LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

JULY 7, 2025

REAL LIFE GAMES:
HOW GAME THEORY SHAPES HUMAN DECISIONS

THE GAME THEORY OF COOPERATION

PUBLIC GOODS GAMES & SOCIAL NORMS

Adrian Haret
a.haret@lmu.de

Let's play a game!

A PUBLIC GOODS GAME

Players start out with an *endowment*, or *budget*.

Each contributes a fraction of their endowment.

Contributions are pooled, multiplied by 1.5, and the result is divided equally among the players.

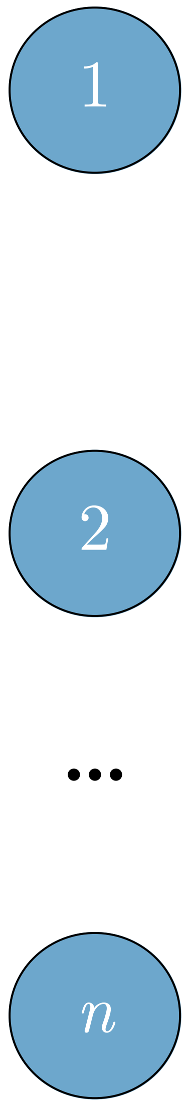
The original endowment, minus the contribution, plus the amount earned becomes the budget for the next round.

Starting with round 3, players have the option of imposing penalties: they give up x points to penalize a player of their choice by $2x$ points.

Player who finishes the game with the most points is the winner.

PUBLIC GOODS GAMES: THE MODEL

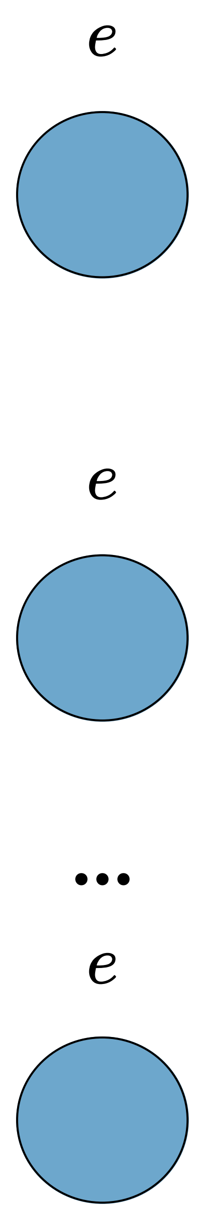
There is a set $N = \{1, \ldots, n\}$ of *players*.



PUBLIC GOODS GAMES: THE MODEL

There is a set $N = \{1, \ldots, n\}$ of *players*.

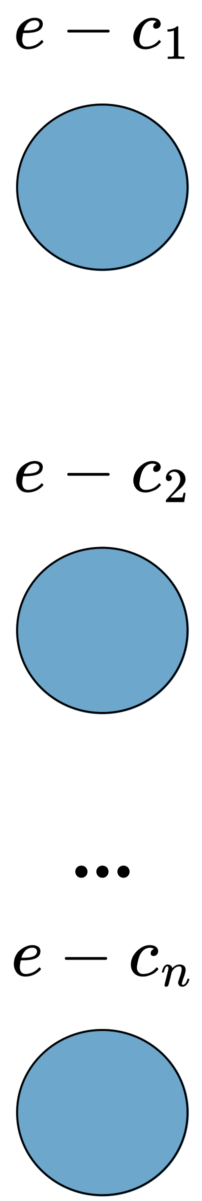
Each player i starts with an *endowment* e



PUBLIC GOODS GAMES: THE MODEL

There is a set $N = \{1, \ldots, n\}$ of *players*.

Each player i starts with an *endowment* e and makes a *contribution* $c_i \in [0, e]$.

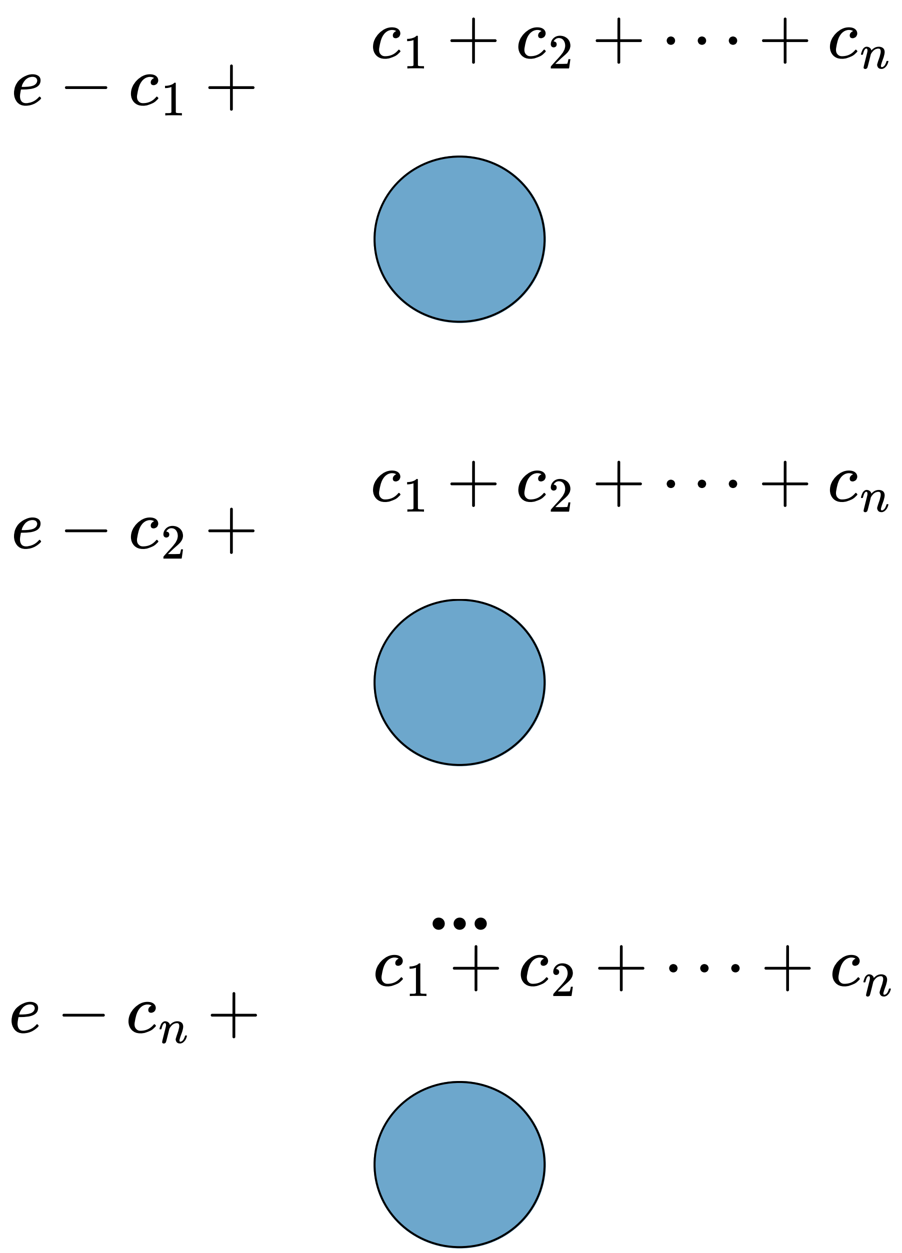


PUBLIC GOODS GAMES: THE MODEL

There is a set $N = \{1, \dots, n\}$ of *players*.

Each player i starts with an *endowment* e and makes a *contribution* $c_i \in [0, e]$.

The *public good* is obtained by pooling all contributions

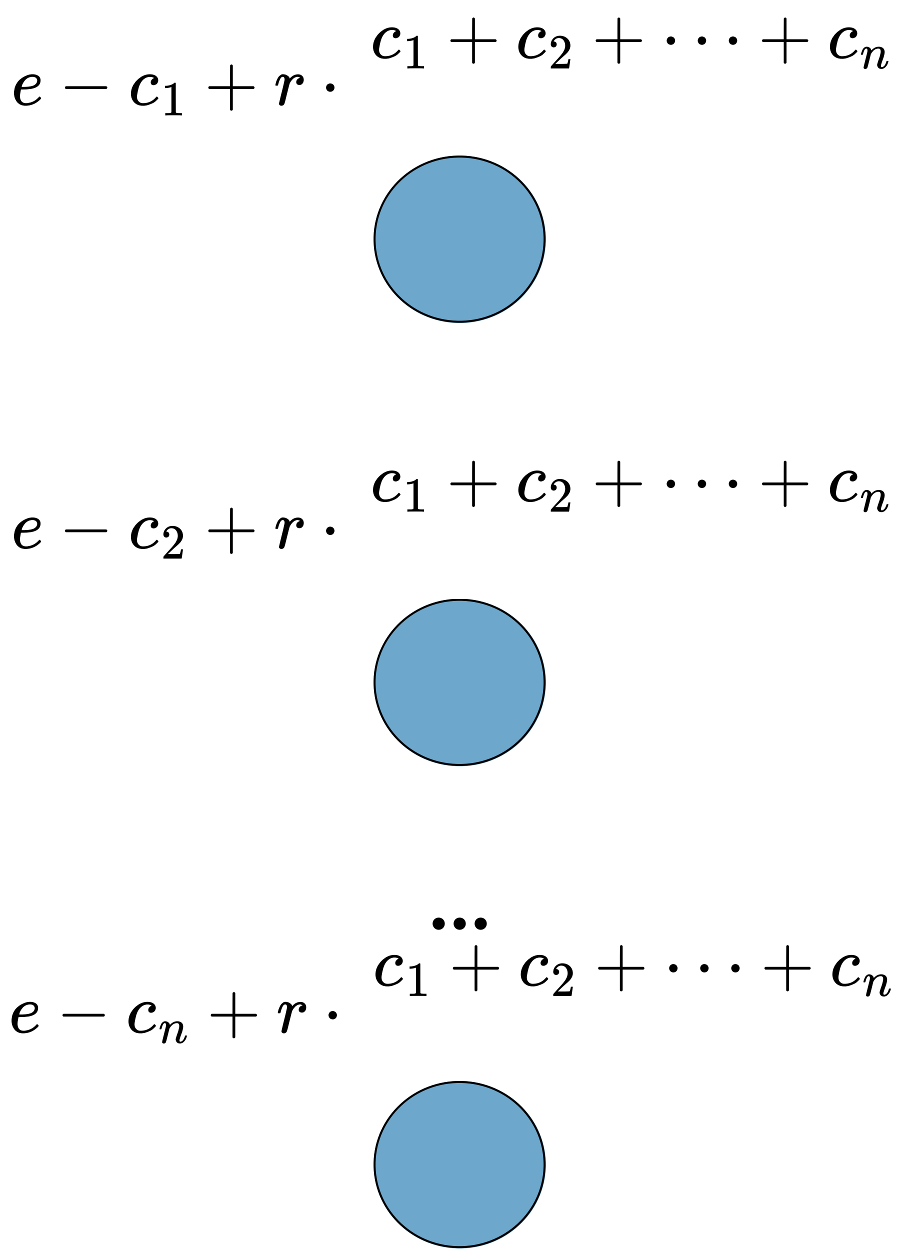


PUBLIC GOODS GAMES: THE MODEL

There is a set $N = \{1, \ldots, n\}$ of *players*.

Each player i starts with an *endowment* e and makes a *contribution* $c_i \in [0, e]$.

The *public good* is obtained by pooling all contributions, and multiplying them by a constant $r > 1$.



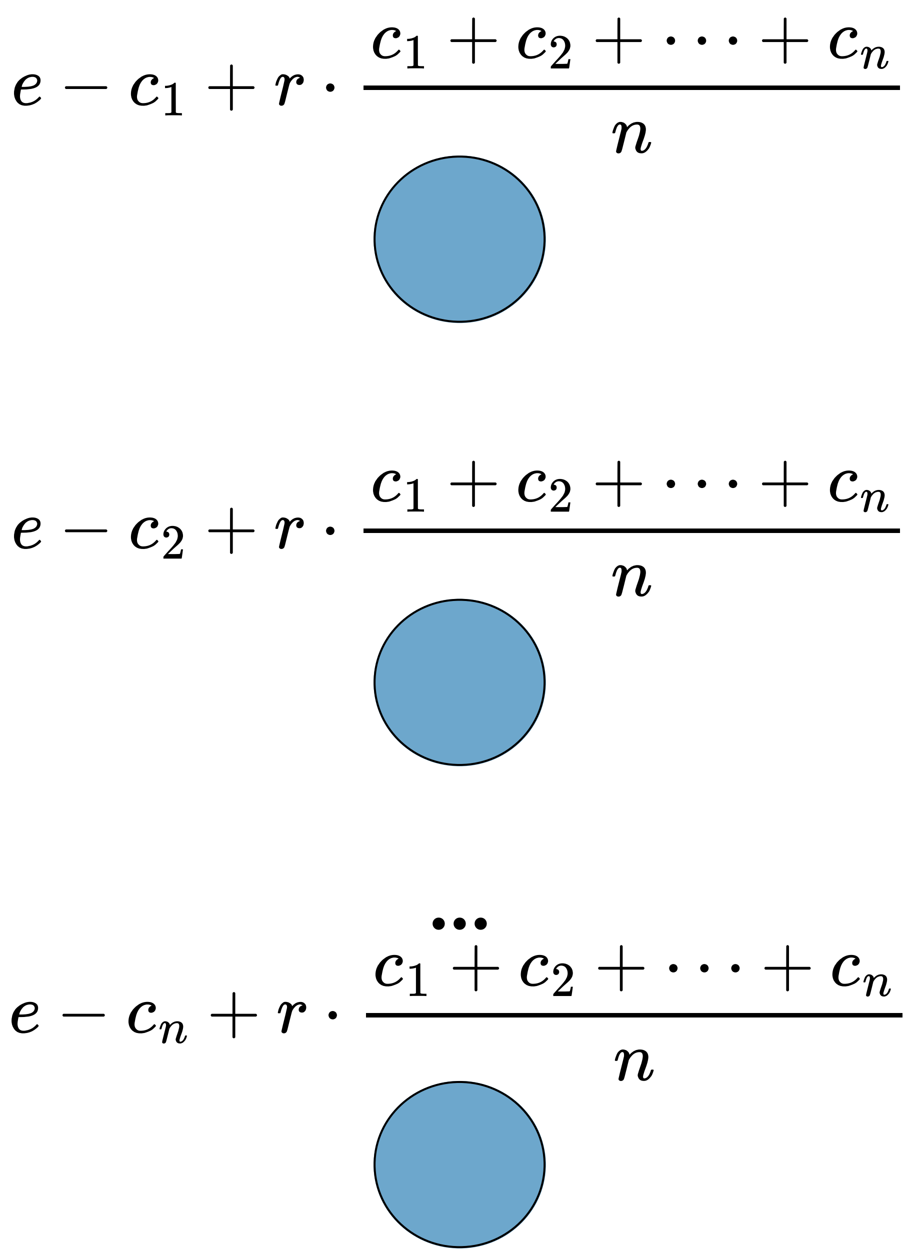
PUBLIC GOODS GAMES: THE MODEL

There is a set $N = \{1, \ldots, n\}$ of *players*.

Each player i starts with an *endowment* e and makes a *contribution* $c_i \in [0, e]$.

The *public good* is obtained by pooling all contributions, and multiplying them by a constant $r > 1$.

Each player receives an equal share of the public good.



$$e - c_1 + r \cdot \frac{c_1 + c_2 + \cdots + c_n}{n}$$

$$e - c_2 + r \cdot \frac{c_1 + c_2 + \cdots + c_n}{n}$$

...

$$e - c_n + r \cdot \frac{c_1 + c_2 + \cdots + c_n}{n}$$

What are the equilibria?

PUBLIC GOODS GAMES: BEST RESPONSES

If player i contributes $c_i > 0$, they make:

$$e - c_i + r \cdot \frac{\sum_{j \neq i} c_j + c_i}{n}.$$

$$e - c_1 + r \cdot \frac{c_1 + c_2 + \cdots + c_n}{n}$$

$$e - c_2 + r \cdot \frac{c_1 + c_2 + \cdots + c_n}{n}$$

...

$$e - c_n + r \cdot \frac{c_1 + c_2 + \cdots + c_n}{n}$$

PUBLIC GOODS GAMES: BEST RESPONSES

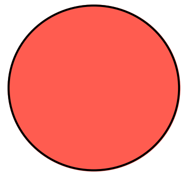
If player i contributes $c_i > 0$, they make:

$$e - c_i + r \cdot \frac{\sum_{j \neq i} c_j + c_i}{n}.$$

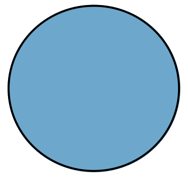
If player i lowers their contribution to $c_i' < c_i$, while everyone else keeps their contribution constant, they make:

$$e - c_i' + r \cdot \frac{\sum_{j \neq i} c_j + c_i'}{n}.$$

$$e - c_i' + r \cdot \frac{c_1' + c_2 + \cdots + c_n}{n}$$

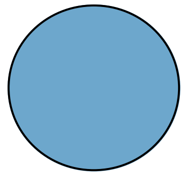


$$e - c_2 + r \cdot \frac{0 + c_2 + \cdots + c_n}{n}$$



...

$$e - c_n + r \cdot \frac{0 + c_2 + \cdots + c_n}{n}$$



PUBLIC GOODS GAMES: BEST RESPONSES

If player i contributes $c_i > 0$, they make:

$$e - c_i + r \cdot \frac{\sum_{j \neq i} c_j + c_i}{n}.$$

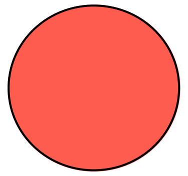
If player i lowers their contribution to $c_i' < c_i$, while everyone else keeps their contribution constant, they make:

$$e - c_i' + r \cdot \frac{\sum_{j \neq i} c_j + c_i'}{n}.$$

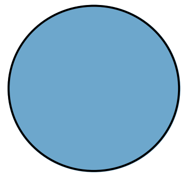
Lowering the contribution is worth it just in case:

$$e - c_i' + r \cdot \frac{\sum_{j \neq i} c_j + c_i'}{n} > e - c_i + r \cdot \frac{\sum_{j \neq i} c_j + c_i}{n} \qquad \text{iff}$$

$$e - c_i' + r \cdot \frac{c_1' + c_2 + \cdots + c_n}{n}$$

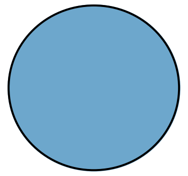


$$e - c_2 + r \cdot \frac{0 + c_2 + \cdots + c_n}{n}$$



...

$$e - c_n + r \cdot \frac{0 + c_2 + \cdots + c_n}{n}$$



PUBLIC GOODS GAMES: BEST RESPONSES

If player i contributes $c_i > 0$, they make:

$$e - c_i + r \cdot \frac{\sum_{j \neq i} c_j + c_i}{n}.$$

If player i lowers their contribution to $c_i' < c_i$, while everyone else keeps their contribution constant, they make:

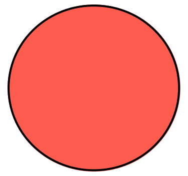
$$e - c_i' + r \cdot \frac{\sum_{j \neq i} c_j + c_i'}{n}.$$

Lowering the contribution is worth it just in case:

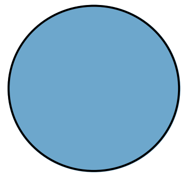
$$e - c_i' + r \cdot \frac{\sum_{j \neq i} c_j + c_i'}{n} > e - c_i + r \cdot \frac{\sum_{j \neq i} c_j + c_i}{n} \quad \text{iff}$$

$$c_i - c_i' > r \cdot \frac{c_i - c_i'}{n} \quad \text{iff}$$

$$e - c_i' + r \cdot \frac{c_1' + c_2 + \cdots + c_n}{n}$$

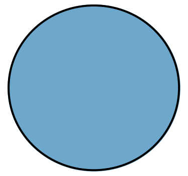


$$e - c_2 + r \cdot \frac{0 + c_2 + \cdots + c_n}{n}$$



…

$$e - c_n + r \cdot \frac{0 + c_2 + \cdots + c_n}{n}$$



PUBLIC GOODS GAMES: BEST RESPONSES

If player i contributes $c_i > 0$, they make:

$$e - c_i + r \cdot \frac{\sum_{j \neq i} c_j + c_i}{n}.$$

If player i lowers their contribution to $c_i' < c_i$, while everyone else keeps their contribution constant, they make:

$$e - c_i' + r \cdot \frac{\sum_{j \neq i} c_j + c_i'}{n}.$$

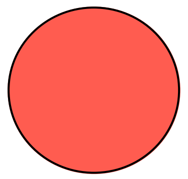
Lowering the contribution is worth it just in case:

$$e - c_i' + r \cdot \frac{\sum_{j \neq i} c_j + c_i'}{n} > e - c_i + r \cdot \frac{\sum_{j \neq i} c_j + c_i}{n} \quad \text{iff}$$

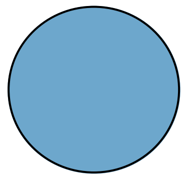
$$c_i - c_i' > r \cdot \frac{c_i - c_i'}{n} \quad \text{iff}$$

$$n > r.$$

$$e - c_i' + r \cdot \frac{c_1' + c_2 + \cdots + c_n}{n}$$

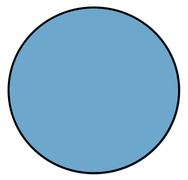


$$e - c_2 + r \cdot \frac{0 + c_2 + \cdots + c_n}{n}$$



...

$$e - c_n + r \cdot \frac{0 + c_2 + \cdots + c_n}{n}$$



So as long as r is not larger than n, a player has an incentive to lower their contribution.

So as long as r is not larger than n, a player has an incentive to lower their contribution. The lowest they can go is 0…

PUBLIC GOODS GAMES: BEST RESPONSES

If everyone contributes 0, there is no public good and each player is stuck with their initial endowment e.

$$e - 0 + r \cdot \frac{0 + 0 + \cdots + 0}{n}$$

$$e - 0 + r \cdot \frac{0 + 0 + \cdots + 0}{n}$$

...

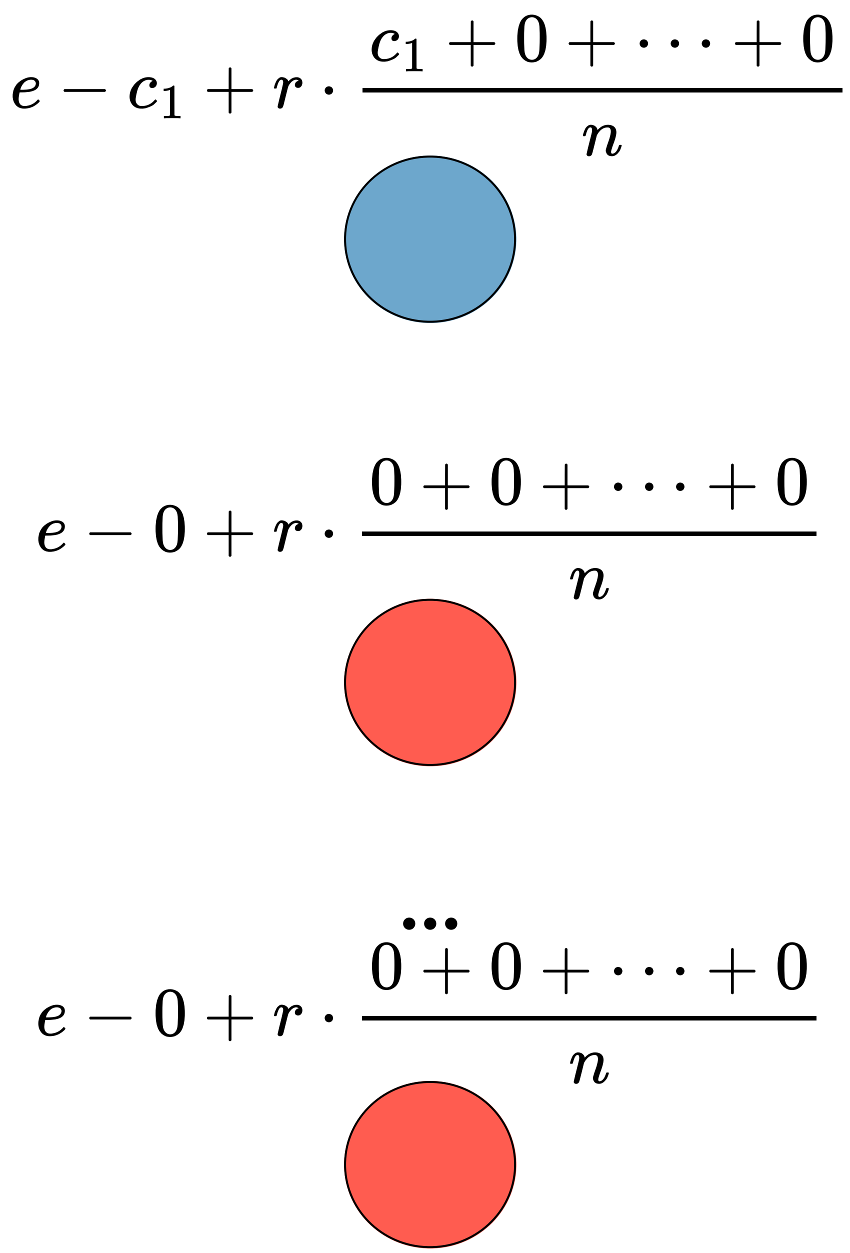
$$e - 0 + r \cdot \frac{0 + 0 + \cdots + 0}{n}$$

PUBLIC GOODS GAMES: BEST RESPONSES

If everyone contributes 0, there is no public good and each player is stuck with their initial endowment e.

In this situation, if a player i decides to deviate and contribute $c_i > 0$, they make:

$$e - c_i + r \cdot \frac{c_i}{n}$$



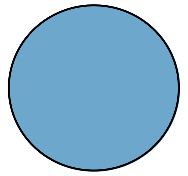
PUBLIC GOODS GAMES: BEST RESPONSES

If everyone contributes 0, there is no public good and each player is stuck with their initial endowment e.

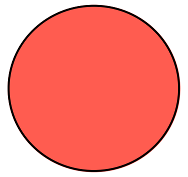
In this situation, if a player i decides to deviate and contribute $c_i > 0$, they make:

$$e - c_i + r \cdot \frac{c_i}{n} < e, \qquad \text{as long as } r < n.$$

$$e - c_1 + r \cdot \frac{c_1 + 0 + \cdots + 0}{n}$$

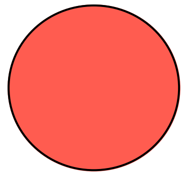


$$e - 0 + r \cdot \frac{0 + 0 + \cdots + 0}{n}$$



...

$$e - 0 + r \cdot \frac{0 + 0 + \cdots + 0}{n}$$

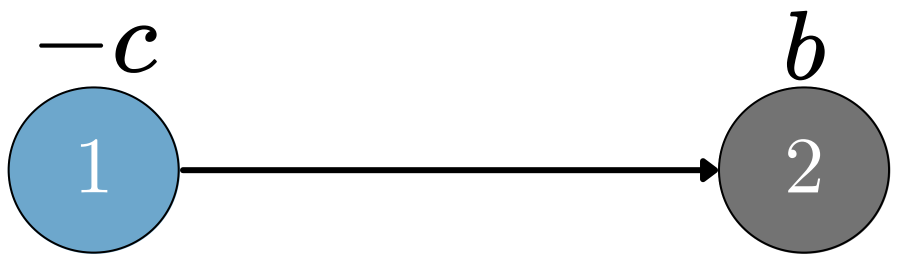


So there is one equilibrium, where no one contributes and the public good is not provisioned.

So there is one equilibrium, where no one contributes and the public good is not provisioned. Despite the fact that everyone would be better off if players contributed!

PUBLIC GOODS GAMES & PRISONER'S DILEMMAS

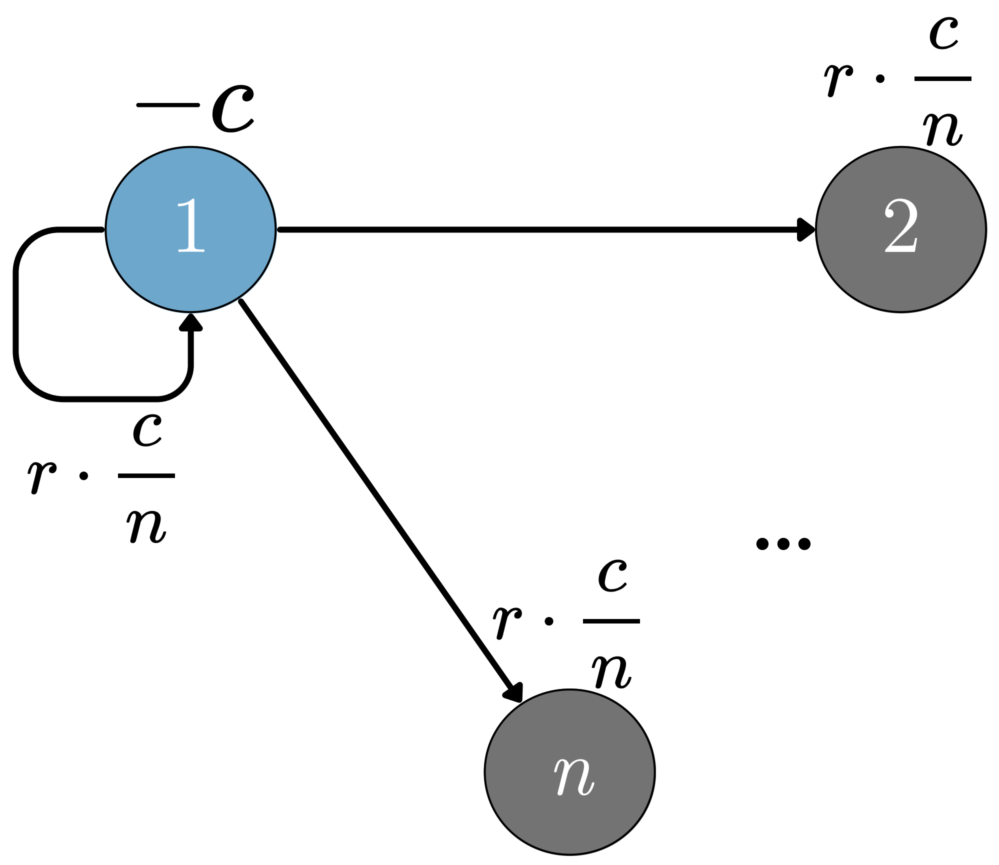
In a Prisoner's Dilemma a cooperator pays a cost to provide a benefit to the other player.



PUBLIC GOODS GAMES & PRISONER'S DILEMMAS

In a Prisoner's Dilemma a cooperator pays a cost to provide a benefit to the other player.

In a Public Goods Game, everyone gets a share of the cooperator's contribution, including the focal player themselves.

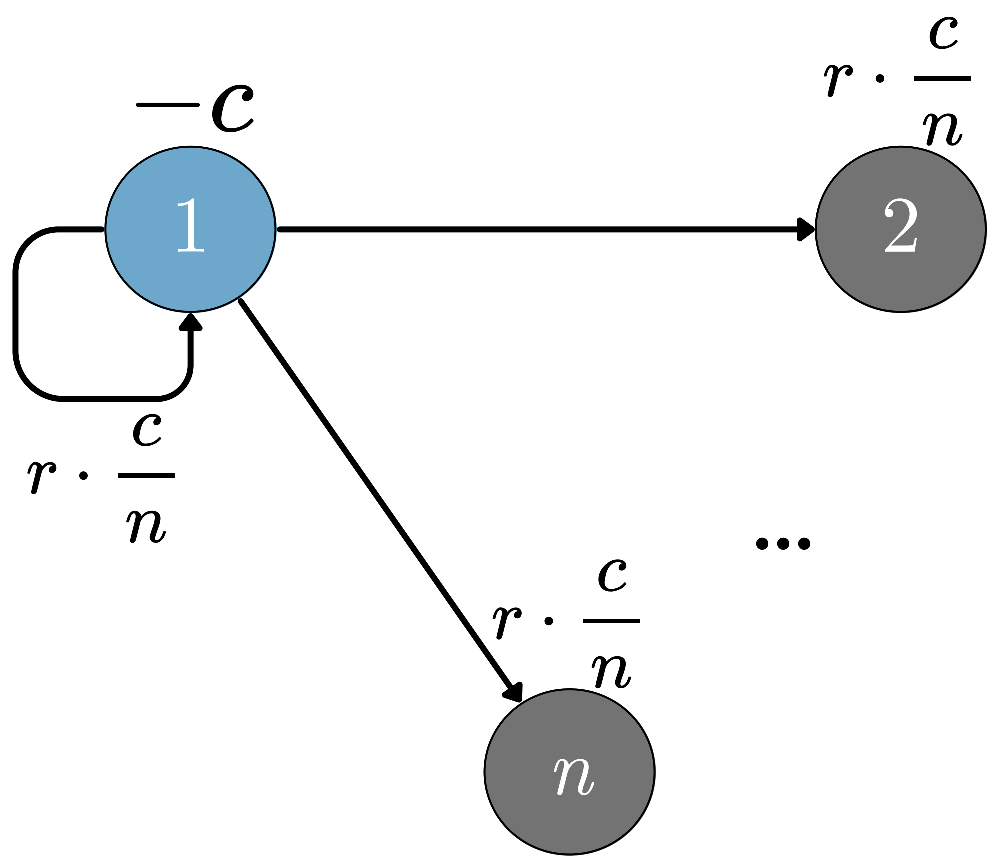


PUBLIC GOODS GAMES & PRISONER’S DILEMMAS

In a Prisoner’s Dilemma a cooperator pays a cost to provide a benefit to the other player.

In a Public Goods Game, everyone gets a share of the cooperator’s contribution, including the focal player themselves.

In this sense, a Public Goods Game are a more general version of a Prisoner’s Dilemma.



Public goods games show up everywhere… but do we see robust contributions?

Public goods games show up everywhere… but do we see robust contributions? It’s tricky…

Chuck Feeney

1931 - 2023

Chuck Feeney

1931 - 2023

Co-founder of Duty Free Shoppers Group.

Chuck Feeney

1931 - 2023

Co-founder of Duty Free Shoppers Group.

Gave away most of his wealth, in the billions.

Would also pick up rubbish from the street.

If everybody picked up trash, there would be no trash on the streets.

At the same time…

In January 2025, the US withdrew (again) from the Paris Climate Accords.

In January 2025, the US withdrew (again) from the Paris Climate Accords. We don't know what the text chain was.

In January 2025, the US withdrew (again) from the Paris Climate Accords. We don't know what the text chain was. But it might have looked like this…

VP

Shall we commit to the goals of the Paris Agreement and work towards reducing emissions?

VP

POTUS

•••

P

VP

VP

Shall we commit to the goals of the Paris Agreement and work towards reducing emissions?

POTUS

I don’t see the point, if it hurts our industry. Especially while there are other large polluters out there...

P

Noor, D. (2025, January 20). Trump signs order to withdraw US from Paris climate agreement for second time. *The Guardian*.

The dilemma of public goods. We all prefer the good to be provided.

The dilemma of public goods. We all prefer the good to be provided. But we'd like someone else to pay for it.

The dilemma of public goods. We all prefer the good to be provided. But we'd like someone else to pay for it. Widespread freeriding, though, can be disastrous…


THE TRAGEDY OF THE COMMONS
A pasture, open to all: a *commons*.


THE TRAGEDY OF THE COMMONS
A pasture, open to all: a *commons*.
Unrestricted grazing leads to a depleted pasture, and everyone suffers.

THE TRAGEDY OF THE COMMONS

A pasture, open to all: a *commons*.

Unrestricted grazing leads to a depleted pasture, and everyone suffers.

But each herder wants to give their *own* livestock as much grazing time as possible.

GARRETT HARDIN

Therein is the tragedy.

Hardin, G. (1968). The Tragedy of the Commons. *Science*, 162(3859), 1243–1248.

GARRETT HARDIN

Therein is the tragedy.

Each man is locked into a system that compels him to increase his herd without limit—in a world that is limited.

Hardin, G. (1968). The Tragedy of the Commons. *Science*, 162(3859), 1243–1248.

GARRETT HARDIN

Therein is the tragedy.

Each man is locked into a system that compels him to increase his herd without limit—in a world that is limited.

Ruin is the destination toward which all men rush, each pursuing his own best interest in a society that believes in the freedom of the commons.

Hardin, G. (1968). The Tragedy of the Commons. *Science*, 162(3859), 1243–1248.

GARRETT HARDIN

Therein is the tragedy.

Each man is locked into a system that compels him to increase his herd without limit—in a world that is limited.

Ruin is the destination toward which all men rush, each pursuing his own best interest in a society that believes in the freedom of the commons.

Freedom in a commons brings ruin to all.

Hardin, G. (1968). The Tragedy of the Commons. *Science*, 162(3859), 1243–1248.

How do we avoid ‘ruin’? How do we get less freeriding and more Chuck Feeney-ing?

MANCUR OLSON

Unless … there is coercion…, rational, self-interested individuals will not act to achieve their common or group interests.

Olson, M. (1971). *The Logic of Collective Action: Public Goods and the Theory of Groups*. Harvard University Press.

MANCUR OLSON

Unless … there is coercion…, rational, self-interested individuals will not act to achieve their common or group interests.

Olson, M. (1971). *The Logic of Collective Action: Public Goods and the Theory of Groups*. Harvard University Press.

THOMAS HOBBES

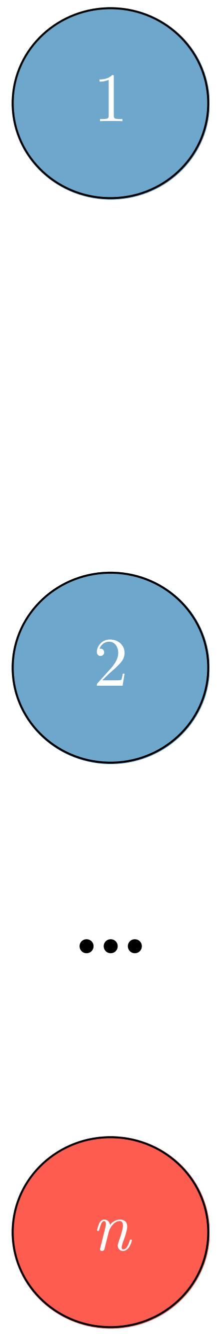
We need an external authority, a *Leviathan*, to keep people in check.

The most obvious mechanism of coercion is *punishment*.

The most obvious mechanism of coercion is *punishment*. Let’s model it using a simplified version of a public goods game.

PUBLIC GOODS GAMES WITH PUNISHMENT

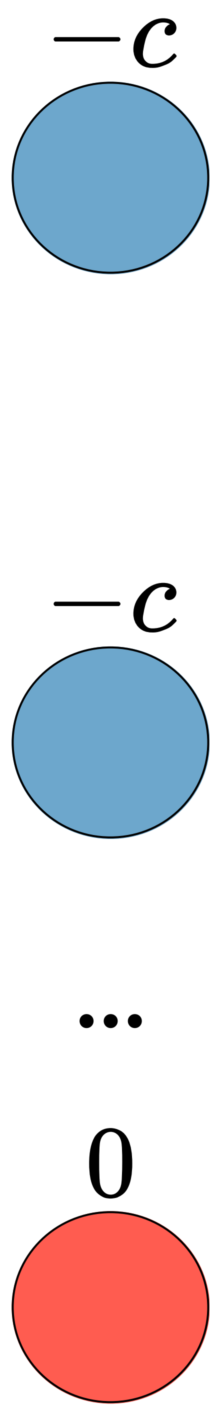
There is a set $N = \{1, \ldots, n\}$ of *players*.



PUBLIC GOODS GAMES WITH PUNISHMENT

There is a set $N = \{1, \ldots, n\}$ of *players*.

At the first round players either make a *contribution* of $c > 0$, or *freeride* by contributing 0.

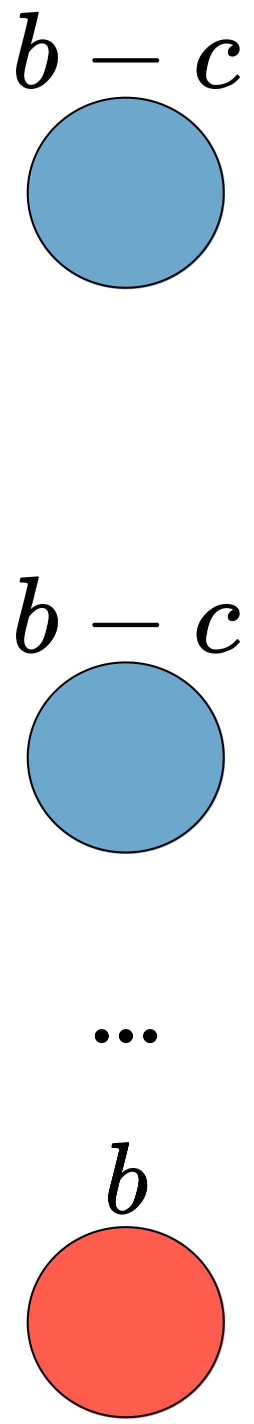


PUBLIC GOODS GAMES WITH PUNISHMENT

There is a set $N = \{1, \ldots, n\}$ of *players*.

At the first round players either make a *contribution* of $c > 0$, or *freeride* by contributing 0.

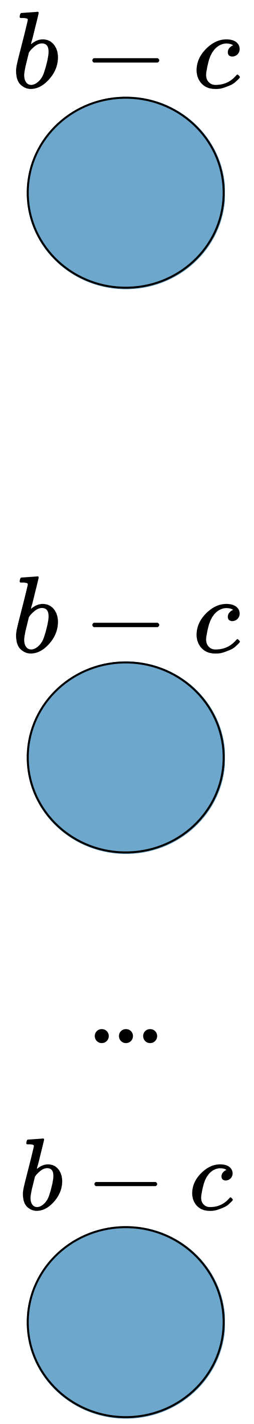
If there are enough* contributions the public good is provided and everyone receives a benefit $b > c$.



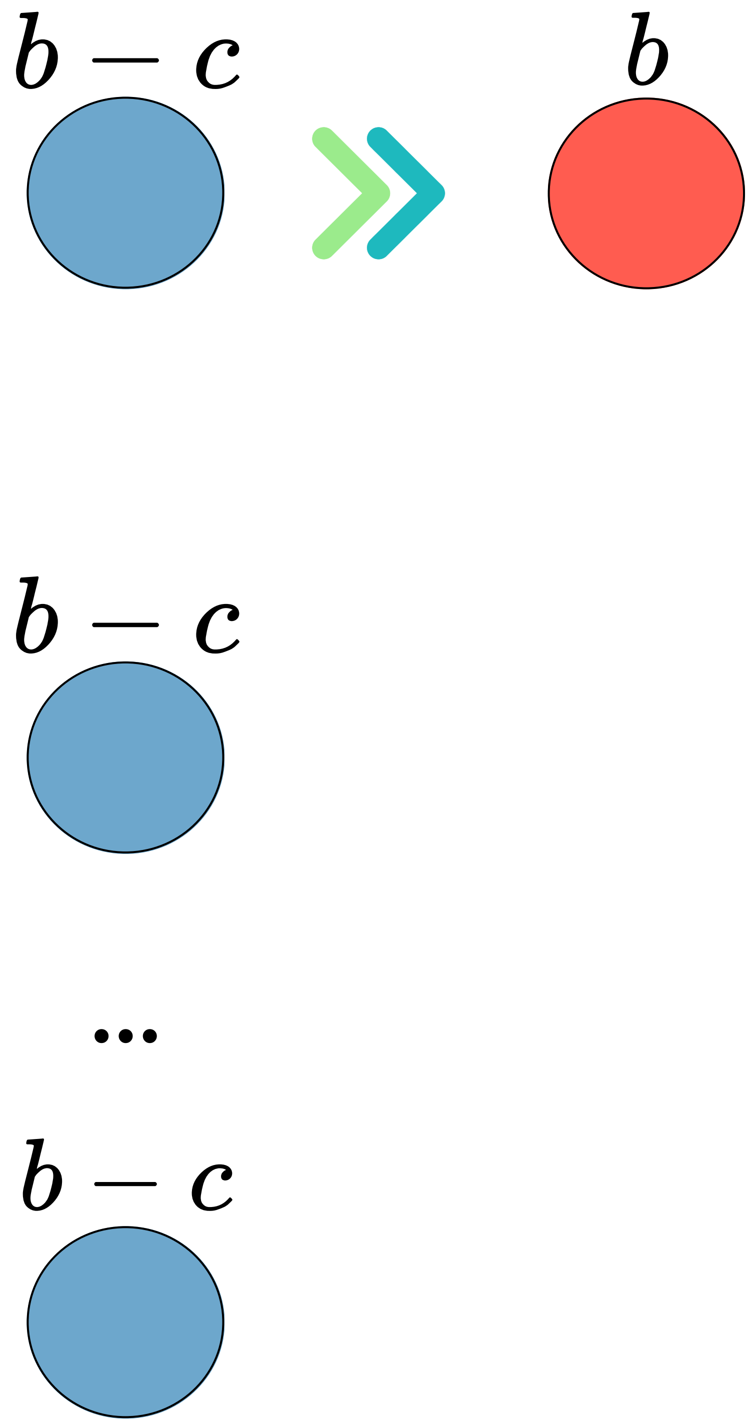
*We leave it a big vague as to how many are needed, but more than one and less than all.

SIDENOTE

As long as the public good is provided, at this point each player has an incentive to freeride.



SIDENOTE



As long as the public good is provided, at this point each player has an incentive to freeride.

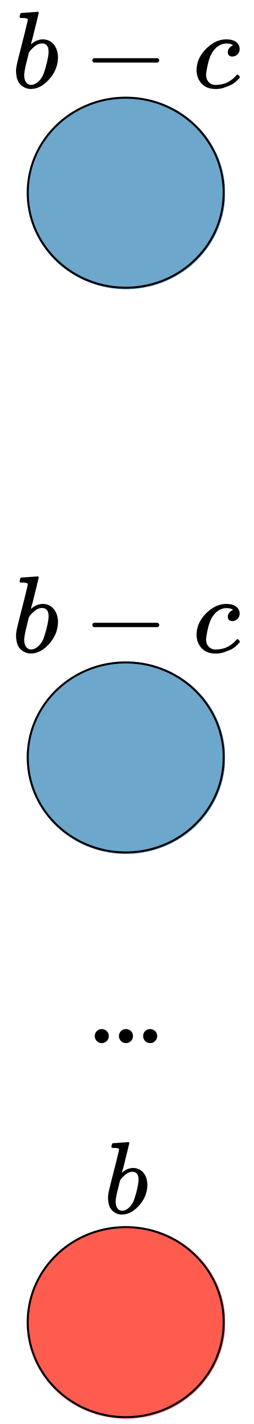
Widespread cooperation is, again, not an equilibrium.

PUBLIC GOODS GAMES WITH PUNISHMENT

There is a set $N = \{1, \ldots, n\}$ of *players*.

At the first round players either make a *contribution* of $c > 0$, or *freeride* by contributing 0.

If there are enough* contributions the public good is provided and everyone receives a benefit $b > c$.



*We leave it a big vague as to how many are needed, but more than one and less than all.

PUBLIC GOODS GAMES WITH PUNISHMENT

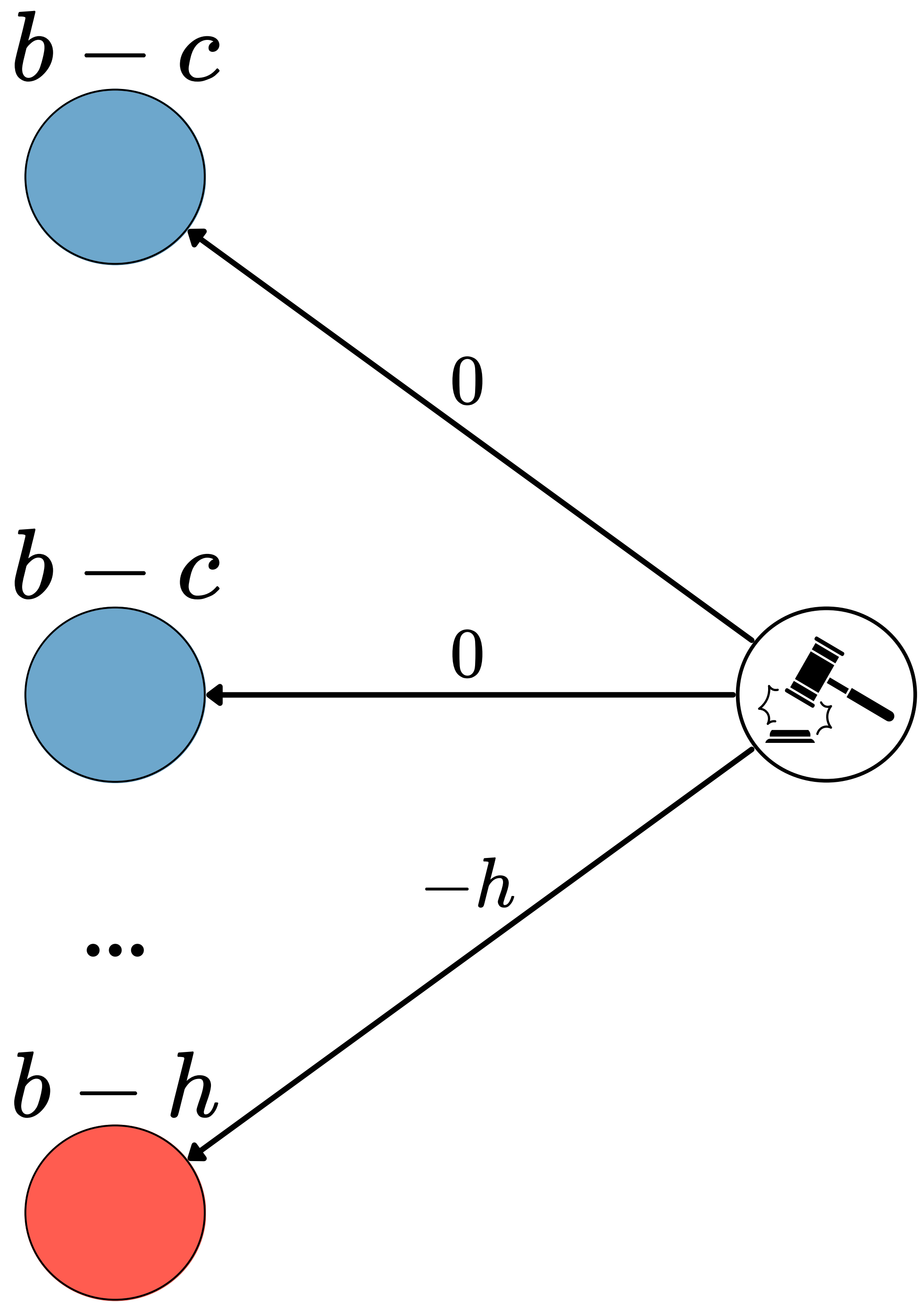
There is a set $N = \{1, \ldots, n\}$ of *players*.

At the first round players either make a *contribution* of $c > 0$, or *freeride* by contributing 0.

If there are enough* contributions the public good is provided and everyone receives a benefit $b > c$.

At the second round an authority imposes a penalty h to non-contributors.

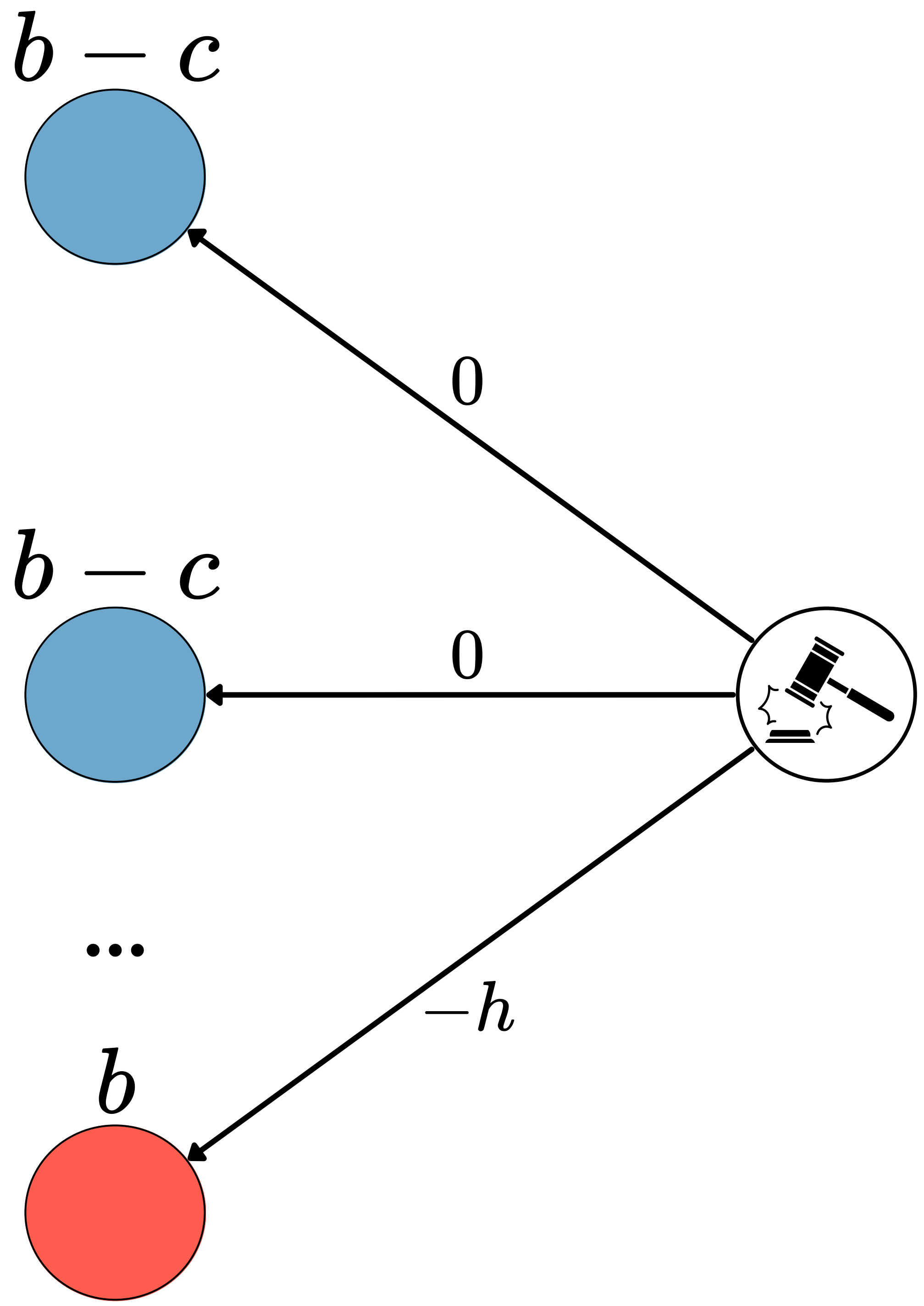
*We leave it a big vague as to how many are needed, but more than one and less than all.



Now universal cooperation can be an equilibrium!

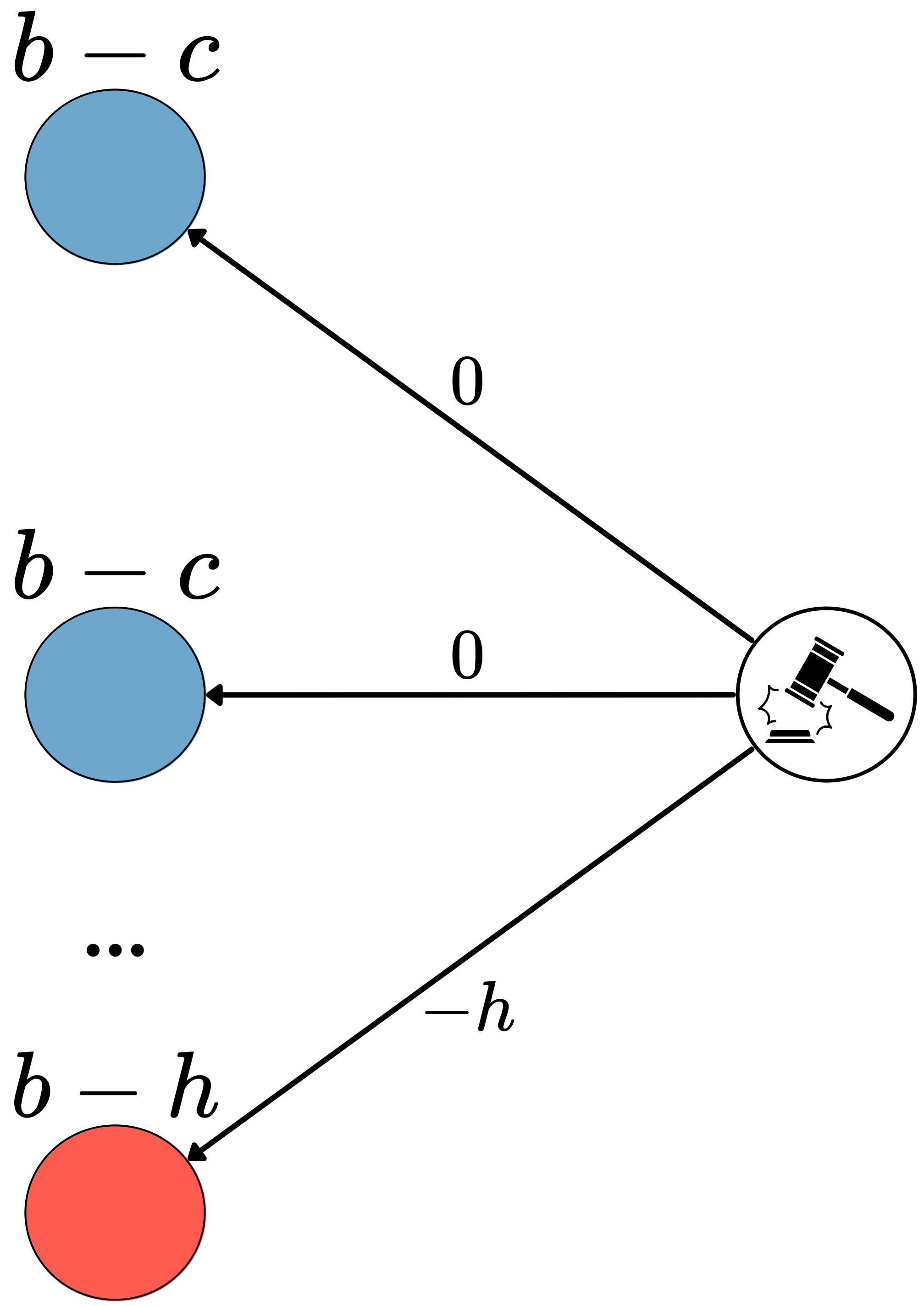
PUNISHMENT STABILIZES COOPERATION

If a player fails to cooperate, they get a penalty.



PUNISHMENT STABILIZES COOPERATION

If a player fails to cooperate, they get a penalty.

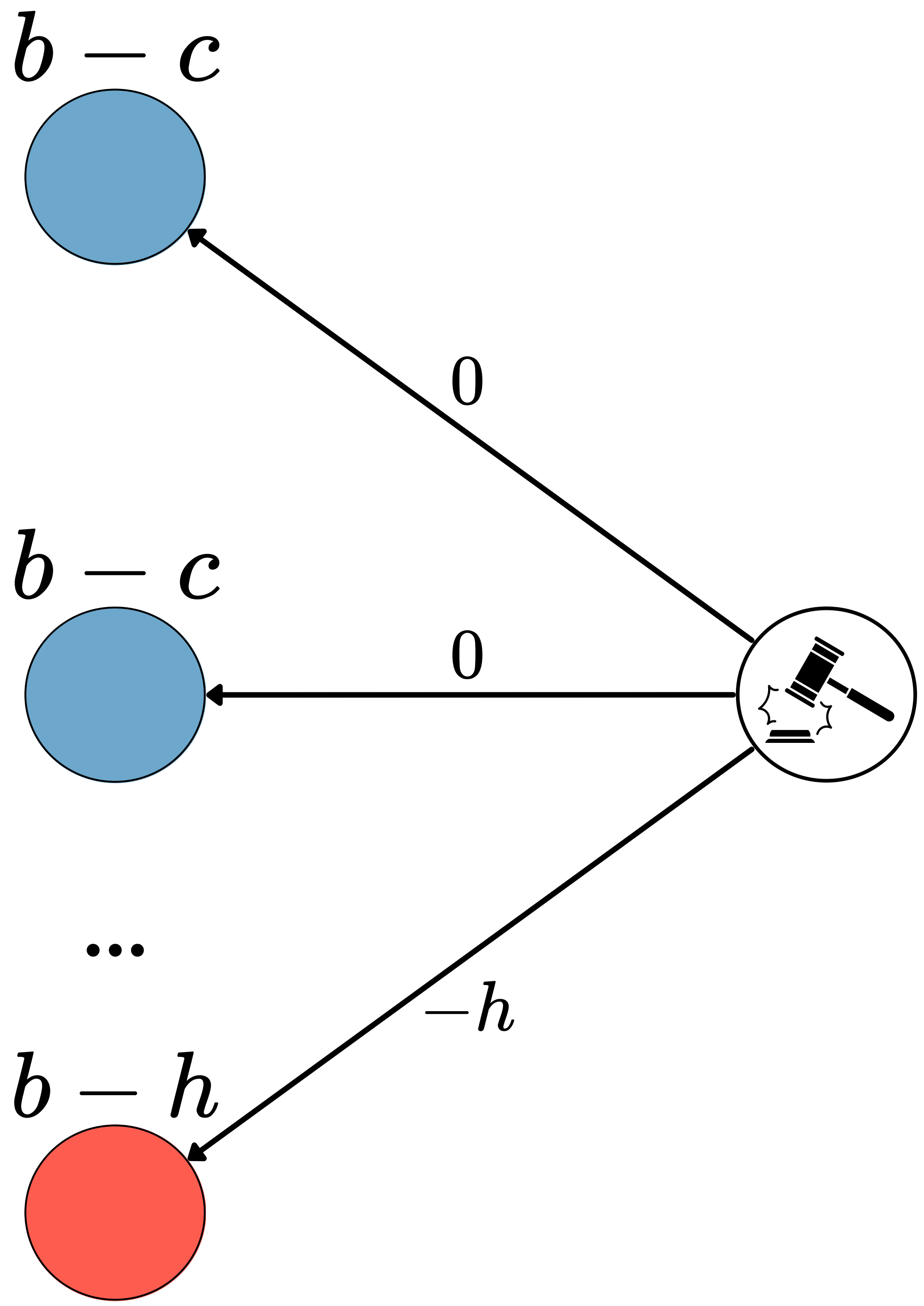


PUNISHMENT STABILIZES COOPERATION

If a player fails to cooperate, they get a penalty.

Freeriding doesn't pay off as long as:

$$h > c.$$



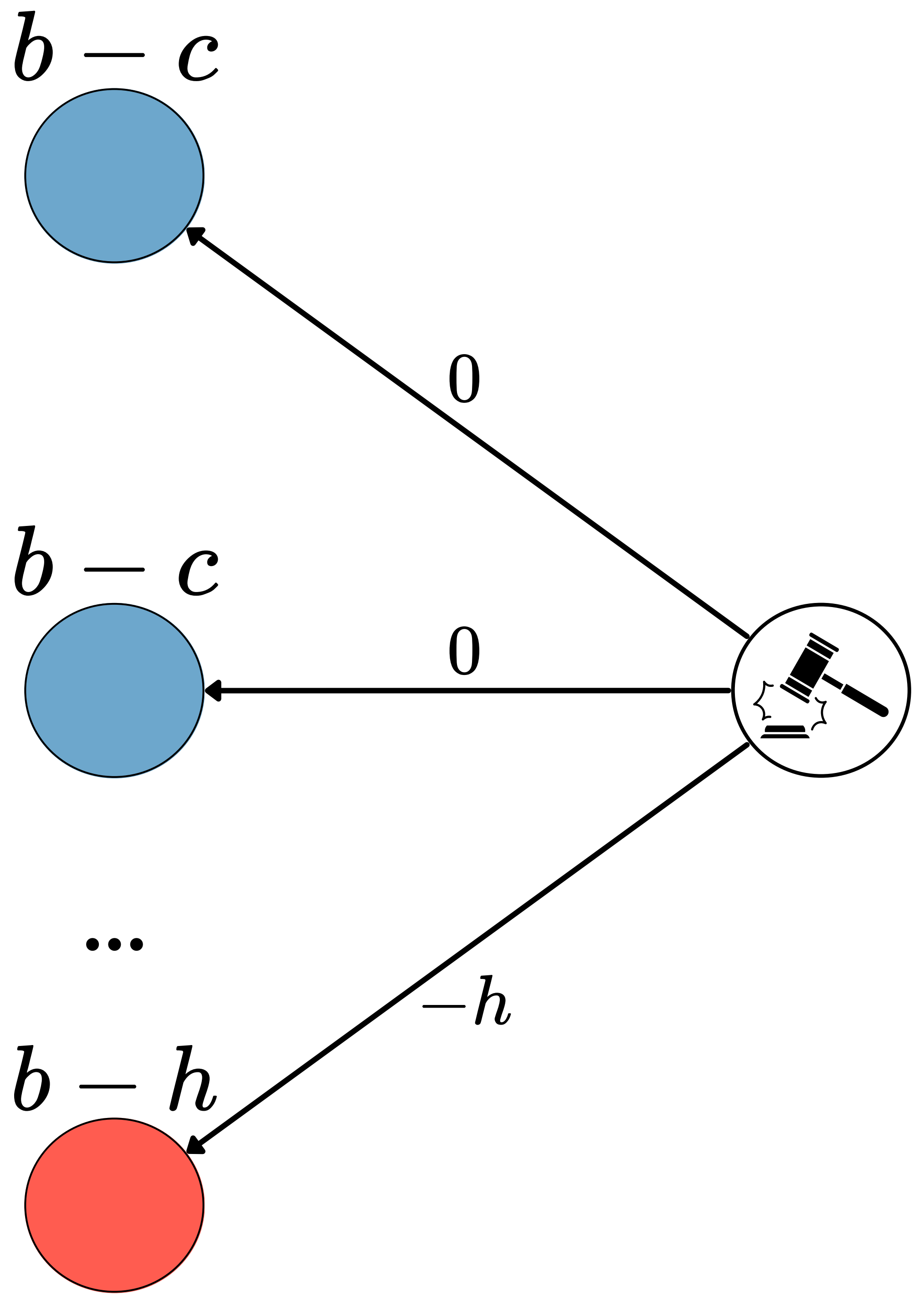
PUNISHMENT STABILIZES COOPERATION

If a player fails to cooperate, they get a penalty.

Freeriding doesn't pay off as long as:

$$h > c.$$

That is, as long as not cooperating is more expensive than cooperating.



PUNISHMENT STABILIZES COOPERATION

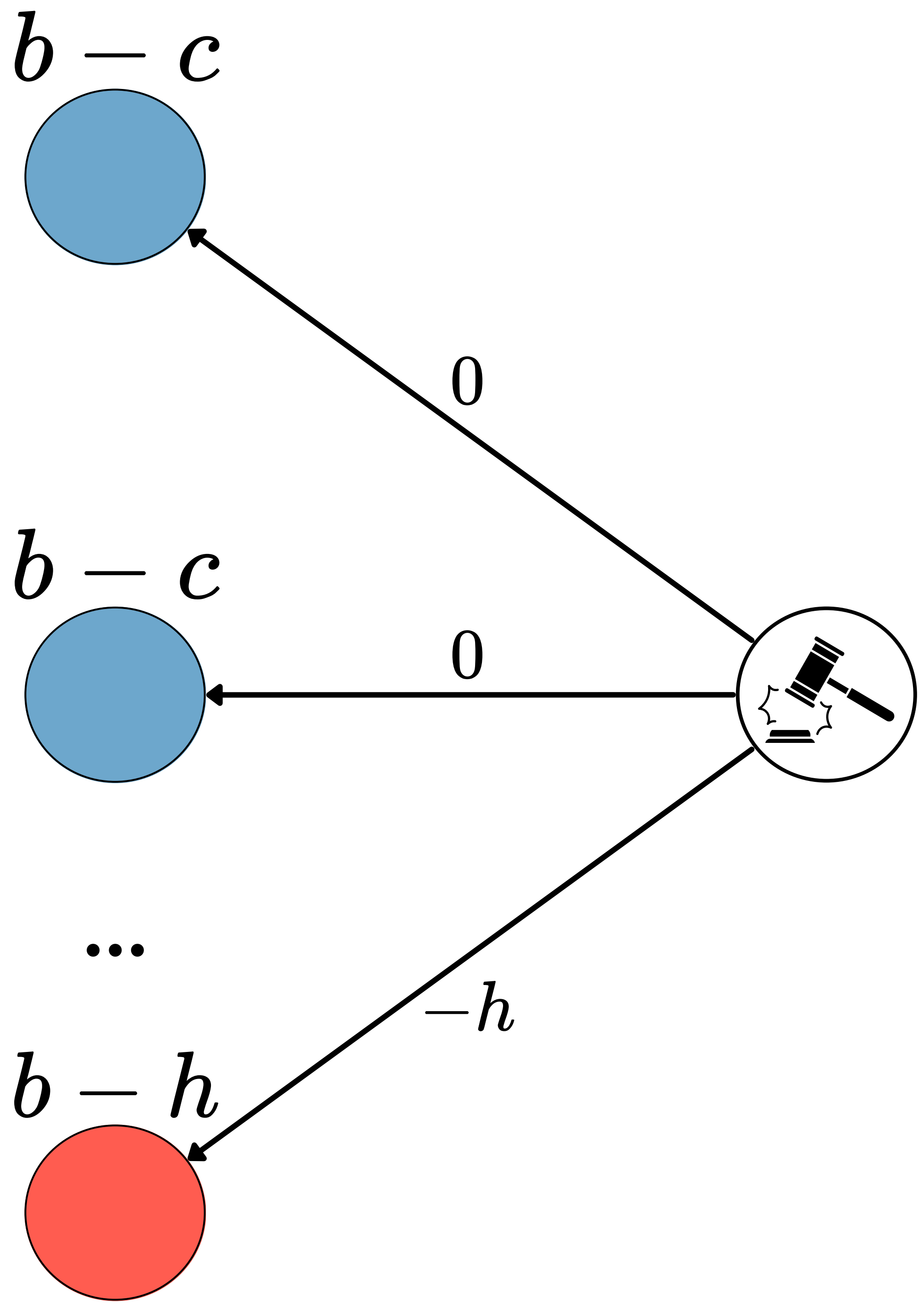
If a player fails to cooperate, they get a penalty.

Freeriding doesn't pay off as long as:

$$h > c.$$

That is, as long as not cooperating is more expensive than cooperating.

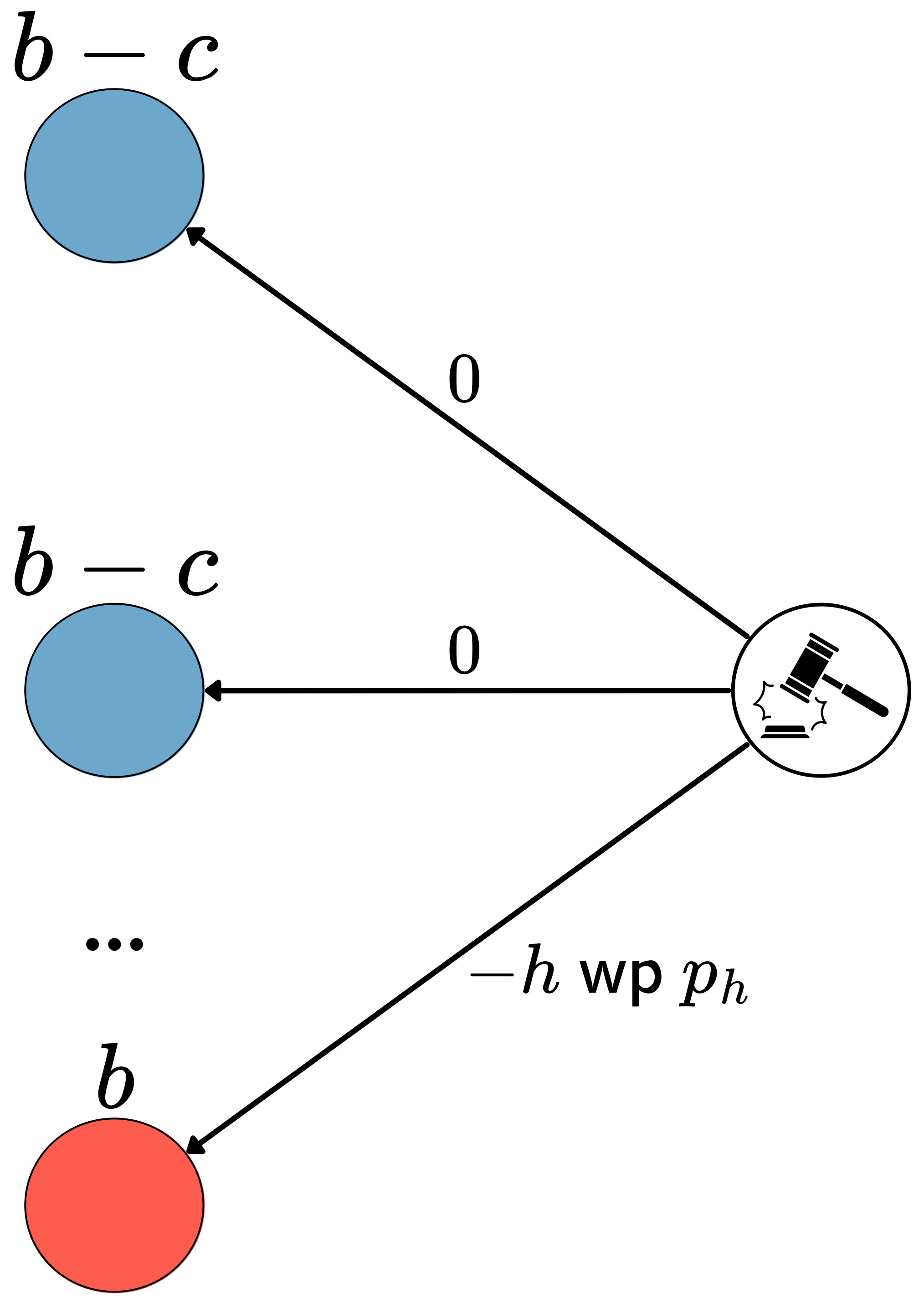
With punishment, the only equilibrium is with universal cooperation.



But this model is rather unrealistic. What if, for instance, the Leviathan is not perfectly efficient in enacting punishment?

IMPERFECT PUNISHMENT

Suppose the Leviathan punishes a defection with probability p_h.

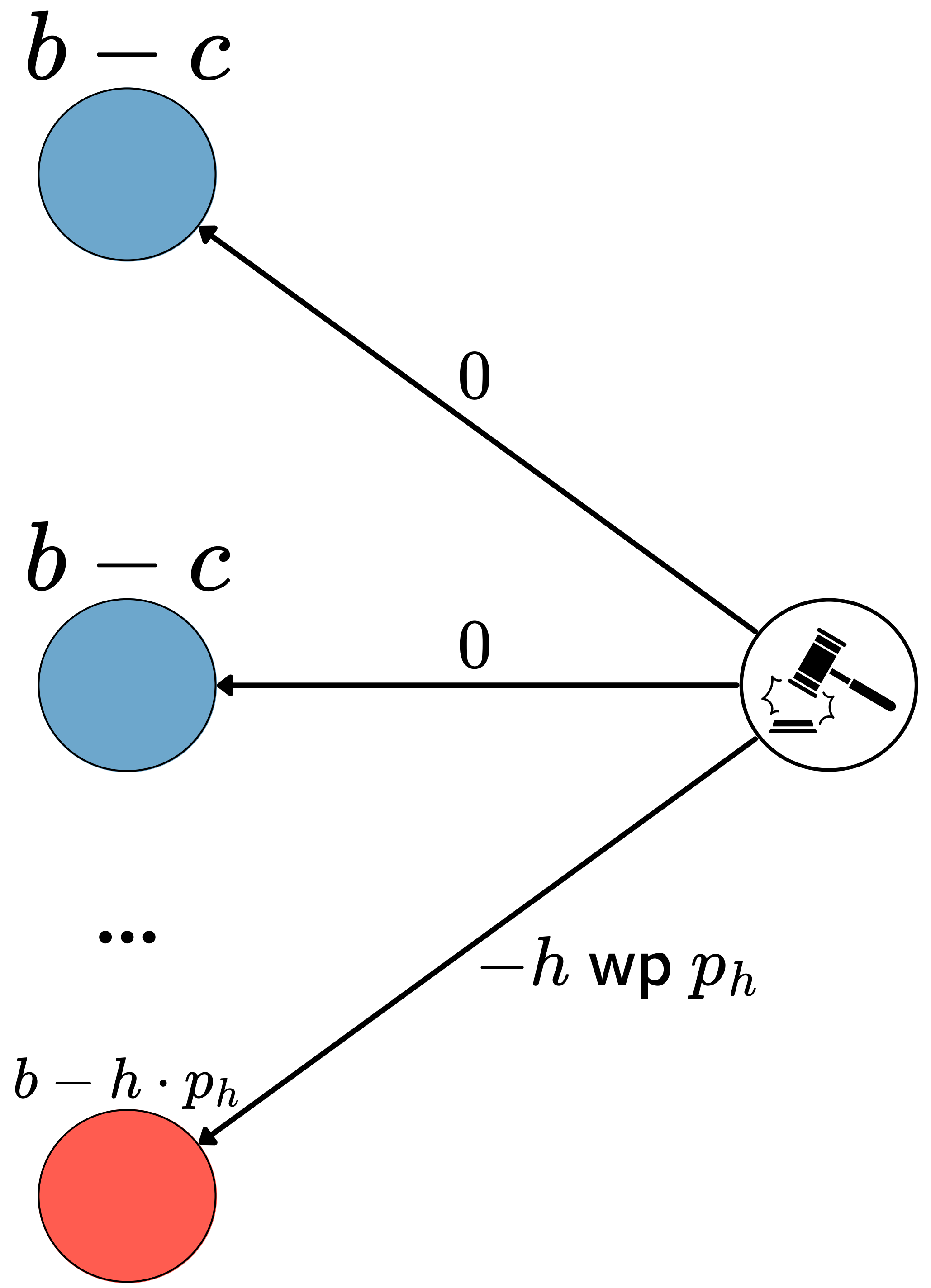


IMPERFECT PUNISHMENT

Suppose the Leviathan punishes a defection with probability p_h.

So a defector's expected payoff is:

$$b - h \cdot p_h.$$



IMPERFECT PUNISHMENT

Suppose the Leviathan punishes a defection with probability p_h.

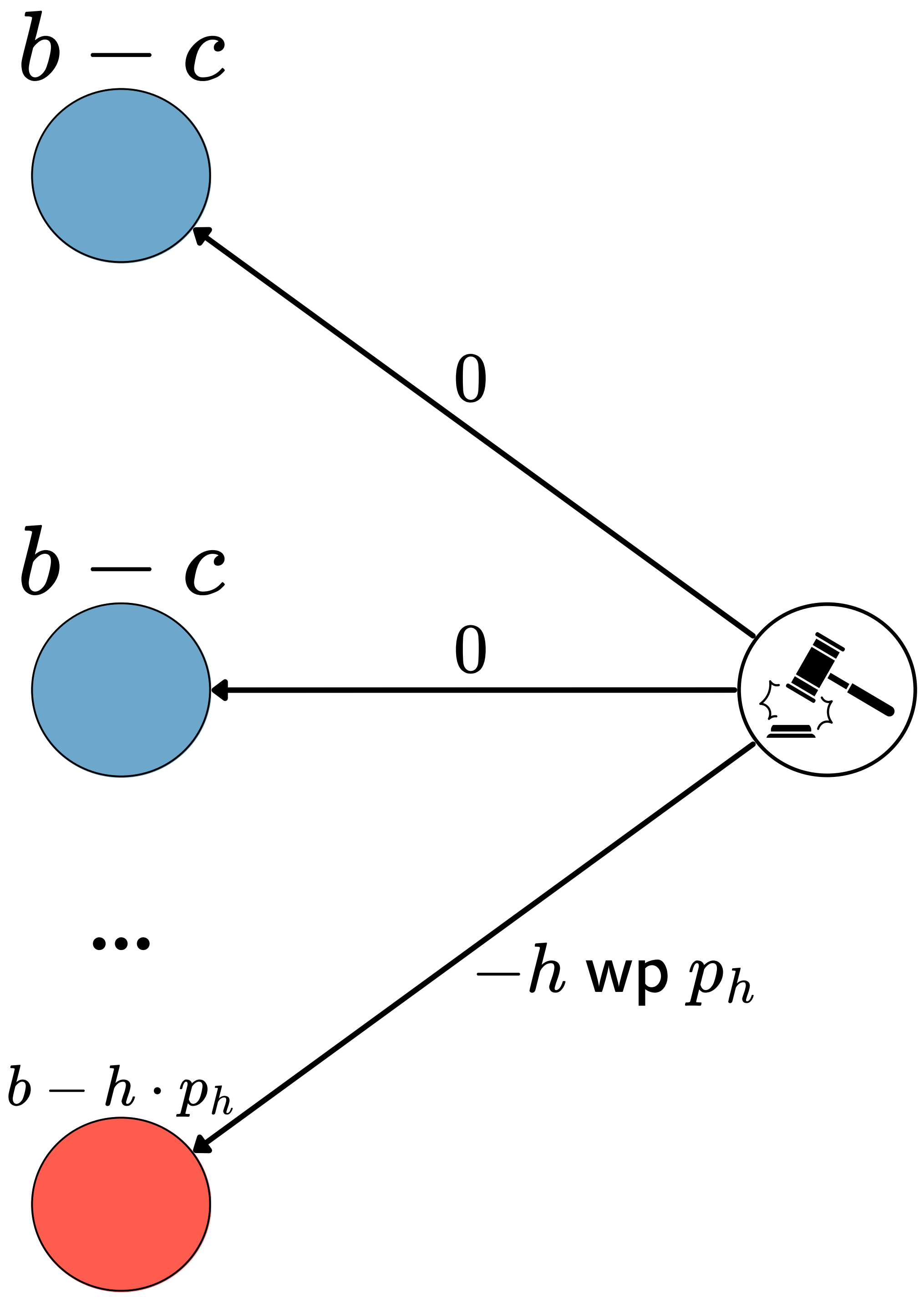
So a defector's expected payoff is:

$$b - h \cdot p_h.$$

Cooperating pays off when:

$$b - c > b - h \cdot p_h \qquad \text{iff}$$

$$h \cdot p_h > c.$$



IMPERFECT PUNISHMENT

Suppose the Leviathan punishes a defection with probability p_h.

So a defector's expected payoff is:

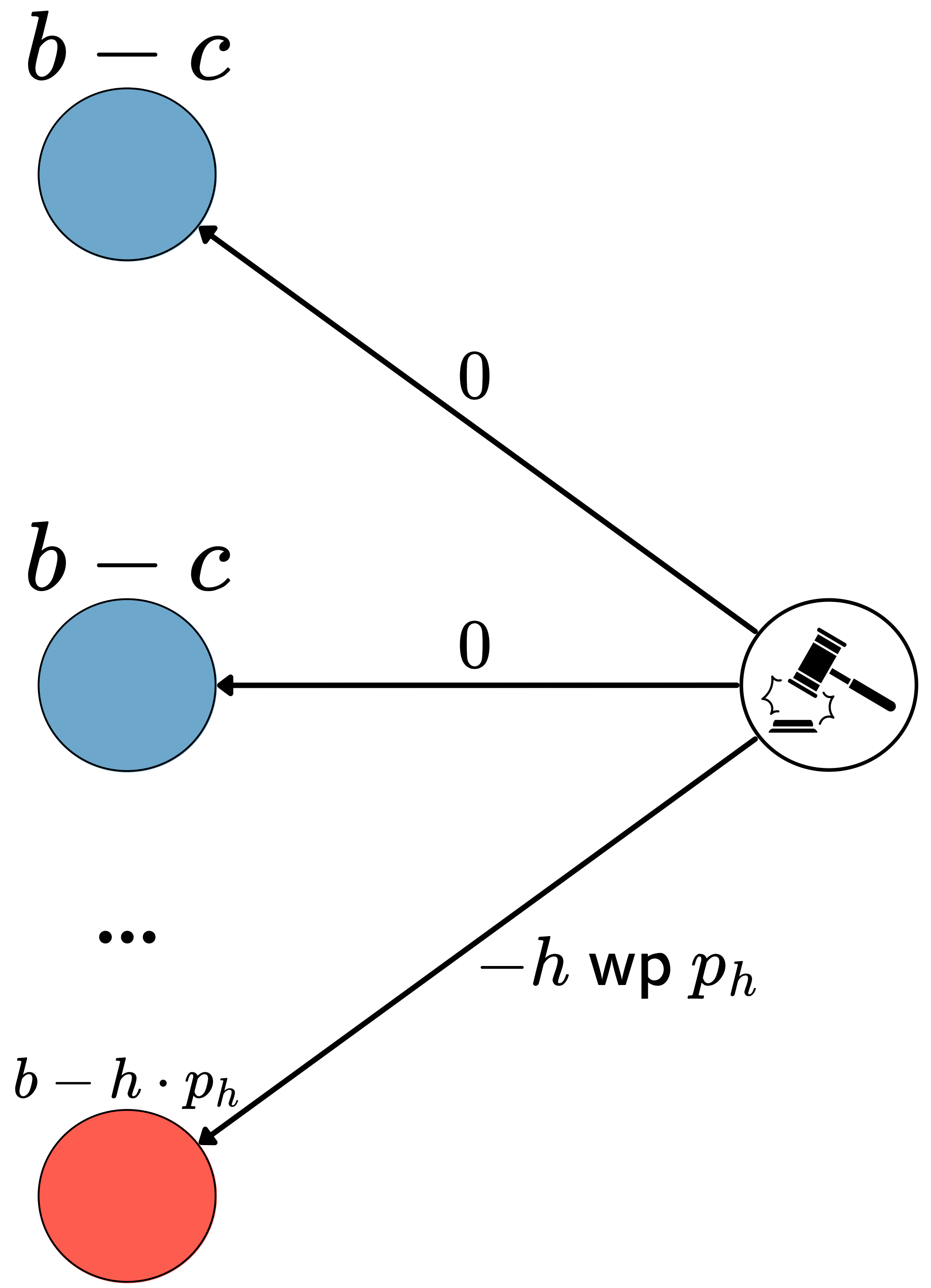
$$b - h \cdot p_h.$$

Cooperating pays off when:

$$b - c > b - h \cdot p_h \qquad \text{iff}$$

$$h \cdot p_h > c.$$

One immediate lesson is that as the authority gets less effective at enacting punishment (i.e., as $p_h \to 0$), the penalty (h) has to grow…



EREZ YOELI

The Leviathan can be an institution, like the state.

EREZ YOELI

The Leviathan can be an institution, like the state.

MOSHE HOFFMAN

But it can also be a person, like the captain of a ship.

Yoeli, E., & Hoffman, M. (2022). *Hidden Games: The Surprising Power of Game Theory to Explain Irrational Human Behavior*. Basic Books.

PIRATE CODES

Pirate crews had very strict codes of conduct, with high penalties.

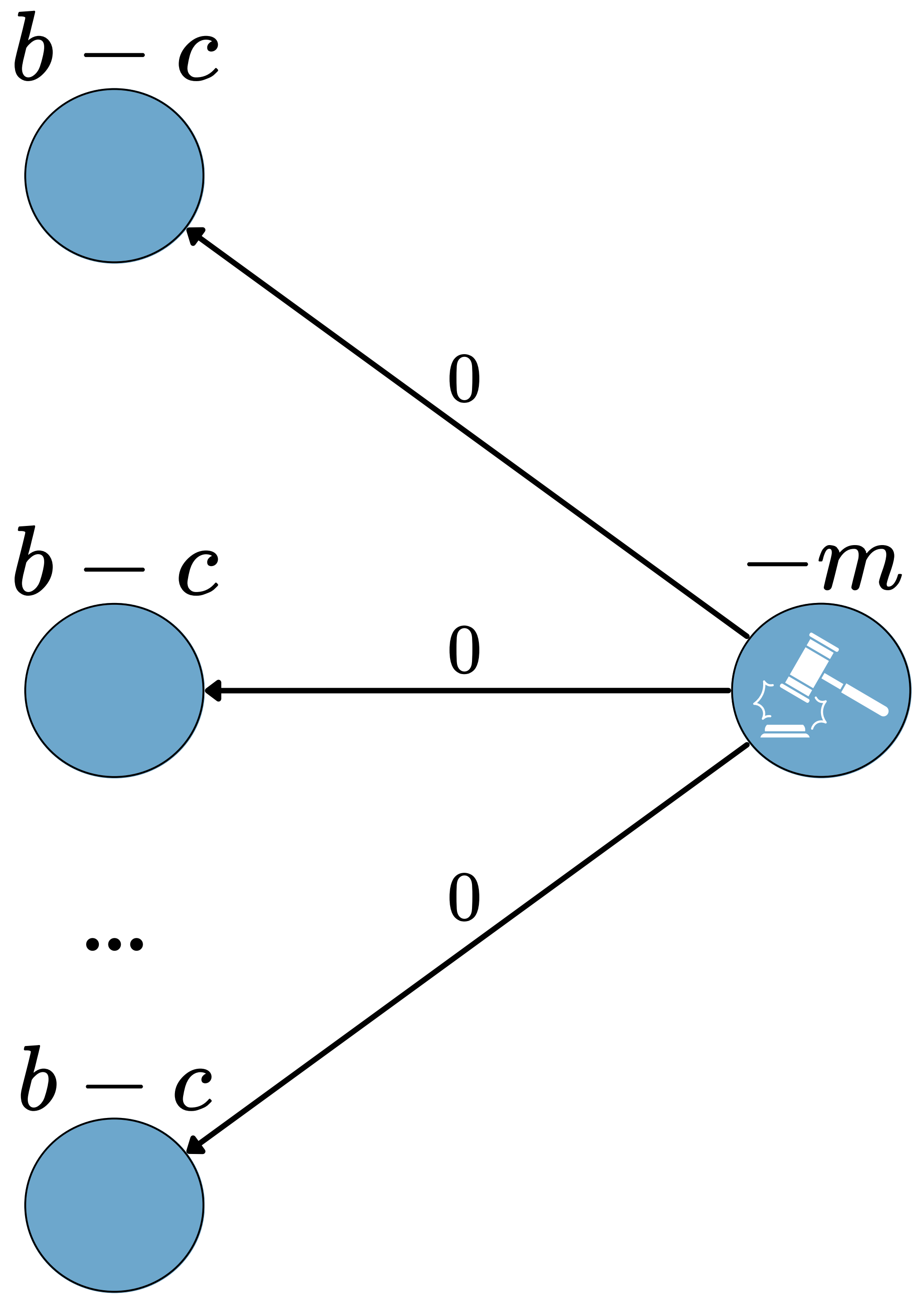
Punishment was generally decided by officers.

Yoeli, E., & Hoffman, M. (2022). *Hidden Games: The Surprising Power of Game Theory to Explain Irrational Human Behavior*. Basic Books.

What of the Leviathan itself? Punishment is costly…

COSTLY PUNISHMENT

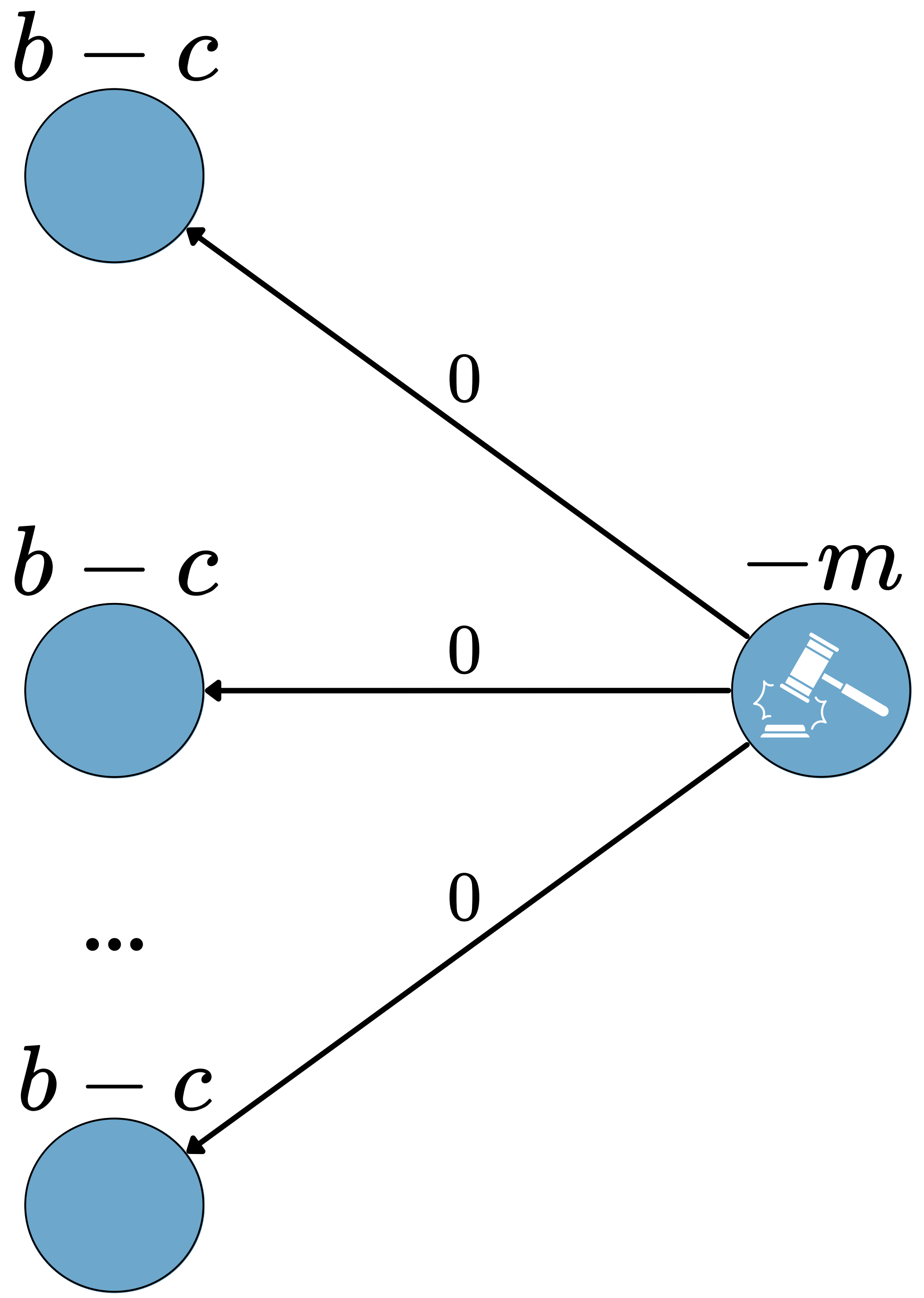
Suppose punishment incurs monitoring costs m > 0.



COSTLY PUNISHMENT

Suppose punishment incurs monitoring costs $m > 0$.

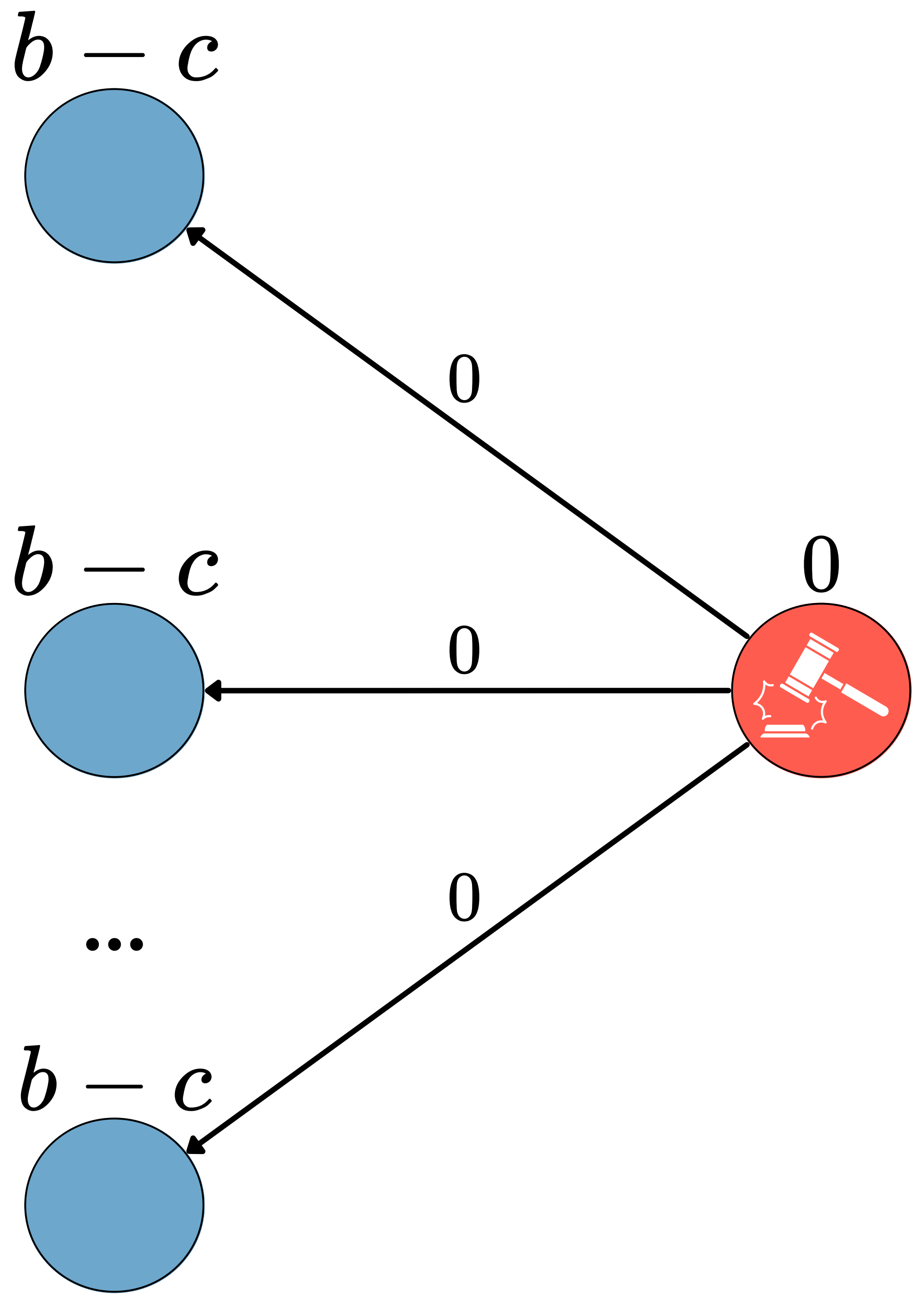
This creates a problem, because now the Leviathan has an incentive to avoid punishing.



COSTLY PUNISHMENT

Suppose punishment incurs monitoring costs $m > 0$.

This creates a problem, because now the Leviathan has an incentive to avoid punishing.

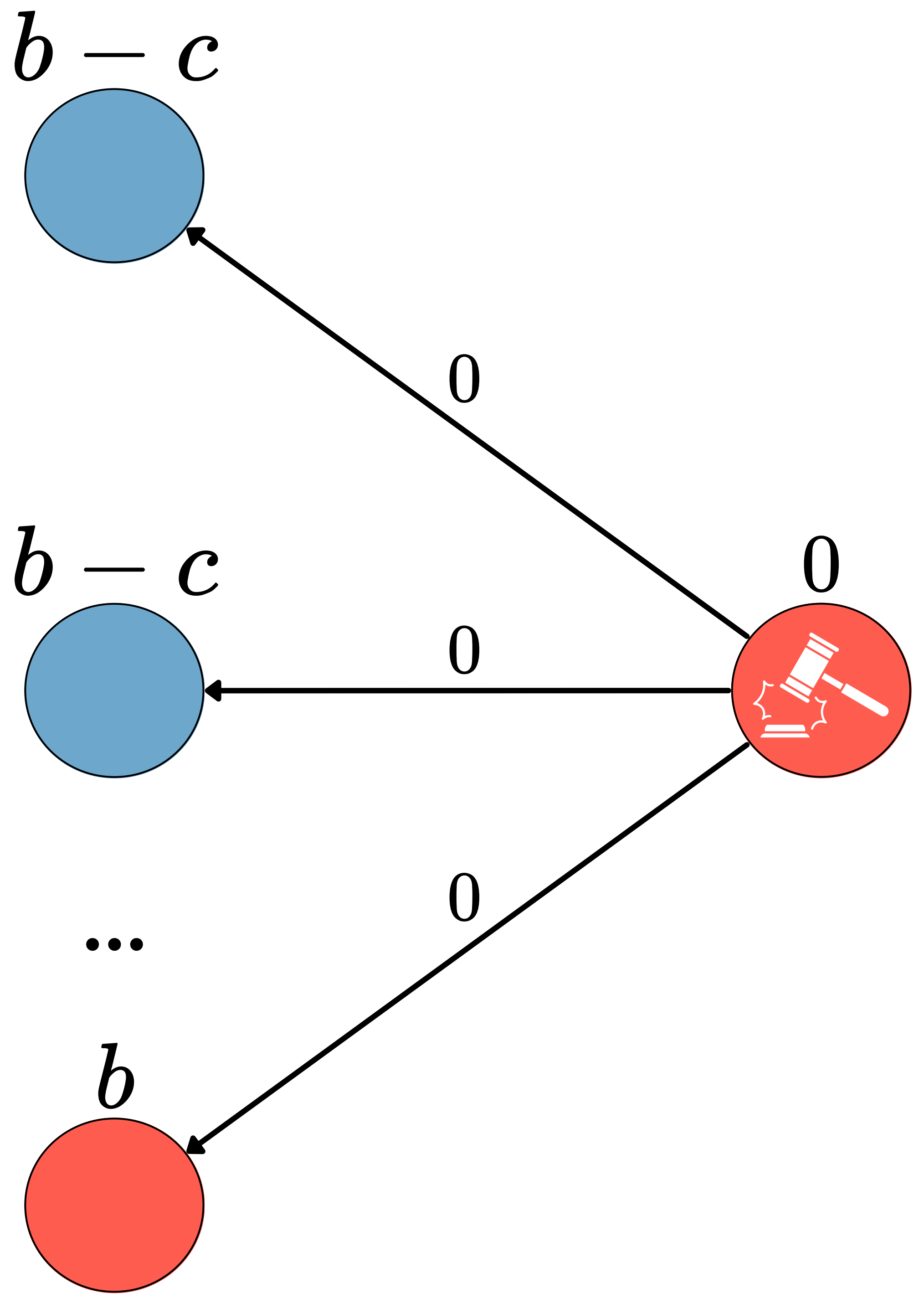


COSTLY PUNISHMENT

Suppose punishment incurs monitoring costs $m > 0$.

This creates a problem, because now the Leviathan has an incentive to avoid punishing.

Which means regular players can freeride without consequences, and cooperation unravels.

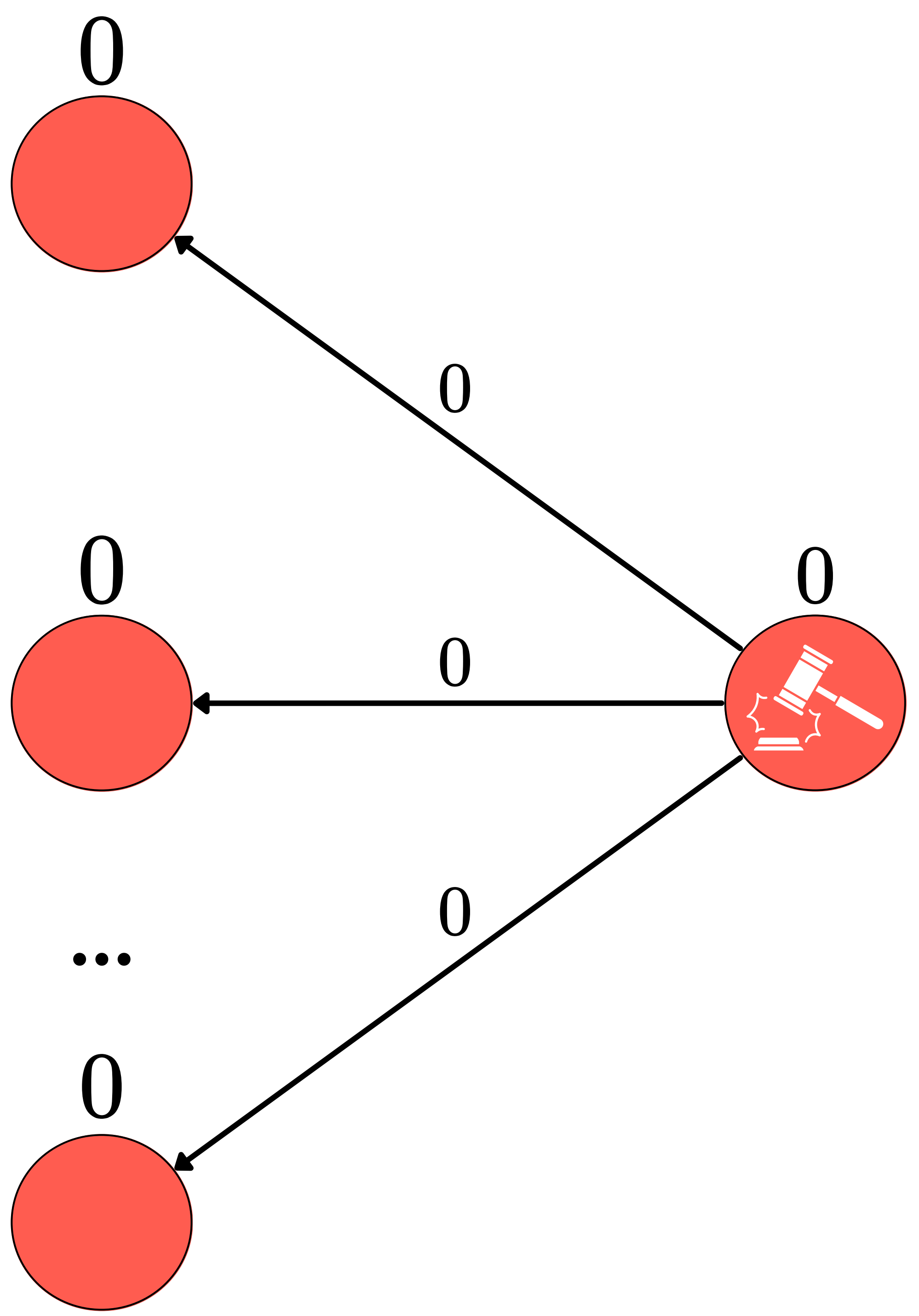


COSTLY PUNISHMENT

Suppose punishment incurs monitoring costs $m > 0$.

This creates a problem, because now the Leviathan has an incentive to avoid punishing.

Which means regular players can freeride without consequences, and cooperation unravels.

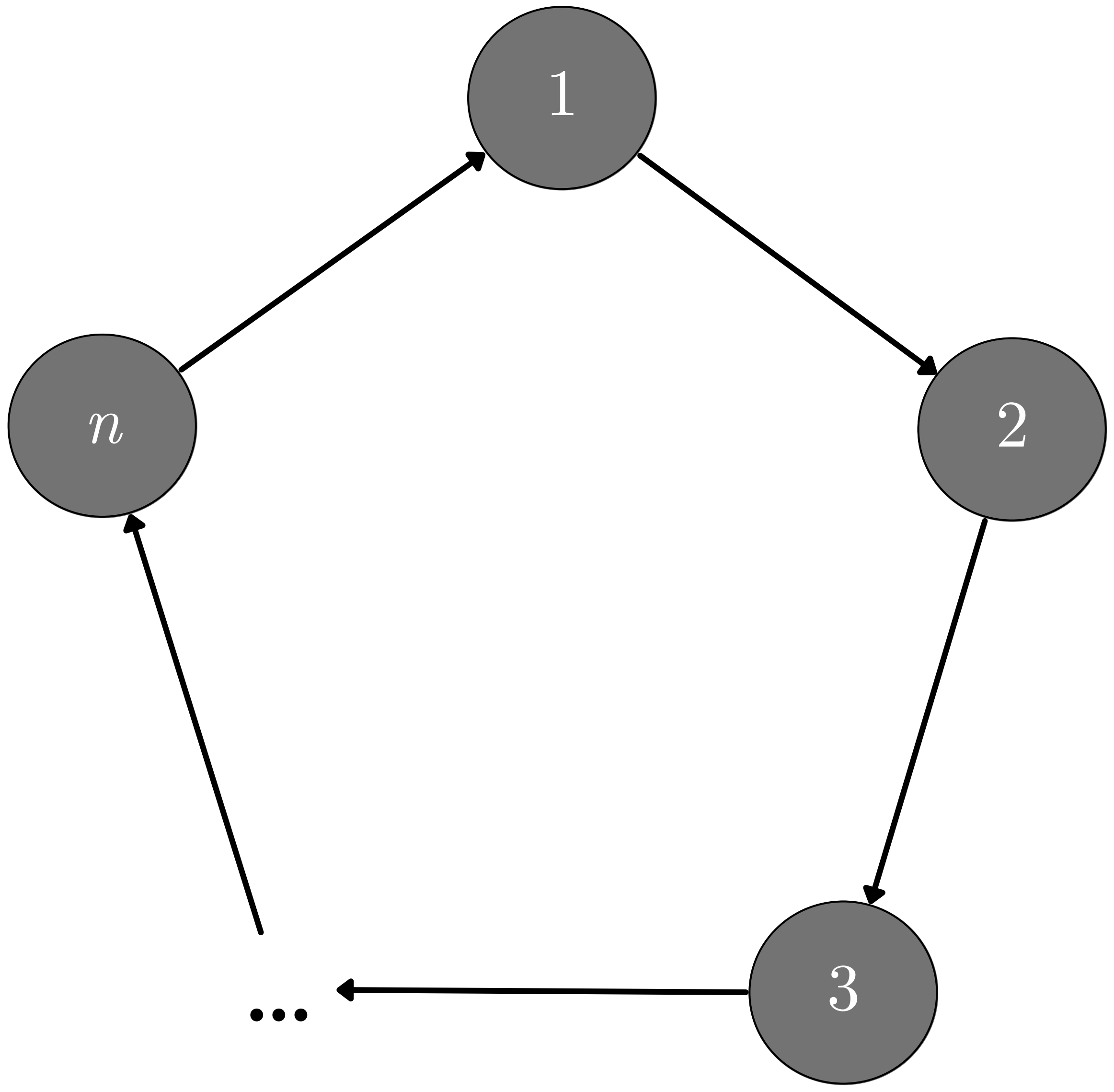


Punishing is like contributing towards a public good in its own right.

Punishing is like contributing towards a public good in its own right. And is also subject to the free rider problem…

PUBLIC GOODS WITH THIRD PARTY PUNISHMENT

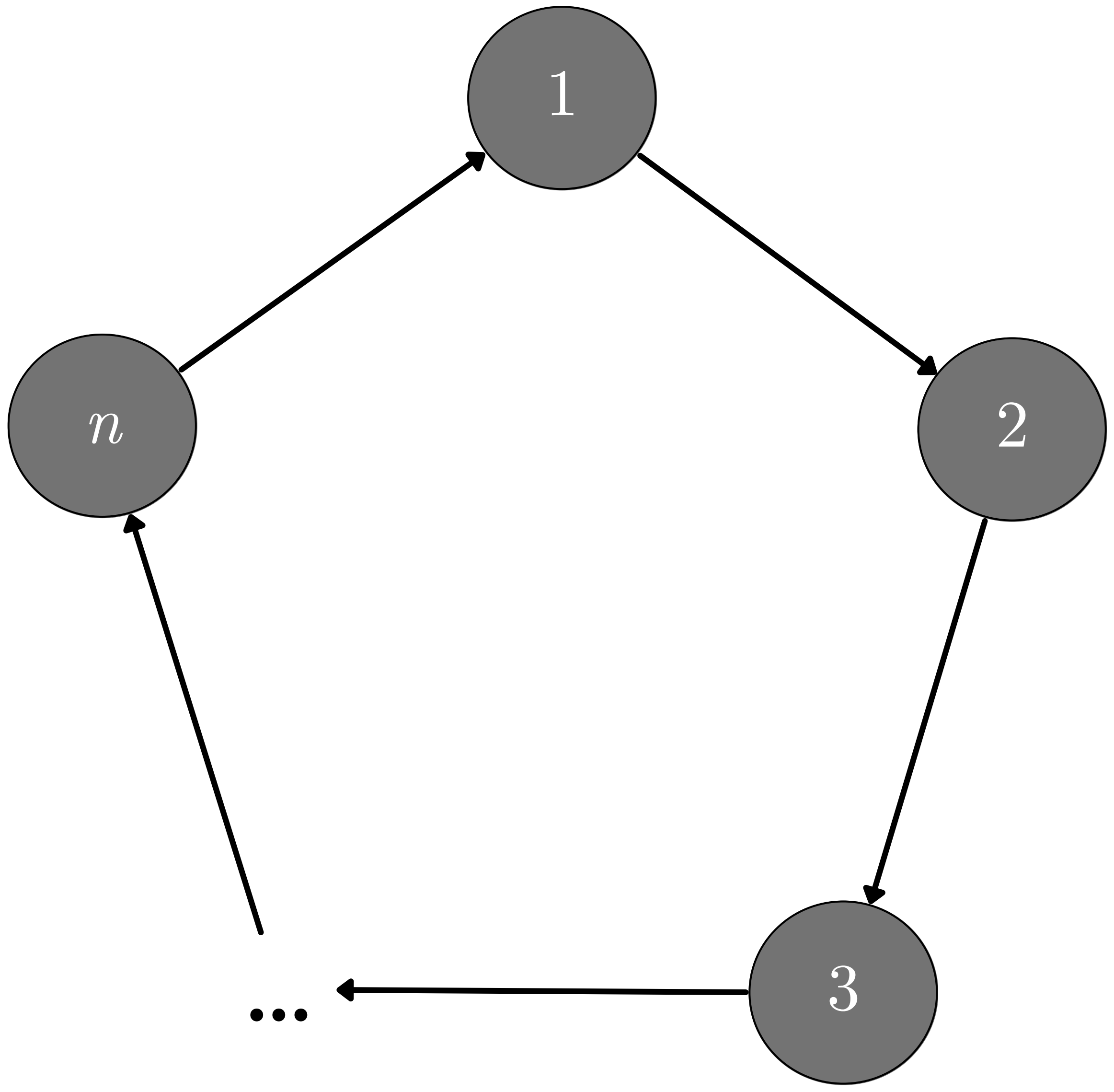
Suppose players monitor each other in a ring,
such that i monitors $i + 1$.



PUBLIC GOODS WITH THIRD PARTY PUNISHMENT

Suppose players monitor each other in a ring, such that i monitors $i+1$.

The game occurs in rounds, with a probability δ of a new round.

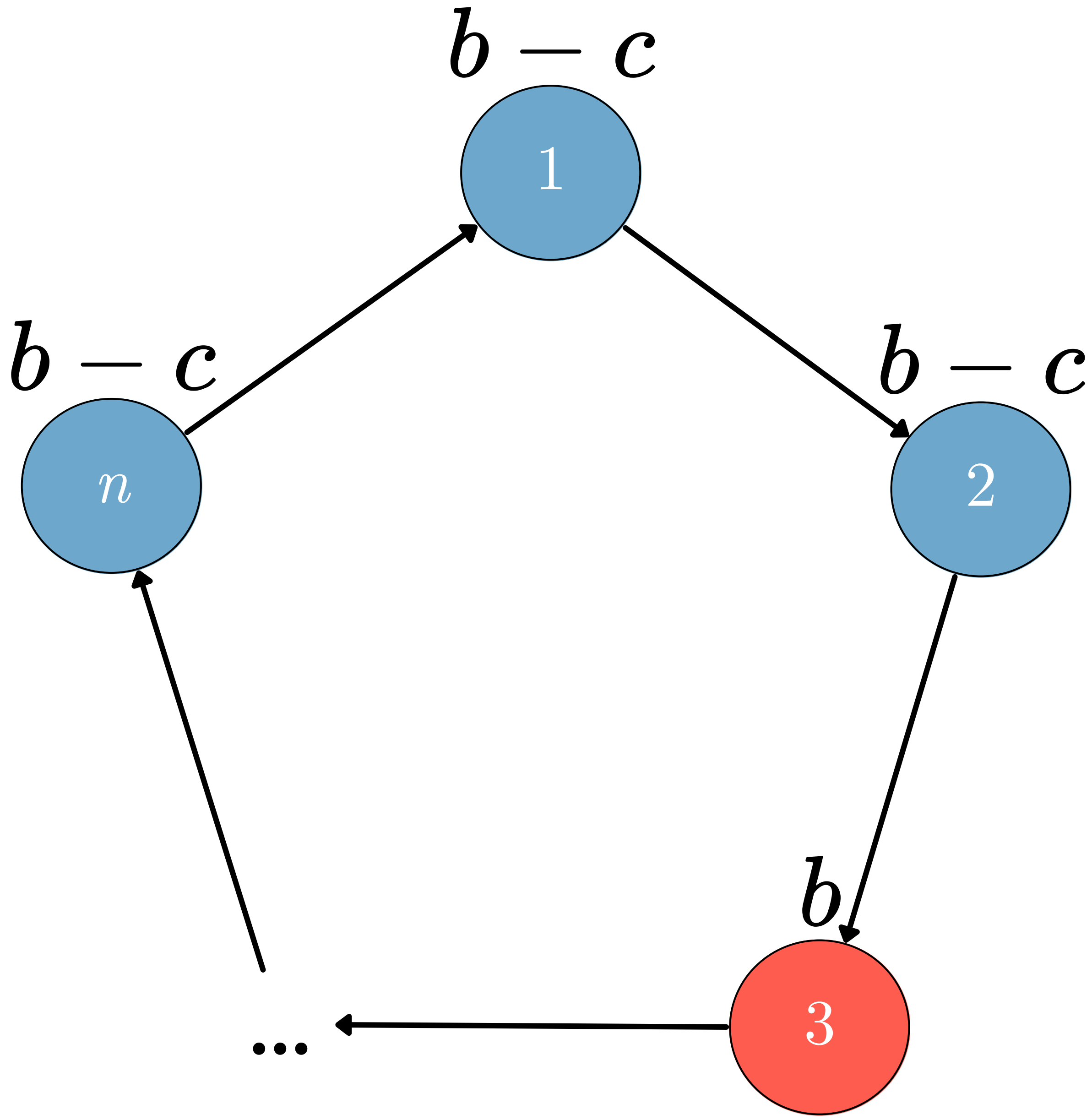


PUBLIC GOODS WITH THIRD PARTY PUNISHMENT

Suppose players monitor each other in a ring, such that i monitors $i+1$.

The game occurs in rounds, with a probability δ of a new round.

At the first round, players submit their contributions.



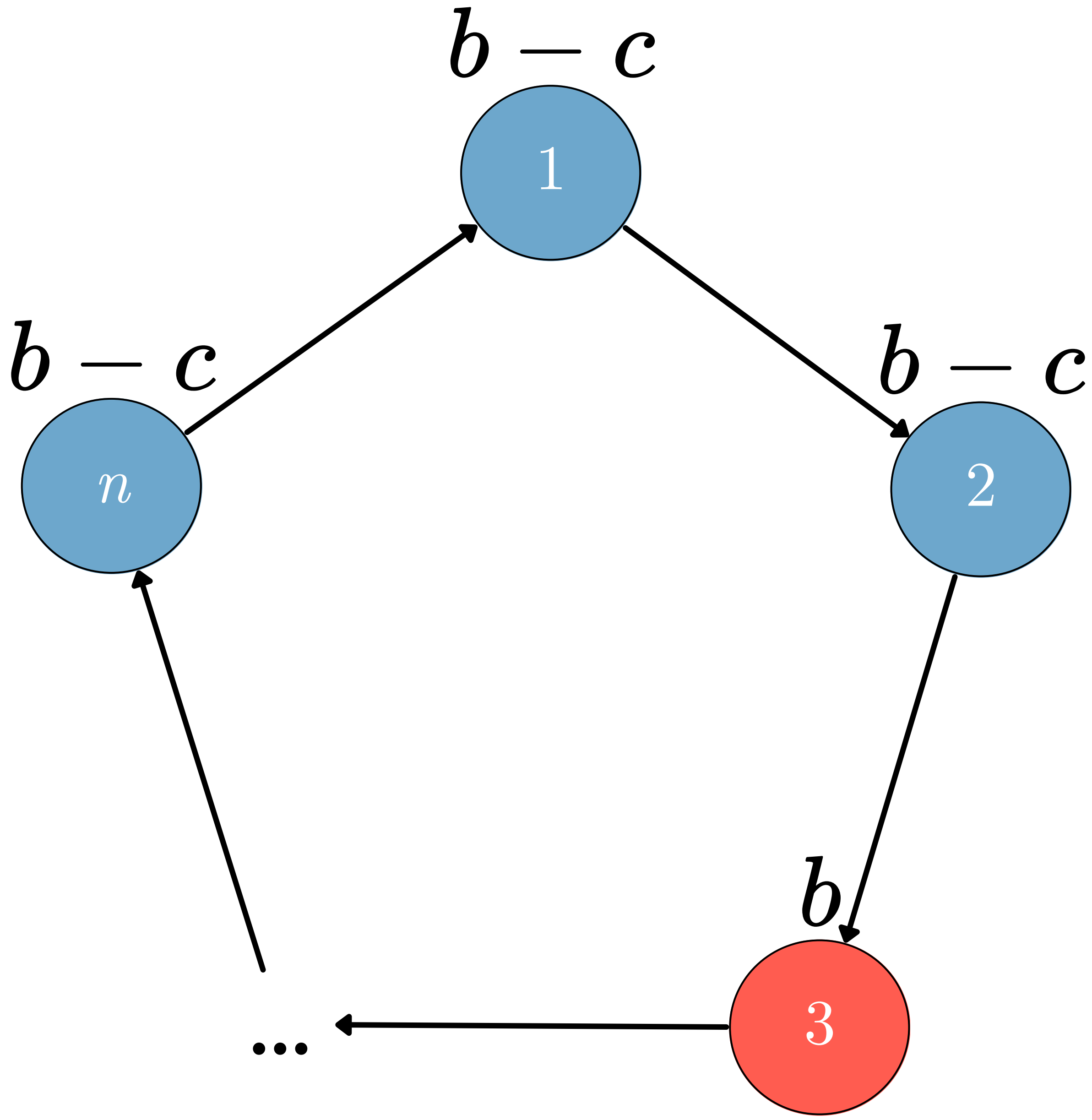
PUBLIC GOODS WITH THIRD PARTY PUNISHMENT

Suppose players monitor each other in a ring, such that i monitors $i+1$.

The game occurs in rounds, with a probability δ of a new round.

At the first round, players submit their contributions.

At subsequent rounds, players can punish the players they monitor.



PUBLIC GOODS WITH THIRD PARTY PUNISHMENT

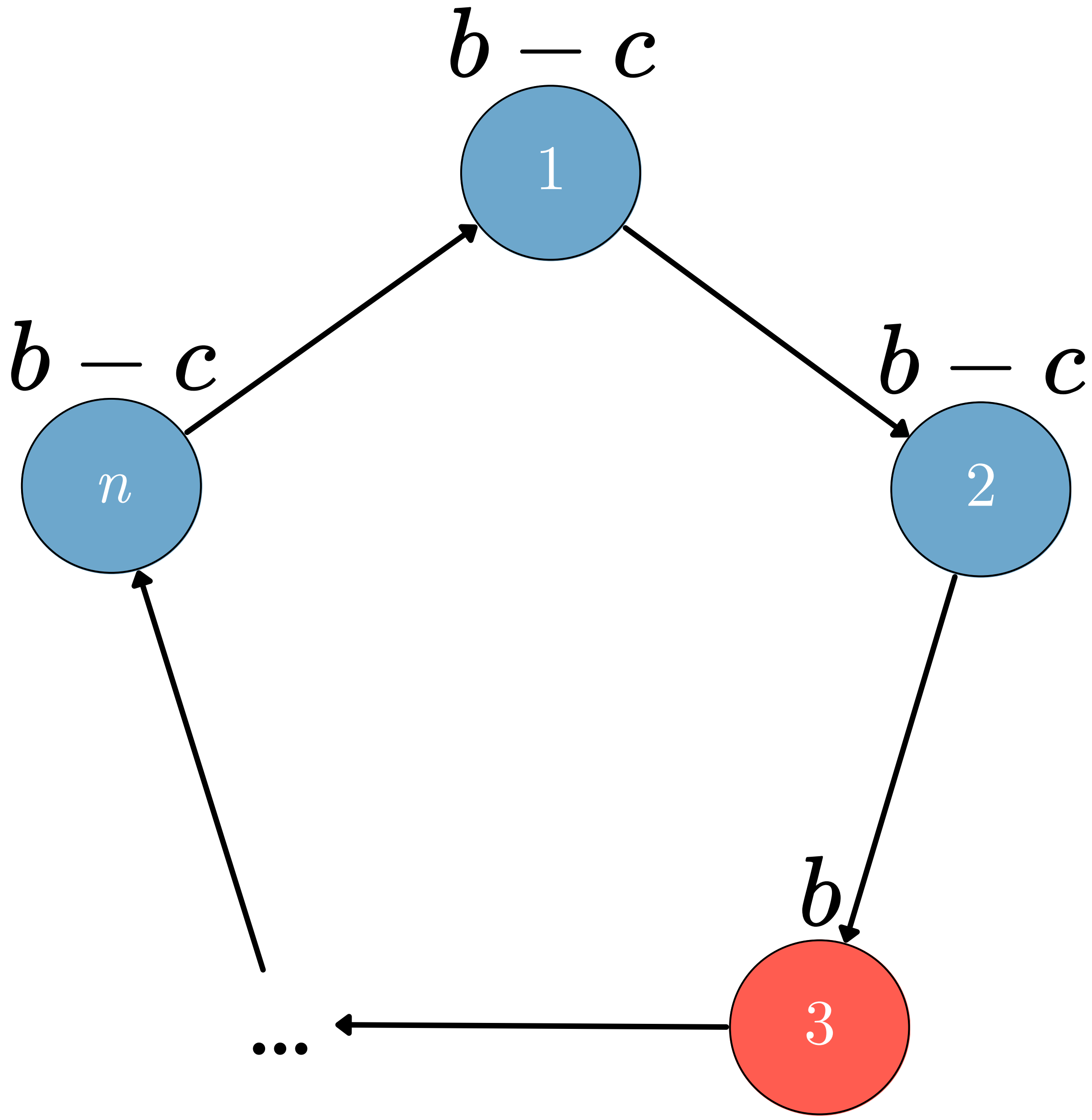
Suppose players monitor each other in a ring, such that i monitors $i+1$.

The game occurs in rounds, with a probability δ of a new round.

At the first round, players submit their contributions.

At subsequent rounds, players can punish the players they monitor.

Punishment incurs cost m to the punishing agent and h to the punished agent.



PUBLIC GOODS WITH THIRD PARTY PUNISHMENT

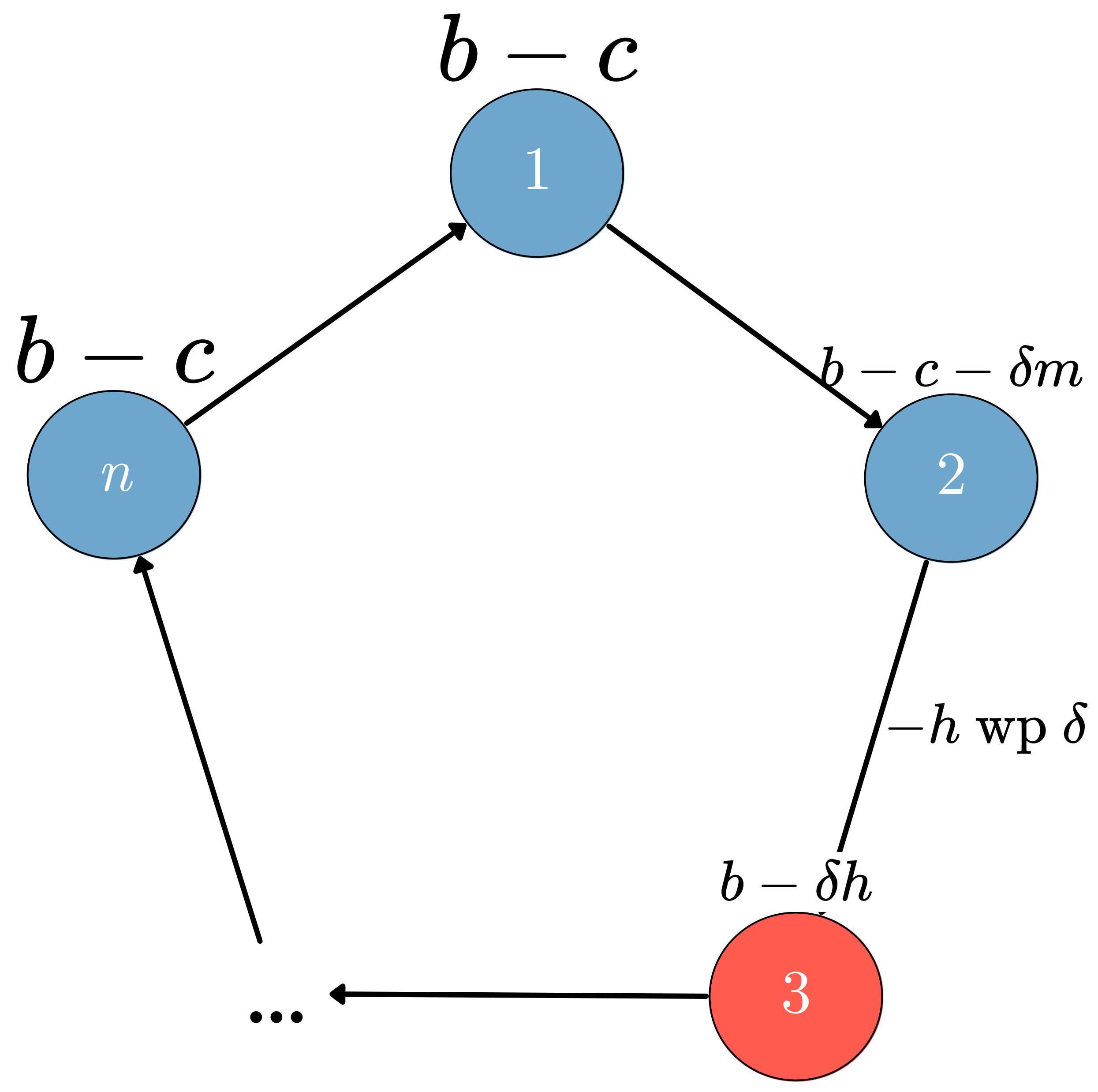
Suppose players monitor each other in a ring, such that i monitors $i+1$.

The game occurs in rounds, with a probability δ of a new round.

At the first round, players submit their contributions.

At subsequent rounds, players can punish the players they monitor.

Punishment incurs cost m to the punishing agent and h to the punished agent.



Player 2 punishes player 3 at round 2

PUBLIC GOODS WITH THIRD PARTY PUNISHMENT

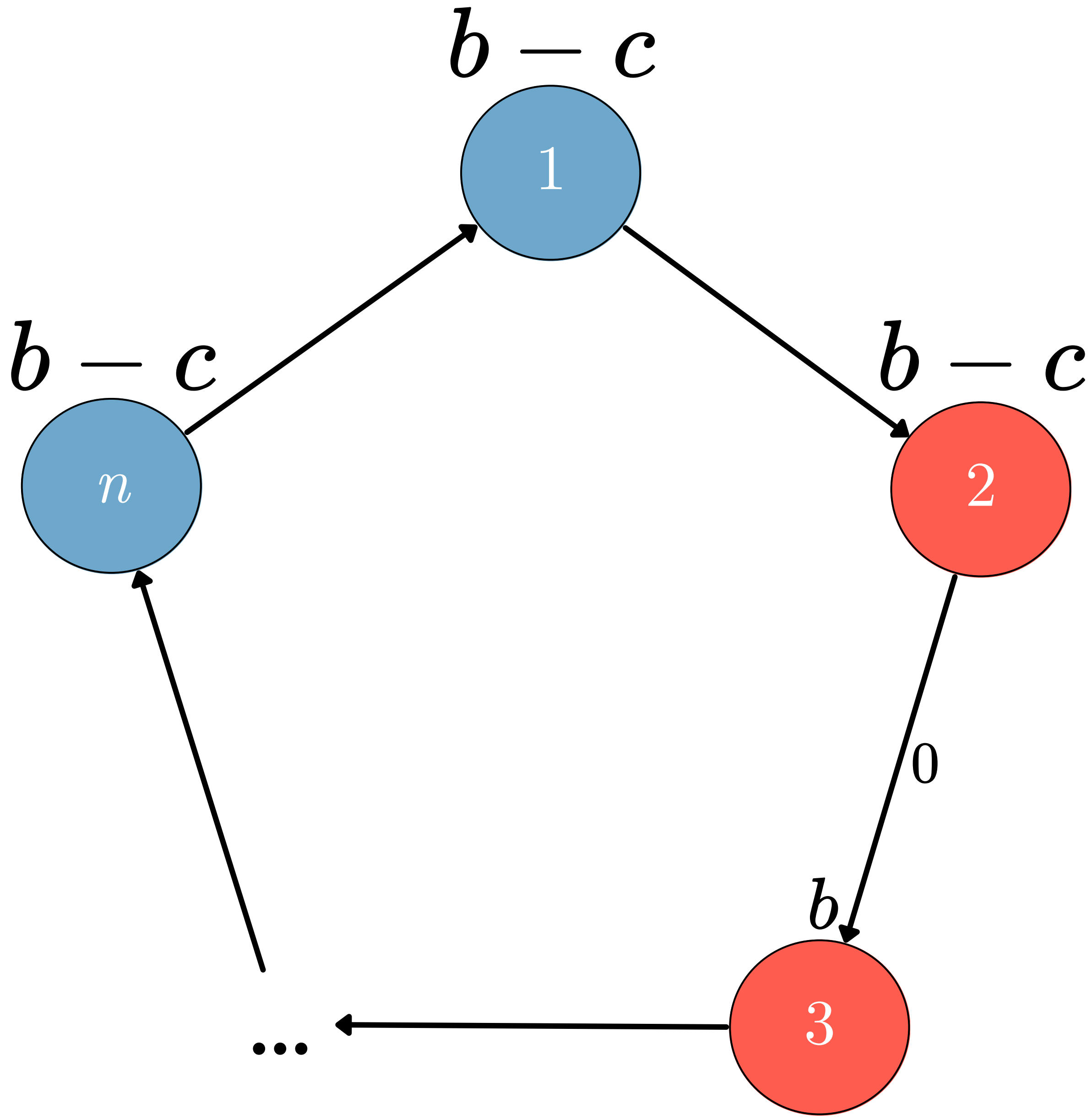
Suppose players monitor each other in a ring, such that i monitors $i+1$.

The game occurs in rounds, with a probability δ of a new round.

At the first round, players submit their contributions.

At subsequent rounds, players can punish the players they monitor.

Punishment incurs cost m to the punishing agent and h to the punished agent.



Player 2 doesn't punish player 3 at round 2

PUBLIC GOODS WITH THIRD PARTY PUNISHMENT

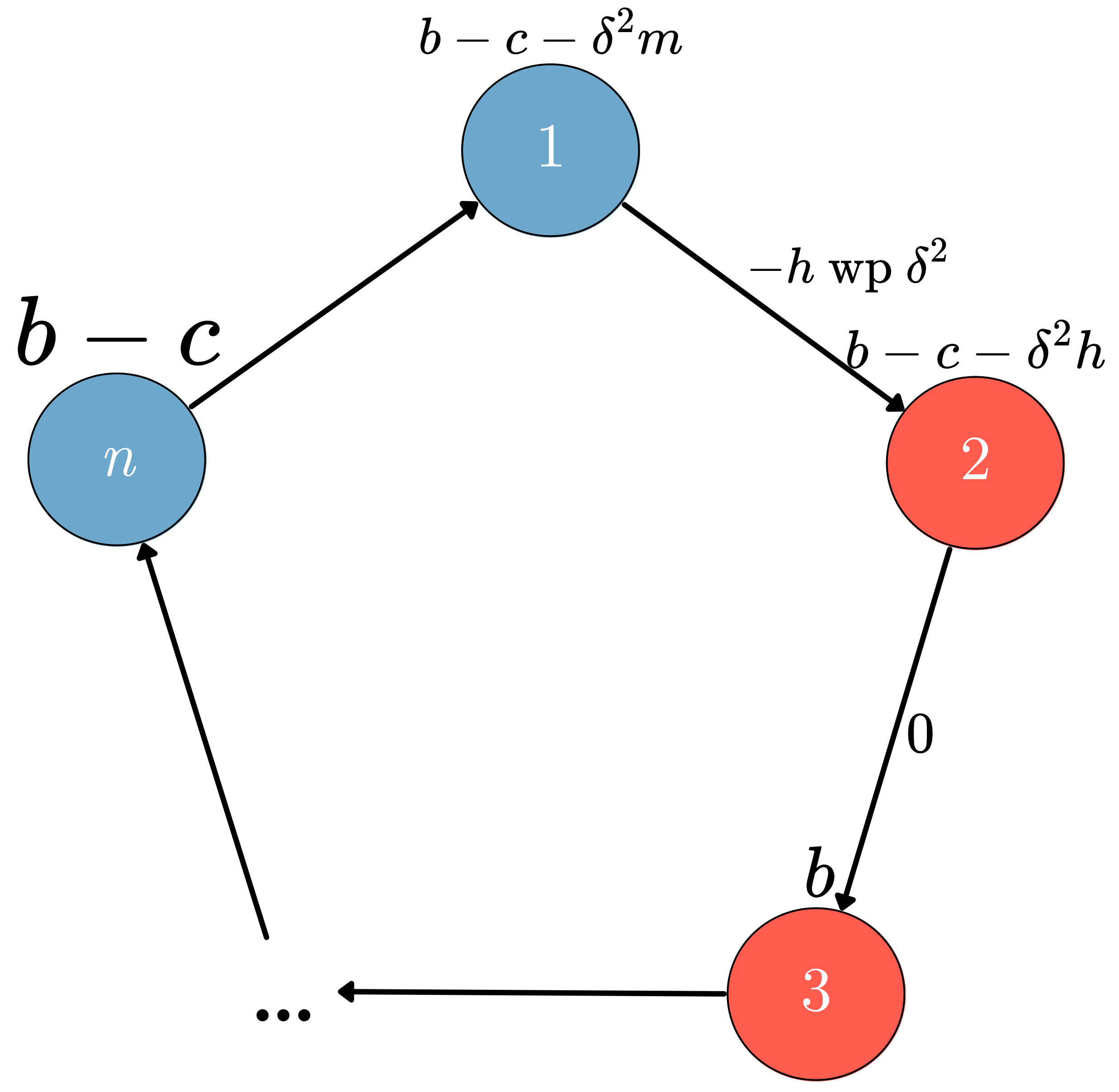
Suppose players monitor each other in a ring, such that i monitors $i+1$.

The game occurs in rounds, with a probability δ of a new round.

At the first round, players submit their contributions.

At subsequent rounds, players can punish the players they monitor.

Punishment incurs cost m to the punishing agent and h to the punished agent.



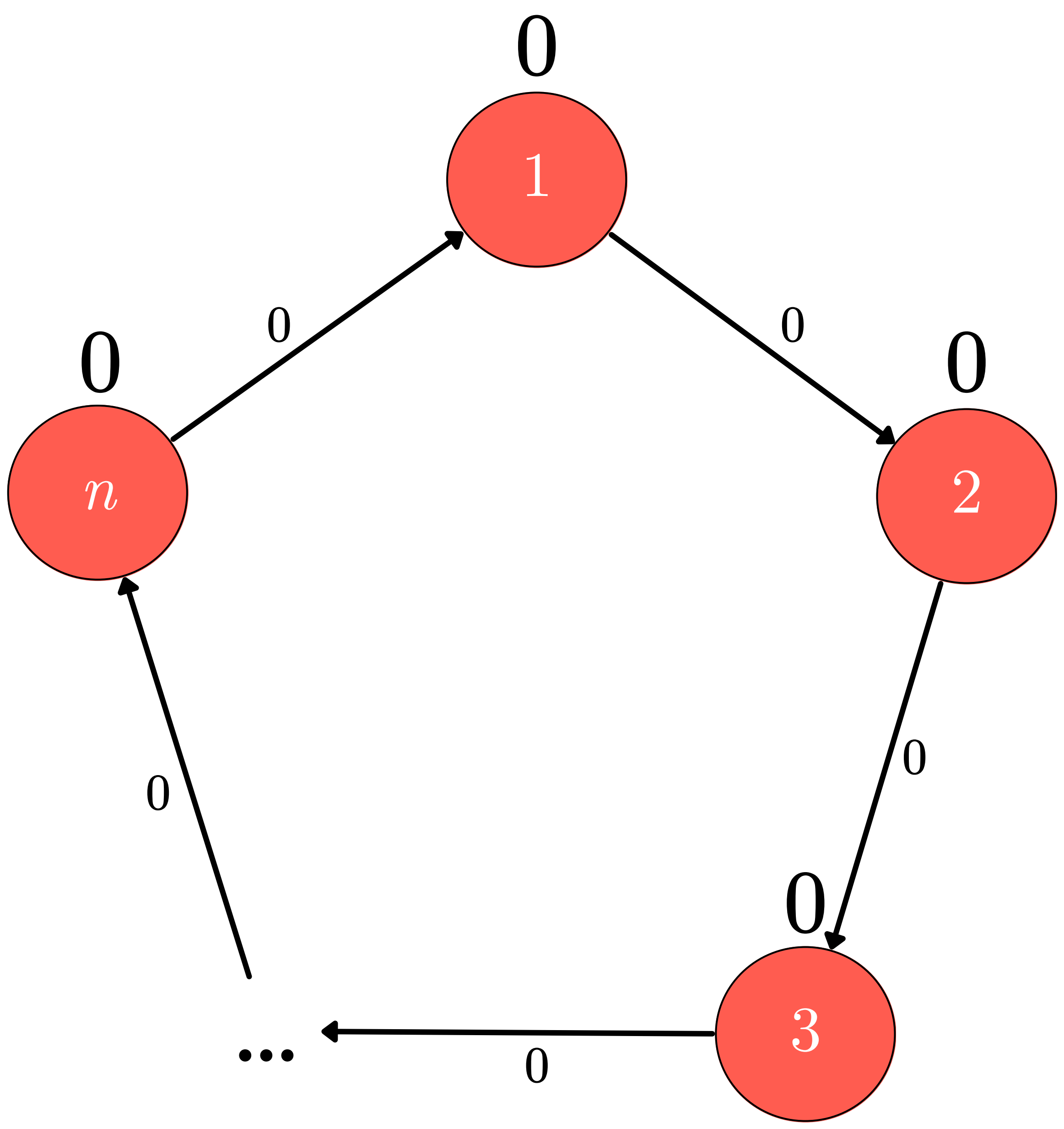
Player 2 doesn't punish player 3 at round 2, but gets punished by player 1 at round 3

Equilibria?

NO CONTRIBUTION, NO PUNISHMENT

Suppose no one contributes, and no one punishes.

This is an equilibrium, since any deviation incurs costs and no benefits.*

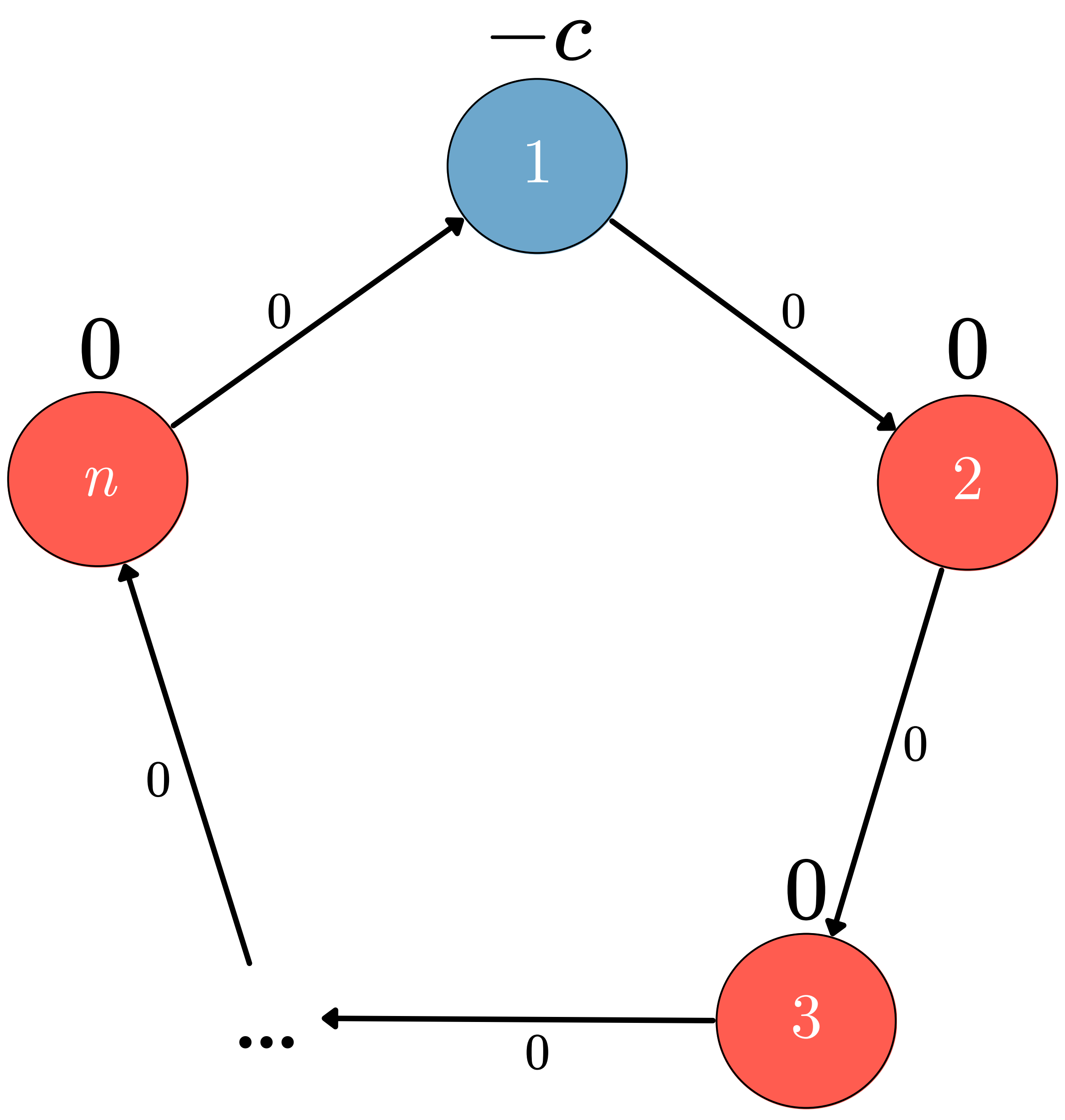


*Assuming one contribution is not sufficient to generate the public good.

NO CONTRIBUTION, NO PUNISHMENT

Suppose no one contributes, and no one punishes.

This is an equilibrium, since any deviation incurs costs and no benefits.*



Player 1 switches to contributing

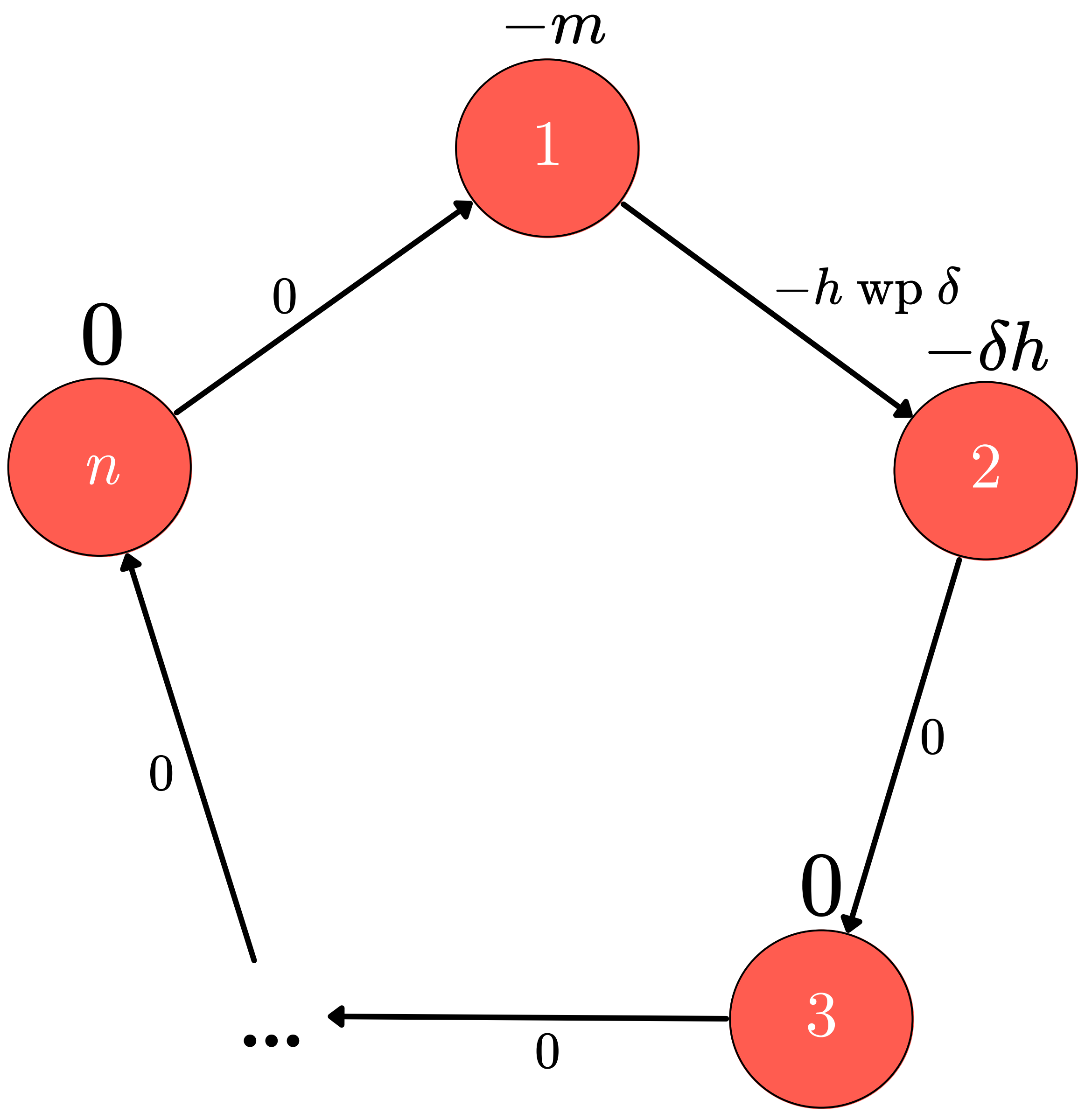
*Assuming one contribution is not sufficient to generate the public good.

NO CONTRIBUTION, NO PUNISHMENT

Suppose no one contributes, and no one punishes.

This is an equilibrium, since any deviation incurs costs and no benefits.*

*Assuming one contribution is not sufficient to generate the public good.



Player 1 switches to punishing player 2

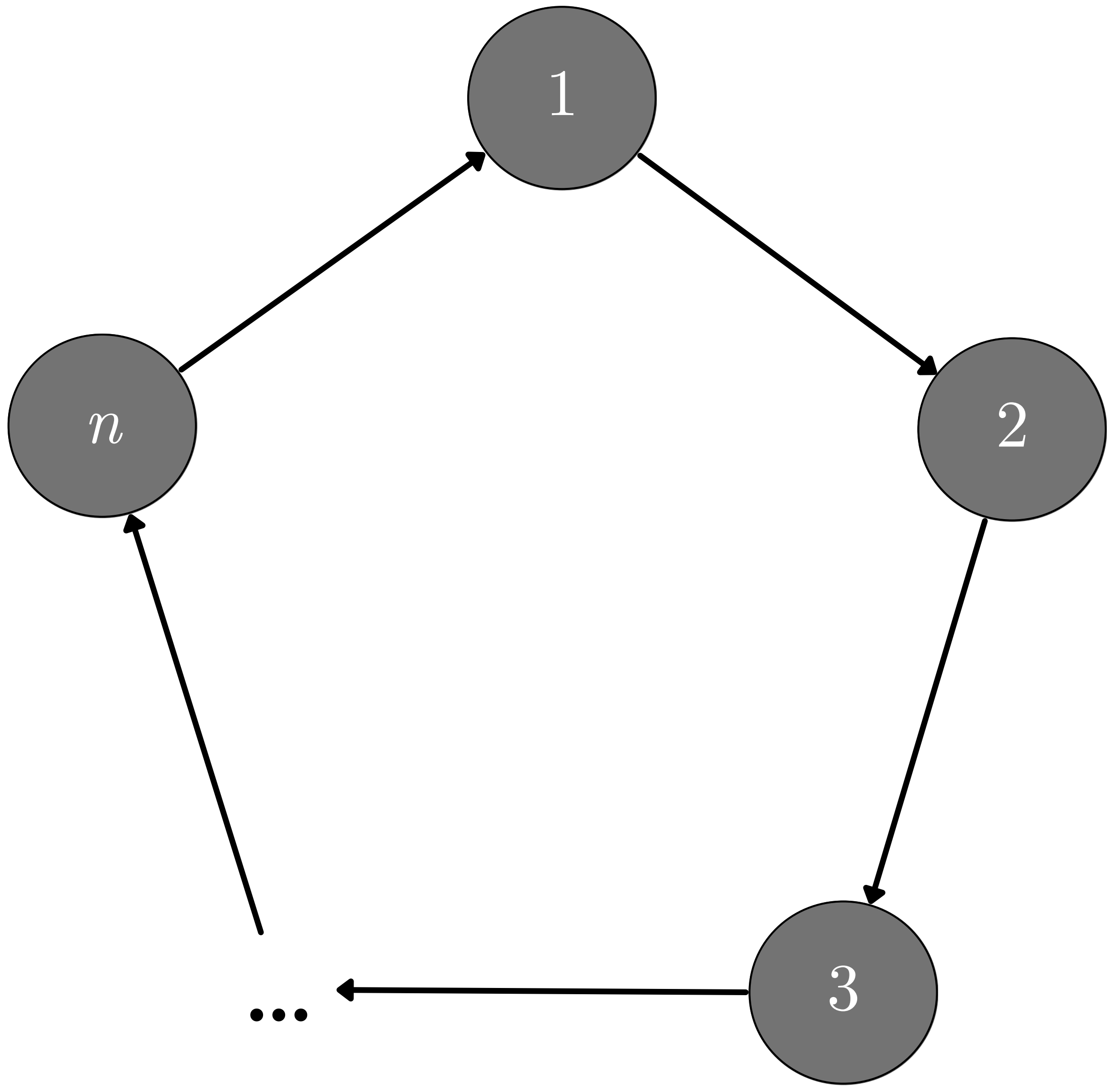
This is a very bad equilibrium. Anything else?

CONTRIBUTE AND PUNISH

Suppose players have different strategies: start by contributing, and punish any transgressions of the monitored agent from the previous round.

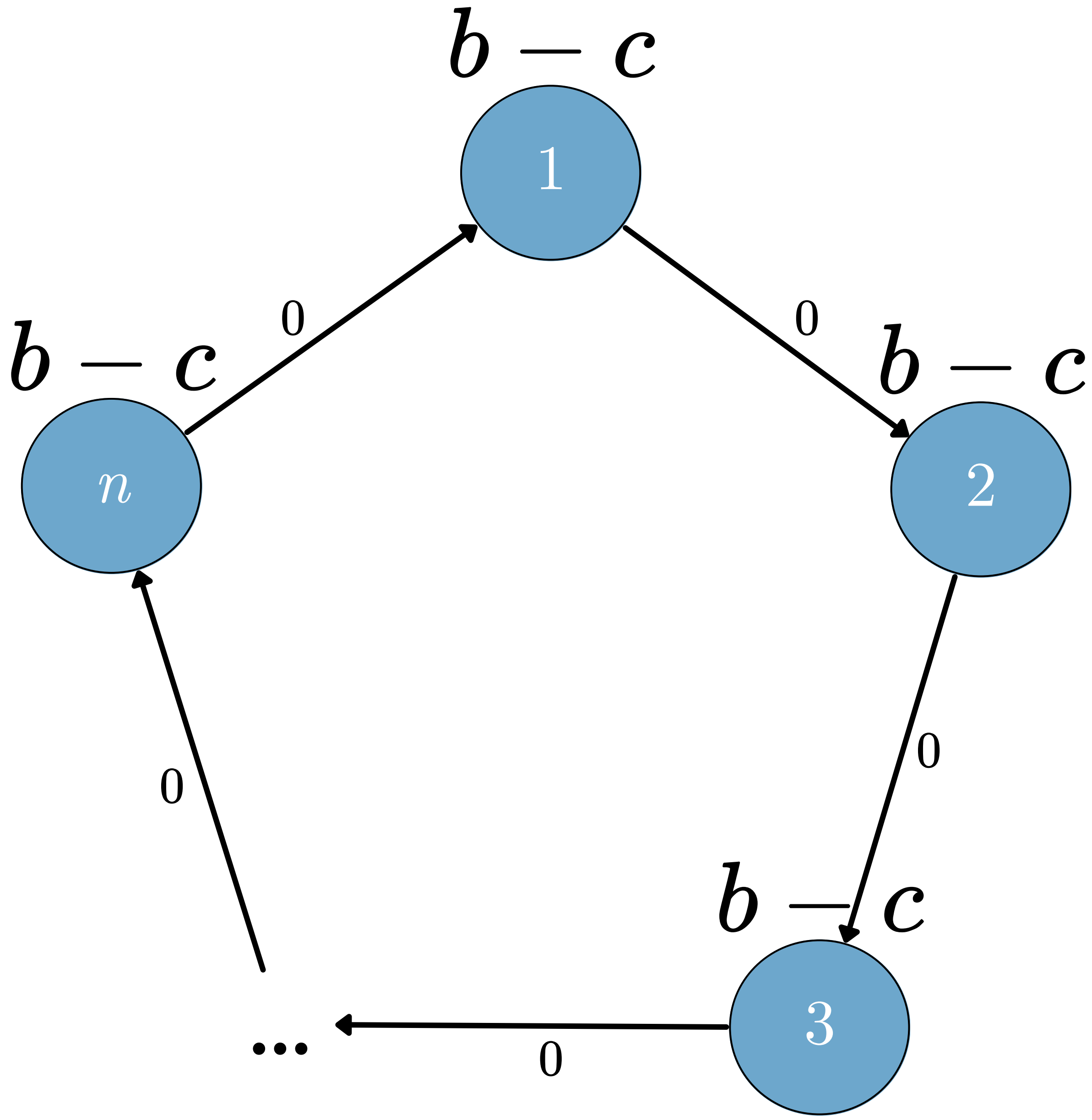
A transgression at round 1 is to not contribute.

A transgression at round *t*+1 is failure to punish transgressions at round *t*.



CONTRIBUTE AND PUNISH

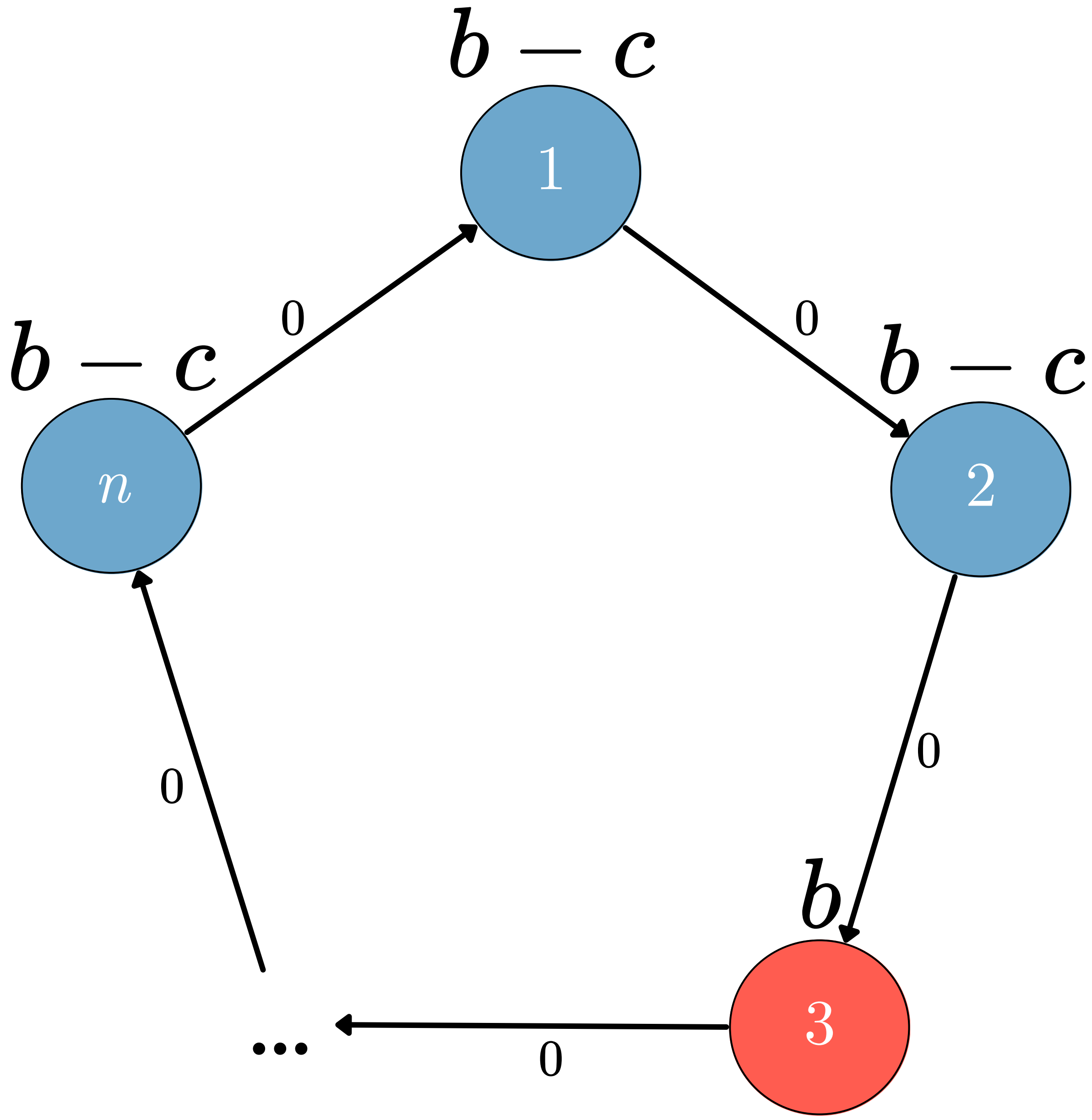
With on-strategy play, everyone contributes (and no punishment necessary).



CONTRIBUTE AND PUNISH

With on-strategy play, everyone contributes (and no punishment necessary).

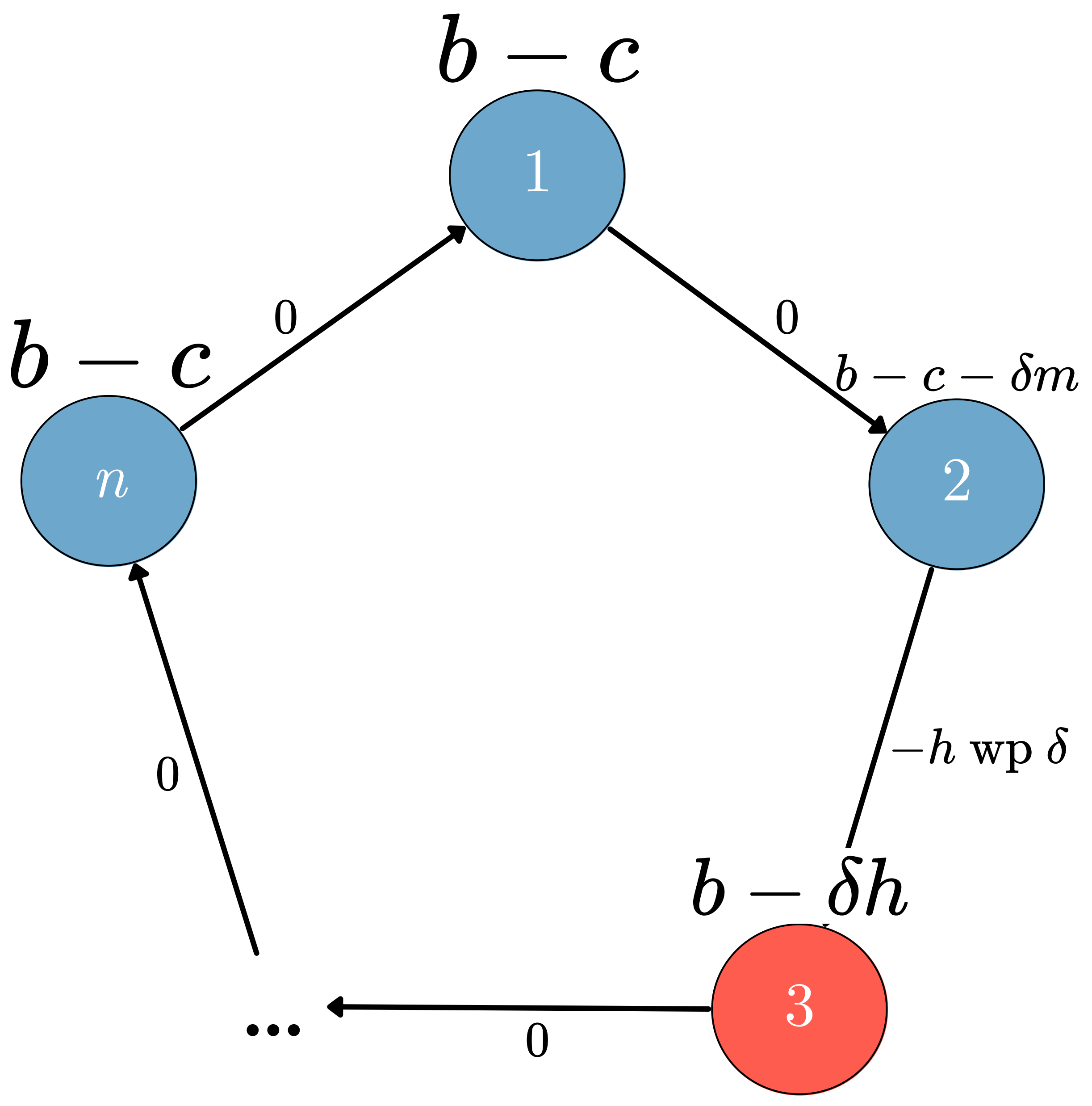
If one player deviates at the first round by not contributing, they get punished by the player upstream.



CONTRIBUTE AND PUNISH

With on-strategy play, everyone contributes (and no punishment necessary).

If one player deviates at the first round by not contributing, they get punished by the player upstream.



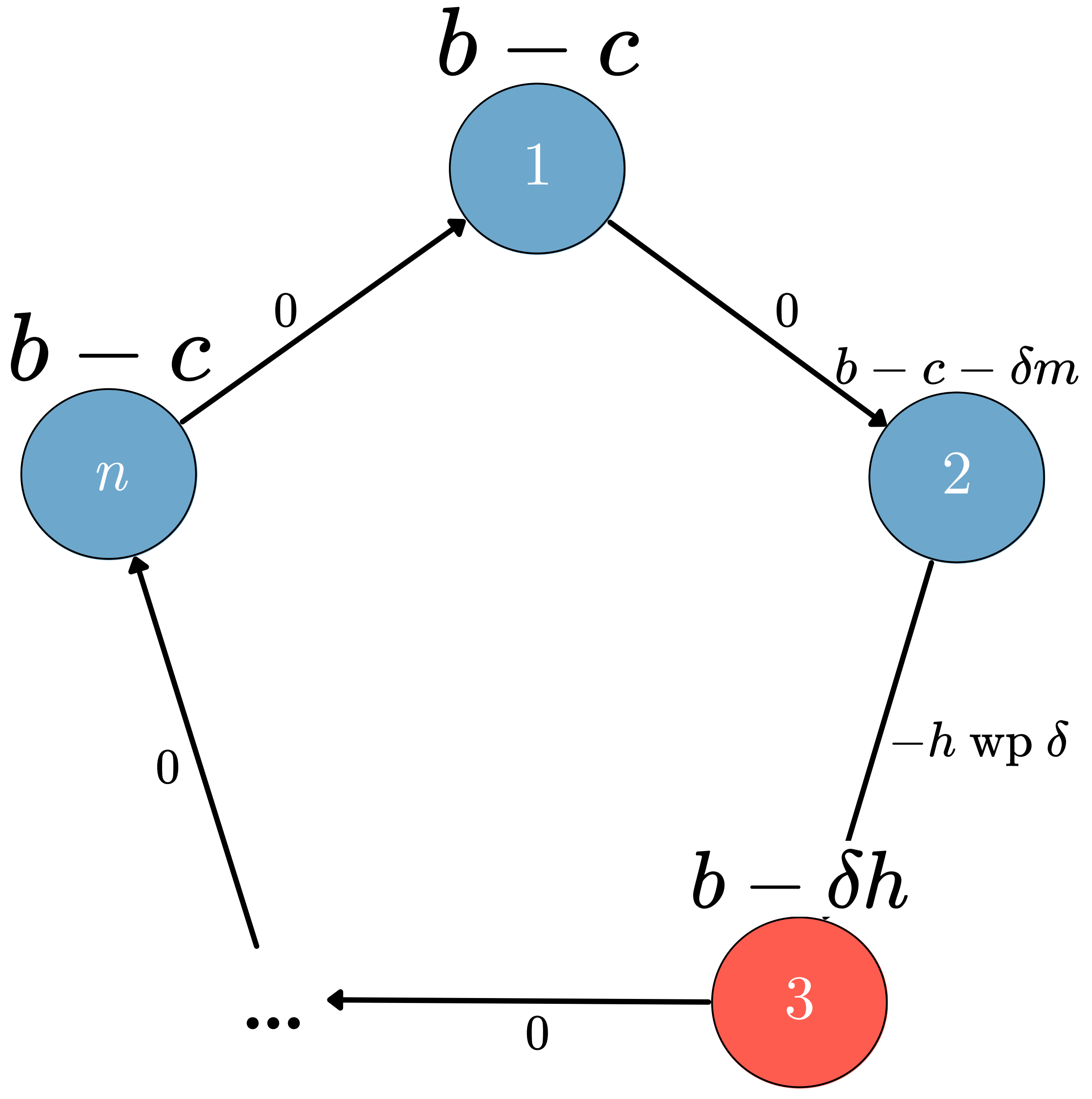
Player 2 punishes player 3 at round 2

CONTRIBUTE AND PUNISH

With on-strategy play, everyone contributes (and no punishment necessary).

If one player deviates at the first round by not contributing, they get punished by the player upstream. This does not pay off as long as:

$$b - c > b - \delta h \quad \text{iff}$$
$$\delta h > c.$$



Player 2 punishes player 3 at round 2

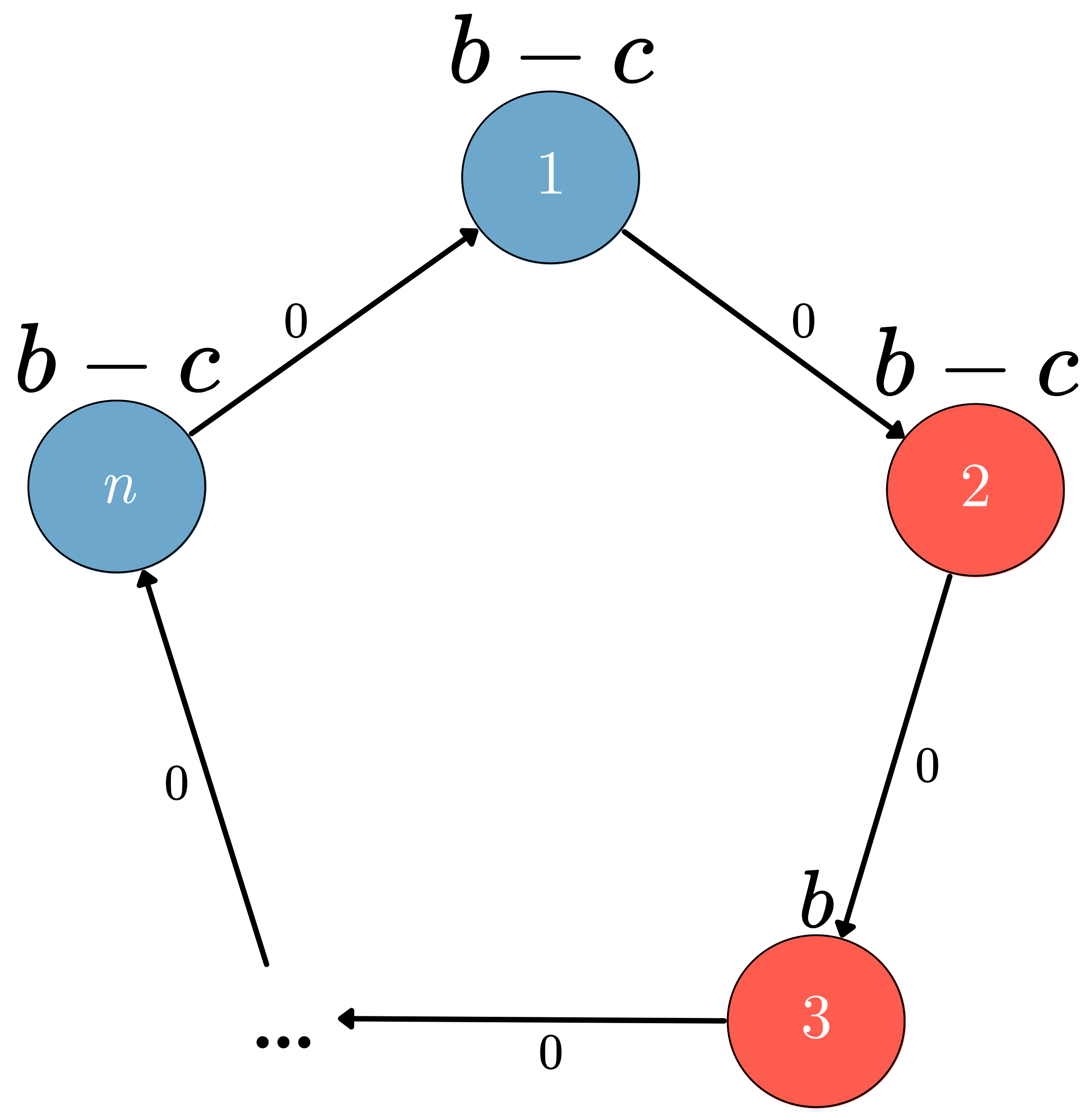
CONTRIBUTE AND PUNISH

With on-strategy play, everyone contributes (and no punishment necessary).

If one player deviates at the first round by not contributing, they get punished by the player upstream. This does not pay off as long as:

$$b - c > b - \delta h \quad \text{iff}$$
$$\delta h > c.$$

If some player deviates at a round i by not punishing a defector, they get punished at the next round by the player upstream.



Player 2 doesn't punish player 3 at round 2

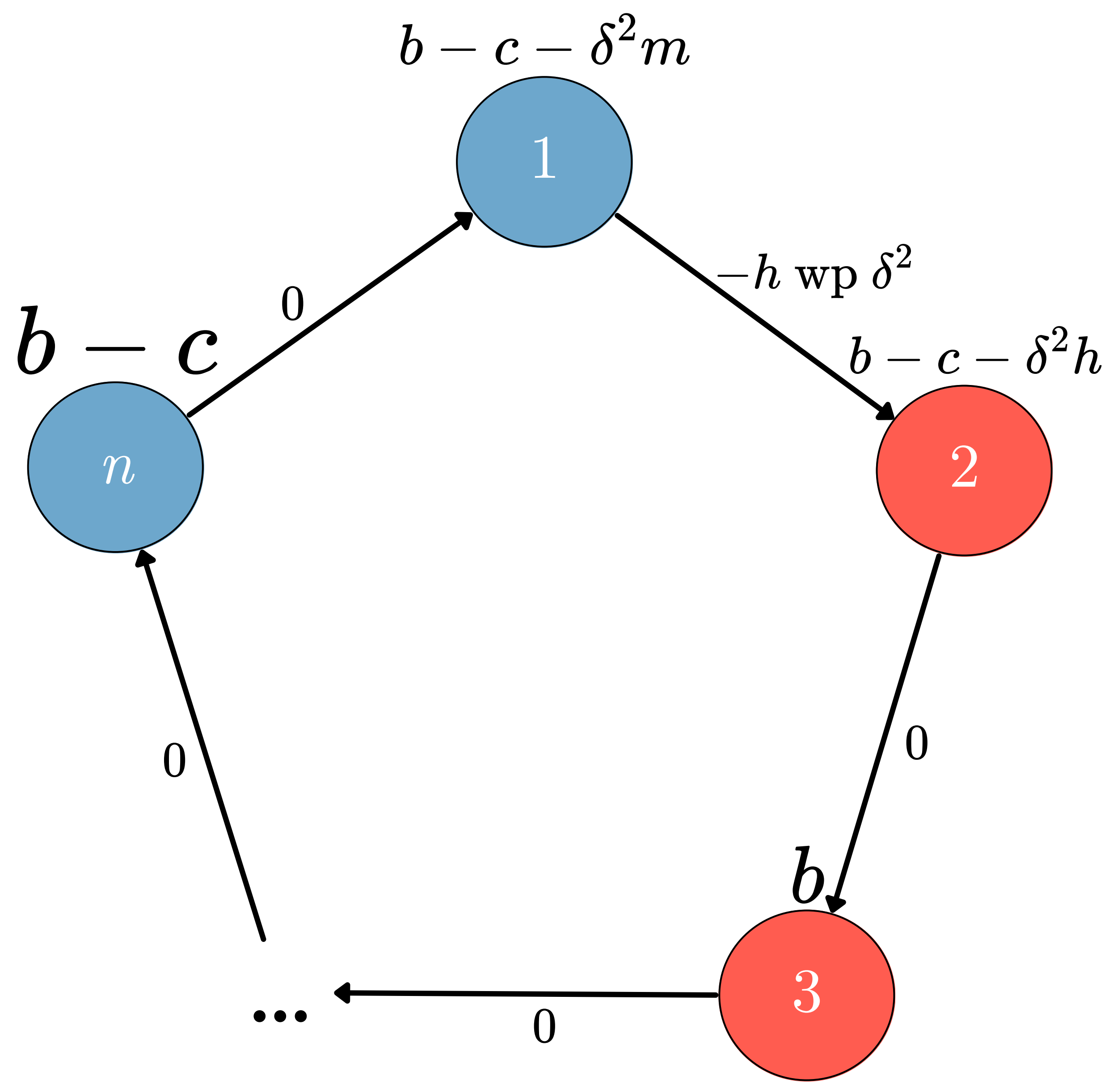
CONTRIBUTE AND PUNISH

With on-strategy play, everyone contributes (and no punishment necessary).

If one player deviates at the first round by not contributing, they get punished by the player upstream. This does not pay off as long as:

$$b - c > b - \delta h \quad \text{iff}$$
$$\delta h > c.$$

If some player deviates at a round i by not punishing a defector, they get punished at the next round by the player upstream.



Player 2 doesn't punish player 3 at round 2, but gets punished by player 1 at round 3

CONTRIBUTE AND PUNISH

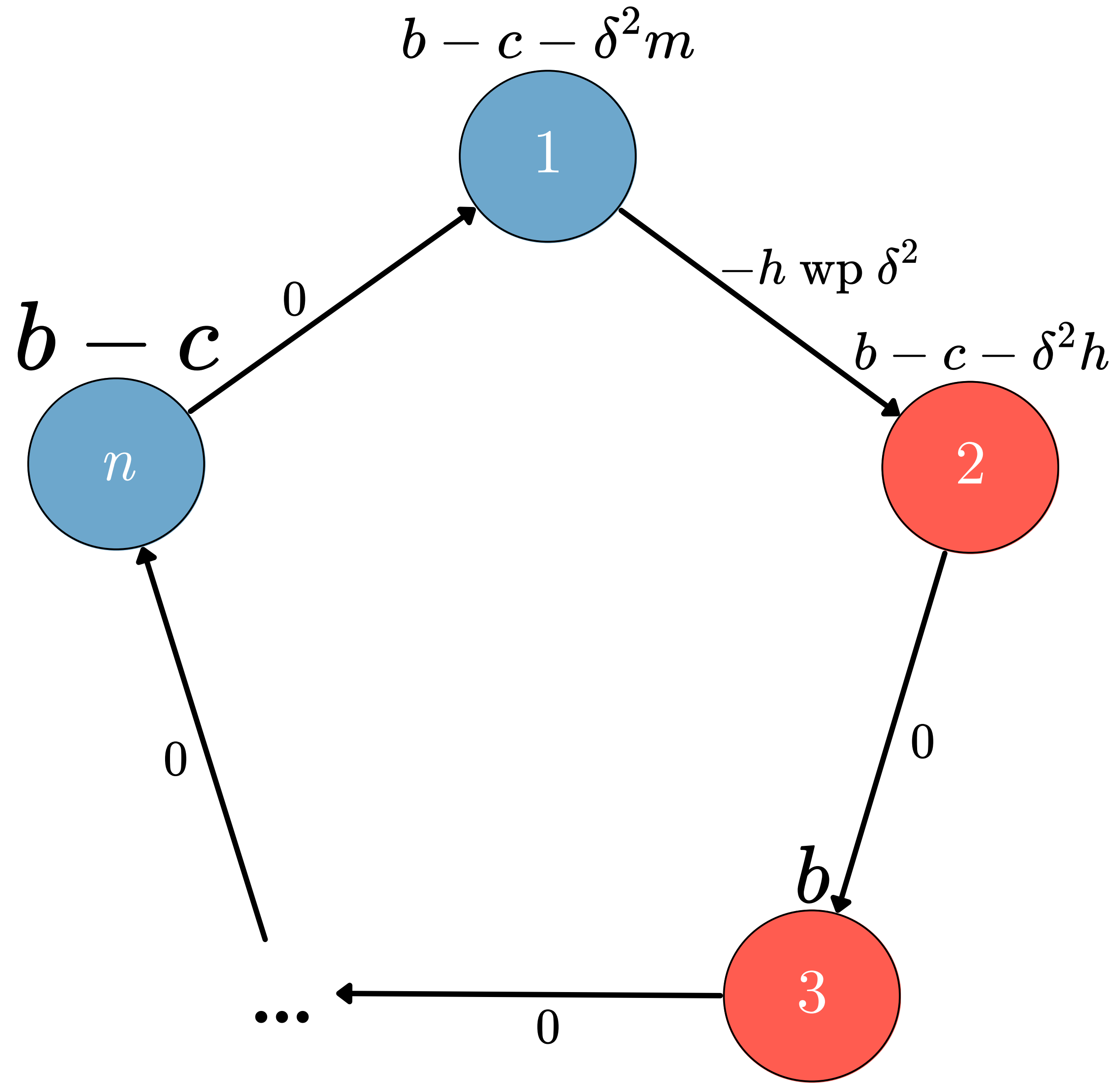
With on-strategy play, everyone contributes (and no punishment necessary).

If one player deviates at the first round by not contributing, they get punished by the player upstream. This does not pay off as long as:

$$\begin{aligned} b - c > b - \delta h \quad &\text{iff} \\ \delta h > c. \end{aligned}$$

If some player deviates at a round i by not punishing a defector, they get punished at the next round by the player upstream. This is not worth it as long as:

$$\begin{aligned} b - c - \delta^i m > b - c - \delta^{i+1} h \quad &\text{iff} \\ \delta h > m. \end{aligned}$$



Player 2 doesn't punish player 3 at round 2, but gets punished by player 1 at round 3

So, if players are willing to monitor and punish each other, if the penalties are high enough, and if the game goes on long enough, cooperation can be maintained.

So, if players are willing to monitor and punish each other, if the penalties are high enough, and if the game goes on long enough, cooperation can be maintained. Lots of ifs.

ERBEZ YOELI

So we have two different equilibria, one good and one bad.

EREZ YOELI
So we have two different equilibria, one good and one bad.

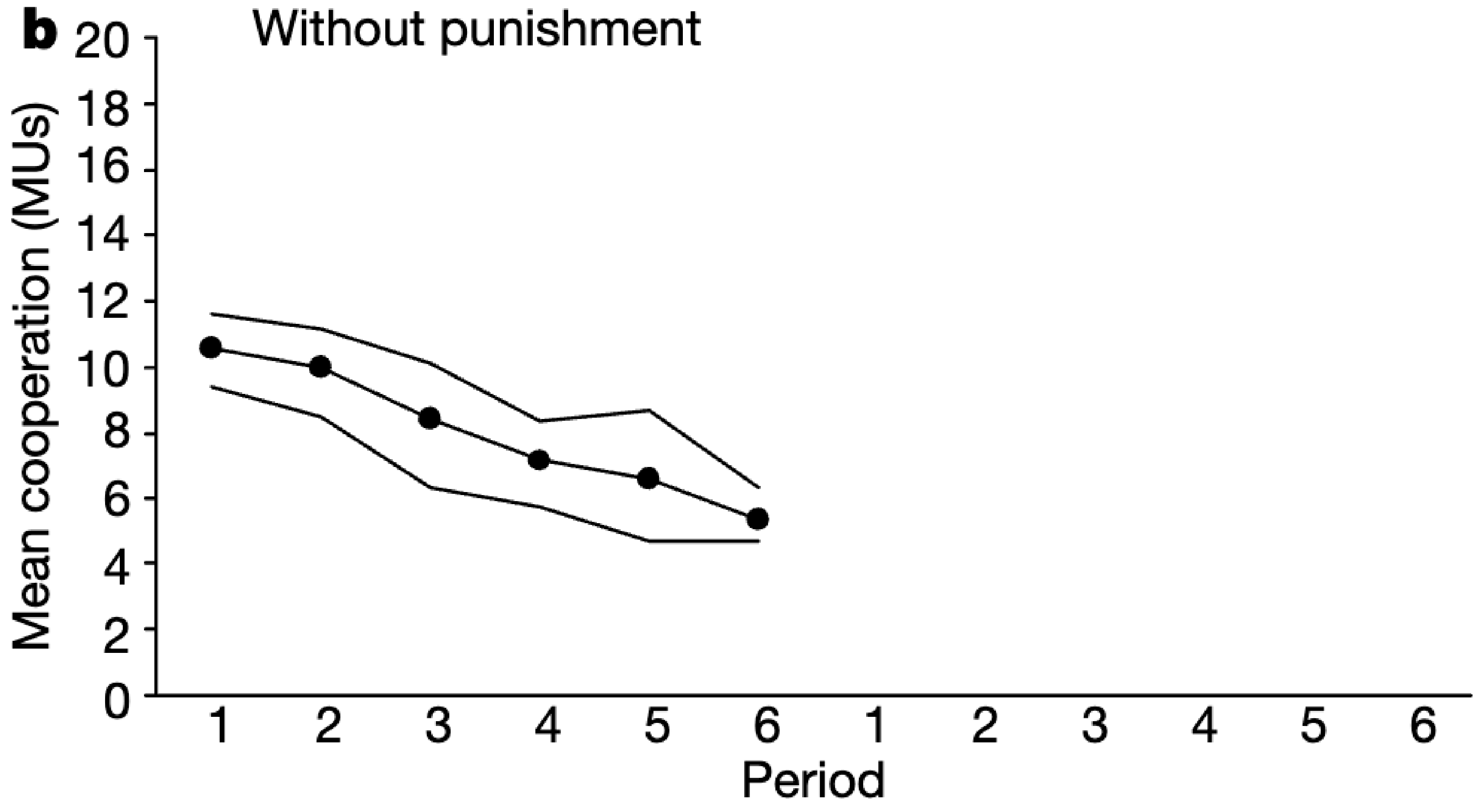
MOSHE HOFFMAN
What do people do?

In the lab people generally contribute above the Nash equilibrium.

PUBLIC GOODS GAMES: EXPERIMENTS

Contributions typically decline over time (are people learning the Nash equilibrium?...)

But generally they start above 0.



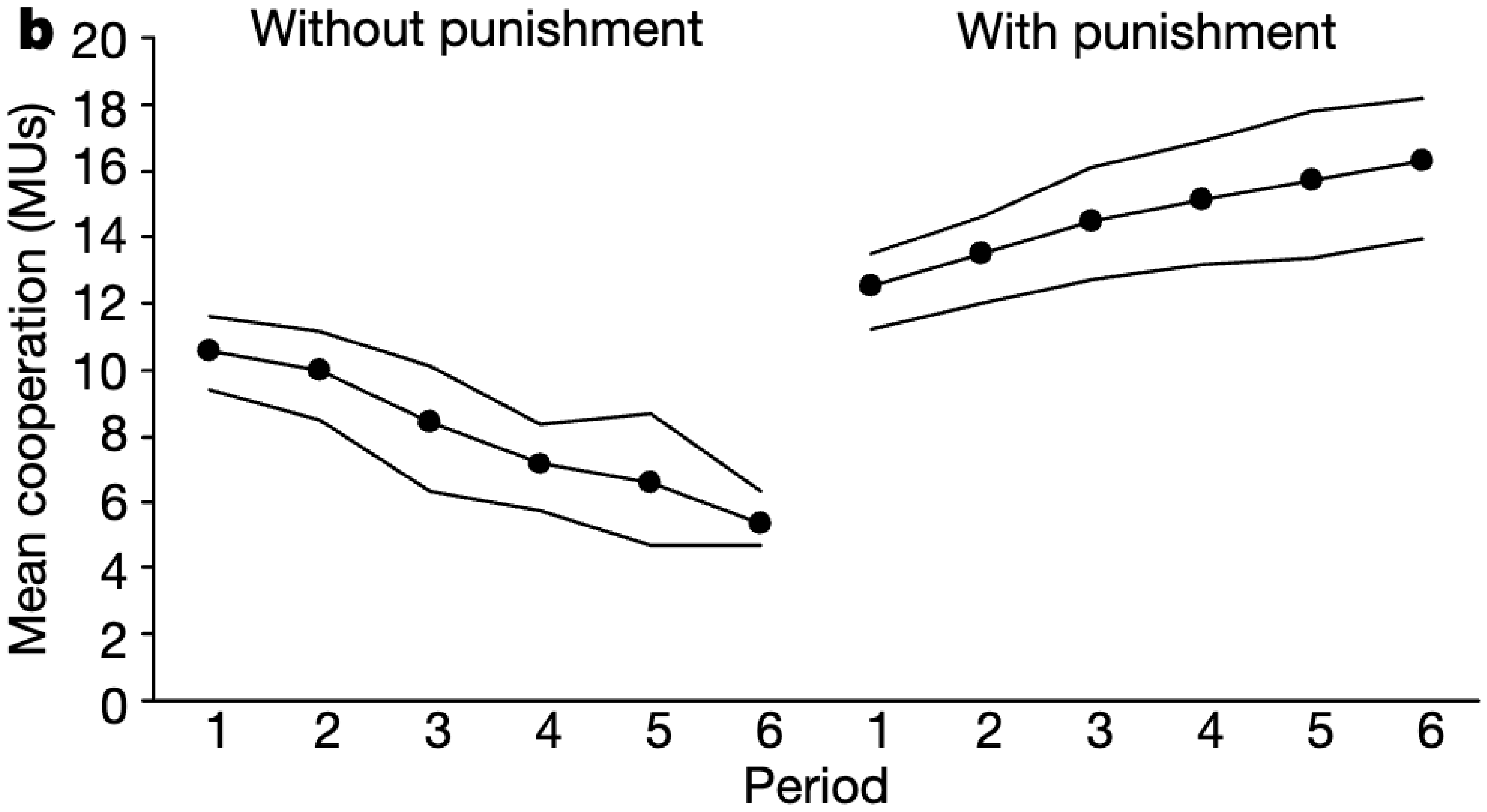
Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868), 137–140.

PUBLIC GOODS GAMES: EXPERIMENTS

Contributions typically decline over time (are people learning the Nash equilibrium?...)

But generally they start above 0.

Punishment helps (!).*



*Interesting, considering that people could default to the other equilibrium.

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868), 137–140.

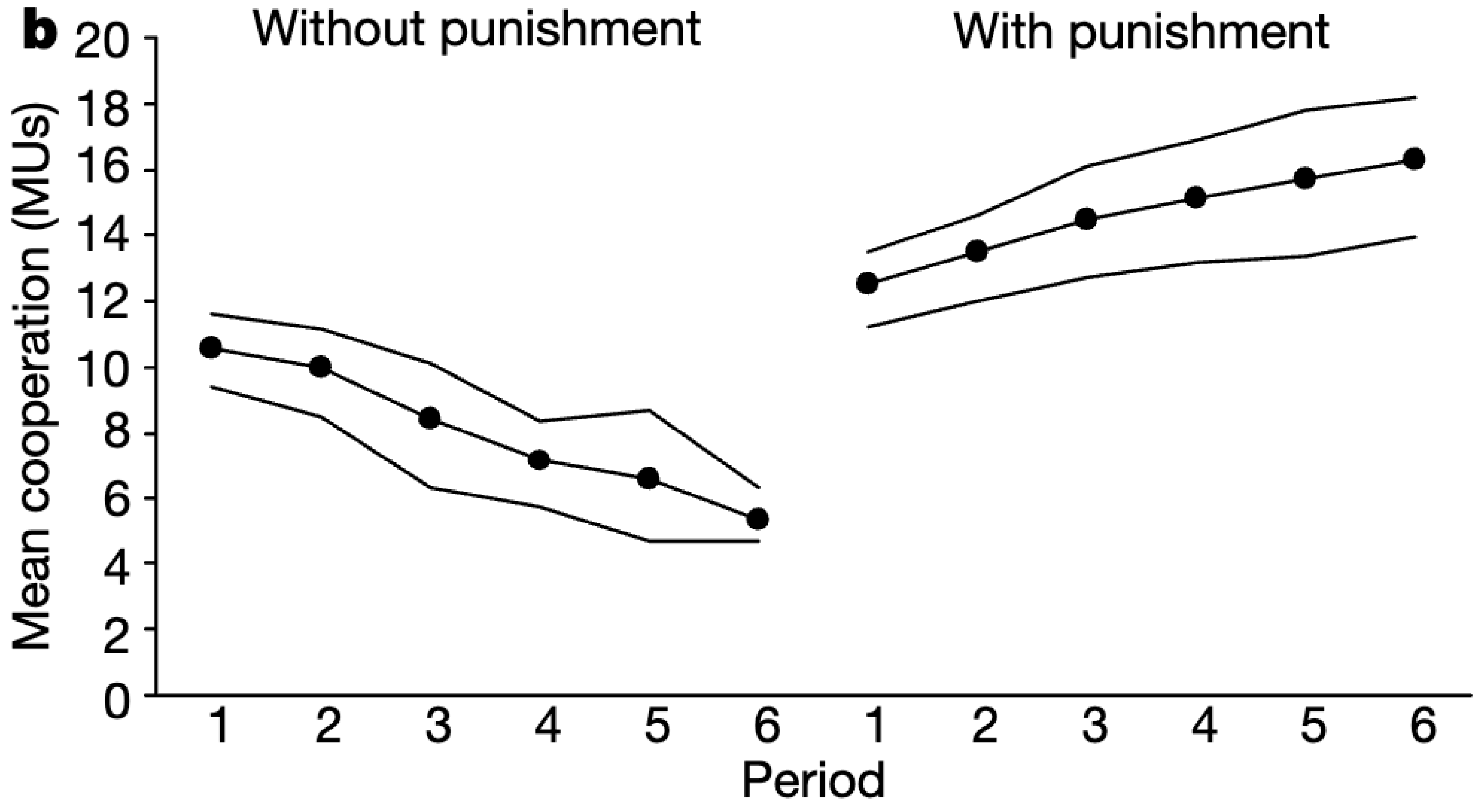
PUBLIC GOODS GAMES: EXPERIMENTS

Contributions typically decline over time (are people learning the Nash equilibrium?...)

But generally they start above 0.

Punishment helps (!).*

Explaining these trends is an ongoing field of research.



*Interesting, considering that people could default to the other equilibrium.

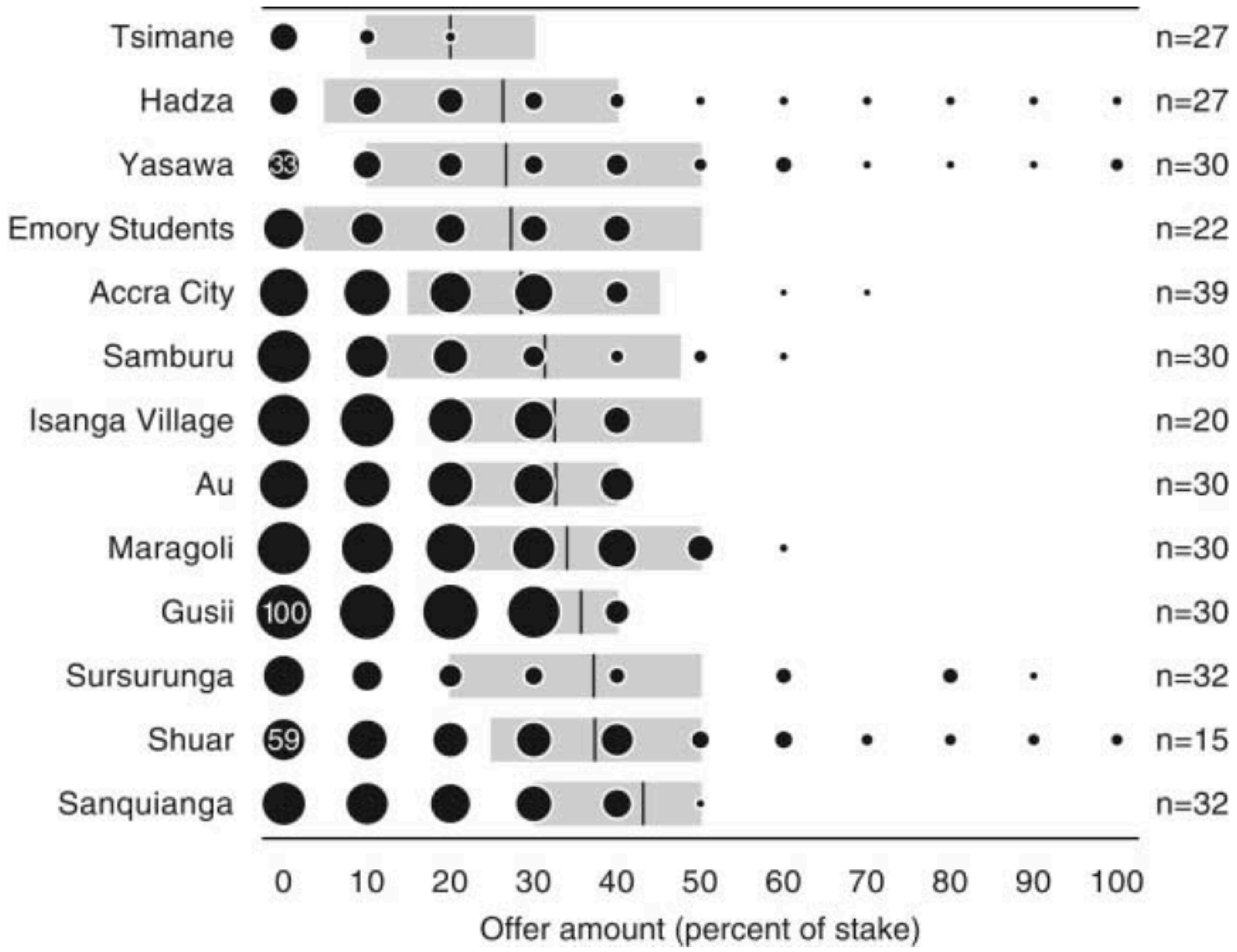
Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868), 137–140.

Interestingly, people seem willing to punish even when the transgression does not affect them personally (third-party punishment).

THIRD PARTY PUNISHMENT GAME

A third party can punish unfair splits in a Dictator game, at a personal cost.

The size of the bubble shows the fraction of people willing to punish at that split.

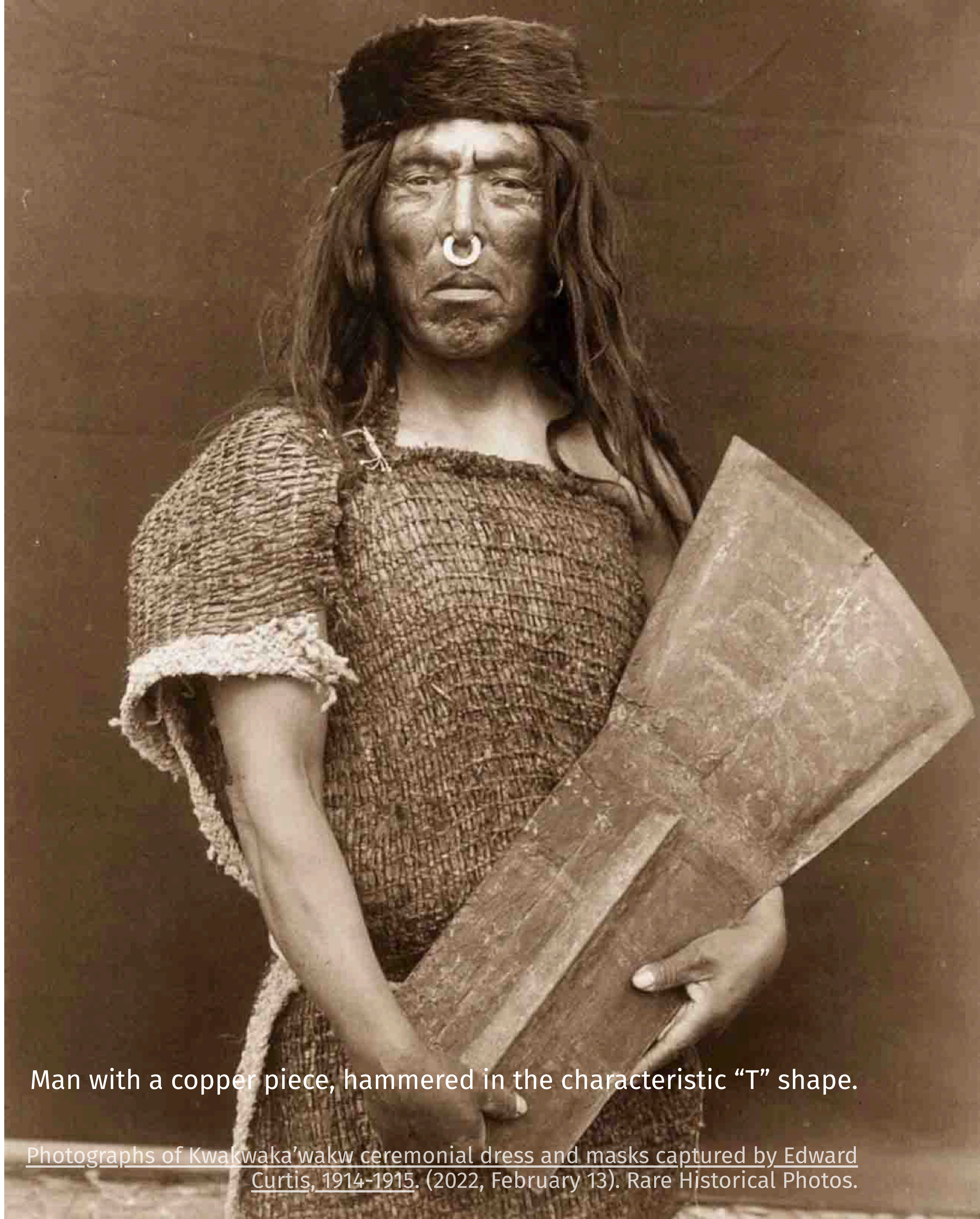


Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., & Ziker, J. (2006). Costly punishment across human societies. *Science*, 312(5781), 1767–1770.

Anecdotally, virtuous chains of contributions are also a mainstay of many cultures.

POTLATCH

The Kwakwa̱ka̱’wakw are an indigenous group of the Pacific Northwest Coast, in southwestern Canada.



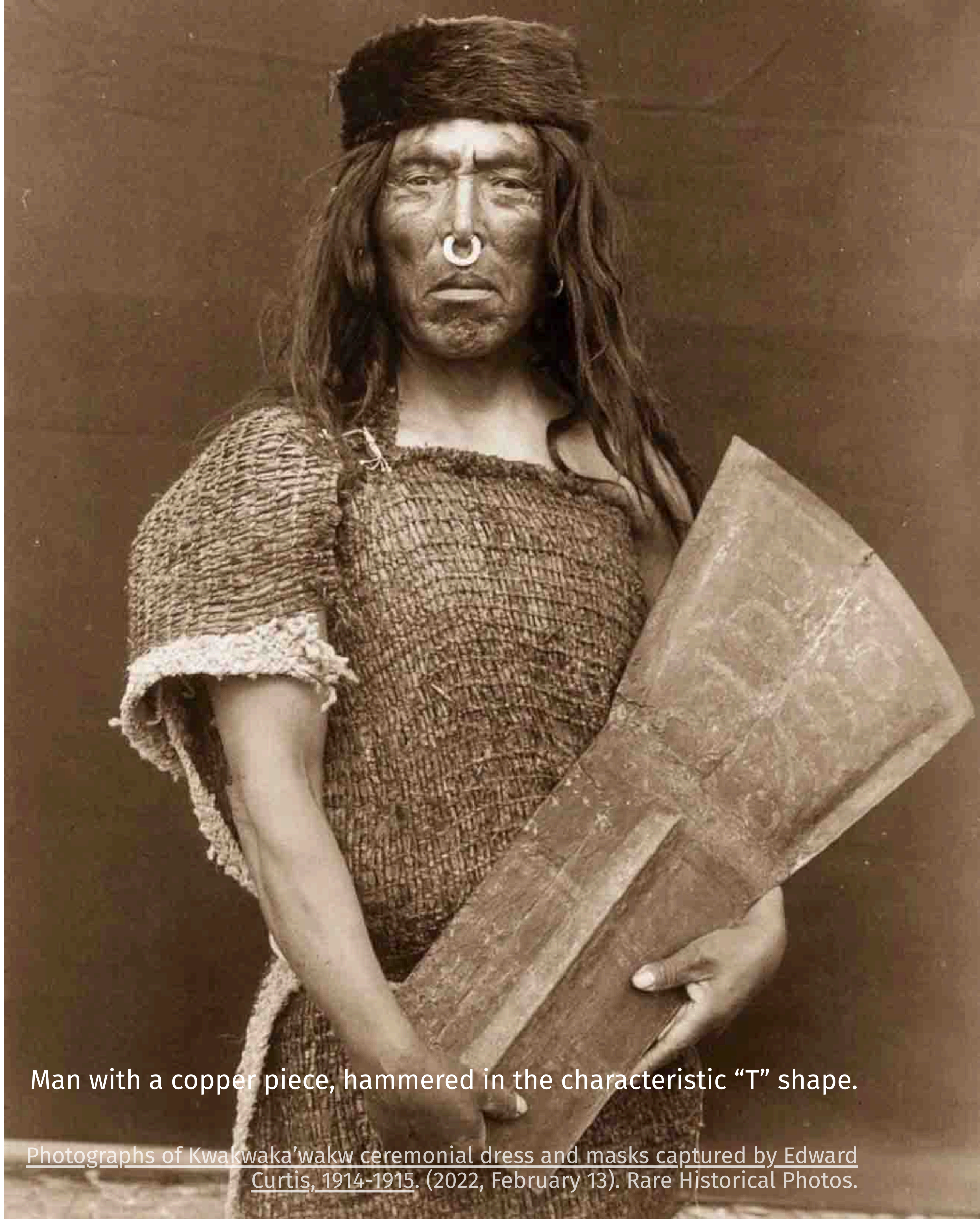
Man with a copper piece, hammered in the characteristic “T” shape.

Photographs of Kwakwaka’wakw ceremonial dress and masks captured by Edward Curtis, 1914-1915. (2022, February 13). Rare Historical Photos.

POTLATCH

The Kwakwa̱ka̱'wakw are an indigenous group of the Pacific Northwest Coast, in southwestern Canada.

Clan chiefs stage winter feasts (*potlatches*) where they shower rivals with blankets, carved copper 'shields,' today even washing-machines.



Man with a copper piece, hammered in the characteristic "T" shape.

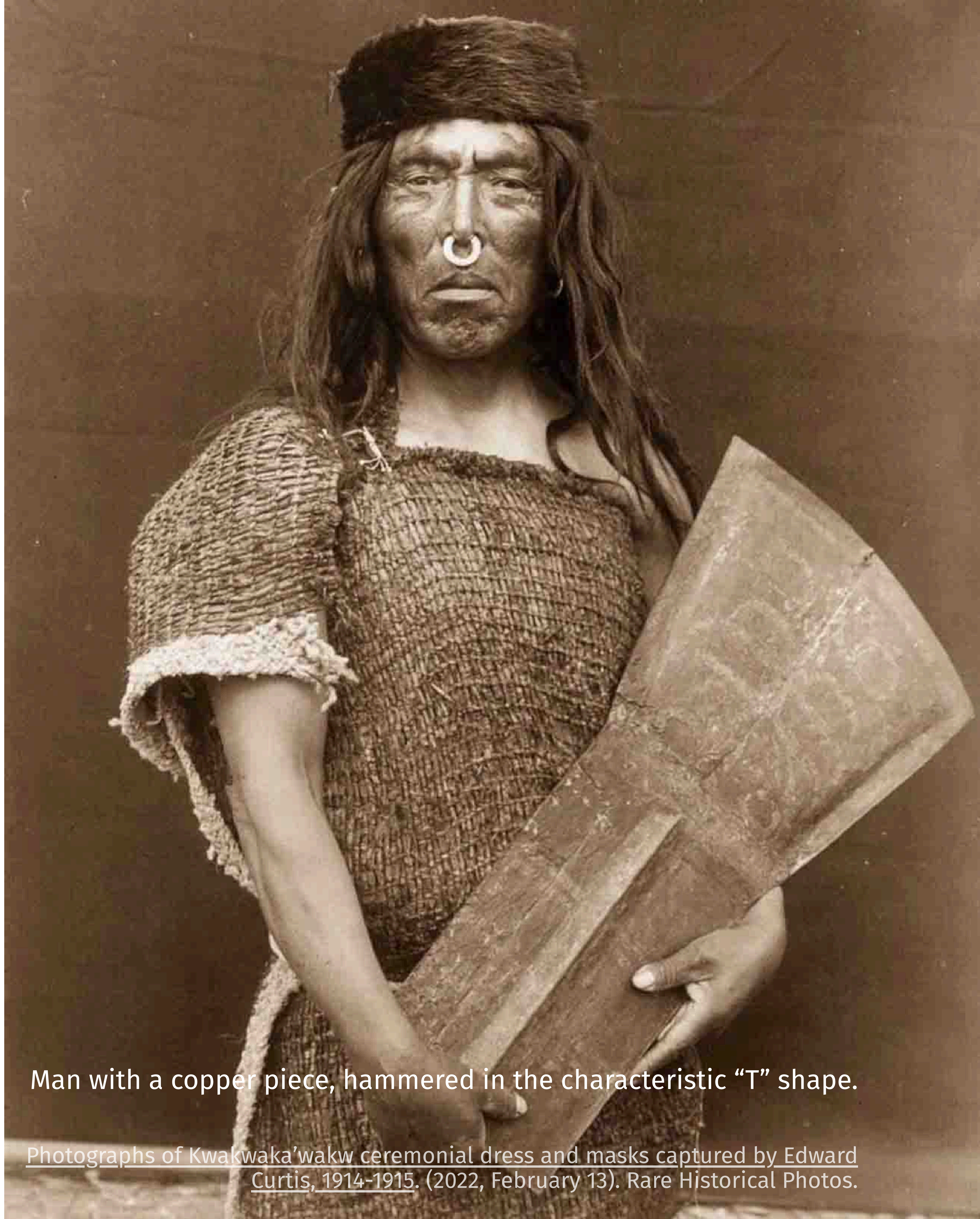
Photographs of Kwakwaka'wakw ceremonial dress and masks captured by Edward Curtis, 1914-1915. (2022, February 13). Rare Historical Photos.

POTLATCH

The Kwakwa̱ka̱'wakw are an indigenous group of the Pacific Northwest Coast, in southwestern Canada.

Clan chiefs stage winter feasts (*potlatches*) where they shower rivals with blankets, carved copper 'shields,' today even washing-machines.

To keep rank, the honoured guests must later exceed the donation.



Man with a copper piece, hammered in the characteristic "T" shape.

Photographs of Kwakwaka'wakw ceremonial dress and masks captured by Edward Curtis, 1914-1915. (2022, February 13). Rare Historical Photos.

People also seem willing to engage in *higher-order punishment*: punishing those who do not punish.

TURKANA RAIDS

The Turkana are semi-nomadic pastoralists based in North-Western Kenya.

They periodically organize large-scale raids against neighboring ethnic groups to acquire cattle, and gain access to pasture and watering sites.

Mathew, S., & Boyd, R. (2014). The cost of cowardice: punitive sentiments towards free riders in Turkana raids. *Evolution and Human Behavior*, 35(1), 58–64.

TURKANA RAIDS

The Turkana are semi-nomadic pastoralists based in North-Western Kenya.

They periodically organize large-scale raids against neighboring ethnic groups to acquire cattle, and gain access to pasture and watering sites.

Freeriders are frowned upon and punished.

Mathew, S., & Boyd, R. (2014). The cost of cowardice: punitive sentiments towards free riders in Turkana raids. *Evolution and Human Behavior*, 35(1), 58–64.

TURKANA RAIDS

The Turkana are semi-nomadic pastoralists based in North-Western Kenya.

They periodically organize large-scale raids against neighboring ethnic groups to acquire cattle, and gain access to pasture and watering sites.

Freeriders are frowned upon and punished.

So are people who do not punish freeriders.

Mathew, S., & Boyd, R. (2014). The cost of cowardice: punitive sentiments towards free riders in Turkana raids. *Evolution and Human Behavior*, 35(1), 58–64.

ERNST FEHR

It seems that nature has endowed us with a disposition to punish wrongdoers.

ERNST FEHR

It seems that nature has endowed us with a disposition to punish wrongdoers.

SIMON GÄCHTER

Punishment, in this sense, is an altruistic act. And demands explanation...

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868), 137–140.