A Risky Problem for Permutation Invariance, Improved

Jacob M. Nebel

November 14, 2026

Abstract

This note identifies a tension between the following principles in risky infinite-population ethics. First, betterness is invariant to permutations on the set of individuals. Second, a prospect that gives each person the same, sufficiently high probability of being happy, otherwise sad, is better than a prospect that gives each person that same probability of being sad, otherwise happy. Third, a prospect is better than another only if some outcome of the one is better than some outcome of the other. This result unifies and improves upon two impossibility results due to Nebel (2026a): it entails one of them, and relies on more plausible assumptions than the other.

1 Puzzle

Consider an infinite population of people, made up of nine subpopulations of the same cardinality. The subpopulations are laid out in a three-by-three grid, indexed by row a and column b. There are three equiprobable events. The fate of each subpopulation in each event depends on our choice. Prospect f ensures that the individuals in row a are sad in event a; otherwise they are happy. Prospect h ensures that the individuals in column b are happy in event b; otherwise they are sad. The case is depicted in Table 1.

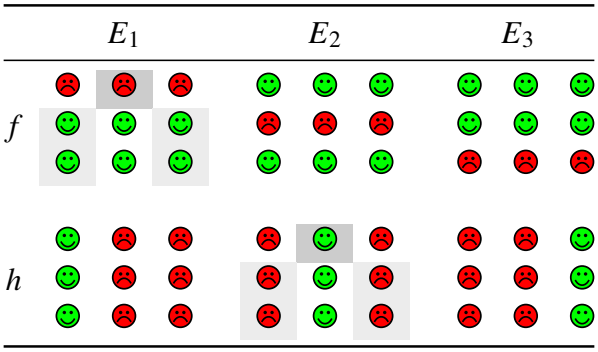


Table 1: f makes the individuals in row a sad in event a, otherwise happy; h makes the individuals in column b happy in event b, otherwise sad. The shaded cells indicate a swap relating f's outcome under E_1 to h's outcome under E_2.

We can argue as follows. In f, each person has a two-thirds chance of being happy, otherwise sad. In h, each person has a one-third chance of being happy, otherwise sad. Intuitively, f must be better than h. (The particular choice of probabilities here is inessential; we just need it to be better to give each person an arbitrarily high chance of being happy than to instead give them that same chance of being sad.)

On the other hand, each outcome of f is related to each outcome of h by simply swapping which of two disjoint infinite sets of people are happy, leaving others unchanged. (Take the intersection of the sad row in f and the happy column in h; this is a single infinite subpopulation, which gets swapped with the union of the subpopulations outside that row and column. An example is indicated by the shaded cells in f's outcome under E_1 and h's outcome under E_2.) It is hard to see how this could possibly make things better or worse. But if no outcome of f is any better than any outcome of h, it is hard to see how f could be better than h.

The puzzle is formulated more precisely in section 2. It unifies and improves upon two related results in Nebel (2026a); I explain how in section 3.

2 Principles and Theorem

We have an infinite set $\mathcal{I}$ of individuals. There are at least two welfare levels, happy (☺) and sad (☹). A welfare distribution assigns a welfare level to each individual. For any distribution w and permutation $\pi : \mathcal{I} \to \mathcal{I}$, πw is the distribution that assigns welfare level $w_{\pi(i)}$ to individual i.

There is a set $\mathcal{S}$ of states, which for every natural number n can be partitioned into n equiprobable events, relative to a probability measure P. A prospect assigns a welfare distribution to each state; we restrict our attention to prospects that have only finitely many possible outcomes. For any prospect f, we let $f_i(s)$ denote the welfare level assigned to individual i by the welfare distribution that f yields in state s.

Let $\succ$ denote the overall betterness relation over prospects. This relation is assumed to be asymmetric. We do not assume that it is transitive. (The corresponding relation over welfare distributions can be understood as the restriction of $\succ$ to constant prospects.)

The first argument in section 1, for f being better than h, appeals to the following Pareto principle:

Minimal Stochastic Pareto For some probability $q \in (1/2, 1)$, any prospects f and h, and any $p \in [q, 1)$, if

1. $P(\{ s \in \mathcal{S} \mid f_i(s) = ☺ \}) = p$ and $P(\{ s \in \mathcal{S} \mid f_i(s) = ☹ \}) = 1 - p$, and
2. $P(\{ s \in \mathcal{S} \mid h_i(s) = ☹ \}) = p$ and $P(\{ s \in \mathcal{S} \mid h_i(s) = ☺ \}) = 1 - p$,

for every individual i, then $f \succ h$.

This principle says that a prospect that gives everyone a sufficiently high (q-or-greater) chance of being happy, otherwise sad, is better than one that gives everyone that same chance of being sad, otherwise happy. (We require q to be greater than one-half to ensure that the first prospect stochastically dominates the second for every individual.)

The second argument in section 1, against f being better than h, appeals to two more premises:

Permutation Invariance For any distributions w and v, and any permutation $\pi : \mathcal{I} \to \mathcal{I}$, $w \succ v$ if and only if $\pi w \succ \pi v$.

Negative Dominance For any prospects f and h, if $f \succ h$ then there exist states $s, s' \in \mathcal{S}$ such that $f(s) \succ h(s')$.

Permutation Invariance says that betterness comparisons are preserved by permutations of individuals, including infinitary ones. Negative Dominance (due to Lederman, forthcoming) says that a prospect can be better than another only if some outcome of the one is better than some outcome of the other.

Our result is that these principles are together inconsistent:[1]

Theorem 1. *Minimal Stochastic Pareto, Negative Dominance, and Permutation Invariance are inconsistent.*

Proof. Pick a natural number n so that $\frac{n-1}{n} \geq q$, where q satisfies Minimal Stochastic Pareto. We partition the set of states $\mathcal{S}$ into n equiprobable events $E_1, \ldots, E_n$ and our population $\mathcal{I}$ into n^2 infinite subpopulations of the same cardinality, indexed by ordered pairs $a, b \in \{1, \ldots, n\}$—i.e., $I_{1,1}, \ldots, I_{1,n}, \ldots, I_{n,1}, \ldots, I_{n,n}$. Define prospects f and h as follows: for each $k = 1, \ldots, n$ and state $s \in E_k$, let $f_i(s) =$ ☹ for every individual $i \in I_{k,b}$ ($b = 1, \ldots, n$), $f_i(s) =$ ☺ otherwise; and let $h_i(s) =$ ☺ for every $i \in I_{a,k}$ ($a = 1, \ldots, n$), $h_i(s) =$ ☹ otherwise. Thus, each individual $i \in \mathcal{I}$ is happy in $n - 1$ out of n events in f, otherwise sad; and she is happy in 1 out of n events in h, otherwise sad. So f is better than h, by Minimal Stochastic Pareto.

Notice, however, that Permutation Invariance rules out any strict comparison between $f(s)$ and $h(s')$ for all states $s, s' \in \mathcal{S}$. To see this, take any $j, k \in \{1, \ldots, n\}$ and consider the events E_j and E_k. For each state $s \in E_j$ and $s' \in E_k$, exactly one infinite subpopulation ($I_{j,k}$) is sad in $f(s)$ and happy in $h(s')$, while $(n-1)^2$ subpopulations ($I_{a,b}$ for $a \neq j$, $b \neq k$) are happy in $f(s)$ and sad in $h(s')$. Thus there is a permutation $\pi : \mathcal{I} \to \mathcal{I}$ such that $f_i(s) = h_{\pi(i)}(s')$ and $f_{\pi(i)}(s) = h_i(s')$ for every individual i. So, by Permutation Invariance, $f(s) \succ h(s')$ if and only if $h(s') \succ f(s)$. Therefore, $f(s) \not\succ h(s')$ for all states s, s', and so $f \not\succ h$ by Negative Dominance, contrary to what we showed above. □

1. The construction used in the proof was proposed by Claude (Anthropic, Opus 4.8).

The argument of section 1, concerning the prospects in Table 1, illustrates the proof on the assumption that $q = 2/3$ is a sufficiently high probability.

3 Comparison

Theorem 1 unifies and improves upon two similar impossibility results in Nebel (2026a).

The first (Proposition 1 in Nebel 2026a) uses a *Statewise* Negative Dominance principle. This principle says that a prospect can be better than another only if it yields a better outcome in some state of the world. The puzzle is roughly that, in Table 2 below, f seems better than h, because it gives each person the same, higher chance of being happy. And yet it is not better, according to Permutation Invariance, in any state of the world. As explained in my earlier paper, however, many philosophers are willing to reject Statewise Negative Dominance for reasons having nothing to do with infinite populations (Hare 2010; Bader 2018; Rabinowicz 2021; Francis and Gustafsson 2025). But since Negative Dominance is strictly weaker than Statewise Negative Dominance, this result is a corollary of Theorem 1 above.

	E_1			E_2			E_3		
	I_1	I_2	I_3	I_1	I_2	I_3	I_1	I_2	I_3
f	☹	☺	☺	☺	☹	☺	☺	☺	☹
g	😐	😐	😐	😐	😐	😐	😐	😐	😐
h	☺	☹	☹	☹	☺	☹	☹	☹	☺

Table 2: Puzzle from Nebel 2026a

The second (Theorem 2 in Nebel 2026a) uses Negative Dominance. But it appeals to additional assumptions and comparisons that obscure the central tension. First, it assumes the transitivity of *better than*. Second, it requires there to be some intermediate welfare level, in between happy and sad, such that giving everyone a sufficiently high chance of being happy (otherwise sad) is better than giving everyone the middle level for sure, which is in turn better than giving everyone a correspondingly high chance of being sad (otherwise happy).

With these assumptions in place, the argument goes roughly like this. In Table 2, f is better than g, because it gives everyone a uniformly high chance of being happy, as opposed to the middle level for sure. Some outcome of f must be better than the sole outcome of g, given Negative Dominance. By Permutation Invariance, every outcome of f must then be better than the outcome of g. By similar reasoning, the sole outcome of g must be better than every outcome of h. Thus, by transitivity, every outcome of f must be better than every outcome of h. But this violates Permutation Invariance.

The additional assumptions of this argument are plausible, but not unquestionable. Rachels (1998) and Temkin (2012) influentially question the transitivity of *better than* (see Nebel 2018 for a response). But although my original proof invokes transitivity, it is in fact inessential to the result.[2] Since each outcome of f is better than the sole outcome of g, we can apply a permutation to those outcomes to transform an f-outcome into an h-outcome, leaving g's outcome fixed; we can then infer, by Permutation Invariance, that each outcome of h is also better than the sole outcome of g. Thus, the outcome of g cannot be better than any outcome of h. But since g is supposed to be better than h, we have a violation of Negative Dominance.

Alternatively, one could deny the requisite comparisons to the middle prospect g by appealing to a highly permissive view about the value of risk. On such a view, it might not be irrational or unreasonable for someone to prefer a certainly middling outcome to a prospect that gives her even an arbitrarily high chance of a good outcome and an arbitrarily low chance of a bad outcome, reducing the pressure to judge that f is better than g. Tarsney (2020), for example, suggests an approach to rational choice on which stochastic dominance is the only substantive requirement; Buchak and Nebel (forthcoming) mention a "super-permissive" conception of individual betterness along these lines, though they do not endorse it. The present puzzle, however, shows that we do not need any such intermediate prospect, or any comparisons that go beyond Minimal Stochastic Pareto. (Nonetheless, my earlier Theorem 2 is not a corollary of the present result.)

A further advantage of the present result is that, because all of the outcomes involved are related by permutations on the set of individuals, it makes certain *ex post* considerations irrelevant to the puzzle. As observed in my earlier paper, one might deny that f is better than g because f guarantees an unequal outcome, whereas g guarantees perfect equality. Alternatively, one might deny that g is better than h because one wants to ensure that there are some (perhaps infinitely many) people who have happy lives. Neither response is germane to the present puzzle.

In this way, I think the present argument most cleanly illustrates that there is a real and very general tension between Permutation Invariance, negative dominance reasoning, and *ex ante* Pareto principles. Though the construction is more complex, the result itself is simpler in terms of the axioms required, and avoids any contentious assumptions about risk or inequality. It shows that the tension cannot be solved by appealing to a highly permissive view about risk or to distributive features of outcomes that might justify violations of *ex ante* Pareto. Nebel (2026a, 2026b) recommends that we reject Permutation Invariance, but I leave that question open here.

2. This was pointed out by Refine (refine.ink), an AI reviewing tool, in a report on an earlier draft of this note.

References

Bader, Ralf M. 2018. "Stochastic Dominance and Opaque Sweetening." *Australasian Journal of Philosophy* 96 (3): 498–507. https://doi.org/10.1080/00048402.2017.1362566.

Buchak, Lara, and Jacob M. Nebel. Forthcoming. "The Value of Risky Prospects." *The Journal of Philosophy.*

Francis, Tomi, and Johan E Gustafsson. 2025. "Stochastic Dominance for Incomplete Preferences." *Analysis,* anaf046. https://doi.org/10.1093/analys/anaf046.

Hare, Caspar. 2010. "Take the Sugar." *Analysis* 70 (2): 237–247. https://doi.org/10.1093/analys/anp174.

Lederman, Harvey. Forthcoming. "Of Marbles and Matchsticks." In *Oxford Studies in Epistemology.*

Nebel, Jacob M. 2018. "The Good, the Bad, and the Transitivity of Better Than." *Noûs* 52 (4): 874–899. https://doi.org/10.1111/nous.12198.

———. 2026a. "A Risky Problem for Permutation Invariance." *Analysis,* anag012. https://doi.org/10.1093/analys/anag012.

———. 2026b. "Infinite Ethics and the Limits of Impartiality." *Noûs* 60 (2): 433–453. https://doi.org/10.1111/nous.70010.

Rabinowicz, Wlodek. 2021. "Incommensurability Meets Risk." In *Value Incommensurability: Ethics, Risk, and Decision-Making,* edited by Henrik Andersson and Anders Herlitz, 201–230. Routledge.

Rachels, Stuart. 1998. "Counterexamples to the Transitivity of *Better Than.*" *Australasian Journal of Philosophy* 76 (1): 71–83. https://doi.org/10.1080/00048409812348201.

Tarsney, Christian. 2020. "Exceeding Expectations: Stochastic Dominance as a General Decision Theory." GPI Working Paper No. 3-2020, Global Priorities Institute, Oxford.

Temkin, Larry S. 2012. *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning.* Oxford: Oxford University Press.