LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

STRATEGIC MINDS: THE GAME THEORY OF COOPERATION, COORDINATION AND COLLABORATION

GAME THEORY 101

SOCIAL DILEMMAS AND EQUILIBRIA

Adrian Haret
a.haret@lmu.de

April 22, 2024

Let’s play a game!

The Trust Game

There are two players with initial endowment of 1 each.

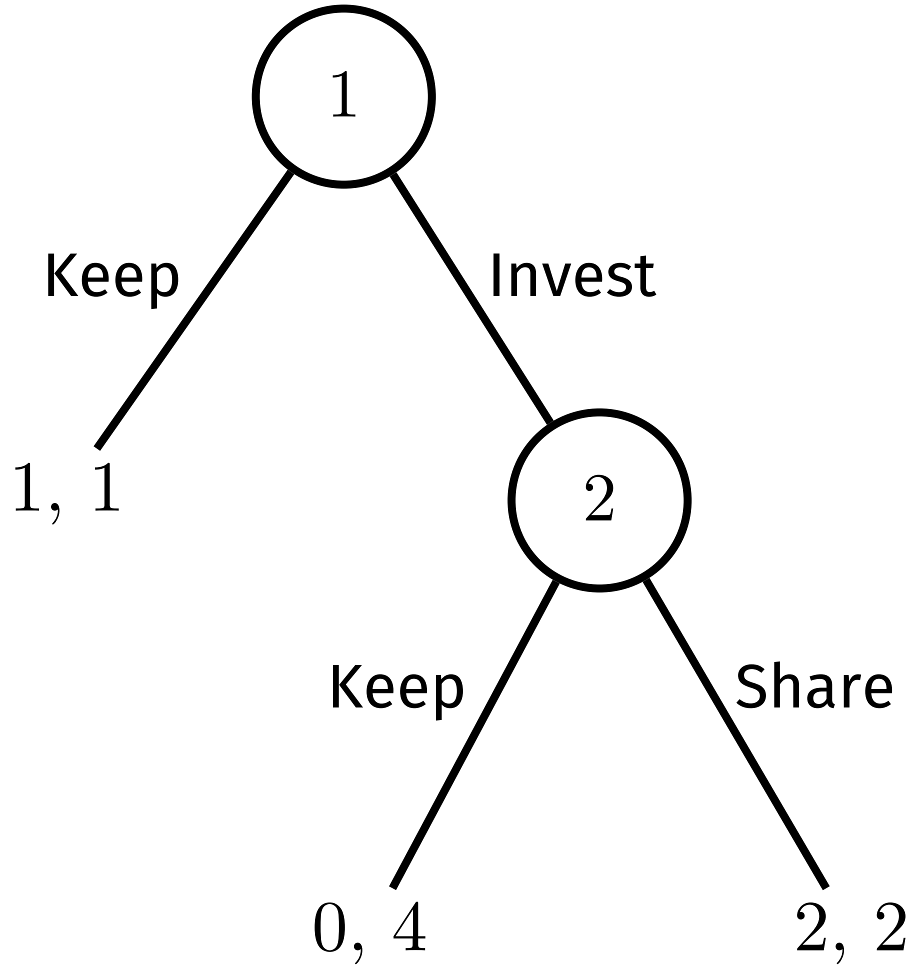
Player 1 makes the first move, by deciding whether to invest in a joint venture.

If Player 1 makes no investment, the game is over and both players retain their endowments.

If Player 1 invests, the 1 generates a surplus and turns into 3.

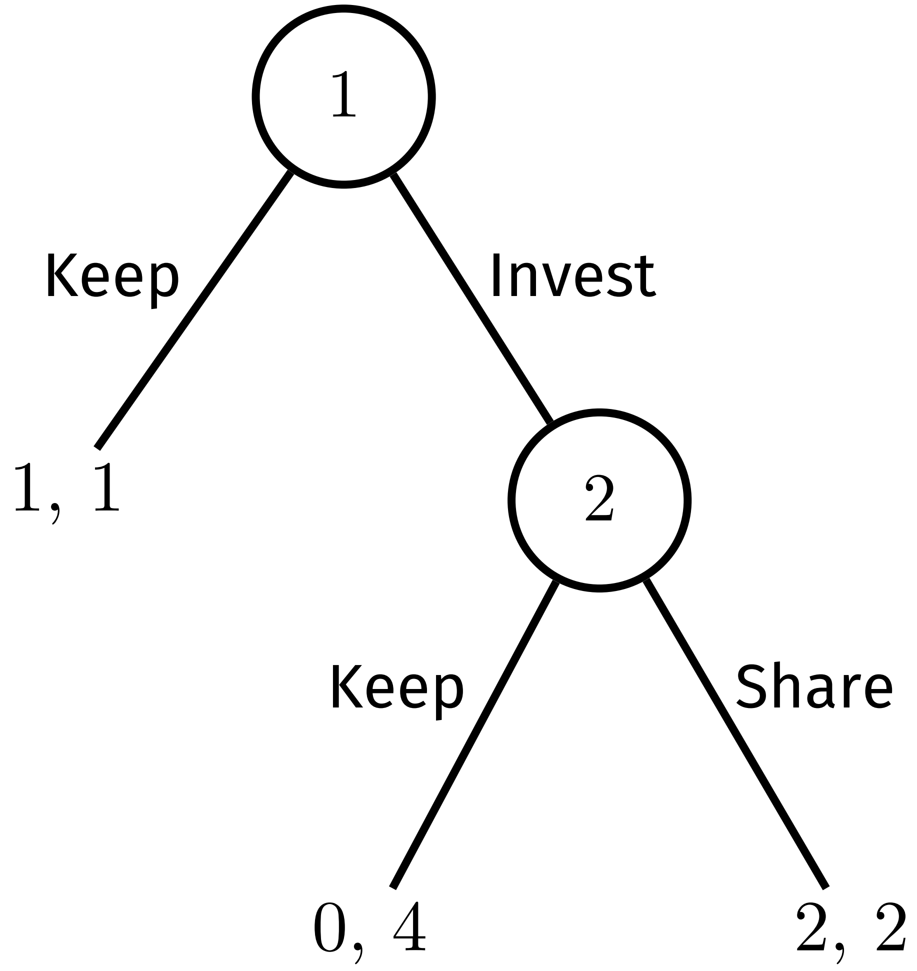
Player 2 now has to decide how to allocate the available sum of 1 + 3 = 4 among the players.

Player 2 can either divide the sum equally, or keep everything.



Suppose you are an individually rational economic agent, i.e., aiming to maximize your own payoff.

How would you play this game?



The Trust Game

There are two players with initial endowment of 1 each.

Player 1 makes the first move, by deciding whether to invest in a joint venture.

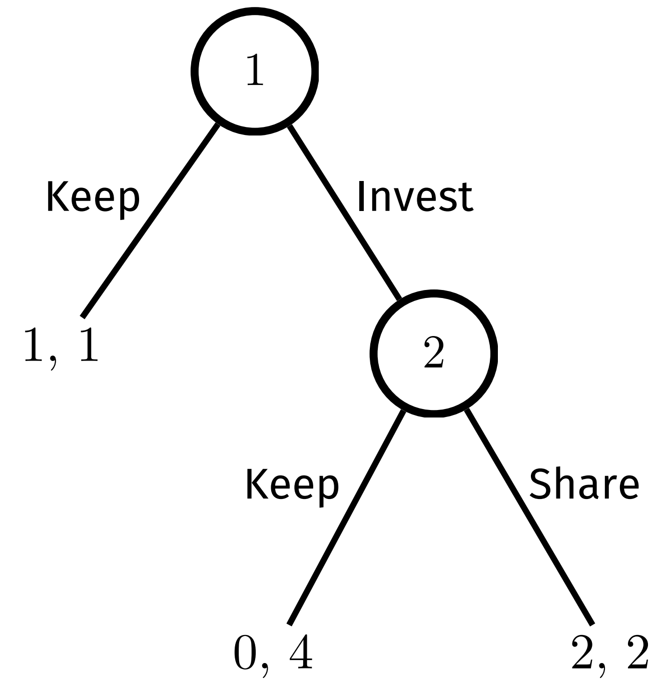
If Player 1 makes no investment, the game is over and both players retain their endowments.

If Player 1 invests, the 1 generates a surplus and turns into 3.

Player 2 now has to decide how to allocate the available sum of 1 + 3 = 4 among the players.

Player 2 can either divide the sum equally, or keep everything.

payoffs



predicted play with self-interested players

If Player 2 is in the position of allocating the sum of 4, they will keep the entire sum (duh).

Player 1, knowing this, realizes there is no point in investing, and keeps the money.

Both players end up with 1 each.

How do people generally play this game?

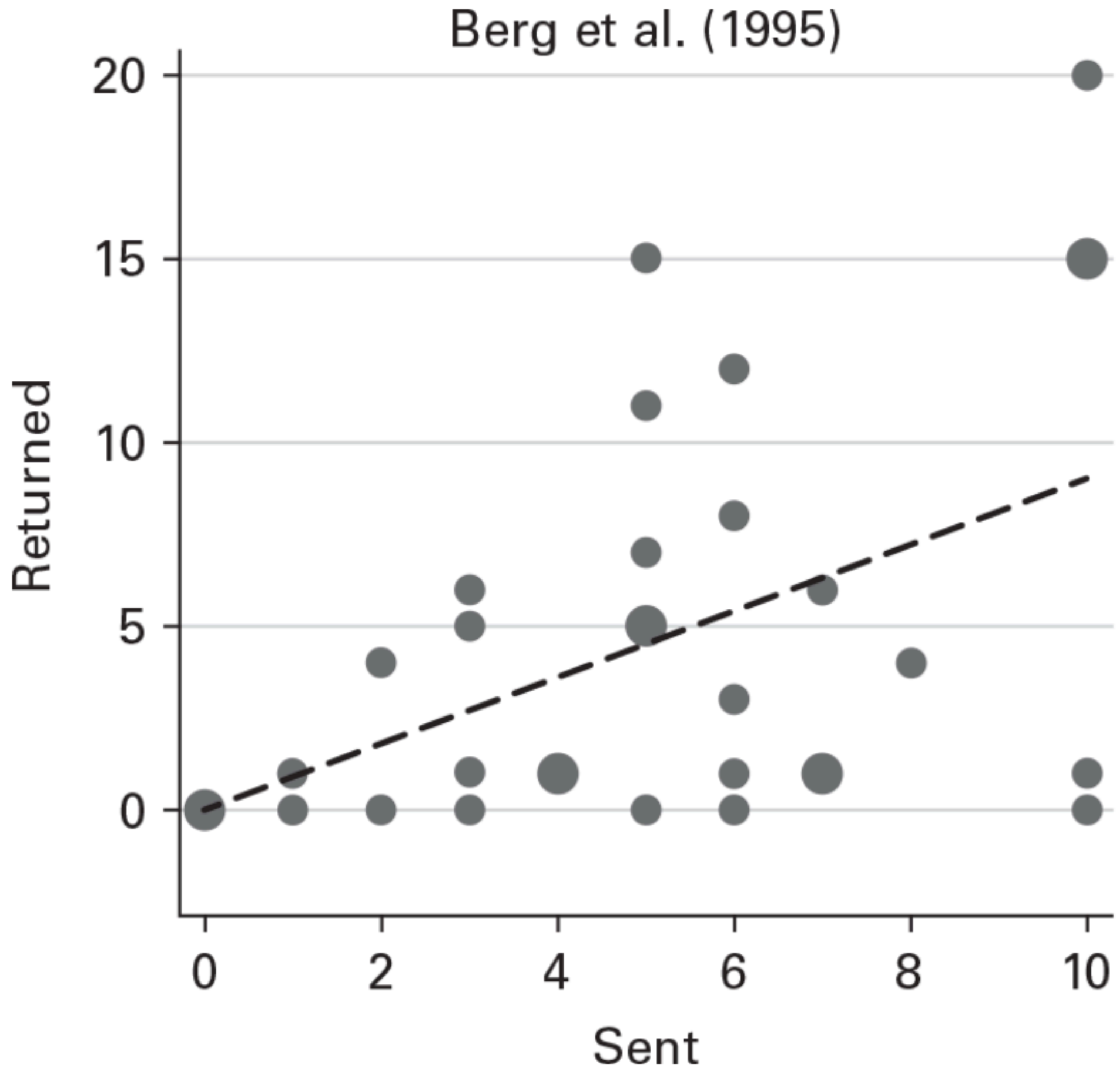
EXPERIMENTAL RESULTS IN THE TRUST GAME

The original experiment features 32 participants from the University of Minnesota.

Player 1 could send any amount between \$0 and \$10, and Player 2 could return anything between \$0 and \$20.

Average amount sent by Player 1 was \$5,16.

Average amount returned by Player 2 was \$4,66.



Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior*, 10(1), 122–142.

RESULTS FROM A META-STUDY

These results have been replicated across many other instances and cultures.

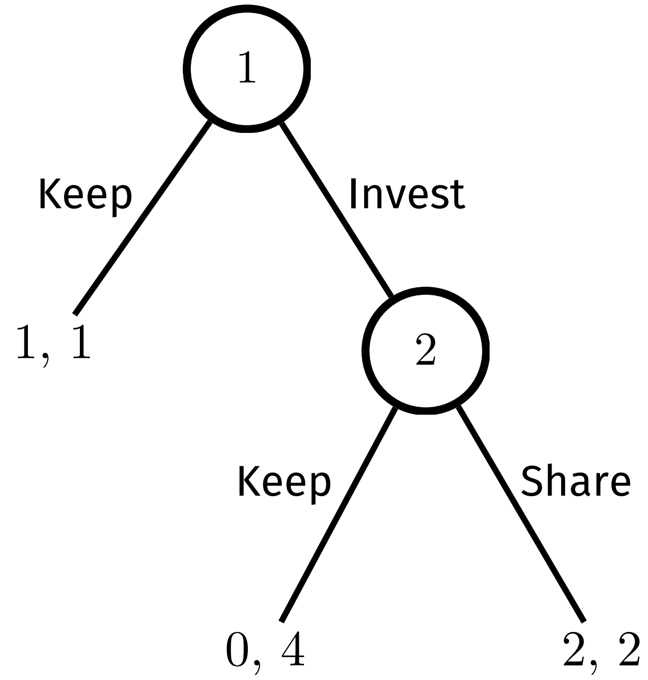
Variable name	Obs.	Sum N	Mean
Panel A: Sent fraction (trust)			
All regions	161	23,900	0.502
North America	46	4579	0.517
Europe	64	9030	0.537
Asia	23	3043	0.482
South America	13	4733	0.458
Africa	15	2515	0.456
Panel B: Proportion returned (trustworthiness)			
All regions	137	21,529	0.372
North America	41	4324	0.340
Europe	53	7596	0.382
Asia	15	2361	0.460
South America	13	4733	0.369
Africa	15	2515	0.319

Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal Of Economic Psychology*, 32(5), 865–889.

Note that by acting in according to their self interest, players are leaving money (or chocolate) on the table.

Money that could be gotten if Player 2 could muster up some self-restraint (or gratitude), and Player 1 could trust Player 2 to do so.

payoffs



predicted play with self-interested players

If Player 2 is in the position of allocating the sum of 4, they will keep the entire sum (duh).

Player 1, knowing this, realizes there is no point in investing, and keeps the money.

Both players end up with 1 each.

This is an example of a *social dilemma*.

SOCIAL DILEMMAS

DEFINITION (PRELIMINARY)

A social dilemma is a situation in which individual incentives are at odds with group incentives. Individual rationality leads members of a group to an outcome that is suboptimal.

Carpenter, J., & Robbett, A. (2022). *Game Theory and Behavior*. MIT Press.

Dawes, R. M. (1980). Social Dilemmas. *Annual Review of Psychology*, 31 (80), 169–193.

How to get out of it?

If the two players could write a *contract*, to be enforced by a strong party, like a scary leviathan, the dilemma is solved.

THOMAS HOBBES

Yassss!

Humans in their natural state are subject to a social dilemma.

They can’t trust each other, so nothing ever gets done.

We need a strong government to intervene, establish the rule of law, punish knaves, and enforce contracts.

IMMANUEL KANT

Alternatively, people should just act in the way they want everyone else to act.

If you don’t want to be taken advantage of, don’t do it to others.

Or, if we look at history, perhaps it was the civilizing effect of markets that drew us out of social dilemmas.

MONTESQUIEU

Commerce cures destructive prejudices, and it is an almost general rule that everywhere there are gentle mores, there is commerce and that everywhere there is commerce, there are gentle mores.

de Montesquieu, C. (1989). *The Spirit of the Laws*. Cambridge University Press.

For economic activity to thrive, you need people to trust each other.

People may have the knowhow to make things, but if they fear that they will be confiscated by the lord, or stolen by thieves, they produce little.

KENNETH ARROW

Virtually every commercial transaction has within itself an element of *trust*.

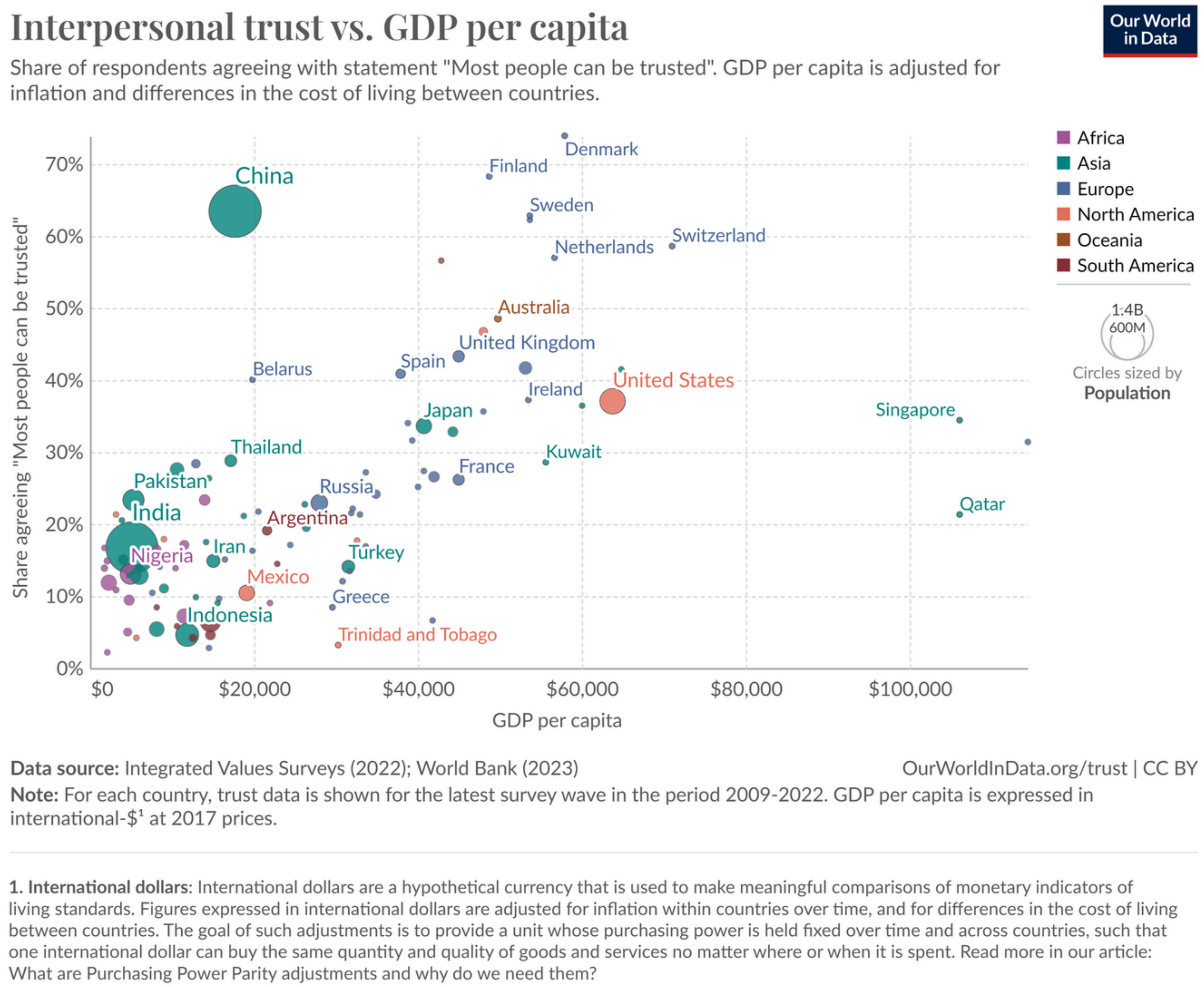
It can be plausibly argued that much of the economic backwardness in the world can be explained by the lack of mutual confidence.

Arrow, K. J. (1972). Gifts and Exchanges. *Philosophy & Public Affairs*, 1(4), 343–362.

CAN MOST PEOPLE BE TRUSTED?

There is a correlation between levels of trust and GDP per capita.

There is a similar correlation with levels of inequality.



Ortiz-Ospina, E., Roser, M., Arriagada, P. (2016). Trust. Published online at OurWorldInData.org

More generally, there are interactions where what is best for you to do depends on what the other does.

And the other way around.

JOHN VON NEUMANN

We should call that *game theory.*

OSKAR MORGENSTERN

JOHN VON NEUMANN
We should call that *game theory*.

OSKAR MORGENSTERN
And write a classic textbook on it!

JOHN VON NEUMANN

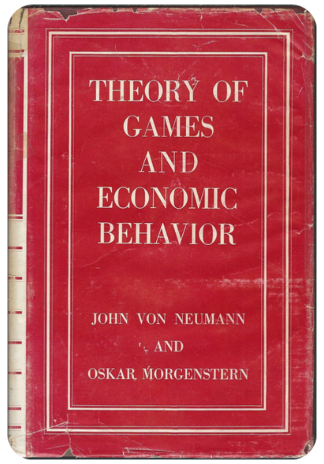
JOHN VON NEUMANN

We should call that *game theory*.

OSKAR MORGENSTERN

And write a classic textbook on it!

JOHN VON NEUMANN



von Neumann, J., & Morgenstern, O. (1953). *Theory of Games and Economic Behavior.* Princeton University Press.

Let’s start with the most basic type of game: games in *normal form*.

A game in normal form consists of *players*, that have *strategies*, based on *actions*, which lead to *payoffs*.

NOTATION

players	$N = \{1, \ldots, n\}$
strategy of player i	s_i
profile of strategies	$\boldsymbol{s} = (s_1, \ldots, s_n)$
utility of player i with strategy profile $\boldsymbol{s}$	$u_i(\boldsymbol{s}) \in \mathbb{R}$
strategy profile $\boldsymbol{s}$ without s_i	$\boldsymbol{s}_{-i} = (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$
$\boldsymbol{s}$, alternatively	$\boldsymbol{s} = (s_i, \boldsymbol{s}_{-i})$

When there are only two players, we can represent the game using a table.

The Trust Game

There are two players with initial endowment of 1 each.

Player 1 makes the first move, by deciding whether to invest in a joint venture.

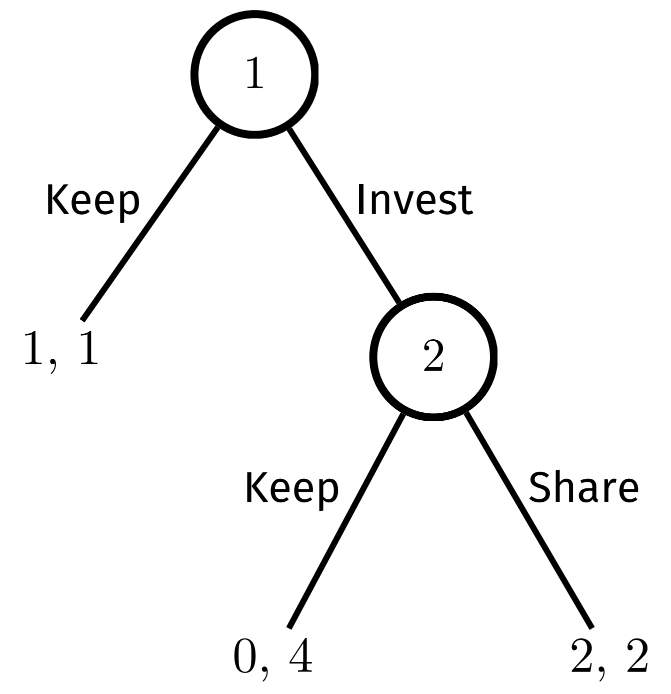
If Player 1 makes no investment, the game is over and both players retain their endowments.

If Player 1 invests, the 1 generates a surplus and turns into 3.

Player 2 now has to decide how to allocate the available sum of 1 + 3 = 4 among the players.

Player 2 can either divide the sum equally, or keep everything.

payoffs



The Trust Game

There are two players with initial endowment of 1 each.

Player 1 makes the first move, by deciding whether to invest in a joint venture.

If Player 1 makes no investment, the game is over and both players retain their endowments.

If Player 1 invests, the 1 generates a surplus and turns into 3.

Player 2 now has to decide how to allocate the available sum of 1 + 3 = 4 among the players.

Player 2 can either divide the sum equally, or keep everything.

payoffs

	Keep	Share
Keep	1, 1	1, 1
Invest	0, 4	2, 2

The Trust Game

There are two players with initial endowment of 1 each.

Player 1 makes the first move, by deciding whether to invest in a joint venture.

If Player 1 makes no investment, the game is over and both players retain their endowments.

If Player 1 invests, the 1 generates a surplus and turns into 3.

Player 2 now has to decide how to allocate the available sum of 1 + 3 = 4 among the players.

Player 2 can either divide the sum equally, or keep everything.

action of column player

payoffs

player 2

action of row player

payoff of row player for this combination of actions

payoff of column player

	Keep	Share
Keep	1, 1	1, 1
Invest	0, 4	2, 2

player 1

The Trust Game

There are two players with initial endowment of 1 each.

Player 1 makes the first move, by deciding whether to invest in a joint venture.

If Player 1 makes no investment, the game is over and both players retain their endowments.

If Player 1 invests, the 1 generates a surplus and turns into 3.

Player 2 now has to decide how to allocate the available sum of 1 + 3 = 4 among the players.

Player 2 can either divide the sum equally, or keep everything.

action of column player

payoffs

action of row player

payoff of column player

payoff of row player for this combination of actions

player 2 → / player 1 ↑	Keep	Share
Keep	1, 1	1, 1
Invest	0, 4	2, 2

elements

The players are 1 and 2.

Possible strategy profiles are (Keep, Keep), (Keep, Share), (Invest, Keep), (Invest, Share).

Payoffs are

$u_1(\text{Keep}, \text{Keep}) = 1$, $u_2(\text{Invest}, \text{Keep}) = 4, \ldots$

We generally assume that Player 1 is the row Player and player 2 is the column player.

And, for now, that a strategy consists in choosing an available action and playing it.

Oh, and players want to maximize their payoffs in the game.

OSKAR MORGENSTERN

If we knew what strategies players would play we could go on and compute their utilities, expected utilities and so on.

JOHN VON NEUMANN

But that's not how rational agents behave: strategies change depending on what others do.

OSKAR MORGENSTERN

Indeed! If Player 1 invests, the best thing for Player 2 to do is to keep. But if Player 2 plays keep, Player 1 also wants to keep...

JOHN VON NEUMANN

We need to reason the other way around: from utilities to strategies.

OSKAR MORGENSTERN

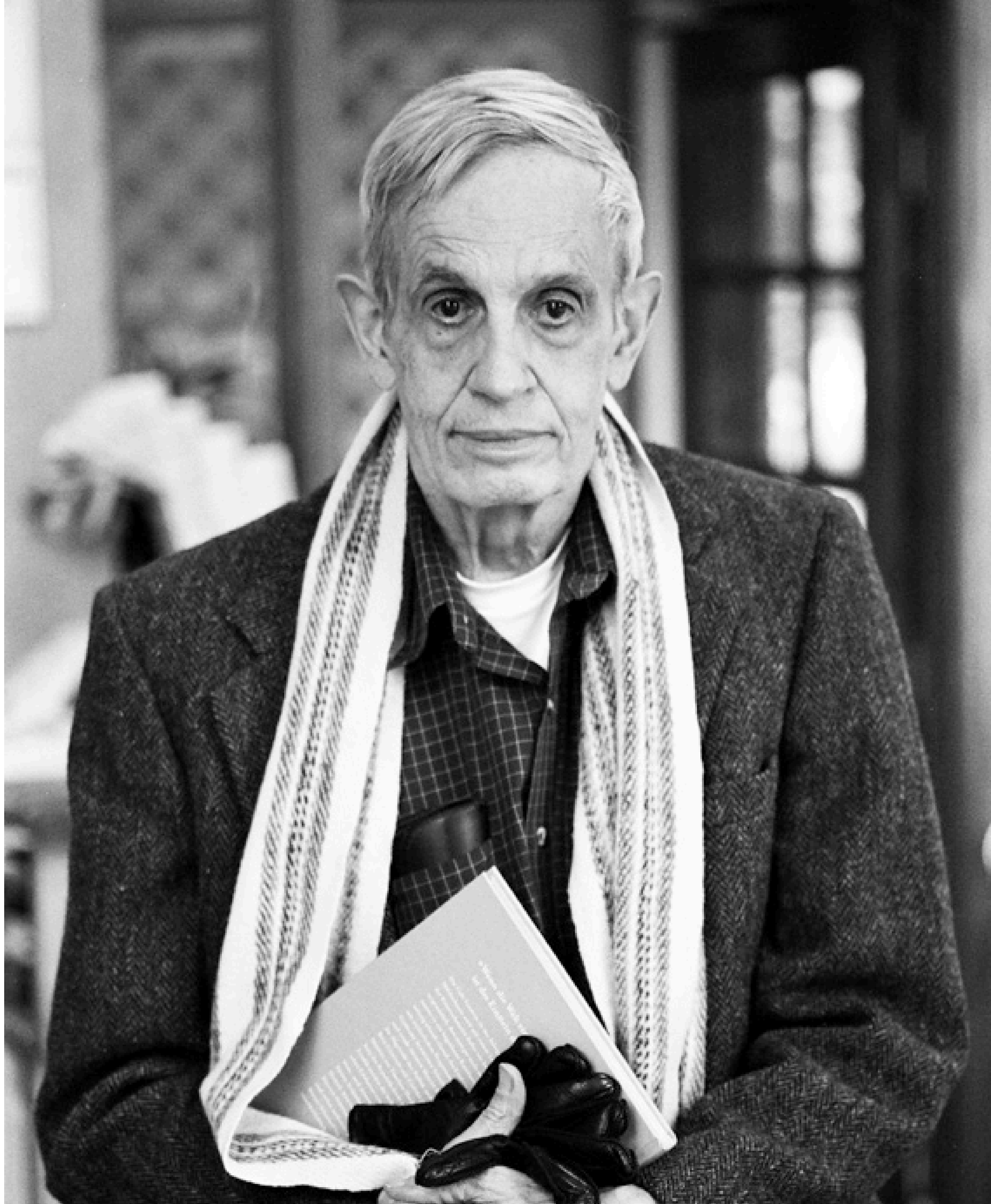
We need to reason about *solution concepts*.

	Keep	Share
Keep	$1, 1$	$1, 1$
Invest	$0, 4$	$2, 2$

A solution concept describes what strategies we expect the players to play.

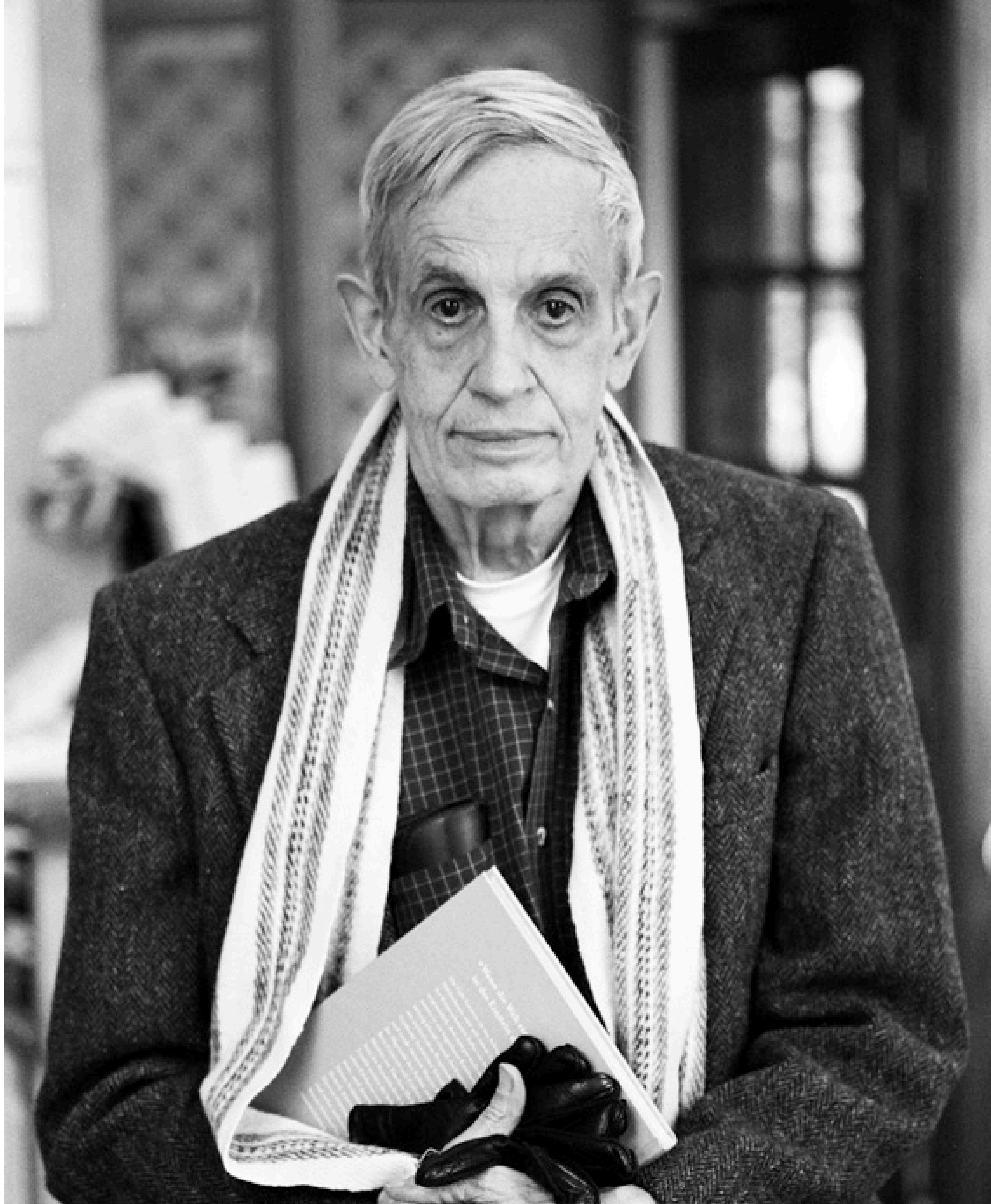
And the outcome of the game.

Enter Nash.



Enter Nash.

John Nash.



JOHN NASH

In a Nash equilibrium no one has an incentive to change their strategy, given the other players' strategies.

BEST RESPONSE & NASH EQUILIBRIUM

DEFINITION (BEST RESPONSE)

Player i's *best response* to the other players' strategies $\boldsymbol{s}_{-i} = (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$ is a strategy s_i^* such that $u_i(s_i^*, \boldsymbol{s}_{-i}) \geq u_i(s_i, \boldsymbol{s}_{-i})$, for any strategy s_i of player i.

BEST RESPONSE & NASH EQUILIBRIUM

DEFINITION (BEST RESPONSE)

Player i's *best response* to the other players' strategies $\boldsymbol{s}_{-i} = (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$ is a strategy s_i^* such that $u_i(s_i^*, \boldsymbol{s}_{-i}) \geq u_i(s_i, \boldsymbol{s}_{-i})$, for any strategy s_i of player i.

DEFINITION (PURE NASH EQUILIBRIUM)

A strategy profile $\boldsymbol{s}^* = (s_1^*, \ldots, s_n^*)$ is a *pure Nash equilibrium* if s_i^* is a best response to $\boldsymbol{s}_{-i}^*$, for every player i.

In other words, $\boldsymbol{s}^*$ is a pure Nash equilibrium if there is no player i and strategy s_i' such that $u_i(s_i', \boldsymbol{s}_{-i}^*) > u_i(s_i^*, \boldsymbol{s}_{-i}^*)$.

profitable deviation

And now for the moment we’ve all been waiting for.

The Prisoner's Dilemma

You and a friend are at the police station. You are the main suspects in a string of Oktoberfest beer thefts.

You are interrogated at the same time, in separate rooms.

If both of you stick to the common story (Cooperate), you get off with a smallish fine.

But if you tell on your friend (Defect) you get off free, while they get a hefty fine.

Your friend faces the same situation.

If you rat each other out, you split the large fine.

payoffs

	Cooperate	Defect
Cooperate	-20, -20	-100, 0
Defect	0, -100	-50, -50

pure Nash equilibria

The Prisoner's Dilemma

You and a friend are at the police station. You are the main suspects in a string of Oktoberfest beer thefts.

You are interrogated at the same time, in separate rooms.

If both of you stick to the common story (Cooperate), you get off with a smallish fine.

But if you tell on your friend (Defect) you get off free, while they get a hefty fine.

Your friend faces the same situation.

If you rat each other out, you split the large fine.

payoffs

	Cooperate	Defect
Cooperate	-20, -20	-100, 0
Defect	0, -100	-50, -50

pure Nash equilibria

(Defect, Defect)

And the Trust Game?

The Trust Game

There are two players with initial endowment of 1 each.

Player 1 makes the first move, by deciding whether to invest in a joint venture.

If Player 1 makes no investment, the game is over and both players retain their endowments.

If Player 1 invests, the 1 generates a surplus and turns into 3.

Player 2 now has to decide how to allocate the available sum of 1 + 3 = 4 among the players.

Player 2 can either divide the sum equally, or keep everything.

payoffs

	Keep	Share
Keep	1, 1	1, 1
Invest	0, 4	2, 2

pure Nash equilibria

The Trust Game

There are two players with initial endowment of 1 each.

Player 1 makes the first move, by deciding whether to invest in a joint venture.

If Player 1 makes no investment, the game is over and both players retain their endowments.

If Player 1 invests, the 1 generates a surplus and turns into 3.

Player 2 now has to decide how to allocate the available sum of 1 + 3 = 4 among the players.

Player 2 can either divide the sum equally, or keep everything.

payoffs

	Keep	Share
Keep	1, 1	1, 1
Invest	0, 4	2, 2

pure Nash equilibria

(Keep, Keep)

Why do women endure the discomfort of high heels?

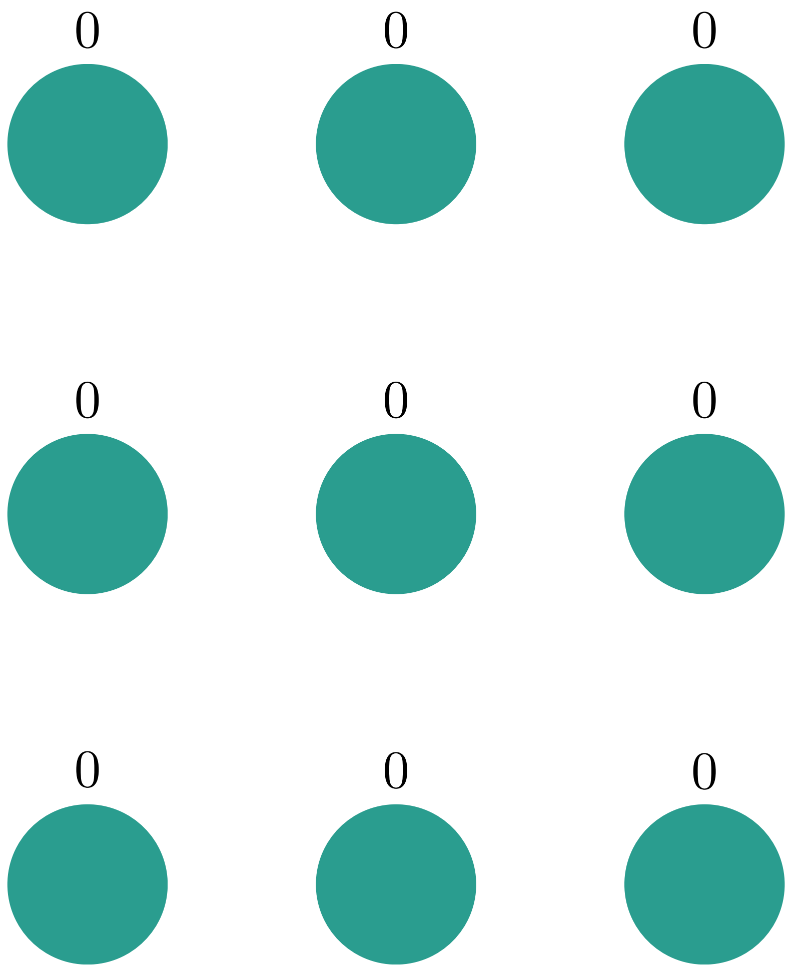
JANE AUSTEN

[Marianne], in having the advantage of height, was more striking [than her sister].

Austen, J. (1811). *Sense and Sensibility.*

no heels heels

Let’s assume that a height advantage makes one more attractive (+3), and a disadvantage is bad (-3).

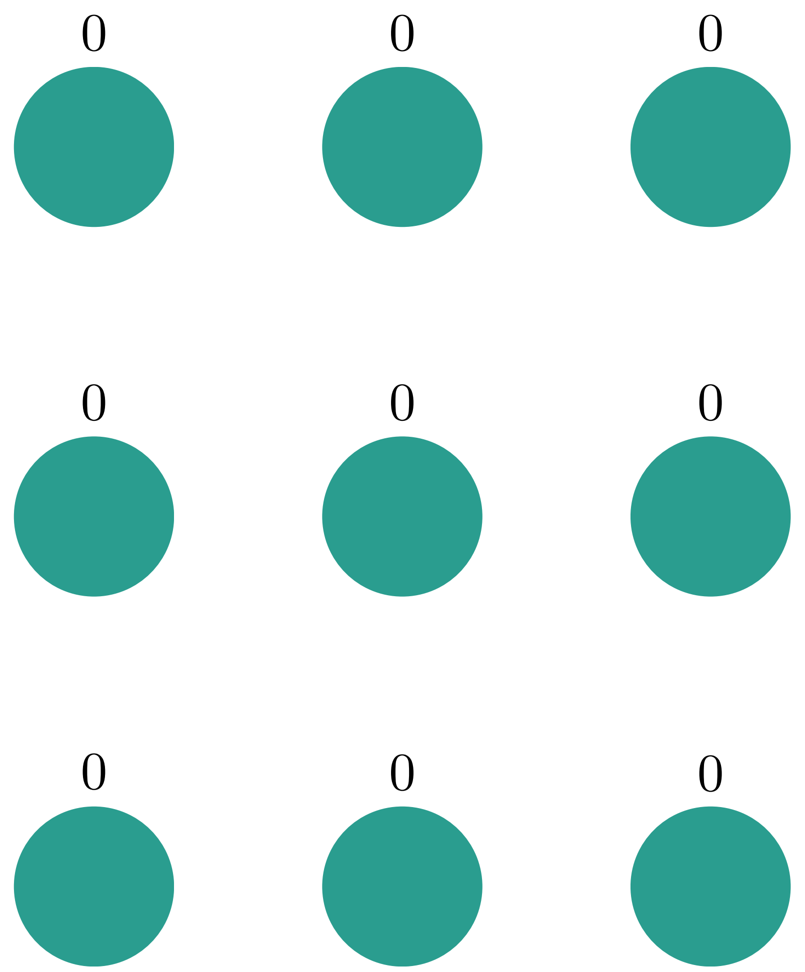


no heels

heels

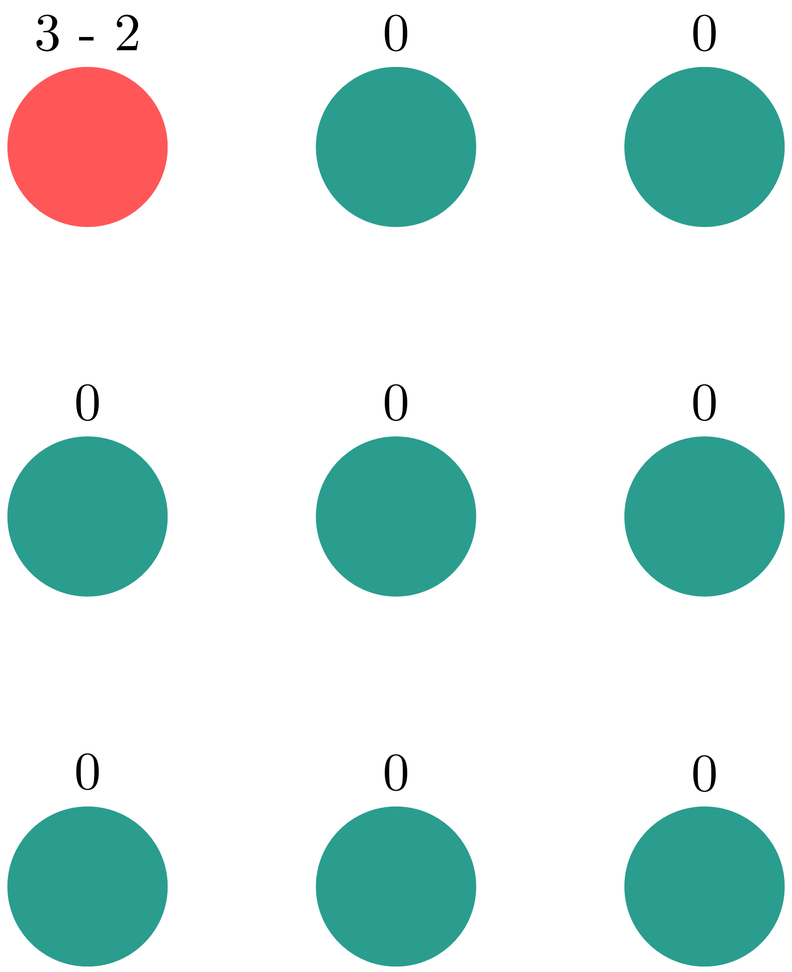
Let’s assume that a height advantage makes one more attractive (+3), and a disadvantage is bad (-3).

And that this boost overweights the discomfort of wearing heels (-2).



Let’s assume that a height advantage makes one more attractive (+3), and a disadvantage is bad (-3).

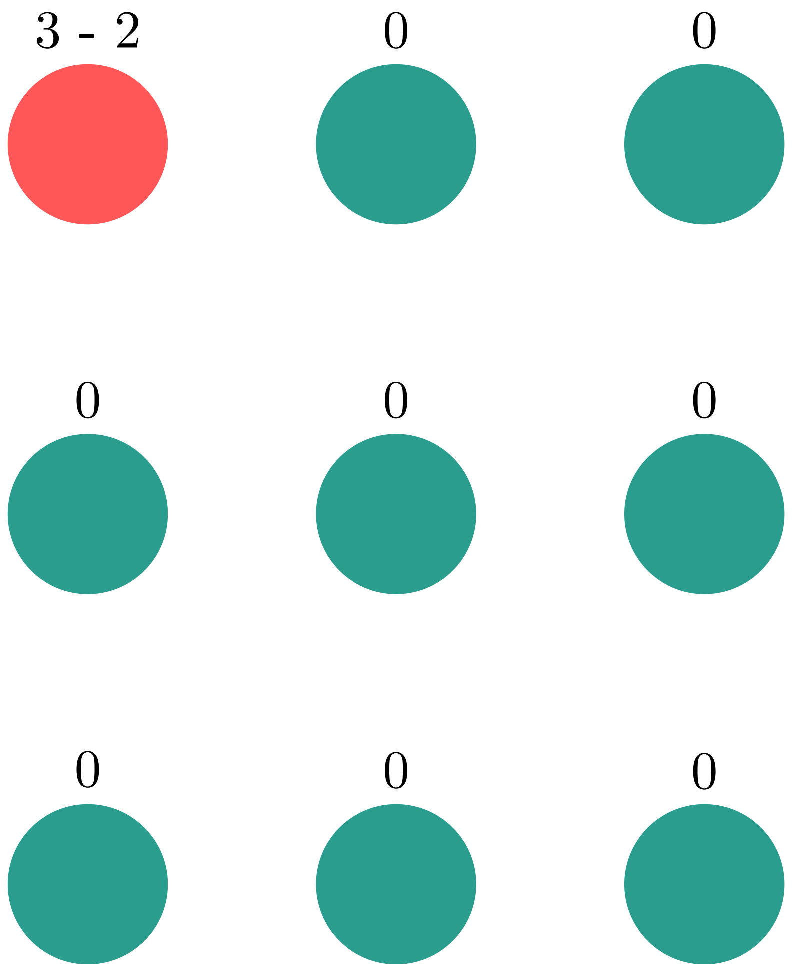
And that this boost overweights the discomfort of wearing heels (-2).



Let's assume that a height advantage makes one more attractive (+3), and a disadvantage is bad (-3).

And that this boost overweights the discomfort of wearing heels (-2).

So everyone adopts high heels.



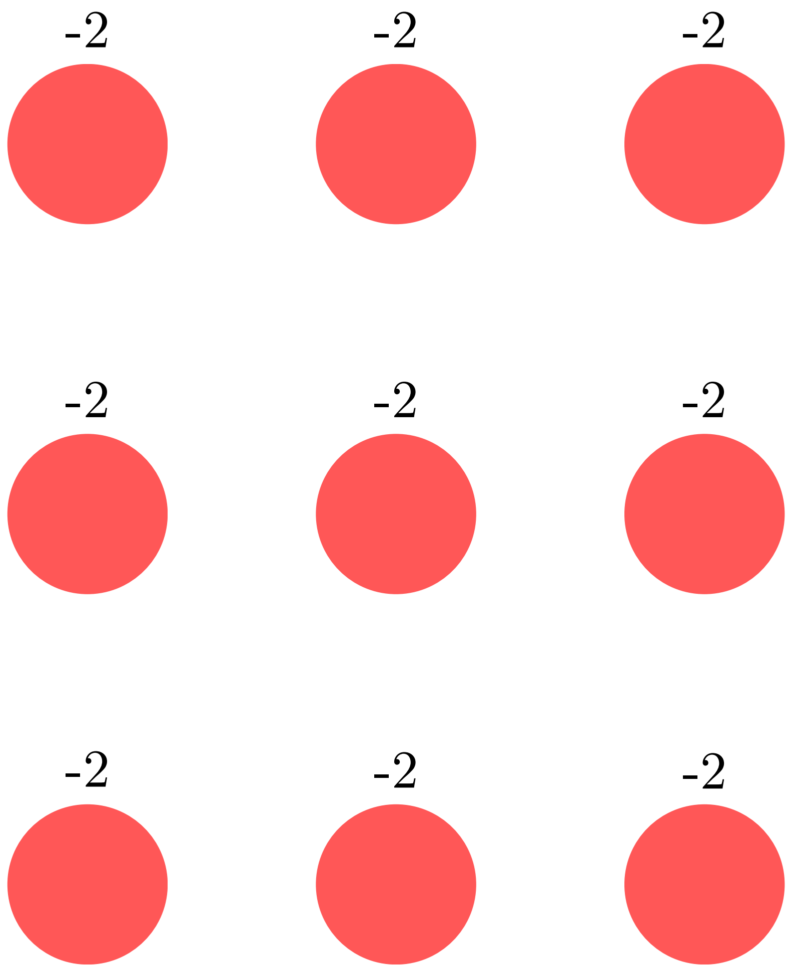
no heels

heels

Let’s assume that a height advantage makes one more attractive (+3), and a disadvantage is bad (-3).

And that this boost overweights the discomfort of wearing heels (-2).

So everyone adopts high heels.



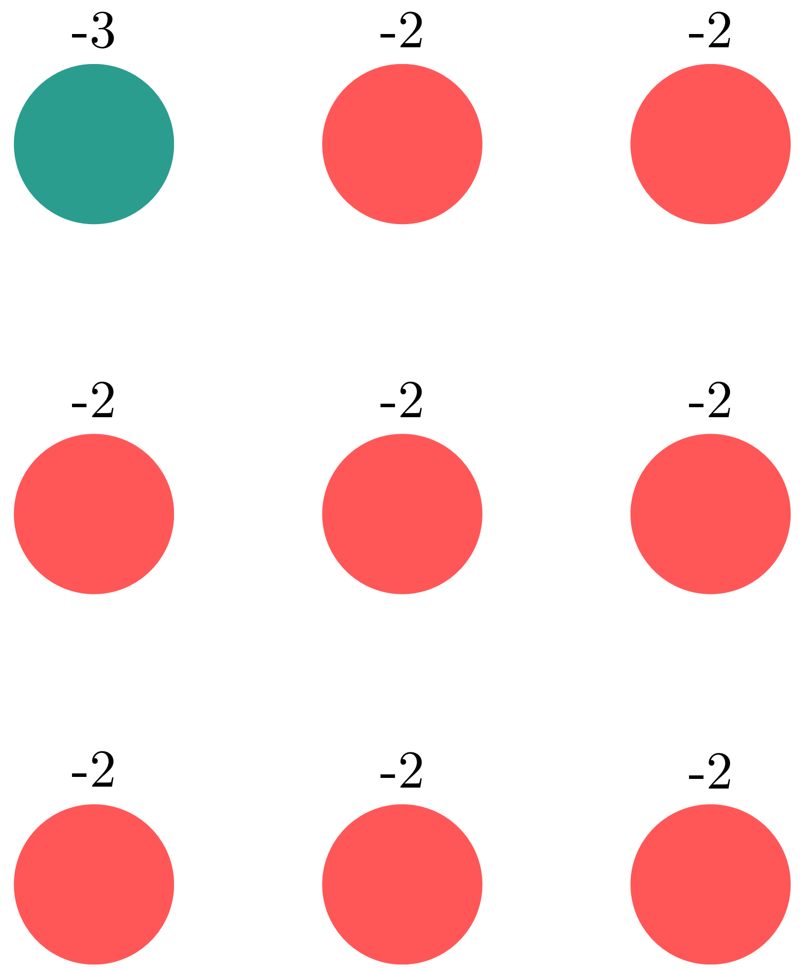
no heels heels

Let's assume that a height advantage makes one more attractive (+3), and a disadvantage is bad (-3).

And that this boost overweights the discomfort of wearing heels (-2).

So everyone adopts high heels.

In a world of high heels, showing up without them puts one at a disadvantage.



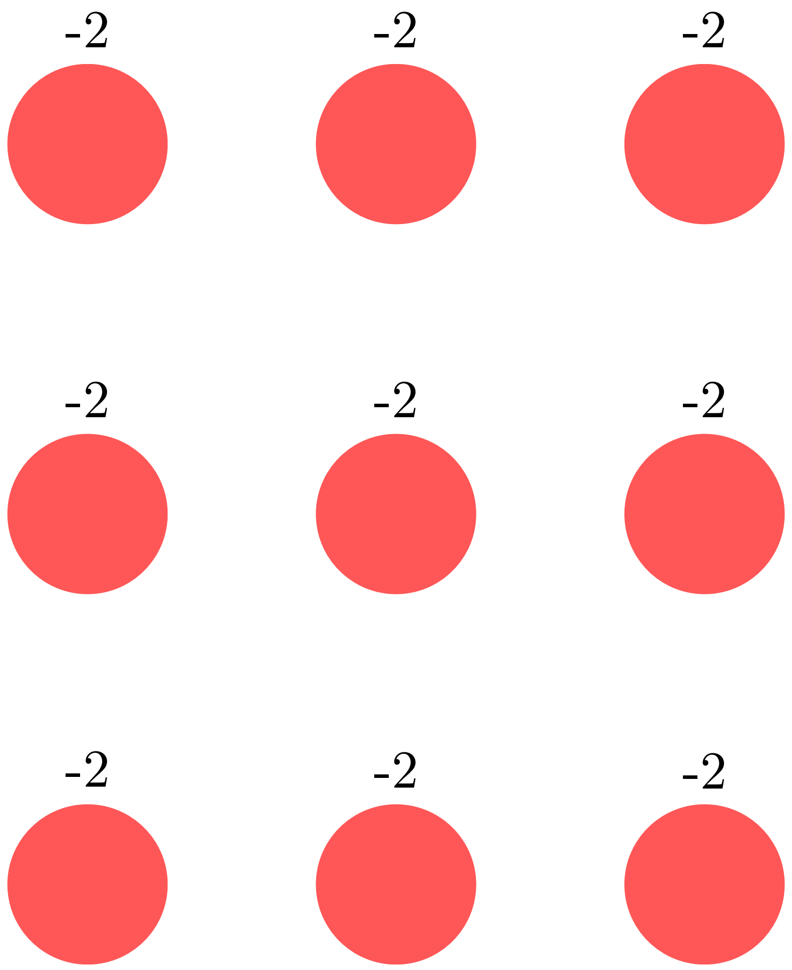
Let’s assume that a height advantage makes one more attractive (+3), and a disadvantage is bad (-3).

And that this boost overweights the discomfort of wearing heels (-2).

So everyone adopts high heels.

In a world of high heels, showing up without them puts one at a disadvantage.

At the Nash equilibrium, everyone puts up with the discomfort… even though the height advantage is gone!



As in the Trust Game, the Nash equilibrium for the Prisoner's Dilemma leaves utility on the table.

payoffs

	Cooperate	Defect
Cooperate	-20, -20	-100, 0
Defect	0, -100	-50, -50

pure Nash equilibria

(Keep, Keep)

As in the Trust Game, the Nash equilibrium for the Prisoner's Dilemma leaves utility on the table.

Can we make this more precise?

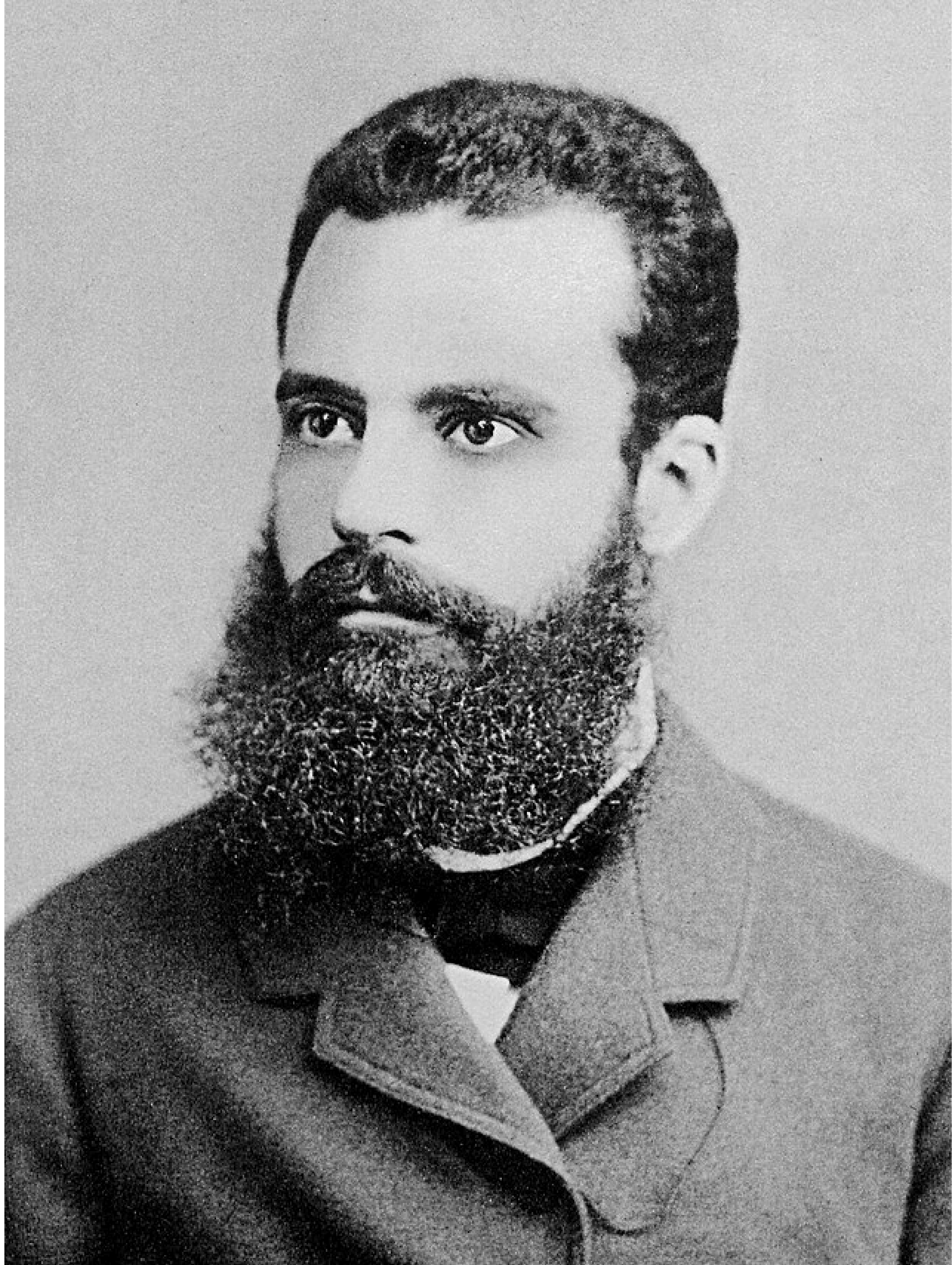
payoffs

	Cooperate	Defect
Cooperate	-20, -20	-100, 0
Defect	0, -100	-50, -50

pure Nash equilibria

(Keep, Keep)

Enter Pareto.



VILFREDO PARETO

How about we look at outcomes where people are (jointly) as well-off as they can be.

In a Pareto optimal outcome no one can be made better off without making someone else worse off.

PARETO DOMINATION & OPTIMALITY

DEFINITION (PARETO DOMINATION)

A strategy profile $\boldsymbol{s}$ *Pareto dominates* strategy profile $\boldsymbol{s}'$ if:

(i) $u_i(\boldsymbol{s}) \geq u_i(\boldsymbol{s}')$, for every agent i, and

(ii) there exists an agent j such that $u_j(\boldsymbol{s}) > u_j(\boldsymbol{s}')$.

PARETO DOMINATION & OPTIMALITY

DEFINITION (PARETO DOMINATION)

A strategy profile s *Pareto dominates* strategy profile s' if:

(i) $u_i(\boldsymbol{s}) \geq u_i(\boldsymbol{s}')$, for every agent i, and

(ii) there exists an agent j such that $u_j(\boldsymbol{s}) > u_j(\boldsymbol{s}')$.

DEFINITION (PARETO OPTIMALITY)

A strategy profile s is *Pareto optimal* if there is no (other) strategy profile s' that Pareto dominates s.

What dominates what in the Trust Game?

payoffs

	Keep	Share
Keep	1, 1	1, 1
Invest	0, 4	2, 2

pure Nash equilibria

(Keep, Keep)

Pareto Optimal strategies

What dominates what in the Trust Game?

(Keep, Keep) and (Keep, Share) are dominated by (Invest, Share).

(Invest, Keep) and (Invest, Share) are not dominated by anything.

payoffs

	Keep	Share
Keep	1, 1	1, 1
Invest	0, 4	2, 2

pure Nash equilibria

(Keep, Keep)

Pareto Optimal strategies

(Invest, Keep), (Invest, Share)

What about the Prisoner's Dilemma?

payoffs

	Cooperate	Defect
Cooperate	-20, -20	-100, 0
Defect	0, -100	-50, -50

pure Nash equilibria

(Defect, Defect)

Pareto Optimal strategies

What about the Prisoner’s Dilemma?

(Defect, Defect) is Pareto dominated by (Cooperate, Cooperate).

Everything else is optimal.

Everything *but* the Nash equilibrium is Pareto optimal!

payoffs

	Cooperate	Defect
Cooperate	-20, -20	-100, 0
Defect	0, -100	-50, -50

pure Nash equilibria

(Defect, Defect)

Pareto Optimal strategies

(Cooperate, Cooperate), (Cooperate, Defect), (Defect, Cooperate)

We can now be more precise about social dilemmas.

SOCIAL DILEMMAS REVISITED

DEFINITION

A *social dilemma* is a situation in which individual incentives are at odds with group incentives. Individual rationality leads members of a group to an outcome that is suboptimal.

Carpenter, J., & Robbett, A. (2022). *Game Theory and Behavior*. MIT Press.
Dawes, R. M. (1980). Social Dilemmas. *Annual Review of Psychology*, 31 (80), 169–193.

SOCIAL DILEMMAS REVISITED

DEFINITION

A *social dilemma* is a situation in which individual incentives are at odds with group incentives. Individual rationality leads members of a group to an outcome that is suboptimal.

More formally, a *social dilemma* is a game in which the equilibria are Pareto dominated by some other outcome.

Carpenter, J., & Robbett, A. (2022). *Game Theory and Behavior*. MIT Press.
Dawes, R. M. (1980). Social Dilemmas. *Annual Review of Psychology*, 31 (80), 169–193.

Can we just not expect that players will gravitate towards a Pareto-optimal outcome?

PARETO IS FRAGILE

Supposing players end up in a situation where both cooperate, they each have a strong incentive to defect.

Pareto-optimal outcomes may not survive, in the long run.

payoffs

	Cooperate	Defect
Cooperate	-20, -20	-100, 0
Defect	0, -100	-50, -50

pure Nash equilibria

(Defect, Defect)

Pareto Optimal strategies

(Cooperate, Cooperate), (Cooperate, Defect), (Defect, Cooperate)

How is this relevant to the problem of cooperation?

JOHN NASH

Note that the numbers in the payoff matrix are not *per se* relevant.

What's important is the *relationship* between them.

The Prisoner's Dilemma

GENERAL VERSION

There are two players, each with two actions: Cooperate or Defect.

If they both cooperate they both get a payoff of R (the *reward*).

If they both defect, they each get a payoff of P (the *punishment*).

In the case of defection with cooperation, the defector gets T (the *temptation*), while the cooperator gets S (the *sucker's payoff*).

The relationship between the payoffs is T > R > P > S.

payoffs

	Cooperate	Defect
Cooperate	R, R	S, T
Defect	T, S	P, P

pure Nash equilibria

(Defect, Defect)

Pareto Optimal strategies

(Cooperate, Cooperate), (Cooperate, Defect), (Defect, Cooperate)

MARTIN NOWAK

Things become even clearer when considering a simplified version of the Prisoner's Dilemma: the *Donation Game*.

Nowak, M.A. (2006). *Evolutionary Dynamics*. Belknap Press

The Donation Game

SPECIAL CASE OF PRISONER'S DILEMMA

There are two players, each with two actions: Cooperate or Defect.

A cooperator pays a cost c for the other player to receive a benefit b, with b > c > 0.

A defector does not pay any cost, and provides no benefit.

Nowak, M.A. (2006). *Evolutionary Dynamics*. Belknap Press

payoffs

	Cooperate	Defect
Cooperate	$b - c, b - c$	$-c, b$
Defect	$b, -c$	$0, 0$

pure Nash equilibria

(Defect, Defect)

Pareto Optimal strategies

(Cooperate, Cooperate), (Cooperate, Defect), (Defect, Cooperate)

A lot of social dilemmas have the structure of a Prisoner’s Dilemma.

VAMPIRE BAT ELDER

Vampire bats face a prisoner's dilemma when having to decide whether to feed their hungry colleagues.

LANCE ARMSTRONG

Sports people too, when deciding whether to take performance enhancing drugs.

Schneier, B. (2006, August 10). Drugs: Sports' Prisoner's Dilemma. *Wired*.

THE UN

Or countries deciding whether to cut down carbon emissions.

MARTIN NOWAK

Indeed, the Prisoner's Dilemma is the paradigmatic game used to study the evolution of cooperation.

Nowak, M.A. (2006). *Evolutionary Dynamics*. Belknap Press.

We can make the problem of cooperation more precise now.

How can we manage to avoid bad equilibria in social dilemmas?