

Climate Change Cooperation: the Themis Mechanism

Carl Edward Rasmussen

Cambridge University

Pioneer Centre for AI, Copenhagen, April 22 2026

Outline

- Purpose
- Cooperation: reciprocity and common commitments
 - three illustrative games
- Climate cooperation properties
- The Themis Mechanism
 - annual cycle
 - elicitation process
- Conclusions

Three illustrative games

The independent commitment game:

- ① players can invest an amount of their choice in a communal pot
- ② the pot doubles in size
- ③ the pot is shared equally between all players

Let's say there are 100 players, each with €1.00. How would you play?

Investment is good for everyone: if all 100 players invested €1.00, everyone would end up with €2.00.

But what would be your return on investment? Compare not investing with investing €1.00. The difference (irrespective of what others do) in payout is €0.02: that's **-98% return, terrible!**

It leads to defection or free-riding.

Three illustrative games

The fixed budget game:

- ① players can invest an amount of their choice in a communal pot, subject to the condition that the total budget must be met
- ② the pot doubles in size
- ③ the pot is shared equally between all players

There are 100 players, each with €1.00. The budget is €50.00. How would you play?

Because the budget is fixed, the payout is fixed at €1.00. Players become **adversarial**. It's a zero sum game. No cooperation can take place.

Three illustrative games

The **common commitment game**:

- ① players can pledge an amount of their choice towards a communal pot; only the part of the pledge which is *common to all* is invested (the rest is retained)
- ② the pot doubles in size
- ③ the pot is shared equally between all players

There are 100 players, each with €1.00. How would you play?

There is no reason not to pledge the full amount. You cannot be exploited by free-riders.

The common commitment is key.

Three game conclusions

The games are instructive ...

- the "pot is shared equally" is the crucial aspect of "common resource pool" (CRP) problems
- "independent commitment" games get sunk by free-riders
- "fixed budget games" are zero-sum. They turn players against each other (adversarial)
- "common commitment" games changes everything: encourages cooperation, or alignment of incentives

... but simplistic ...

- there is no notion of time
- it assumes that agents are equivalent. But real nations differ. So how should we interpret "common commitment"?
- it is not practical that a single nation could prevent a successful coalition

Pricing carbon

The essence of the problem is that we currently treat emissions as if they have no (negative) value – what economists call *unpriced negative externalities*. Key idea: remove externalities by pricing carbon.

The price of carbon must be identical everywhere, otherwise it distorts competitiveness.

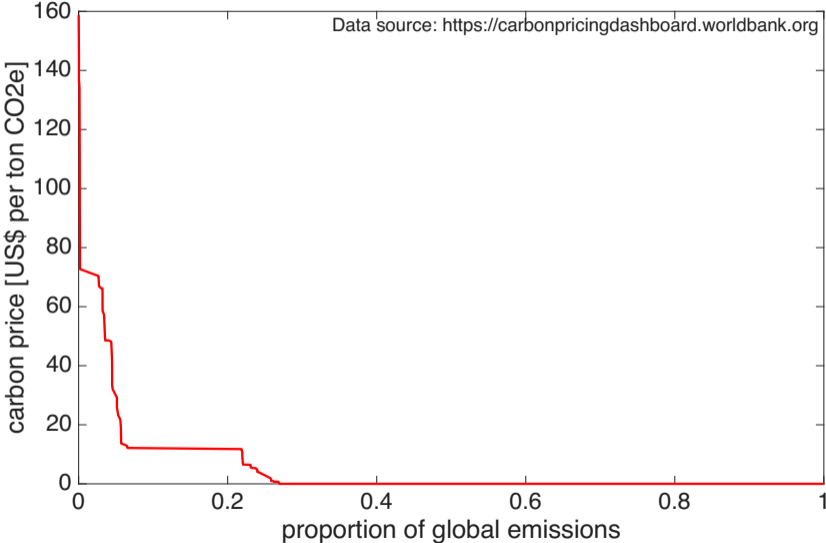
There are two distinct types of carbon pricing:

- In *national* carbon pricing, a minimum price is placed carbon, revenues are retained nationally.
- In *global* carbon pricing, proceeds are collected globally, and divided to coalition members according to population.

Both pure national and pure global carbon pricing won't work well in practice, for opposite reasons. We need something in between: a price p and a global fraction g .

Carbon pricing is often criticised for being socially regressive. That depends entirely on how the proceeds are used.

Current carbon prices



Cooperation in practice

The political economist Elinor Ostrom in her book *Governing the Commons*¹ identified the following conditions for cooperation:

- commitments
- responsibility
- mutual trust
- transparency
- reciprocity²
- ability to sanction non-cooperators

The Paris Agreement is not a cooperative agreement, and shares **none** of Ostrom's properties.

¹Elinor Ostrom: *Governing the Commons*, CUP (1990)

²MacKay, Cramton, Ockenfels and Stoft: *Price Carbon – I will if you will*, Nature (2015)

The Themis Mechanism: foundations

The Themis mechanism is built on four foundations

- ① our atmosphere is a shared resource
- ② we must price carbon
that means **being paid** if you use less than your fair share, and paying if you use more
- ③ urgency: we need to talk about this year, not promises for the future
- ④ balancing fairness and benefits

Themis coalition members sanction non-members with trade barriers.

The Themis Mechanism

The *Coalition Authority* (CA) administers the Themis Mechanism in an annual cycle:

- ① all nations are invited to join the coalition at predefined price p and global fraction g
- ② members report their emissions (using the established UNFCCC reporting rules), and pay the fraction g of price times emissions to the coalition (and retain the remainder)
- ③ the coalition immediately pays out all proceeds in proportion to member population
- ④ next year's p and g are determined by a member elicitation process

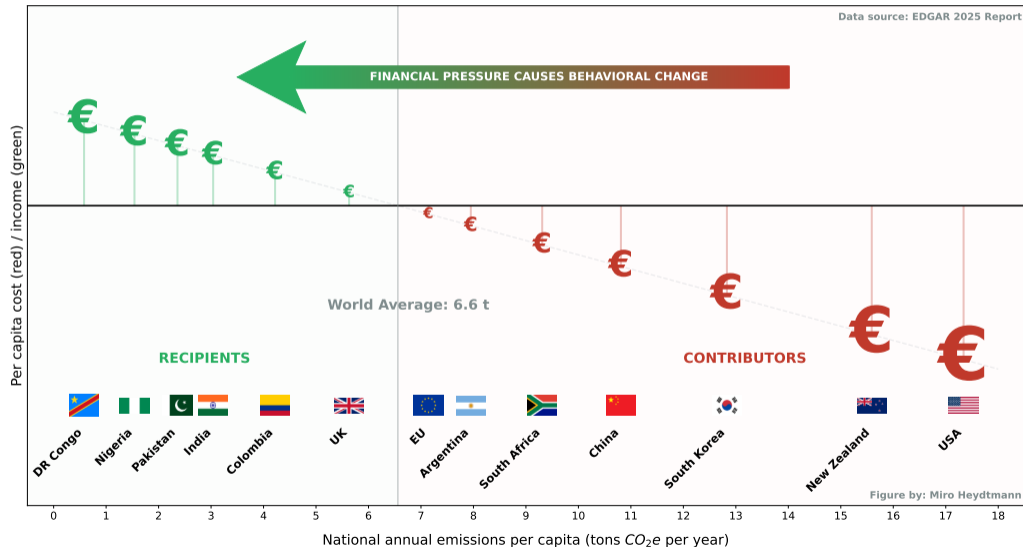
In the inaugural year, the price is set to $p = €0$.

Only per capita emissions are relevant for transfers, reflecting equity. Smaller than average per capita emitters get paid, larger than average emitters pay.

Themis immediately creates economic pressure on all nations to emit less.

The Themis Mechanism

The Themis Mechanism: wealth redistribution assuming universal membership



The Themis elicitation process

The Coalition Authority (CA) uses an *elicitation process* to fix the governing parameters, p and g – not a democratic vote with majority decision and compulsion to comply.

The *coverage* c is the fraction of emissions covered by the coalition.

The CA asks every nation for their preferred value of p and uses these preferences to determine p and c .

The CA chooses the price to maximise the impact of the coalition; the *total price paid* is a proxy:

$$p_{\text{total}} = p \times c(p), \quad \hat{p} = \operatorname{argmax}_p p_{\text{total}}. \quad (1)$$

Elicitation properties

Unanimity is a bad principle, it would give individual nations too much power.

Two important implications of this elicitation procedure:

- no nation is ever asked to do anything it hasn't already volunteered to
- the determination of price depends only on *quantiles*. Quantiles are Strategy Proof (SP). You cannot manipulate the price outcome in your favour: you're always best off by **telling the truth**.

These properties are essential to guarantee a fruitful process: In unstructured negotiations, information is power. This stalls the process. In an SP process, no-one can exploit you expressing your preferences.

The same process can be extended to allow price preferences *conditional* on coverage c and global fraction g , to also determine these.

Next steps

What do we need to *do*?

Understanding cooperation is the missing piece:

- it's not complicated
- but it's not conventional. New ideas take time
- tell your friends, your colleagues and your politicians about it.

We need nations to express their price preferences – this is actually not a big ask. It's not a commitment.

The question is not: what are you willing to do – it's what will you do if others do the same.

Conclusions

Solving our climate crisis requires cooperation: reciprocal agreements and common commitments.

It's not about altruism. It's about enabling you to act in your self interest.

The Themis Mechanism is a simple, concrete and detailed proposal for how to implement climate cooperation.

People and policy makers are not yet thinking about this in the right way.

I need your help to engage people. We can solve the problem. It won't solve itself.

Climate change is not an intractable problem.