

# Infinite Worlds and the Two Envelope Paradox

Brian Weatherson

2025-10-14

There have been a lot of discussions recently about how versions of the St Petersburg paradox affect theories about welfare in infinite worlds. This note is about how the two envelope paradox affects things. The tentative suggestion is that it gives us a reason to give up on an otherwise very plausible looking principle of state-level strict dominance.

There have been a flurry of interesting recent papers on welfare aggregation principles in infinite worlds. A small sample includes:

- Hayden Wilkinson, *Infinite Aggregation and Risk* (Wilkinson 2023)
- Frank Hong and Jeffrey Sanford Russell, *Paradoxes of Infinite Aggregation* (Hong and Russell 2025)
- Jake Nebel, *Infinite Ethics and the Limits of Impartiality* (Nebel 2025)
- Jeremy Goodman, *Permutation-Invariant Social Welfare Orders Are Anonymous* (Goodman 2025)
- Jeremy Goodman and Harvey Lederman, *Maximal Social Welfare Relations on Infinite Populations Satisfying Permutation Invariance* (Goodman and Lederman 2024)

I'm going to be drawing on all of these here, but especially the last one.<sup>1</sup>

Hong and Russell (2025) use variants of the St Petersburg paradox to show that the following four principles are inconsistent. Notably, they do so even given the extra assumption that welfare levels for individuals can only take one of two possible values.<sup>2</sup> I'll adopt that assumption too.<sup>3</sup>

**(Weak) Ex Ante Pareto** If  $X$  is at least as good as  $Y$  for every individual, then  $X$  is at least as good as  $Y$  overall.

---

<sup>1</sup>And it should be noted that a lot of these works are downstream of Amanda Askill's PhD Thesis *Pareto Principles in Infinite Ethics* (Askill 2018). The positive point of this note is to provide another argument for using what Goodman and Lederman call the *Sum Preorder* for comparing infinite worlds. But I'll leave most of that argument implicit; hopefully I'll return to it in a later post.

<sup>2</sup>It's much easier to derive contradictions if we allow that individuals can have unbounded utility.

<sup>3</sup>The following principles are mostly direct quotes from Hong and Russell's paper, but I've expanded the notation for **Stochastic Compensation** for clarity.

**Stochastic Compensation** For any good  $x$  for an individual  $i$ , and any events  $E$  and  $F$  such that  $\Pr(E) > \Pr(F)$ , there is some good  $y$  for  $i$  such that the gamble  $y$  if  $F$ , 0 otherwise is at least as good for  $i$  than the gamble  $x$  if  $E$ , 0 otherwise.

**Interpersonal Compensation** For any allocation  $x$  for a finite set of individuals  $I$ , there is an allocation  $y$  for some finite set of individuals  $J$  disjoint from  $I$  such that  $y$  is better overall than  $x$ .

**(Strict) Statewise Dominance** If the outcome that results from a prospect  $X$  is strictly better than the outcome that results from a prospect  $Y$  in every possible state, then  $X$  is strictly better than  $Y$ .

The first point I'll make is that slight generalisations of the last two on their own suffice for a contradiction. That makes me rather sceptical that the first two principles are the problems. Here are the two principles that I need.

**Risk Spreading** For some integer  $m$ , and some probability  $p < 1/2$ , the following holds. It is better that  $m^n$  people get 1 for certain, and  $m^{n+2} - m^n$  people have chance  $p$  of 1 (and chance  $1-p$  of 0), than that  $m^{n+1}$  people get 1 and everyone else gets 0.

**(Strict) Partition Dominance** If a prospect  $X$  is strictly better than a prospect  $Y$  conditional on each member of a countable partition, then  $X$  is strictly better than  $Y$ .

I take **Risk Spreading** to be a probabilistic form of **Interpersonal Compensation**. It's better that a much larger group of people have a good chance (nearly  $1/2$ ) of the good, than that a much smaller group have the good for sure, and everyone misses out. This would hold given any plausible kind of Additivity principle, and it's if anything a fairer distribution of goods.

And I take **(Strict) Partition Dominance** to be motivated by the same things that motivate **(Strict) Statewise Dominance**. When we talk about states in this bit of philosophy/economics, that's always shorthand. We never really know in real life what the full outcome of any action will be. We can save the child from the pond, but they might grow up to be a mass murderer. The typical, and I think correct, practice is to simply bracket off those considerations. The relevant 'states' are that the child lives, or the child drowns. In reality, each of those states is another lottery; it's just that the first is a much much better one. Any principle that reasons from properties of states to properties of prospects is really reasoning from properties of more fine-grained prospects to properties of less fine-grained prospects. From that perspective, **(Strict) Partition Dominance** is really just a reformulation of **(Strict) Statewise Dominance**.

Given this, that these are inconsistent is quite simple, and just follows the familiar pattern from the literature on the Two Envelope Paradox.<sup>4</sup> There are two coins, a **green** coin and a fair **brown** coin. Unlike the **brown** coin, the **green** coin is not fair; **it** has chance  $p$  of landing **heads**. The **green** coin will be flipped repeatedly until **it** lands **heads**. Let  $n$  be the number of times **it** lands **tails**; unless **it** never lands **tails**, in which case  $n = 0$ . Then the **brown** coin is flipped once. Assume (as is normal in these problems) that we have an infinite population,  $p_1$ ,

---

<sup>4</sup>For a primer on that, see Jackson, Menzies, and Oppy (1994).

$p_2, \dots$ , with the numbers being arbitrary. And then assume we have two prospects, which we'll call **Blue** and **Red**. (Assume  $m$  is the same large number that witnesses **Risk Spreading**, and that  $m > 2$ .)

**Blue** If the **brown** coin lands **heads**, the first  $m^n$  people get 1, everyone else gets 0.

If the **brown** coin lands **tails**, the first  $m^{n+1}$  people get 1, everyone else gets 0.

**Red** If the **brown** coin lands **heads**, the first  $m^{n+1}$  people get 1, everyone else gets 0.

If the **brown** coin lands **tails**, the first  $m^n$  people get 1, everyone else gets 0.

Let  $B$  be a random variable equalling the number of people who will get 1 if **Blue** is chosen, and  $R$  be a random variable equalling the number of people who will get 1 if **Red** is chosen. Just look at the possible values for  $B$ ; what happens with  $R$  will be entirely parallel. If  $B = 1$ , then  $R = m$ , so **Red** is clearly better. If  $B = m^{n+1}$  for  $n \geq 0$ , then  $R = m^n$  with probability  $1/(2-)$ , and  $R = m^{n+2}$  with probability  $(1-)/(2-)$ . Provided  $m$  is small enough, that will be greater than the value  $p$  in **Risk Spreading**, so again **Red** will be better.

So conditional on any possible value of  $B$ , **Red** is better than **Blue**. So by **(Strict) Statewise Dominance**, **Red** is better than **Blue**. But an exactly parallel argument shows that conditional on any possible value of  $R$ , **Blue** is better than **Red**. So by **(Strict) Statewise Dominance**, **Blue** is better than **Red**. Contradiction.

We could probably tighten this up, but I think this shows that fairly weak dominance principles, plus fairly weak compensation principles, are inconsistent in infinite worlds. This makes me think that the guilty principle in Hong and Russell's tetralemma will be one of **(Strict) Statewise Dominance** and **Interpersonal Compensation**. I'm inclined to give up **(Strict) Statewise Dominance**, though it's possible that the right lesson to draw from the two envelope cases is that there is an argument in the style of Nebel (2025) against impartiality which does not rely on completeness.

Askill, Amanda. 2018. "Pareto Principles in Infinite Ethics." PhD thesis, New York University.

Goodman, Jeremy. 2025. "Permutation-Invariant Social Welfare Orders Are Anonymous." *Journal of Mathematical Economics* 120.

Goodman, Jeremy, and Harvey Lederman. 2024. "Maximal Social Welfare Relations on Infinite Populations Satisfying Permutation Invariance." <https://arxiv.org/abs/2408.05851v2>. arXiv preprint.

Hong, Frank, and Jeffrey Sanford Russell. 2025. "Paradoxes of Infinite Aggregation." *Nous* 59 (3): 809–27. doi: [10.1111/nous.12535](https://doi.org/10.1111/nous.12535).

Jackson, Frank, Peter Menzies, and Graham Oppy. 1994. "The Two Envelope 'Paradox'." *Analysis* 54 (1): 43–45. doi: [10.1093/analys/54.1.43](https://doi.org/10.1093/analys/54.1.43).

Nebel, Jacob M. 2025. "Infinite Ethics and the Limits of Impartiality." *Noûs*. 2025. doi: [10.1111/nous.70010](https://doi.org/10.1111/nous.70010).

Wilkinson, Hayden. 2023. "Infinite Aggregation and Risk." *Australasian Journal of Philosophy* 101 (2): 340–59. doi: [10.1080/00048402.2021.2013265](https://doi.org/10.1080/00048402.2021.2013265).