* (pp. 297–302). Springer. https://doi.org/10.1007/978-94-010-1141-9$_1$9.

Sen, A. (1979). Utilitarianism and welfarism. *Journal of Philosophy*, *76*, 463–489.

Sen, A. (1986). Social choice theory. In K. J. Arrow & M. D. Intriligator (Eds.), *Handbook of Mathematical Economics* (p. 1073-1181). Elsevier. https://doi.org/10.1016/S1573-4382(86)03004-7.

Sen, A. (2017). *Collective Choice and Social Welfare: An Expanded Edition*. Harvard University Press.

Stefánsson, H. O., & Bradley, R. (2015). How valuable are chances? *Philosophy of Science*, *82*, 602–625. https://doi.org/10.1086/682915.

Voorhoeve, A. (2021). Equality for prospective people: A novel statement and defence. *Utilitas*, *33*, 304–320. https://doi.org/10.1017/S0953820821000017.

Voorhoeve, A., & Fleurbaey, M. (2016). Priority or equality for possible people? *Ethics*, *126*, 929–954. https://doi.org/10.1086/686000.

Weymark, J. (2005). Measurement theory and the foundations of utilitarianism. *Social Choice and Welfare*, *25*, 527–555. https://doi.org/10.1007/s00355-005-0017-7.

Weymark, J. A. (1991). A reconsideration of the Harsanyi-Sen debate on utilitarianism. In J. Elster & J. E. Roemer (Eds.), *Interpersonal Comparisons of Well-Being* (pp. 255–320). Cambridge University Press. https://doi.org/10.1017/CBO9781139172387.009.

Wilkinson, H. (2022). Can risk aversion survive the long run? *Philosophical Quarterly*, *73*, 625–647. https://doi.org/10.1093/pq/pqac056.

Zhao, M. (2021). Ignore risk; maximize expected moral value. *Noûs*, *57*, 144–161. https://doi.org/10.1111/nous.12398.