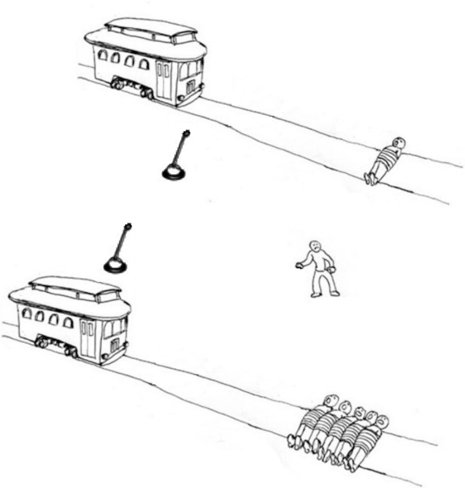
Doing good in an infinite, chaotic world

Hayden Wilkinson

Australian National University

EA Global 2019

Slides available at: haydenwilkinson.co.uk/slides

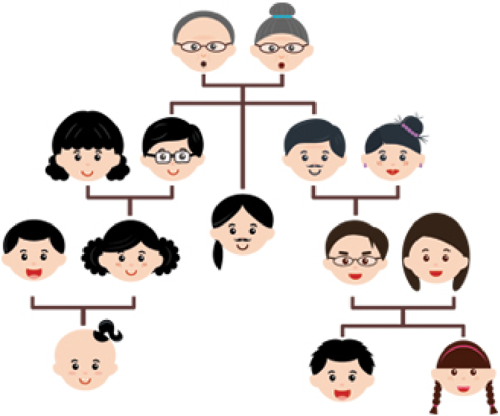


Chaos

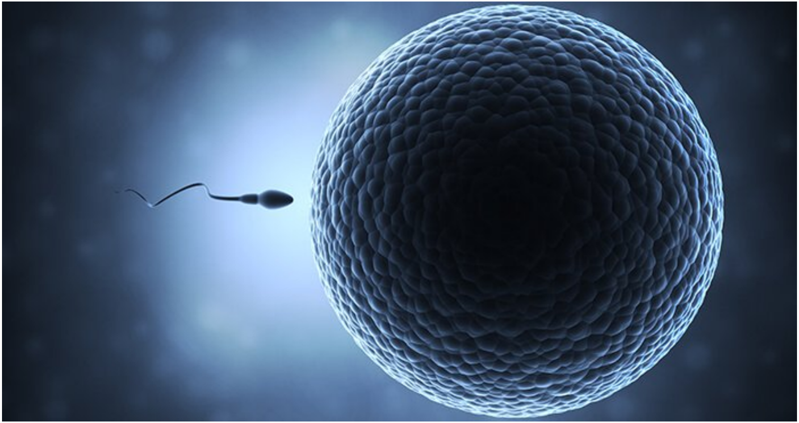
Chaos

- Identity effects

Identity effects



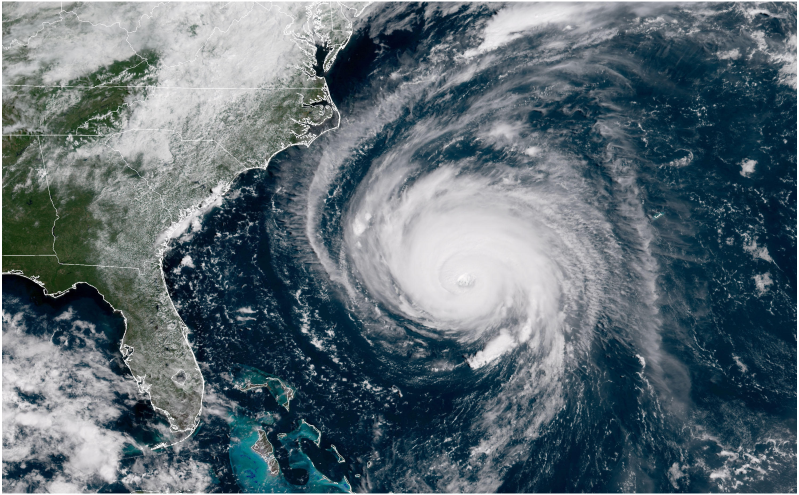
Identity effects



Chaos

- Identity effects
- Climatic effects

Climatic effects

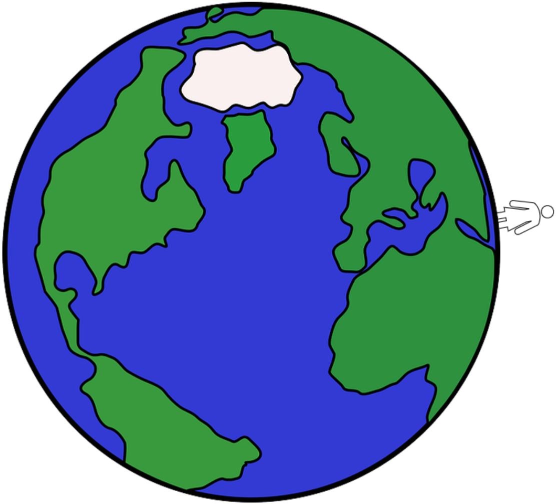


Climatic effects

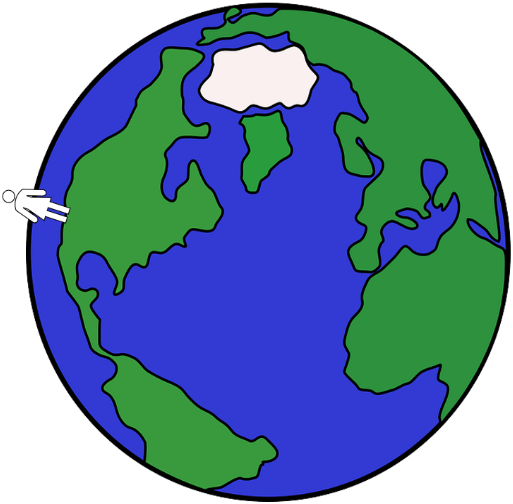
Chaos

- Identity effects
- Climatic effects
- Gravitational effects

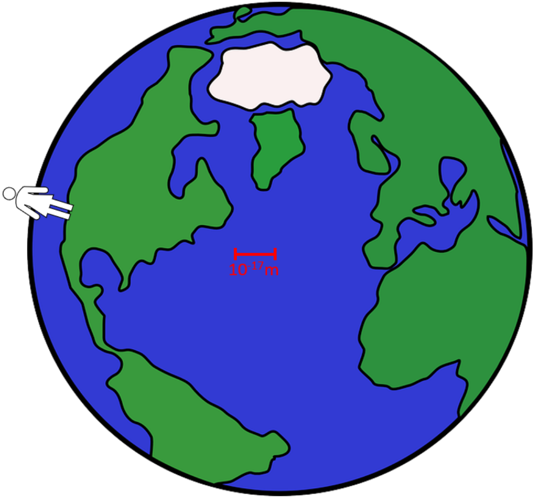
Gravitational effects



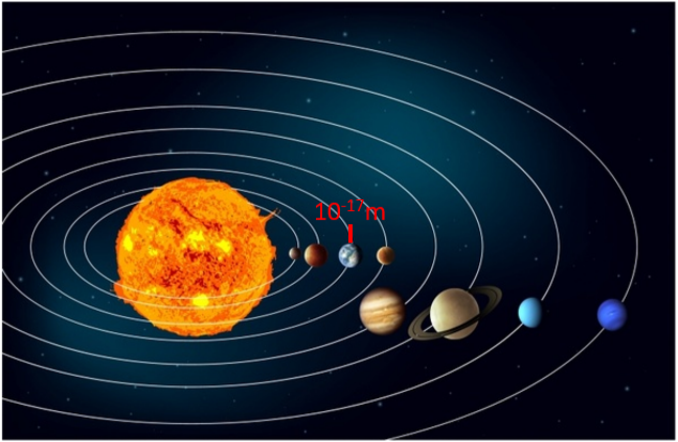
Gravitational effects



Gravitational effects

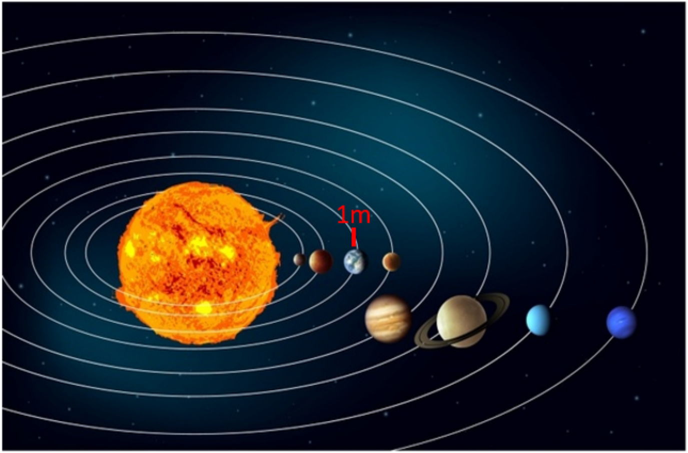


Gravitational effects



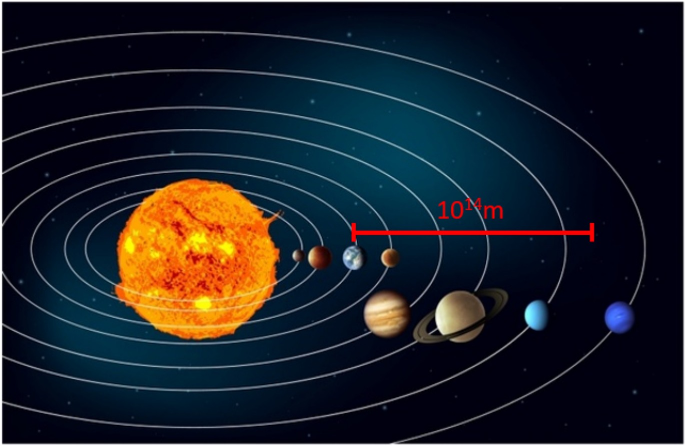
Present day

Gravitational effects



After 170 million years

Gravitational effects



After 310 million years

Chaos

- Identity effects
- Climatic effects
- Gravitational effects
- And others...

Cluelessness

Cluelessness

$$\begin{array}{rc} & t_1 \\ W_1 : & 1 \\ W_2 : & 5 \end{array}$$

Cluelessness

	t_1	t_2	t_3	t_4	t_5	t_6	...	t_N
W_1 :	1							
W_2 :	5							

Cluelessness

$$
\begin{array}{lcccccccc}
 & t_1 & t_2 & t_3 & t_4 & t_5 & t_6 & \ldots & t_N \\
W_1: & 1 & X_2 & X_3 & X_4 & X_5 & X_6 & \ldots & X_N \\
W_2: & 5 & Y_2 & Y_3 & Y_4 & Y_5 & Y_6 & \ldots & Y_N
\end{array}
$$

Cluelessness

	t_1	t_2	t_3	t_4	t_5	t_6	...	t_N
W_1 :	1	X_2	X_3	X_4	X_5	X_6	...	X_N
W_2 :	5	Y_2	Y_3	Y_4	Y_5	Y_6	...	Y_N

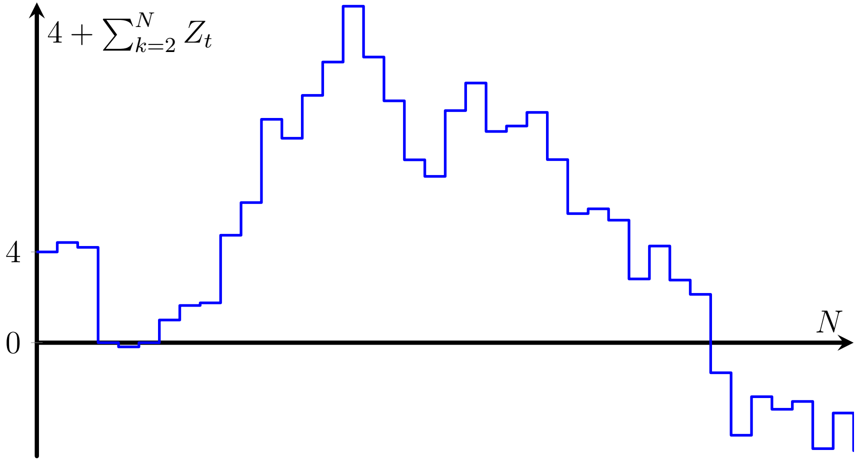
$$V(W_2) - V(W_1) = (5 - 1) + \sum_{k=2}^{N} (Y_k - X_k)$$

Cluelessness

$$\begin{array}{lcccccccc} & t_1 & t_2 & t_3 & t_4 & t_5 & t_6 & \ldots & t_N \\ W_1: & 1 & X_2 & X_3 & X_4 & X_5 & X_6 & \ldots & X_N \\ W_2: & 5 & Y_2 & Y_3 & Y_4 & Y_5 & Y_6 & \ldots & Y_N \end{array}$$

$$V(W_2) - V(W_1) = 4 + \sum_{k=2}^{N} Z_k$$

Cluelessness



Cluelessness

The cluelessness worry (for objective betterness) (Greaves 2016)

For any pair of distinct acts (A_1, A_2) ever available to us, we can never have even the faintest idea which will have the better outcome.

Cluelessness - subjective betterness

Cluelessness - subjective betterness

Total difference:

$$V(W_2) - V(W_1) = 4 + \sum_{k=2}^{N} Z_k$$

Cluelessness - subjective betterness

Total difference:

$$V(W_2) - V(W_1) = 4 + \sum_{k=2}^{N} Z_k$$

Expected total difference:

$$\begin{aligned} EV(W_2) - EV(W_1) &= 4 + EV(\sum_{k=2}^{N} Z_k) \\ &= 4 + 0 \end{aligned}$$

Infinite worlds

Infinite worlds

- Several leading cosmological theories predict an infinite future, containing infinitely many instances of every physically possible phenomenon (see Gott 2008; Carroll 2017; Rauer *et al.* 2018).

Infinite worlds

- Several leading cosmological theories predict an infinite future, containing infinitely many instances of every physically possible phenomenon (see Gott 2008; Carroll 2017; Rauer *et al.* 2018).
- We have a few ways of comparing infinite worlds.

Infinite worlds

- Several leading cosmological theories predict an infinite future, containing infinitely many instances of every physically possible phenomenon (see Gott 2008; Carroll 2017; Rauer *et al.* 2018).
- We have a few ways of comparing infinite worlds.
  - Strongly impartial views

Infinite worlds

- Several leading cosmological theories predict an infinite future, containing infinitely many instances of every physically possible phenomenon (see Gott 2008; Carroll 2017; Rauer *et al.* 2018).
- We have a few ways of comparing infinite worlds.
    - Strongly impartial views
    - Weakly impartial views

Infinite worlds

- Several leading cosmological theories predict an infinite future, containing infinitely many instances of every physically possible phenomenon (see Gott 2008; Carroll 2017; Rauer *et al.* 2018).
- We have a few ways of comparing infinite worlds.
  - Strongly impartial views
  - Weakly impartial views
  - Position-dependent views

Definition

- **Strong impartiality**: Rankings of outcomess are independent of *which* persons obtain value in each outcome, and independent of which times and places those persons are.

Definition

- **Strong impartiality**: Rankings of outcomess are independent of *which* persons obtain value in each outcome, and independent of which times and places those persons are.
  - Implies that, if W_1 and W_2 contain the same number of persons at each level of value, then $W_1 \simeq W_2$.

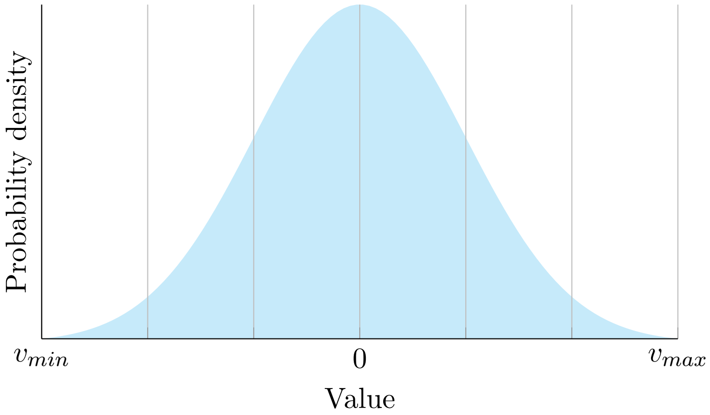
Definition

- **Strong impartiality**: Rankings of outcomess are independent of *which* persons obtain value in each outcome, and independent of which times and places those persons are.
  - Implies that, if W_1 and W_2 contain the same number of persons at each level of value, then $W_1 \simeq W_2$.
- e.g., Bader MS; Clark MS

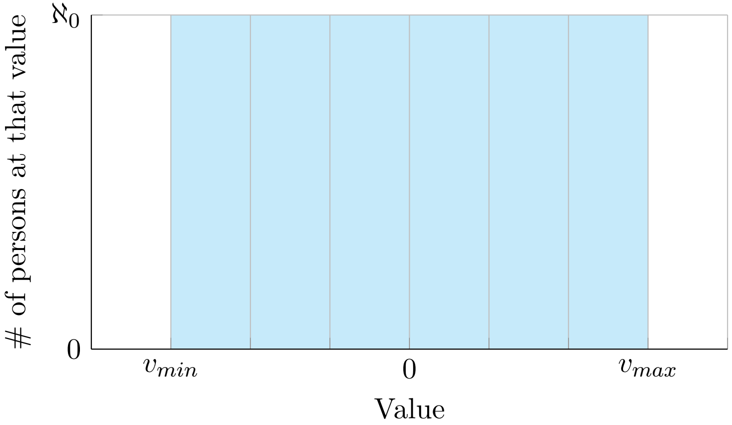
Chaotic outcomes

$$
\begin{array}{lccccccccccc}
 & p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & p_a & p_b & \cdots & p_\alpha & p_\beta & \cdots \\
W_1: & 1 & 0 & 0 & 0 & 0 & 0 & X_a & X_b & \ldots & - & - & \ldots \\
W_2: & 0 & 1 & 1 & 1 & 1 & 1 & - & - & \ldots & X_\alpha & X_\beta & \ldots
\end{array}
$$

Chaotic outcomes - probability distribution for some X_i



Chaotic outcomes



Chaotic outcomes

$$
\begin{array}{lccccccccccccc}
 & p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & p_a & p_b & \dots & p_\alpha & p_\beta & \dots \\
W_1: & 1 & 0 & 0 & 0 & 0 & 0 & X_a & X_b & \dots & - & - & \dots \\
W_2: & 0 & 1 & 1 & 1 & 1 & 1 & - & - & \dots & X_\alpha & X_\beta & \dots
\end{array}
$$

$W_1 \simeq W_2$ by any strongly impartial view.

Definition

- **Weak impartiality**: Rankings of outcomes are independent of which persons obtain value across each *outcome pair*, and independent of the times and places those persons are in each outcome.

Definition

- **Weak impartiality**: Rankings of outcomes are independent of which persons obtain value across each *outcome pair*, and independent of the times and places those persons are in each outcome.
  - Implies that, if the pairs of outcomes (W_1, W_2) and (W_3, W_4) contain the same number of persons obtaining values (a, b) in the respective outcomes, then $W_1 \succcurlyeq W_2$ iff $W_3 \succcurlyeq W_4$.

Definition

- **Weak impartiality**: Rankings of outcomes are independent of which persons obtain value across each *outcome pair*, and independent of the times and places those persons are in each outcome.
  - Implies that, if the pairs of outcomes (W_1, W_2) and (W_3, W_4) contain the same number of persons obtaining values (a, b) in the respective outcomes, then $W_1 \succcurlyeq W_2$ iff $W_3 \succcurlyeq W_4$.
- e.g., Vallentyne & Kagan 1997: 11; Lauwers & Vallentyne 2004; see also Askell 2018

Pareto

Pareto (over persons): *If outcomes W_1 and W_2 contain the same persons, and every person obtains as much value in W_1 as in W_2, then $W_1 \succcurlyeq W_2$. And if some p_i obtains strictly more value in W_1, then $W_1 \succ W_2$.*

Example

Take the simplified case of W_1 and W_2, where each $X_{p_i} \in \{0, 1\}$:

$$\begin{array}{lcccccc} & p_1 & p_2 & A & B & C & D \\ W_1 : & 1 & 0 & 1 & - & 0 & - \\ W_2 : & 0 & 1 & - & 1 & - & 0 \end{array}$$

Example

We can construct W_3 and W_4 as follows. (n.b., $B_1 \cup B_2 = B$, $C_1 \cup C_2 = C$)

	p_1	p_2	A	B_1	B_2	C_1	C_2	D
W_1 :	1	0	1	–	–	0	0	–
W_4 :	1	0	1	–	–	1	0	–

Example

We can construct W_3 and W_4 as follows. (n.b., $B_1 \cup B_2 = B$, $C_1 \cup C_2 = C$)

	p_1	p_2	A	B_1	B_2	C_1	C_2	D
W_1 :	1	0	1	–	–	0	0	–
W_4 :	1	0	1	–	–	1	0	–
W_2 :	0	1	–	1	1	–	–	0
W_3 :	0	1	–	0	1	–	–	0

Example

We can construct W_3 and W_4 as follows. (n.b., $B_1 \cup B_2 = B$, $C_1 \cup C_2 = C$)

	p_1	p_2	A	B_1	B_2	C_1	C_2	D
W_1 :	1	0	1	–	–	0	0	–
W_4 :	1	0	1	–	–	1	0	–
W_2 :	0	1	–	1	1	–	–	0
W_3 :	0	1	–	0	1	–	–	0

$W_4 \succ W_1$ and $W_2 \succ W_3$ by Pareto

Example

But the only difference between the pairs (W_1, W_2) and (W_3, W_4) is the identities of the persons in each pair.

	p_1	p_2	A	B	C	D
W_1 :	1	0	1	–	0	–
W_2 :	0	1	–	1	–	0

	p_2	p_1	B_2	$A \cup C_1$	$B_1 \cup D$	C_2
W_3 :	1	0	1	–	0	–
W_4 :	0	1	–	1	–	0

Example

But the only difference between the pairs (W_1, W_2) and (W_3, W_4) is the identities of the persons in each pair.

	p_1	p_2	A	B	C	D
W_1 :	1	0	1	–	0	–
W_2 :	0	1	–	1	–	0

	p_2	p_1	B_2	$A \cup C_1$	$B_1 \cup D$	C_2
W_3 :	1	0	1	–	0	–
W_4 :	0	1	–	1	–	0

So weakly impartial views must say $W_1 \succcurlyeq W_2$ iff $W_3 \succcurlyeq W_4$.

Example

So weakly impartial views must say $W_1 \succcurlyeq W_2$ iff $W_3 \succcurlyeq W_4$.

Example

So weakly impartial views must say $W_1 \succcurlyeq W_2$ iff $W_3 \succcurlyeq W_4$.

But Pareto says that $W_4 \succ W_1$ and $W_2 \succ W_3$.

Example

So weakly impartial views must say $W_1 \succcurlyeq W_2$ iff $W_3 \succcurlyeq W_4$.

But Pareto says that $W_4 \succ W_1$ and $W_2 \succ W_3$.

If $W_1 \succcurlyeq W_2$, we get a cycle: $W_1 \succcurlyeq W_2 \succ W_3 \succcurlyeq W_4 \succ W_1$.
(And, if $W_2 \succcurlyeq W_1$, we can construct W_5, W_6 for another cycle.)

Example

So weakly impartial views must say $W_1 \succcurlyeq W_2$ iff $W_3 \succcurlyeq W_4$.

But Pareto says that $W_4 \succ W_1$ and $W_2 \succ W_3$.

If $W_1 \succcurlyeq W_2$, we get a cycle: $W_1 \succcurlyeq W_2 \succ W_3 \succcurlyeq W_4 \succ W_1$.
(And, if $W_2 \succcurlyeq W_1$, we can construct W_5, W_6 for another cycle.)

$\therefore W_1$ and W_2 are incomparable, according to any weakly impartial, Paretian view.
(adapted from Askell 2018: ch.3)

- **Position-dependence**: Rankings of outcomes are (at least sometimes) dependent on the positions of value in time and space, even when the outcomes contain the same persons and each obtains the same value.

- **Position-dependence**: Rankings of outcomes are (at least sometimes) dependent on the positions of value in time and space, even when the outcomes contain the same persons and each obtains the same value.
- e.g., Koopmans 1960; Vallentyne & Kagan 1997:19; Bostrom 2011:16; Jonsson & Voorneveld 2018; Wilkinson MS

Overtaking

Overtaking criterion (from von Weizsäcker 1965)

$W_1 \succcurlyeq W_2$ iff there exists $T_0 \in \mathbb{N}$ such that, for all $T > T_0$,

$$\sum_{t=1}^{T} v_{W_1}(t) - v_{W_2}(t) \geq 0$$

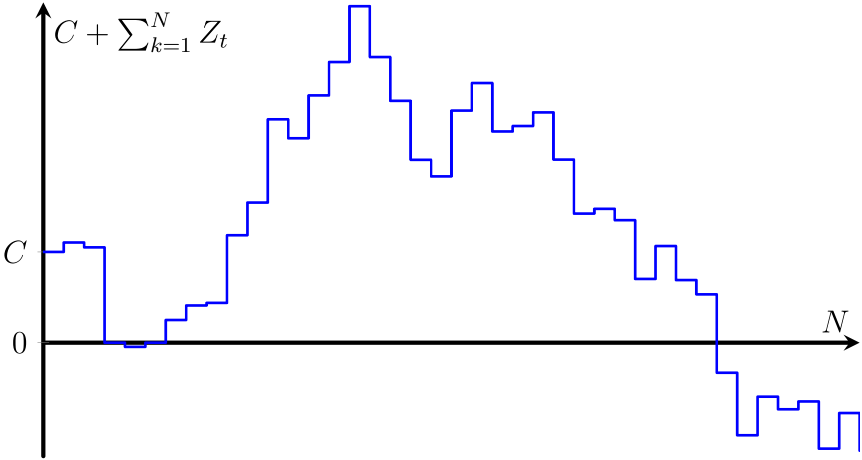
Overtaking for chaotic outcomes

Overtaking criterion (from von Weizsäcker 1965)

$W_1 \succcurlyeq W_2$ iff there exists $T_0 \in \mathbb{N}$ such that, for all $T > T_0$,

$$C + \sum_{t=1}^{T} Z_t \geq 0$$

Overtaking for chaotic outcomes



Overtaking

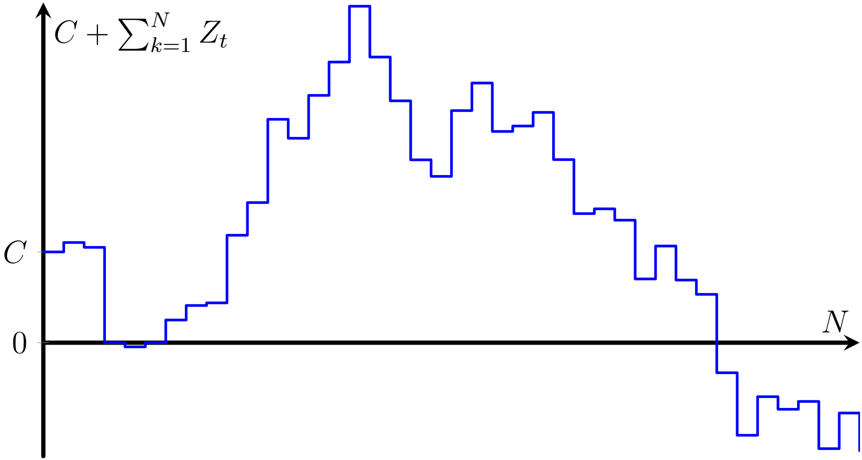
Overtaking criterion (from von Weizsäcker 1965)

$W_1 \succcurlyeq W_2$ iff there exists $T_0 \in \mathbb{N}$ such that, for all $T > T_0$,

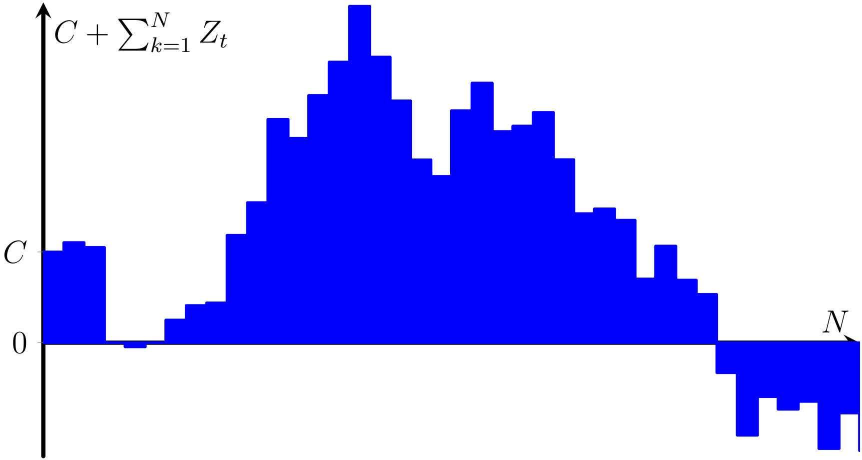
$$\sum_{t=1}^{T} v_{W_1}(t) - v_{W_2}(t) \geq 0$$

A solution

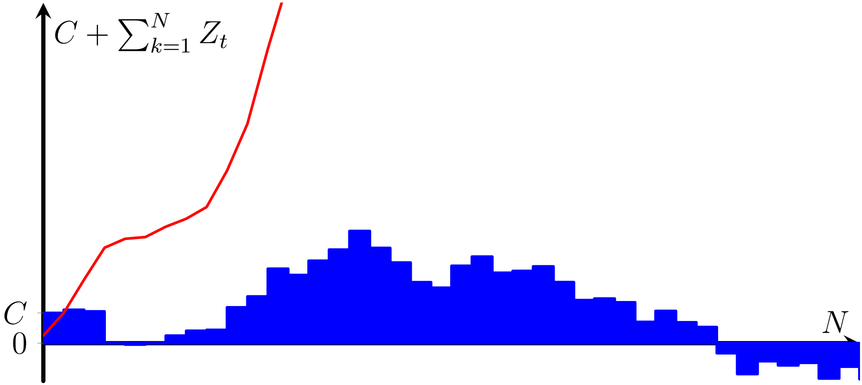
A solution



A solution



A solution



A solution

Integrated catching-up criterion

$W_1 \succ W_2$ iff the following integral approaches $+\infty$.

$$\int_0^\infty \sum_{t=1}^{T} v_{W_1}(t) - v_{W_2}(t) dT$$

$W_1 \simeq W_2$ iff the integral is bounded above and below.

Summary

- In a finite chaotic world, we are *clueless* about which acts have the better outcome.

Summary

- In a finite chaotic world, we are *clueless* about which acts have the better outcome.
  - But we can still say which is better *in expectation*, but only because outcomes are comparable.

Summary

- In a finite chaotic world, we are *clueless* about which acts have the better outcome.
    - But we can still say which is better *in expectation*, but only because outcomes are comparable.
- In an infinite chaotic world:

Summary

- In a finite chaotic world, we are *clueless* about which acts have the better outcome.
  - But we can still say which is better *in expectation*, but only because outcomes are comparable.
- In an infinite chaotic world:
  - *All* strongly impartial views say that all (possible) acts have *equally good* outcomes.

Summary

- In a finite chaotic world, we are *clueless* about which acts have the better outcome.
  - But we can still say which is better *in expectation*, but only because outcomes are comparable.
- In an infinite chaotic world:
  - *All* strongly impartial views say that all (possible) acts have *equally good* outcomes.
  - *All* weakly impartial (Paretian) views say that all acts have *incomparable* outcomes.

Summary

- In a finite chaotic world, we are *clueless* about which acts have the better outcome.
  - But we can still say which is better *in expectation*, but only because outcomes are comparable.
- In an infinite chaotic world:
  - *All* strongly impartial views say that all (possible) acts have *equally good* outcomes.
  - *All* weakly impartial (Paretian) views say that all acts have *incomparable* outcomes.
  - *Many* position-dependent views say that all acts have *incomparable* outcomes.

Summary

- In a finite chaotic world, we are *clueless* about which acts have the better outcome.
  - But we can still say which is better *in expectation*, but only because outcomes are comparable.
- In an infinite chaotic world:
  - *All* strongly impartial views say that all (possible) acts have *equally good* outcomes.
  - *All* weakly impartial (Paretian) views say that all acts have *incomparable* outcomes.
  - *Many* position-dependent views say that all acts have *incomparable* outcomes.
- To say that any available act has a better outcome than another, we must accept position-dependence (or something even less plausible).