



# Frege's puzzle and the *ex ante* Pareto principle

Anna Mahtani<sup>1</sup> 

© The Author(s) 2020

**Abstract** The *ex ante* Pareto principle has an intuitive pull, and it has been a principle of central importance since Harsanyi's defence of utilitarianism (to be found in e.g. Harsanyi, *Rational behaviour and bargaining equilibrium in games and social situations*. CUP, Cambridge, 1977). The principle has been used to criticize and refine a range of positions in welfare economics, including egalitarianism and prioritarianism. But this principle faces a serious problem. I have argued elsewhere (Mahtani, *J Philos* 114(6):303-323 2017) that the concept of *ex ante* Pareto superiority is not well defined, because its application in a choice situation concerning a fixed population can depend on how the members of that population are designated. I show in this paper that in almost all cases of policy choice, there will be numerous sets of rival designators for the same fixed population. I explore two ways that we might complete the definition of *ex ante* Pareto superiority. I call these the 'supervaluationist' reading and the 'subvaluationist' reading. I reject the subvaluationist reading as uncharitable, and argue that the supervaluationist reading is the most promising interpretation of the *ex ante* Pareto principle. I end by exploring some of the implications of this principle for prioritarianism and egalitarianism.

**Keywords** *Ex ante* · Pareto principle · Frege's puzzle · Reference · Prioritarian · Egalitarian · Harsanyi

---

✉ Anna Mahtani  
a.mahtani@lse.ac.uk

<sup>1</sup> Department of Philosophy, Logic and Scientific Method, London School of Economics and Political Science (LSE), London, UK

## 1 Introduction

The *ex ante* Pareto principle—like many other principles in welfare economics—involves the idea of prospects or expected welfare for individuals. I have argued elsewhere (Mahtani 2017) that, for reasons connected with Frege’s puzzle (Frege 1980), the prospects for an agent under a policy can depend on how that agent is designated. As it stands, then, the concept of *ex ante* Pareto superiority is not well defined, because its application in a choice situation concerning a fixed population can depend on how the members of that population are designated. I show in this paper that in almost all cases of policy choice, there will be numerous sets of rival designators for the same fixed population, and so that the problem with the *ex ante* Pareto principle—and indeed with many other principles of welfare economics, including certain fairness principles, and principles concerning competing claims—is widespread.

I explore two ways that we might complete the definition of *ex ante* Pareto superiority. I call these the ‘supervaluationist’ reading and the ‘subvaluationist’ reading. I argue that on the subvaluationist reading, the principle is equivalent to a version of utilitarianism, and so that this reading is an uncharitable interpretation. The supervaluationist reading is much more promising, and leaves us with a coherent principle that is faithful to the underlying rationale. I end by exploring some of the implications of this principle for debates over egalitarianism and prioritarianism.

## 2 Why the *ex ante* Pareto principle is not completely defined

The *ex ante* Pareto principle states that when you have a choice between a range of policies, you should not choose some policy  $P_y$  if some other policy  $P_x$  is *ex ante* Pareto superior. And policy  $P_x$  is *ex ante* Pareto superior to policy  $P_y$  if and only if (a) the prospects for each person under policy  $P_x$  are at least as good as under policy  $P_y$ , and (b) the prospects for at least one person under policy  $P_x$  are better than under policy  $P_y$ . In this paper I understand the prospects for a person under a policy to be the expected welfare for that person under that policy. There are interesting questions about how to assess the welfare for a person at some outcome, but I set these questions aside here and assume that there is no difficulty in assigning a number that gives the welfare for any given person in any given outcome. The more important (but, I think, not controversial) assumption that I make is that the *expected* welfare of some policy is calculated using the decision-maker’s credences.<sup>1</sup>

I have argued (Mahtani 2017) that the concept of *ex ante* Pareto superiority is not fully defined. The reason for this relates to an insight from Frege, which can be illustrated with an example (Frege 1980). The ancient Greeks used the name

---

<sup>1</sup> Or subjective probabilities, where these are the probabilities that a *rational* agent would have given the decision maker’s total evidence.

'Hesperus' for an object that they saw in the sky in the evening, and the name 'Phosphorus' for an object that they saw in the sky in the morning. In fact, the object they saw in the evening was the very same object that they saw in the morning: it was the planet Venus. Because they did not know that the two names co-referred, they could believe different things about Hesperus and Phosphorus. For example, Penelope, pointing to an object in the sky in the evening, could believe that she was pointing at Hesperus without believing that she was pointing at Phosphorus. Hence both (1) and (2) can be true, without Penelope being irrational.

- (1) Penelope believes that Hesperus is over there
- (2) Penelope does not believe that Phosphorus is over there

We might put this by saying that Penelope's beliefs are not really, or at least not directly, about the object Venus itself. Frege claimed that a name has both a reference and a sense, where the sense of a name is some 'mode of presentation' of the reference. The names 'Hesperus' and 'Phosphorus' have the same reference but different senses: they present the same object in different ways. The object of Penelope's belief is sensitive to sense as well as reference. To summarise this thought, I'll say in a rough and ready way that what an agent believes about an object can depend on how that object is designated.<sup>2</sup>

What goes for belief also goes for credences. Thus the following two claims can be true together, without Penelope being irrational:

- (1) Penelope has a high credence that Hesperus is over there
- (2) Penelope has a low credence that Phosphorus is over there

Thus what credence an agent has in some proposition about an object can depend on how that object is designated.

Obviously, this point still stands if we substitute a person for a planet. Let us suppose that Penelope—a doctor, no longer living in ancient times—is expecting two patients, Alice and Belinda, to arrive by ambulance. Reception call to say that her first patient, Ms Smith, has arrived and is waiting in the foyer to be seen. Penelope is sure that Ms Smith must be either Alice or Belinda, but she doesn't know which. Thus the following two claims can be true, even though—let us suppose—Ms Smith is in fact Alice:

- (1) Penelope has a credence greater than 0.9 that Ms Smith is the first patient to arrive.
- (2) Penelope has a credence lower than 0.9 that Alice is the first patient to arrive.

Because Penelope's credences about Ms Smith and Alice are different, the prospects for Ms Smith and Alice (as calculated using Penelope's credences) can be different. To see this, let us suppose that Penelope knows that both Alice and Belinda have an

<sup>2</sup> There are those who deny this. E.g. on Nathan Salmon's view (Salmon 1989) if Penelope believes that Hesperus is over there, then she also believes that Phosphorus is over there, and perhaps we could create an analogous account for credence (Braun 2016). My argument here does not apply directly on such a view, but I think an alternative version—focused on the analogue of the three-place BEL relation – could be constructed.

acute illness, which can be fully cured by a medicine of which Penelope has only a single dose available. She considers the policy (call it ‘first come, first served’) of just assigning it to the first patient of the two to arrive—i.e. to Ms Smith. What are the prospects for Ms Smith and Alice under this policy? The prospects for Ms Smith are excellent: Penelope knows that this policy will result in Ms Smith getting the curative dose which (let’s suppose) Penelope is certain would give Ms Smith a long and healthy life. The prospects for Alice are much less good. Penelope does not know whether Alice is Ms Smith—in which case under this policy Alice will get the curative dose—or whether Alice is not Ms Smith—in which case under this policy Alice will not get the curative dose. Thus the prospects are less good for Alice than for Ms Smith under this policy—even though Alice and Ms Smith are the very same person.

A useful image here comes from Gareth Evans (1982), who imagines us collecting ‘dossiers’ of information on various objects including people. Penelope will have a dossier of information about Alice—perhaps from emails they have exchanged, previous meetings they have had, and so on. Penelope also has a dossier of information about Ms Smith, which will include the information that Ms Smith is the first visitor to arrive. Of course, this dossier could be added to further: the receptionist might tell Penelope more about Ms Smith, and Penelope will eventually see and talk to Ms Smith when she meets her in the foyer. Thus Penelope has a dossier of information about Alice, and a dossier of information about Ms Smith, and she uses these dossiers (at whatever state they have reached at the time) when calculating the prospects for Alice and Ms Smith under a proposed policy. At some point, Penelope may discover that Alice is Ms Smith (perhaps—but not of course necessarily—at the moment when she sees Ms Smith) and Penelope will then merge the dossiers. But until the dossiers are merged, Penelope’s calculations of the prospects for Ms Smith and Alice can diverge because she has different information about each of them.

This shows that the concept of *ex ante* superiority is not fully defined as it stands: it talks about the prospects for people without specifying how these people should be designated, and (as we have seen) how people are designated makes a difference to their prospects. Simply applying the concept without resolving this issue can lead to contradiction, as we can easily see using our example. Let us suppose the receptionist tells Penelope both that Ms Smith is already in the foyer, and that Ms Jones (Penelope’s other patient) is delayed and will arrive a bit later. Penelope is sure that there are only two patients arriving today, so she is sure that either Ms Smith is Alice and Ms Jones is Belinda, or vice versa. We can suppose (just to make the example simple) that Penelope divides her credence equally between these two possibilities. Now Penelope considers two policies she could adopt with the dose: she could follow the policy ‘first come, first served’; or she could split it in two (call this ‘split’), which will render it much less effective for each individual patient. We can suppose that Penelope is sure that either patient would have a large amount of well-being (10) on receiving a full-dose; a much smaller amount of well-being (4) on receiving a half-dose; and a much smaller amount of well-being still (0) on receiving no dose at all. Penelope then calculates the prospects for Alice and Belinda as follows (Table 1).

**Table 1** Prospects for Alice and Belinda

	State $S_1$ : Alice is Ms Smith, and Belinda is Ms Jones $P(S_1) = 0.5$	State $S_2$ : Alice is Ms Jones, and Belinda is Ms Smith $P(S_2) = 0.5$	Prospects
First come, first served	Alice: 10 Belinda: 0	Alice: 0 Belinda: 10	Alice: 5 Belinda: 5
Split	Alice: 4 Belinda: 4	Alice: 4 Belinda: 4	Alice: 4 Belinda: 4

**Table 2** Prospects for Ms Smith and Ms Jones

	State $S_1$ : Alice is Ms Smith, and Belinda is Ms Jones $P(S_1) = 0.5$	State $S_2$ : Alice is Ms Jones, and Belinda is Ms Smith $P(S_2) = 0.5$	Prospects
First come, first served	Ms Smith: 10 Ms Jones: 0	Ms Smith: 10 Ms Jones: 0	Ms Smith: 10 Ms Jones: 0
Split	Ms Smith: 4 Ms Jones: 4	Ms Smith: 4 Ms Jones: 4	Ms Smith: 4 Ms Jones: 4

From Table 1, we can see that *First come, first served* is *ex ante* Pareto superior to *Split*: the prospects are better for both Alice and Belinda, and we know that Alice and Belinda are the only two people who will be affected by the choice. But these two people can be equally well designated as ‘Ms Smith’ and ‘Ms Jones’ (though of course Penelope does not know which is which). We have seen that the prospects for a person depend on how (s)he is designated, so what happens if we calculate the prospects for Ms Smith and Ms Jones under the two possible policies (Table 2)? Now it seems that neither action is Pareto superior: *First come, first served* is better for Ms Smith, but *Split* is better for Ms Jones.

In the example I’ve described, I’ve supposed that the decision-maker is certain of the outcomes for Ms Smith and Ms Jones under each policy, and so under this set of designators the *ex ante* prospects are identical to the *ex post* outcomes, whereas the *ex ante* prospects for Alice and Belinda are different from the *ex post* outcomes. This might suggest that the example merely shows that the *ex ante* prospects for the people concerned (as given by the prospects for Alice and Belinda) are different from the *ex post* outcomes for the people concerned. But this is just a feature of the simplicity of this case: by introducing some uncertainty, we can make it clear that we are calculating *ex ante* prospects (rather than *ex post* outcomes) for Ms Smith and Ms Jones, as well as for Alice and Belinda. We could suppose, for example, that the decision-maker is less than completely certain that the dose of medicine will be curative, in which case the prospects for Ms Smith under *First Come, First Served* are (let’s suppose) very slightly below 10 (and the prospects for Alice and Belinda under this policy will also need to be slightly reduced similarly): if we adjust the

example in this way, we can still get *First come, first served* to come out as *ex ante* Pareto superior when we consider the prospects for Alice and Belinda, but not when we consider the prospects for Ms Smith and Ms Jones. This should make it clear that the issue here is not a mere distinction between *ex ante* prospects and *ex post* outcomes for an individual, but between different calculations of that individual's *ex ante* prospects depending on how that individual is designated.

In this scenario, Penelope is certain that there are only two people in the population to be considered. Yet she has two different ways of designating this pair of people. Designated in one way (as 'Alice' and 'Belinda'), *First come, first served* appears *ex ante* Pareto superior, but designated in another way (as 'Ms Smith' and 'Ms Jones'), *First come, first served* does not appear *ex ante* Pareto superior. Thus in order to determine whether one action is *ex ante* Pareto superior to the other, we need to complete the definition of *ex ante* Pareto superiority. In this paper, I focus on two ways of completing this definition: a supervaluationist reading and a subvaluationist reading. I describe these in the next section.

### 3 The two readings

Let us start by supposing that we have a fixed population of  $n$  people, and that the decision-maker knows this. We consider some set  $D$  of designators for these  $n$  people. Let us say that such a set is 'admissible' iff it meets the following two requirements:

- (a) For the population of  $n$  people, each person has one and only one designator within  $D$ .
- (b) The decision-maker knows that (a) holds. That is, for every designator within  $D$ , the decision-maker knows that that designator uniquely designates one member of the population; and the designator knows that the designators in  $D$  between them designate every member of the population.

The example in the last section shows that in some cases there may be multiple admissible sets of designators: both the sets {'Alice', 'Belinda'} and {'Ms Smith', 'Ms Jones'} meet requirements (a) and (b). Suppose then that the decision-maker is choosing from some set of possible policies  $P = \{P_1, P_2, \dots, P_n\}$ , and also that there are various possible states of the world,  $S = \{S_1, S_2, \dots, S_z\}$ . For any given designator, policy, and state of the world, there will be some welfare outcome: the outcome that the person (so designated) would certainly have at that state under that policy. The decision-maker has some credence in each state of the world. Thus for each designator and policy, there will be some *expected* welfare, which is the sum of the outcomes for that designator under that policy in each state, weighted by credence.<sup>3</sup> This gives us the prospects for that designator under that policy. It may be that for some policies  $P_x$  and  $P_y$ , and some admissible set of designators  $D = \{D_1,$

<sup>3</sup> It is really *people under designators* that have prospects: I write of designators having prospects as a shorthand.

$D_2, \dots, D_n$ }, prospects are at least as good under  $P_x$  as under  $P_y$  for every designator in set  $D$ , and the prospects are better under  $P_x$  than under  $P_y$  for some designator in set  $D$ : in this case, we can say that  $P_x$  is *ex ante* Pareto superior to  $P_y$  relative to the set of designators  $D$ .

But given the very same population of  $n$  people, there may be alternative admissible sets of designators. Let us suppose that  $D^* = \{D_1^*, D_2^*, D_3^*, \dots, D_n^*\}$  is just such another set of designators, meeting the requirements (a) and (b) above. Just to emphasise—there is no variation in the population here: the idea is rather that there can be two different ways of designating the members of the very same population. We can then also consider whether  $P_x$  is *ex ante* Pareto superior to  $P_y$  relative to the set of designators  $D^*$ . Quite generally, we can consider whether  $P_x$  is *ex ante* Pareto superior to  $P_y$  relative to each admissible set of designators.

I have stated what it is for some policy to be *ex ante* Pareto superior to another policy relative to some admissible set of designators. But what is it for some policy to be *ex ante* Pareto superior (given our fixed population) *simpliciter*—i.e. not relative to some particular set of designators? There are two natural ways we might answer this question. Firstly, we might say that some policy  $P_x$  is *ex ante* Pareto superior simpliciter to some policy  $P_y$  iff there is *some* admissible set of designators  $D$  such that  $P_x$  is *ex ante* Pareto superior to  $P_y$  relative to  $D$ . This is what I am calling a subvaluationist reading. Secondly, we might say that some policy  $P_x$  is *ex ante* Pareto superior simpliciter to some policy  $P_y$  iff for *every* admissible set of designators  $D$ ,  $P_x$  is *ex ante* Pareto superior to  $P_y$  relative to  $D$ : this is what I am calling a supervaluationist reading.

An alternative way of reading the *ex ante* Pareto principle deserves a mention—but only to dismiss it at once. This is the option of fixing—for any given choice scenario and fixed population—a particular admissible set of designators as the special ones: if  $P_x$  is *ex ante* Pareto superior to  $P_y$  relative to the special set of designators, then it is *ex ante* Pareto superior simpliciter. But the problem is that the choice of the set of designators is itself an important moral choice. We might say that we should use the designators that pick people out by their most relevant characteristics, but who is to say which characteristics are the relevant ones in a given scenario? The motivation for the *ex ante* Pareto principle was simply a concern for the prospects for each person, considered separately. If it turns out that in applying the *ex ante* Pareto principle we are (explicitly or tacitly) making some further moral choice in our selection of designators, then the principle will have lost its appeal. The subvaluationist and supervaluationist readings do not require us to make such a moral choice, and to this extent at least they are natural and faithful expressions of the spirit behind the *ex ante* Pareto principle.

I discuss both the subvaluationist and supervaluationist readings in detail below. But first I turn to a key question: besides (a) and (b) above, are there any other restrictions on which sets of designators are admissible?

## 4 Admissible sets of designators

Philosophers of language recognize a variety of different types of designators, including proper names, definite descriptions, demonstratives and indexicals. Numerous designators of any or all of these types can apply to a single person. For example, we are supposing that one of Penelope's patients has the name 'Alice' and also the name 'Ms Smith'. She may have other names: perhaps her colleagues know her as 'Professor Randall' and her childhood friends know her as 'Tiddles'. There will also be a multitude of definite descriptions that apply to her: we know that she is the first patient to arrive to see Penelope; perhaps she is also the patient with NHS number 2487341, and no doubt there are numerous other ways to describe her. Then there are demonstratives, such as 'that' and 'this', and indexicals such as 'me' and 'you', that could be used (in the right circumstances) to designate this person.

There are all sorts of theories about these different sorts of designators and how they work. I don't enter into issues over how to analyse these designators here. I will just assume for the present that any of these different types of designators can be members of an admissible set of designators. This is an assumption that I will revisit shortly, but taking it for granted shows us quickly how to generate multiple admissible sets of designators for a given population. The earlier example involving the sets {'Alice', 'Belinda'} and {'Ms Smith', 'Ms Jones'} may have seemed quite contrived, as though I had to carefully craft the scenario in order to get two rival sets of designators, and this might make you think that the problem I have raised for the concept of *ex ante* Pareto superiority is quite obscure. But in fact there will be multiple sets of rival designators in almost all cases where we might want to apply the concept: they may not all have the intuitive force of the sets in the example I gave in the last section, but (I will argue) the intuitive force is not important. Thus the problem for the concept of *ex ante* Pareto superiority is widespread. I turn now to show how we can generate multiple sets of rival designators in almost any scenario.

Let us suppose that the population consists of just two people, and let us begin with one set of designators  $D$  that is admissible—that is, the decision-maker knows that each designator in  $D$  designates one member of the population, and that every member of the population is designated by some designator in  $D$ . Suppose that this set  $D = \{\text{'Chris', 'Dom'}\}$ . We can then use this set  $D$  to generate other sets  $D^*$ ,  $D^{**}$  and so on that are also admissible. To do so, we just need the decision-maker to have some uncertainty—uncertainty over anything will do. We can suppose for example that the decision-maker has been tossing a coin, and has dropped it under the sofa, so that she doesn't know which way up it has landed. She has a credence of 0.5 that it has landed heads (HEADS), and a credence of 0.5 that it has landed tails (TAILS). Then we can coin the following predicate,  $F$ , where a person  $x$  is  $F$  iff  $x$  is either Chris and HEADS obtains, or Dom and TAILS obtains. This gives us the definite description 'the  $F$ '. We can similarly construct the reverse predicate,  $F^*$ , where a person  $x$  is  $F^*$  iff  $x$  is either Dom and HEADS obtains, or Chris and TAILS obtains. This gives us the definite description 'the  $F^*$ '. Thus we have a second set of admissible designators:  $D^* = \{\text{'the } F\text{'}, \text{'the } F^*\text{'}\}$ .



The prospects for the F and the F\* under a given policy may (but need not) be different from the prospects for Dom and Chris. To see how the prospects might differ, suppose we consider the policy whereby Chris gets 10 units of welfare if the coin landed heads, and nothing otherwise, and Dom gets 10 units of welfare if it landed tails and nothing otherwise. Under this policy, the prospects for both Dom and Chris are  $(0.5)(10) = 5$ . In contrast, the prospects for the F and the F\* under this policy are 10 and 0 respectively. Thus the prospects for each of the members of set D under a given policy can be different from the prospects for each of the members of set D\* under the same policy. The total prospects under a given policy will be the same for the members of D as for D\*, but the distribution can vary. For example, we can see that the total prospects under the relevant policy for the members of  $D = \{\text{'Chris'}, \text{'Dom'}\}$  come to  $5 + 5 = 10$ , and the total prospects for the members of  $D^* = \{\text{'the F'}, \text{'the F*'}\}$  also come to  $10 + 0 = 10$ , but the distribution pattern is different.

In the earlier case of  $\{\text{'Alice'}, \text{'Belinda'}\}$  and  $\{\text{'Ms Smith'}, \text{'Ms Jones'}\}$ , the claim that here we had two rival sets of designators seemed quite compelling. Above, in the case of  $\{\text{'Chris'}, \text{'Dom'}\}$  and  $\{\text{'the F'}, \text{'the F*'}\}$  we saw how from a seed set of designators it was easy to generate a further set of designators, given some uncertainty over a partition. From here we can see (and I discuss below) how widespread are cases of multiple sets of designators. But the generated designators  $\{\text{'the F'}, \text{'the F*'}\}$  are rather unintuitive, and so here it may seem less clear that we really do have two rival sets of designators. After all, the generated designators seem to be gerrymandered: should we allow designators that are gerrymandered in this way? I turn to argue now that we must, because there is no good rationale for excluding them.

Let's start by considering what sorts of restrictions we could try placing on sets of designators that would allow us to exclude sets like  $\{\text{'the F'}, \text{'the F*'}\}$ . We might begin by ruling out definite descriptions, and insisting that only proper names are allowed. But a problem with this idea is that proper names are easily produced, for they can be defined by description.<sup>4</sup> Thus the decision-maker can simply state that henceforth the person who is the F, whoever that is, shall be called 'Frank' and whoever is the F\* shall be called 'Fred'. Then we can replace the set  $\{\text{'the F'}, \text{'the F*'}\}$  with the set  $\{\text{'Frank'}, \text{'Fred'}\}$  which consists of proper names as required. The same point applies to the attempt to limit the designators to rigid designators, for of course given that 'Frank' is a proper name, it is a rigid designator (and indeed we could have just made the definite descriptions rigid, by replacing 'the F' with 'the actual F' and so on).

It might be objected that names produced in this way are not the right sorts of names. We don't *normally* name things by description: normally naming something involves standing in some sort of causal relationship with the thing named. Exactly what this involves is a debated question (Kripke 1980, Searle 1983). When I get my new pet cat and say in its presence, 'I hereby call it Felix', then I do stand in a causal relationship to the cat just by standing near it, but so I do to the new basket, the

<sup>4</sup> See Gareth Evans on 'Julius, the inventor of the zip' (Evans, *The Varieties of Reference* 1982).

immunisation certificate and the other things in the room, and I haven't named *those* things. The causal relationship is not enough all by itself: it needs to be backed up with an intention directed towards a particular object. Why then couldn't I stand in front of Chris and Dom and say 'I hereby call him 'Frank''—intending to name the person who is the F?

A different objection is that *intuitively* there is just no obligation to worry about the prospects for the F in the way that we should worry about the prospects for Chris and Dom. The intuition here might be that the F is not a real person, but some sort of gerrymandered figure. Of course, the person 'the F' designates is not a gerrymandered figure: the expression designates an actual living person, just as the names 'Chris' and 'Dom' do.<sup>5</sup> The designator is reasonably described as gerrymandered, but the person designated by the name is not. But perhaps the point is just that because the designator is gerrymandered, there is no intuitive obligation to worry about the F's prospects. To address this worry, we can add some detail to the scenario so that intuitively we *should* be concerned with the prospects for the F. Let's suppose that there is someone else ('the taunter') in the room with the decision-maker. The taunter knows Chris and Dom well (much better than the decision-maker, let's suppose), and has looked under the sofa and seen how the coin has landed, and so knows who the F is. The taunter can then tell the decision-maker lots of information about the F. For example, the taunter can explain that the F is known to his or her friends as Mosschops, show the decision-maker various photos of the F and so on. The taunter could do the same for the F\*. The decision-maker could then end up far more informed about the F and the F\* than (s)he is about Chris and Dom: the decision-maker's dossiers on 'the F' and 'the F\*' are bulging with information, while his or her dossiers on Chris and Dom are rather thin. It is now very natural for the decision-maker to consider the prospects for the F and the F\* under each policy. And the designators 'the F' and 'the F\*' may no longer seem like gerrymandered designators. In fact, perhaps 'Chris' and 'Dom' are the gerrymandered designators: the decision-maker has a wealth of information about the F and the F\*, and really just thinks of 'Chris' as a name for the person who is the F if the coin has landed heads, and the F\* otherwise. Thus by giving the decision-maker more information about 'the F' and 'the F\*', it becomes intuitive to be concerned with their prospects.<sup>6</sup> But it can't be the case that 'the F' is a designator worthy of

<sup>5</sup> It can help here to remember that the coin has already been tossed.

<sup>6</sup> How about a reading on which when applying the *ex ante* pareto principle the relevant set of designators is the admissible set of designators for which the decision-maker has the largest dossiers? There are some technical challenges for this suggestion: there is a question over how we measure the size of a dossier; a further question over how we aggregate the dossier sizes for each designator to give us a ranking of admissible sets of designators (because there is no guarantee that the designators with the largest dossiers will all fall within a single admissible set); and a question over what happens should two rival admissible sets of dossiers both come at the top of this ranking. But a more fundamental problem is that on this reading the commands of the *ex ante* pareto principle become capricious (and open to manipulation). The principle might require a decision maker not to select a particular policy P—but should the decision maker come to learn some irrelevant and frivolous fact about an individual concerned (under some designator), then the relevant set of designators may shift and that policy P may no longer be debared. Under this reading the principle is hardly worthy of the weighty role that it has played in the

concern only if the decision-maker has enough information about 'the F'. The intuition behind the original *ex ante* Pareto principle was that it concerned *all* people: not just people that we felt some kind of connection with or had lots of information about, but all people regardless. Now we have seen that we need to consider designators rather than people, the same analogous principle should apply: the concept concerns all designators, not just designators that we feel some sort of interest in. Thus there is no good rationale for excluding sets of gerrymandered designators.

An alternative objection is to say that we should focus on designators that pick out the same person at every state. And (the objector might say) while 'Chris' and 'Dom' do pick out the same person at every state, 'the F' and 'the F<sup>0</sup>' do not. The thought here may relate to the idea that some but not all designators are rigid, where a rigid designator picks out the same object at all metaphysically possible worlds, whereas a non-rigid designator does not. Thus for example the rigid designator 'George Orwell' picks out the same person at every metaphysically possible world (where he exists), whereas the non-rigid designator 'the winner of the Hugo award', which happens to also pick out George Orwell at the actual world, picks out other authors at other metaphysically possible worlds, because of course it is possible for other authors to have won that award instead. This distinction between rigid and non-rigid designators seems to make sense when we are thinking about metaphysically possible worlds. But the states that form part of the decision theorists' and welfare economists' framework are not metaphysically possible worlds. To see this, consider that the name 'George Orwell' picks out the same person at every metaphysically possible world (where he exists), and so does the name 'Eric Blair'. Given that at the actual world, 'George Orwell' and 'Eric Blair' pick out the very same person ('George Orwell' was the pen-name of Eric Blair), these two names pick out the same person at every metaphysically possible world (where he exists). Thus there are no metaphysically possible worlds where George Orwell is not Eric Blair. But clearly a decision-maker might have a positive credence in the possibility that George Orwell is not Eric Blair. Thus we need a state where George Orwell is not Eric Blair, and as there is no metaphysically possible world where this holds, states cannot be metaphysically possible worlds. We are dealing here with epistemic rather than metaphysical modality.

Can we make sense of the idea of a designator that is rigid across states—interpreted as epistemically possible worlds, rather than metaphysically possible worlds? It is not at all obvious that we can.<sup>7</sup> At any rate, ordinary proper names—though rigid across metaphysically possible worlds—will not be rigid across epistemically possible worlds. We can see this by noting that, in order to allow an agent to be uncertain whether George Orwell is Eric Blair, we need George Orwell to be Eric Blair in some epistemically possible worlds but not others, so it cannot be

---

Footnote 6 continued

literature, and can be rejected as unpersuasive. Many thanks to an anonymous reviewer for alerting me to this possible reading.

<sup>7</sup> If we can, I think they might be what Russell called 'logically proper names' (Russell 1912).

the case that *both* names refer rigidly or they would refer to the same person at all worlds. Clearly then not all proper names refer rigidly across epistemically possible worlds, and we cannot class a designator as rigid (in this sense) just in virtue of its logical form. We might hope instead to class some designators as rigid (in this sense) in virtue of the information connected with those designators. If our agent has a lot of information about Eric Blair, then we might think that ‘Eric Blair’ should count as a rigid designator; but our agent might also have a lot of information about George Orwell, and as we have seen we can’t have both ‘Eric Blair’ and ‘George Orwell’ counting as rigid designators over epistemically possible worlds. The same holds for our designators ‘Chris’ and ‘the F’: these cannot both be rigid designators (over epistemically possible worlds) because there will be some epistemically possible worlds where Chris is the F, and some where he isn’t; we should not privilege ‘Chris’ as a rigid designator just because it is a proper name (recall that it is easy enough to coin a proper name to designate whoever is the F); and there may be no good way to discriminate between the two designators based on the decision maker’s information. Indeed, the very idea of rigidity—once it is recognised that we are dealing with epistemically rather than metaphysically possible worlds—needs to be rethought.<sup>8</sup> Thus we cannot sensibly restrict the admissible sets of designators to those which contain only designators which are rigid across states, and so the sets {‘Chris’, ‘Dom’} and {‘the F’, ‘the F\*’} are on a par.

We generated the set of designators  $D^* = \{‘the F’, ‘the F^*’\}$  from the original set of designators  $D = \{‘Chris’, ‘Dom’\}$  in the following way. We found some partition over which the decision-maker was uncertain (in this case the events HEADS and TAILS). We then defined each new designator  $D_k^*$  in the set  $D^*$  by stating, for each event, identity between  $D_k^*$  and some member of  $D$ . Thus for example, we defined ‘the F’ in set  $D^*$  by stating that at HEADS the F is Chris, and at TAILS the F is Dom. We defined each designator in  $D^*$  in this way, ensuring that at each event each member of  $D^*$  was paired one-to-one with a member of  $D$ . In this way we arrived at the new set  $D^* = \{‘the F’, ‘the F^*’\}$ . Call this process of moving from the set  $D$  to the set  $D^*$  ‘gerrymandering’. The strategy can be repeated. Suppose for example that the decision-maker is uncertain how his next die roll will land. Then we can define a set of new designators  $D^{**} = \{‘the G’, ‘the G^*’\}$ , where the G is Chris if the die lands on 1, but Dom otherwise, and the  $G^*$  is Dom if the die lands on 1, but Chris otherwise. And we can define another set of new designators  $D^{***} = \{‘the H’, ‘the H^*’\}$ , where the H is Chris if the die lands on 1 or 2, but Dom otherwise, and the  $H^*$  is Dom if the die lands on 1 or 2, but Chris otherwise. And many other sets of designators are possible.

Thus for almost any decision situation involving uncertainty, for a fixed population there will be multiple admissible sets of designators.<sup>9</sup> Starting from an

<sup>8</sup> There are some interesting relations between this line of thinking and the work of several recent philosophers of language on epistemic modality (Ninan 2018; Yalcin 2015).

<sup>9</sup> There may well be situations in which there are an *infinite* number of admissible sets of designators: the subvaluationist and supervaluationist readings can still apply in this case, just as subvaluationist and supervaluationist accounts of vagueness can apply even under the assumption that the number of admissible precisifications is infinite.

admissible set of designators, we can generate more sets of designators by the process that I am calling 'gerrymandering'. Let us say then that the collection of admissible sets of designators should be closed under gerrymandering—meaning that if there is any admissible set  $D$  of designators which can be converted into the set  $D^*$  by the process I am calling 'gerrymandering', then  $D^*$  is also admissible. Wherever there is a population of more than one person and uncertainty over the state of the world, there will be more than one admissible set of designators. There are just three sorts of cases in which for a fixed population, you will have effectively only one set of admissible designators: first, cases where the relevant population is empty; second, cases where the relevant population contains just one person; and third, cases where the decision-maker has no uncertainty about the state of the world. The concept of *ex ante* Pareto superiority would be of limited interest in these sorts of cases. In all other cases, there will be multiple admissible sets of designators. Thus the problem with the concept of *ex ante* Pareto does not just arise for a few contrived examples, but infects almost all cases where we might wish to use the concept. If we want to carry on using the concept then, we will need to figure out how it should be read given that we have these rival sets of designators. I turn now to the first of two ways of reading the *ex ante* Pareto principle—the subvaluationist reading.

## 5 The subvaluationist reading

We already know what it is for one policy  $P_x$  to be *ex ante* Pareto superior to another policy  $P_y$  relative to some admissible set of designators  $D$ . But what is it for a policy  $P_x$  to be *ex ante* Pareto superior simpliciter (given a particular fixed population)? In this section I consider one answer to this question: the subvaluationist reading. On this view,  $P_x$  is *ex ante* Pareto superior<sub>sub</sub> to  $P_y$  provided that there is some admissible set of designators  $D$  such that  $P_x$  is *ex ante* Pareto superior to  $P_y$  relative to  $D$ .

To see how this works, consider again the case of Penelope, deciding what policy to follow with the dose of medicine. Recall that she considers two policies: *First come first served*, and *Split*. Under the policy *First come first served*, the prospects for both Alice and Belinda are  $(0.5)(10) + (0.5)(0) = 5$ . Under the policy *Split*, the prospects for both Alice and Belinda are  $(0.5)(4) + (0.5)(4) = 4$ . Thus relative to the set of designators {'Alice', 'Belinda'}, *First come first served* is *ex ante* Pareto superior to *Split*. And because there is some set of designators relative to which *First come first served* is *ex ante* Pareto superior to *Split*, it follows that it is *ex ante* Pareto superior<sub>sub</sub>. Relative to another set of designators, {'Ms Smith', 'Ms Jones'}, *First come first served* is not *ex ante* Pareto superior to *Split*, but this does not prevent *First come first served* from being *ex ante* Pareto superior<sub>sub</sub> to *Split*, because all that matters is that there is some set of designators relative to which the relation holds.

On this way of reading *ex ante* Pareto superiority, the *ex ante* Pareto principle effectively collapses into utilitarianism. This is because in almost all cases:

Policy  $P_x$  is *ex ante* Pareto superior<sub>sub</sub> to policy  $P_y$   
iff

The total prospects under Policy  $P_x$  are greater than the total prospects under  $P_y$

To see that this holds, consider first that the first claim entails the second. If  $P_x$  is *ex ante* Pareto superior<sub>sub</sub> to policy  $P_y$ , then there must be some set of designators  $D$  relative to which  $P_x$  is *ex ante* Pareto superior to  $P_y$ . The prospects for each designator in  $D$  under  $P_x$  must be at least as great as the prospects for each designator in  $D$  under  $P_y$ , and for some designator in  $D$  the prospects under  $P_x$  must be better than the prospects under  $P_y$ . Thus the total prospects for the designators in  $D$  must be greater under  $P_x$  than under  $P_y$ . The total prospects under a policy for a fixed population is the same under any admissible set of designators, so we can say simply that if policy  $P_x$  is *ex ante* Pareto superior<sub>sub</sub> to policy  $P_y$ , then the total prospects under  $P_x$  are greater than the total prospects under  $P_y$ .

To show that the second claim entails the first, we need to start by making the assumption that the decision-maker has uncertainty across some partition with the following features:

1. The number of events in the partition is the same as the number of people in the population.
2. The decision-maker's credence is distributed equally across the events in the partition.
3. For some admissible set of designators  $D$ , the distribution of welfare across these designators is the same at each event in this partition.

Provided that these conditions are met (as they will be in nearly all policy choice situations),<sup>10</sup> then there will be an admissible set of designators across which the total prospects will be evenly distributed. I illustrate this with an example.

We can suppose that the fixed population contains 6 people, and one admissible set of designators is {'A', 'B', 'C', 'D', 'E', 'F'}. Perhaps under some policy, the total welfare is distributed in some uneven pattern amongst these designators, with A getting 6, B getting 2, and each of C-F each getting 1. Now we suppose that—quite unrelatedly—a dice has been tossed, without the decision-maker knowing the outcome. Thus we have the partition of events {1, 2, 3, 4, 5, 6} which meets the requirements above: there are 6 events and a population of 6; the decision-maker

<sup>10</sup> In the example I assume that a dice has been tossed, and it might be thought that there will be plenty of policy situations where no such relevant random event has occurred to give us the necessary partition. In fact in everyday policy situations a partition fulfilling conditions 1–3 is bound to be available. One rich source of useful partitions involves claims about the people in the population: in what order did each person cut his or her first tooth? Who is currently positioned the closest to a badger sett?

It is a challenge to produce a case in which there is no partition available fulfilling conditions 1–3. One such case would be where the decision-maker is omniscient—but of course the concept of *ex ante* Pareto superiority is not useful in such a case. Another case is one where the decision-maker is very nearly omniscient: perhaps (s)he is uncertain how a particular coin will land, but has no other uncertainty whatsoever. If his or her policy choice concerns more than two people, then there will be no partition available fulfilling conditions 1–3. These are the sorts of (very unusual) cases where the *ex ante* Pareto principle<sub>sub</sub> and utilitarianism might give different verdicts, and it is only because of these sorts of cases that I qualify the claim that the *ex ante* Pareto principle collapses into utilitarianism.

takes each event to be equally likely; the distribution of prospects amongst the designators {‘A’, ‘B’, ‘C’, ‘D’, ‘E’, ‘F’} is independent of the events in this partition—so for example under the supposition that the die lands on 3 it is still the case that the prospects for A are 6, the prospects for B are 2 and so on. We can now coin the predicates ‘G’, ‘G\*’ and so on, where to be G is to be A if the die lands on 1, B if the die lands on 2, and so on; and to be G\* is to be B if the die lands on 1, C if the die lands on 2, and so on. This gives us an admissible set of designators {‘the G’, ‘the G\*...’}. The prospects for the G can then be calculated by summing the prospects for A–F, weighted by the credence that the G is identical to each of A–F. Thus for example, the prospects for A (6) will be weighted by the credence that the G is identical to the A (1/6); and the prospects for B (2) will be weighted by the credence that the G is identical to the B (1/6); and so on. The prospects for the G work out at  $(6 + 2 + 1 + 1 + 1 + 1)/6 = 2$  which of course is 1/6 of the total prospects. The prospects for the G\* and so on can be calculated in a similar way. Thus the total prospects are distributed evenly amongst the designators in {‘the G’, ‘the G\*...’}. If the total prospects under some policy  $P_x$  are greater than the total prospects under some policy  $P_y$ , then every member of {‘the G’, ‘the G\*...’} will have better prospects under  $P_x$  than under  $P_y$ , and so  $P_x$  will be *ex ante* Pareto superior to  $P_y$  relative to this set of designators. And given that  $P_x$  is *ex ante* Pareto superior to  $P_y$  under *some* admissible set of designators, it is *ex ante* Pareto superior<sub>sub</sub>.

Thus Policy  $P_x$  is *ex ante* Pareto superior<sub>sub</sub> to policy  $P_y$  iff the total prospects under Policy  $P_x$  are greater than the total prospects under  $P_y$ . The *ex ante* Pareto principle<sub>sub</sub> is thus equivalent to utilitarianism: to say that you must not choose some policy  $P_x$  if some other available policy  $P_y$  is *ex ante* Pareto superior<sub>sub</sub>, is just to say that you must not choose  $P_x$  if some other policy  $P_y$  has greater total prospects. On this reading, the *ex ante* Pareto principle<sub>sub</sub> is not concerned in any way with the distribution of prospects, but only with the total sum.

This argument has some relation to one of John Harsanyi's argument for utilitarianism.<sup>11</sup> According to Harsanyi, our ‘moral preferences’ are our preferences that are ‘impartial’ and ‘impersonal’. He writes:

“Individual *I*'s choice among alternative social situations would certainly satisfy this requirement of impartiality and impersonality, if he... thought he would have an *equal probability* of being *put in the place* of any one among the *n* individual members of society” (Harsanyi 1977, 49–50).

Thus a person's moral preferences amongst a range of policies are those preferences (s)he would have were (s)he to have an equal credence that (s)he is any member of the relevant population. Harsanyi shows that this would give preferences for policies that have the greatest expected welfare—i.e. the greatest total prospects. We can see a similarity here with my argument above: ‘the G’ and ‘the G\*’ and so on are what we might call anonymous designators, equally likely to be any of A–F, and so when we consider the prospects on behalf of the G we are effectively considering the prospects from the perspective of an agent who doesn't know who

<sup>11</sup> See also a recent criticism of egalitarianism in that vein (McCarthy 2015).

in the population (s)he is, which is just how Harsanyi claims we arrive at moral preferences.

There is a key difference though. Harsanyi's argument faces an important challenge: why should we accept that what is morally right is determined by our preferences in a state of ignorance? As Brian Barry writes: 'No adequate reason has ever been given (by Harsanyi or anybody else) for identifying moral judgments with those made by someone trying to maximise his own prospects from behind a veil of ignorance' (Barry 1989, 78–79). My argument does not face this challenge. If in some choice situation conditions 1–3 above are all met, then there just will be an admissible set of designators that is anonymous in the required way, and the total prospects will be distributed evenly amongst these designators. And so given the subvaluationist reading of the *ex ante* Pareto principle, it simply follows that a policy is *ex ante* Pareto superior<sub>sub</sub> iff it has greater total prospects.

The big question here is whether to accept a subvaluationist reading of the *ex ante* Pareto principle—given that to do so is to accept a version of utilitarianism. Dialectically, nobody would accept the *ex ante* Pareto principle under its subvaluationist reading unless they were already committed to utilitarianism, given that (as we shall see in the next section) an alternative reading of the *ex ante* Pareto principle is available. Furthermore, the subvaluationist reading of the *ex ante* Pareto principle is an uncharitable interpretation. It is a concern with the prospects for each person considered separately that motivates the *ex ante* Pareto principle, whereas whether one policy is *ex ante* Pareto superior<sub>sub</sub> to another turns out to depend just on the total massed prospects for the whole population. For these reasons, I lay the subvaluationist reading to one side here, and turn to consider the alternative.

## 6 The supervaluationist reading

On the supervaluationist reading,  $P_x$  is *ex ante* Pareto superior<sub>sup</sub> to  $P_y$  iff  $P_x$  is *ex ante* Pareto superior to  $P_y$  relative to *all* admissible sets of designators. Once one has grasped the problem with the *ex ante* Pareto principle as it has previously been stated, this reading seems like the natural way to capture its spirit. On this reading, the criteria for *ex ante* Pareto superiority is quite demanding. Instead of just considering one set of admissible designators, to establish *ex ante* Pareto superiority we must consider *all* sets of admissible designators. This means that in many cases where it appeared as though one policy was *ex ante* Pareto superior to another, that does not hold on this reading. As the criteria for *ex ante* Pareto superiority becomes more demanding, the *ex ante* Pareto principle becomes weaker, because it places a restriction on the decision-maker in fewer choice situations. Nevertheless the *ex ante* Pareto principle<sub>sup</sub> is not completely toothless: it still imposes some restrictions, and I explore some of the boundaries of these restrictions below.

### 6.1 Greater expected welfare.

- (i) If policy  $P_x$  is *ex ante* Pareto superior<sub>sup</sub> to policy  $P_y$ , then the total prospects under  $P_x$  will be greater than the total prospects under  $P_y$ .



**Table 3** Prospects for Chris and Dom

	State $S_1$ : HEADS $P(S_1) = 0.5$	State $S_2$ : TAILS $P(S_2) = 0.5$	Prospects
Large gift to Chris if HEADS	Chris: 30 Dom: 0	Chris: 0 Dom: 0	Chris: 15 Dom: 0
Small gift to both	Chris: 3 Dom: 3	Chris: 3 Dom: 3	Chris: 3 Dom: 3

To see why (i) holds, consider that if policy  $P_x$  is *ex ante* Pareto superior<sub>sup</sub> to policy  $P_y$ , then  $P_x$  is certainly *ex ante* Pareto superior<sub>sub</sub> to policy  $P_y$ .<sup>12</sup> And it has already been shown that if  $P_x$  is *ex ante* Pareto superior<sub>sub</sub> to  $P_y$ , then the total prospects under  $P_x$  are greater than the total prospects under  $P_y$ .

The reverse claim does not hold, or in other words:

- (ii) From the claim that the total prospects under  $P_x$  is greater than the total prospects under  $P_y$ , it does not follow that  $P_x$  is *ex ante* Pareto superior<sub>sup</sub> to  $P_y$ .

This is easily proved with an example. In this example, we have a population of just two people, and they can be designated using {Chris, Dom} (Table 3). Here we can see that *Large gift to Chris if HEADS* has greater total prospects ( $15 + 0 = 15$ ) than *Small gift to both* ( $3 + 3 = 6$ ) for the designator set {'Chris', 'Dom'}. And yet *Large gift to Chris if HEADS* is not *ex ante* Pareto superior<sub>sup</sub> to *Small gift to both*, because it is not *ex ante* Pareto superior relative to the set {'Chris', 'Dom'}, and given that we are using the supervaluationist reading, it must be *ex ante* Pareto superior to every set in C to count as *ex ante* Pareto superior<sub>sup</sub>. Thus—unlike on the subvaluationist reading—on the supervaluationist reading greater total prospects does not guarantee that a policy is *ex ante* Pareto superior.

### 6.2 Ex post Pareto

Policy  $P_x$  is *ex post* Pareto superior to  $P_y$  iff for every person the actual (rather than expected) welfare under  $P_x$  is at least as good as the actual welfare under  $P_y$ , and for at least one person the actual welfare under  $P_x$  is better than the actual welfare under  $P_y$ . We do not need to worry about rival sets of designators when we are dealing with *ex post* Pareto superiority, because a person's *actual* (as opposed to expected) welfare does not depend on how (s)he is designated: if Alice is Ms Smith, then obviously whatever welfare Alice ends up with, Ms Smith will end up with the very same.

The relation of *ex ante* Pareto superiority<sub>sup</sub> neither guarantees nor is guaranteed by a relation of *ex post* Pareto superiority. That is:

<sup>12</sup> This follows provided that there is at least one set of admissible designators.

**Table 4** Prospects for Ethan

	State $S_1$ : HEADS $P(S_1) = 0.5$	State $S_2$ : TAILS $P(S_2) = 0.5$	Prospects
Risk	Ethan: 7	Ethan: 0	Ethan: 0.35
Safe	Ethan: 4	Ethan: 4	Ethan: 4

- (iii) From the claim that a policy  $P_x$  is *ex post* Pareto superior to  $P_y$ , it does not follow that policy  $P_x$  is *ex ante* Pareto superior<sub>sup</sub> to  $P_y$ .
- (iv) From the claim that a policy  $P_x$  is *ex ante* Pareto superior<sub>sup</sub> to  $P_y$ , it does not follow that policy  $P_x$  is *ex post* Pareto superior  $P_y$ .

Here is an illustration with a population of just one person, Ethan (Table 4). In this example, we have to decide between taking a risk (giving Ethan 7 units of welfare if HEADS and nothing otherwise), or playing it safe (giving Ethan 4 units of welfare for sure). The prospects for Ethan are better under *Safe* than under *Risk*, and Ethan is the only person involved, so *Safe* is *ex ante* Pareto superior to *Risk*, relative to the set of designators {‘Ethan’}. Because Ethan is the only member of the population, there are no other rival sets of designators which distribute the total prospects differently, and so *Safe* is *ex ante* Pareto superior<sub>sup</sub> to *Risk*. Now assume that the coin has already been tossed (out of sight of the decision-maker) and has landed heads. Then in fact choosing *Risk* would give Ethan 7, whereas choosing *Safe* would give him 4, so *Risk* is *ex post* Pareto superior to *Safe*. This illustrates claims (3) and (4): a policy can be *ex ante* Pareto superior<sub>sup</sub> without being *ex post* Pareto superior, and vice versa.

## 7 Identically situated individuals

Some choices—such as the case of Ethan above—are known to concern only one individual. In these cases, though the decision-maker might have a variety of different ways of designating that individual, the distribution of total prospects will obviously be the same whichever designator is considered. Thus for cases where a policy choice is known to concern just one individual, a policy  $P_x$  is *ex ante* Pareto superior<sub>sup</sub> to a policy  $P_y$  iff  $P_x$  is *ex ante* Pareto superior to  $P_y$  relative to any admissible set of designators—in other words, iff  $P_x$  is *ex ante* Pareto superior<sub>sub</sub> to  $P_y$ . The same holds where a policy choice concerns any number of identically situated individuals. Whatever designators are applied to these individuals, their welfare at each state, and so the calculation of their prospects will be identical. This gives us the following result:

- (v) Where all individuals are identically situated,  $P_x$  is *ex ante* Pareto superior<sub>sup</sub> to  $P_y$  iff  $P_x$  is *ex ante* Pareto superior<sub>sub</sub> to  $P_y$ .

Throughout the literature on the *ex ante* Pareto principle, various results have been proved, and while we cannot now straightforwardly endorse those proofs (for the *ex*

**Table 5** Prospects for Harry and Ian

	State $S_1$ $P(S_1) = 0.5$	State $S_2$ $P(S_2) = 0.5$	Prospects
Risk	Harry: 9	Harry: 0	Harry: 4.5
	Ian: 9	Ian: 0	Ian: 4.5
Safe	Harry: 4	Harry: 4	Harry: 4
	Ian: 4	Ian: 4	Ian: 4

*ante* principle has until now been incompletely defined and so not fit to feature in these proofs), we can see that various analogous such proofs will go through on a supervaluationist reading of the *ex ante* Pareto principle. In particular, proofs that concern cases of identically situated individuals will go through. Here is one such example.

Prioritarians place greater priority on improving the welfare of an individual the worse off that individual is. There are various different sorts of prioritarians, but one sort—the ‘continuous prioritarian’—captures the moral value of an agent’s welfare at an outcome by applying a transformation function such as the square root function on his or her welfare to give that individual’s ‘transformed well-being’. The rationale for this is that it is morally more important to give a fixed-sized well-being gain to a poorly-off individual than to give the same sized well-being gain to a better-off individual. We calculate the expected sum of *transformed* well-being (ESTWB) across individuals for a given policy, and choose whichever policy has the greatest ESTBW (Adler 2017).

This has been proven to result in violations of the *ex ante* Pareto principle (Otsuka and Voorhoeve 2009, 2018) even in cases concerning identically situated individuals. As I have argued, the *ex ante* Pareto principle is incompletely defined, and so we cannot endorse this proof, but we can endorse an analogous version involving the *ex ante* Pareto principle<sub>sup</sub>. To see this, consider the scenario above (Table 5).

We can see that *Risk* is *ex ante* Pareto superior to *Safe* relative to the set of designators {‘Harry’, ‘Ian’}, which proves that *Risk* is *ex ante* Pareto superior<sub>sub</sub> to *Safe*, and given that all individuals are identically situated, it follows that *Risk* is *ex ante* Pareto superior<sub>sup</sub> to *Safe*. To confirm this, consider the table below where the same policy choice is laid out relative to the designators {‘the F’, ‘the F\*’}, where the F is Harry if  $S_1$  obtains, and Ian otherwise, and the F\* is Ian if  $S_1$  obtains and Harry otherwise (Table 6).

Relative to all admissible sets of designators, *Risk* is *ex ante* Pareto superior to *Safe*, and so it is *ex ante* Pareto superior<sub>sup</sub>. However the ESTWB is greater under *Safe* than under *Risk*. Both Harry’s and Ian’s transformed wellbeing under *Risk* is  $\sqrt{9} = 3$  at  $S_1$ , and  $\sqrt{0} = 0$  at  $S_2$ , and under *Safe* it is  $\sqrt{4} = 2$  at both  $S_1$  and  $S_2$ . Thus the ESTWB under *Risk* is  $(0.5)(3) + (0.5)(3) + (0.5)(0) + (0.5)(0) = 3$ , and the ESTWB under *Safe* is  $(0.5)(2) + (0.5)(2) + (0.5)(2) + (0.5)(2) = 4$ . We get the same result here, of course, if we focus on the F and the F\* rather than Harry and

**Table 6** Prospects for the F and the F\*

	State $S_1$ $P(S_1) = 0.5$	State $S_2$ $P(S_2) = 0.5$	Prospects
Risk	The F: 9	The F: 0	The F: 4.5
	The F*: 9	The F*: 0	The F*: 4.5
Safe	The F: 4	The F: 4	The F: 4
	The F*: 4	The F*: 4	The F*: 4

Ian: in cases where all agents are identically situated, it doesn't matter which admissible set of designators we choose. Thus ESTWB is greater under *Safe* than under *Risk*, and so continuous-prioritarianism mandates *Safe* over *Risk*, which is in conflict with the *ex ante* Pareto principle<sub>sup</sub>. This gives us the result below, and a possible reason to reject continuous prioritarianism:

- (vi) Continuous prioritarianism is incompatible with the *ex ante* Pareto principle<sub>sup</sub>.

## 7.1 Heartland cases

Matt Adler and Nils Holtung define 'Heartland Cases' as follows:

Let's say that the comparison of two prospects, P and P\*, presents a 'heartland case' for the *ex ante* Pareto principles if the following holds true: (a) some number of individuals (meaning zero or more) are sure to be unaffected by the P/P\* choice and (b) all other individuals are equally situated (each such individual has the very same state-conditional well-being level as every other). (Adler and Holtung 2019, 115)

Theorists who reject the *ex ante* Pareto principle have a particularly hard job justifying their stance where heartland cases are concerned. For, we might think, when choosing between policies we need not consider those for whom the choice of policy will make no difference, and in heartland cases the individuals for whom the choice may make a different all face exactly the same outcome in each state, and so the situation is effectively like that in which we have a choice which concerns just one individual. And when we have a choice which just concerns one individual, it seems compelling that we should choose so as to maximise that person's prospects.<sup>13</sup>

<sup>13</sup> These are some of the arguments that Marc Fleurbaey puts forward in support of his principle of 'Weak Pareto for Subgroup Equal Risk' (Fleurbaey 2010, 665), which is closely connected to the *ex ante* Pareto principle restricted to heartland cases.

**Table 7** Prospects for Martha and Neil

	State $S_1$ : HEADS $P(S_1) = 0.5$	State $S_2$ : TAILS $P(S_2) = 0.5$	Prospects
Risk	Martha: 9	Martha: 0	Martha: 4.5
	Neil: 15	Neil: 15	Neil: 15
Safe	Martha: 4	Martha: 4	Martha: 4
	Neil: 15	Neil: 15	Neil: 15

**Table 8** Prospects for the H and the  $H^*$

	State $S_1$ $P(S_1) = 0.5$	State $S_2$ $P(S_2) = 0.5$	Prospects
Risk	The H: 9	The H: 15	The H: 12
	The $H^*$ : 15	The $H^*$ : 0	The $H^*$ : 7.5
Safe	The H: 4	The H: 15	The H: 9.5
	The $H^*$ : 15	The $H^*$ : 4	The $H^*$ : 9.5

Here is an illustration of a heartland case, where the *ex ante* Pareto principle seems to dictate a particular policy (Table 7). This case seems to meet the definition of a heartland case. Neil is unaffected by the choice between *Risk* and *Safe*, and all other people are identically situated, because the only other person is Martha. And it would appear that *Risk* is *ex ante* Pareto superior to *Safe*, because both Martha and Neil have prospects which are at least as good under *Risk* as under *Safe*, and Martha's prospects are better. However as we have seen, in a case like this with more than one person and some uncertainty there will be rival sets of designators, and so far we have considered only the set { 'Martha', 'Neil' }. Let us then consider another set, by coining the predicates H and  $H^*$ . To be H is to be Martha if state  $S_1$  obtains, and Neil if state  $S_2$  obtains, and to be  $H^*$  is to be Neil if state  $S_1$  obtains, and Martha if state  $S_2$  obtains. The table above gives the welfare at each state and prospects for the H and  $H^*$  (Table 8).

We can see that the prospects for the  $H^*$  are worse under *Risk* (7.5) than they are under *Safe* (9.5). Given that the  $H^*$  has worse prospects under *Risk* than under *Safe*, it follows that *Risk* is not *ex ante* Pareto superior relative to { 'the H', 'the  $H^*$ ' }, and so it is not *ex ante* Pareto superior<sub>sup</sub>. Thus this apparent heartland case is not a case of *ex ante* Pareto superiority<sub>sup</sub>, and so is obviously not a case where the *ex ante* Pareto principle<sub>sup</sub> is especially compelling.

In fact, although I skimmed over this at the start of this sub-section, the definition of a heartland case is itself incomplete. A case may be a heartland case relative to one admissible set of designators, but not relative to another: thus in the example above we had a heartland case relative to the designators { 'Martha', 'Neil' }, but not relative to the designators { 'the H', 'the  $H^*$ ' }. We might say then that a heartland case<sub>sup</sub> is a heartland case relative to all admissible sets of designators. Only a very special subset of cases will be heartland cases<sub>sup</sub>. Cases where all individuals are

identically situated are heartland cases<sub>sup</sub>; cases where all individuals are unaffected by the choice of policy are heartland cases<sub>sup</sub>, and cases where all individuals are either identically situated or unaffected by the choice of policy, and the decision-maker has no uncertainty whatsoever are also heartland cases<sub>sup</sub>. There is a small category of more interesting cases which also qualify as heartland cases<sub>sup</sub>, but in these cases the decision maker's uncertainty is very tightly—and entirely unrealistically—restricted.<sup>14</sup> Thus we have the following result:

- (vii) The only heartland cases<sub>sup</sub> that will occur in realistic cases are cases where all individuals are identically situated.

This has implications for various positions in welfare economics. Theorists have argued that continuous prioritaris, along with certain types of egalitarians (where egalitarians place value on equality amongst individuals) violate the *ex ante* Pareto principle in some heartland cases (Adler 2017, 121). Though egalitarians and continuous prioritaris can attempt to dismiss these concerns by denying the *ex ante* Pareto principle, this denial looks particularly counter-intuitive where heartland cases are concerned. As we can now see, these debates have been operating with incompletely defined concepts —of both *ex ante* Pareto superiority, and heartland cases. The situation needs to be reassessed now that we have the concepts of *ex ante* Pareto superiority<sub>sup</sub> and heartland cases<sub>sup</sub> to work with. As shown above, there are no realistic heartland cases<sub>sup</sub> of interest to decision theorists over and above those cases where all individuals are identically situated. Thus egalitarians and continuous prioritaris have no significant quarrel with the *ex ante* Pareto principle<sub>sup</sub> over heartland cases<sub>sup</sub>—no quarrel that is over and above the quarrel that prioritaris [though not egalitarians (Otsuka and Voorhoeve 2009, 2018)] have with the *ex ante* Pareto principle<sub>sup</sub> in cases where all individuals are identically situated.

<sup>14</sup> Here is an outline for such a case:

	State S <sub>1</sub>	State S <sub>2</sub>
	P(S <sub>1</sub> ) = p	P(S <sub>2</sub> ) = 1 - p
Risk	Martha: a	Martha: e
	Neil: b	Neil: f
Safe	Martha: c (where c ≠ a)	Martha: e
	Neil: b	Neil: f

This is an interesting case: if we assume that  $0 < p < 1$ , there is some uncertainty; not all individuals are identically situated (for if Martha and Neil were identically situated then it would follow that  $a = b$  and that  $c = b$  but  $a \neq c$ ); and not all individuals are unaffected by the choice of policy, for Martha is affected. Though I have not space to prove it here, it can be shown that if we assume that the decision's maker's only uncertainty is over whether S<sub>1</sub> or S<sub>2</sub> obtains, then this is a heartland case<sub>sup</sub>. However as soon as we introduce any additional uncertainty we can generate a set of designators relative to which this case is not a heartland case.

## 8 Conclusion

Until now, the *ex ante* Pareto principle has been incompletely defined, because of a failure to recognise that the prospects for an agent can depend on how that agent is designated. As I have shown, this problem is widespread, for in almost any policy choice situation there will be rival admissible sets of designators. I have offered two ways of reading the principle: a subvaluationist reading, and a supervaluationist reading. On the subvaluationist reading, the principle collapses into a version of utilitarianism: this reading of the principle is uncharitable when other readings are available. I focused instead on the supervaluationist reading, which is faithful to the rationale underlying the principle.

We cannot really compare the *ex ante* Pareto principle under the supervaluationist reading to the *ex ante* Pareto principle as it was generally understood before, because as it was generally understood before it was incompletely defined. Let us suppose, however, that on the old reading, the idea was that the *ex ante* Pareto principle could be applied in any situation where some particular set of designators seems like the obvious one to focus on, and then on the old reading  $P_x$  counts as *ex ante* Pareto superior to  $P_y$  provided that  $P_x$  is Pareto superior to  $P_y$  relative to this obvious set of designators. It then follows that it is harder (or at least, no easier) for a policy  $P_x$  to be *ex ante* Pareto superior to  $P_y$  on the supervaluationist reading, than it is for a policy to be *ex ante* Pareto superior to  $P_y$  on the old reading—for of course if  $P_x$  is *ex ante* Pareto superior relative to all admissible sets of designators, then it is *ex ante* Pareto superior relative to the obvious set. It follows that the *ex ante* Pareto principle is weaker under the supervaluationist reading than it is under the old reading. Nevertheless under the supervaluationist reading the principle is not entirely without teeth, and moreover it is both coherent and intuitively compelling.

The implications of the *ex ante* Pareto principle<sub>sup</sub> require much more investigation, and here I just began this investigation by exploring some of the implications for the principle for various versions of prioritarianism and egalitarianism. Furthermore, the points made in this paper can be generalised. The problem of giving a complete definition of the *ex ante* Pareto principle has analogues for many other principles that make use of the idea of expected welfare—including for example fairness principles that favour distributions that give people equal expected welfare, all else being equal (Diamond 1967), or that rate outcomes higher when they were chosen by fair selection processes (Broome 1984). The problem also arises for principles that make use of the idea of a hybrid of expected and final welfare (Voorhoeve and Fleurbaey 2016), and for competing claims models and variants on these, whenever *ex ante* claims are part of the model [for discussion on these models, see Frick (2015), Horton (2017)]. Here I have focused on the particular example of the *ex ante* Pareto principle, but the implications for a wide range of principles are far-reaching.

**Acknowledgements** I'm very grateful for all the comments I've received on this paper, in particular from Matt Adler, Nicolas Cote, Alan Hajek, Katie Steele, Bastian Steuwer, Alex Voorhoeve, and from audiences at ANU, LSE, MCMP and the University of Reading.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Adler, M. (2017). *Measuring social welfare: An introduction*. Manuscript.
- Adler, M., & Holtung, N. (2019). Prioritarianism: A response to critics. *Politics, Philosophy and Economics*, 18(2), 101–144.
- Barry, B. (1989). *Theories of Justice*. Hemel Hempstead: Harvester-Wheatsheaf.
- Braun, D. (2016). The objects of belief and credence. *Mind*, 125(498), 469–497.
- Broome, J. (1984). Uncertainty and fairness. *The Economic Journal*, 94(375), 624–632.
- Diamond, P. A. (1967). Cardinal welfare, individualistic ethics, and interpersonal comparison of utility. *Journal of Political Economy*, 75(5), 756–766.
- Evans, G. (1982). *The varieties of reference*. Oxford: OUP.
- Fleurbaey, M. (2010). Risky social situations. *Journal of Political Economy*, 118(4), 649–680.
- Frege, G. (1980). On sense and reference. In P. Geach & M. Black (Eds.), *Translations from the philosophical writings of Gottlob Frege* (pp. 25–50). Oxford: Blackwell.
- Frick, J. (2015). Contractualism and social risk. *Philosophy & Public Affairs*, 43(3), 175–223.
- Harsanyi, J. (1977). *Rational behaviour and bargaining equilibrium in games and social situations*. Cambridge: CUP.
- Horton, J. (2017). Aggregation, complaints, and risk. *Philosophy & Public Affairs*, 45, 54–81.
- Kripke, S. (1980). *Naming and necessity*. Cambridge, MA: Harvard University Press.
- Mahtani. (2017). The ex ante pareto principle. *Journal of Philosophy*, 114(6), 303–323.
- McCarthy, D. (2015). Distributive equality. *Mind*, 124(496), 1045–1109.
- Ninan, D. (2018). Quantification and epistemic modality. *The Philosophical Review*, 127(4), 433–485.
- Otsuka, M., & Voorhoeve, A. (2009). Why it matters that some are worse off than others: An argument against the priority. *Philosophy & Public Affairs*, 37(2), 171–199.
- Otsuka, M., & Voorhoeve, A. (2018). Equality versus priority. In S. Olsaretti (Ed.), *Oxford handbook of distributive justice* (pp. 65–85). Oxford: OUP.
- Russell, B. (1912). Knowledge by acquaintance and knowledge by description. *Proceedings of the Aristotelian Society*, 11, 108–128.
- Salmon, N. (1989). Illogical belief. *Philosophical Perspectives*, 3, 243–285.
- Searle, J. R. (1983). Proper names and intentionality. *Pacific Philosophical Quarterly*, 63(3), 205–225.
- Voorhoeve, A., & Fleurbaey, M. (2016). Priority or equality for possible people? *Ethics*, 126, 929–954.
- Yalcin, S. (2015). Epistemic modality de re. *Ergo*, 2, 475–527.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.