2 – People & Preferences

Extended slides
this is an extended version (much more crowded with text and with additional explanations) of the slides I will project in class. You can use them as lecture notes.

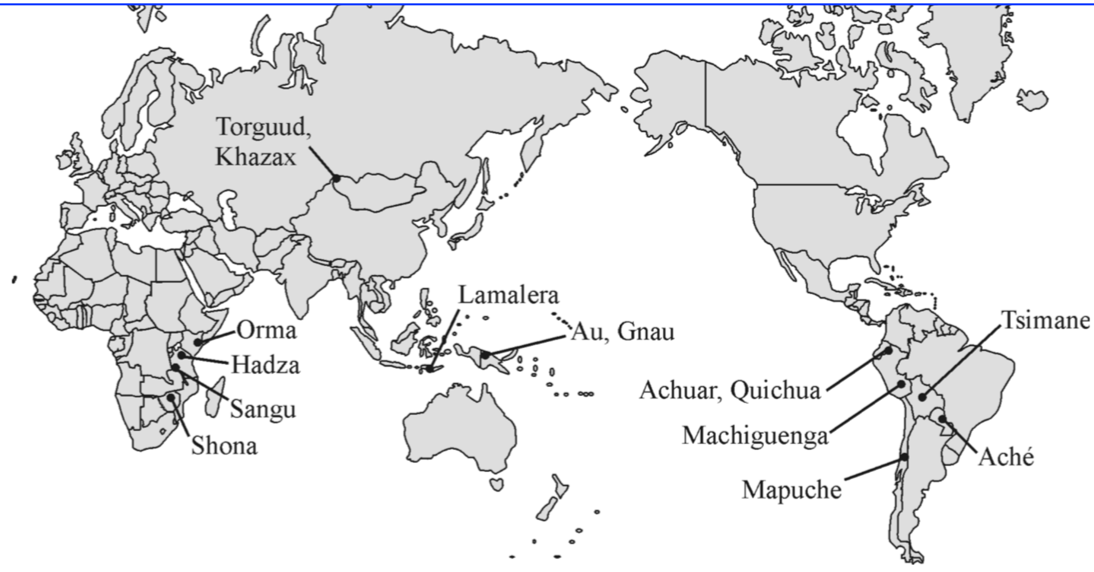


Figure 2.9: A map of the world showing the

Study Materials for this Section:

- Chapter 2 of the textbook ("People: Preferences, Beliefs and Constraints")

2 – People & Preferences

Some questions we ask in this Section:

- How do we study individual behavior in social interactions? How does risk affect people's choices?
- How do we know what will people try to achieve when they make economic choices?
- Do people care only about their material payoffs?
- Do altruism, reciprocity, envy and other 'social preferences' relevantly influence economic behavior?

The main ideas of this Section

- To study people's economic choices, we need to make assumptions about their motives (*preferences*).
- A *rational* person has *complete and consistent* preferences.
- The concept of *expected payoffs* can be used to study decision-making under risk.
- Experiments and case studies show that people are motivated not only by self-interest, but also by altruism, reciprocity, aversion to inequality (*social preferences*).
- Policy-makers must consider this variety of motives when designing optimal rules & incentives.

2 – People & Preferences

The roadmap through Section 2

1. Preferences, beliefs and constraints
2. Risk & expectations
3. Sequential games
4. Social preferences.
5. Experiments on economic behavior.
6. Incentives & social preferences

1 – Preferences, beliefs & constraints

- The ‘preferences, beliefs and constraints’ approach:

  Faced with constraints determined by what is feasible, a person chooses from the set of feasible options the action that she believes will result in the best possible outcome for her.

- *Rational actor* approach

1 – Preferences, beliefs and constraints

PREFERENCES

- Our evaluation of the possible outcomes of our actions.
- Preferences guide people's choices by providing a ranking of the possible outcomes, from most to least preferred.
- They depend on many things: tastes, habits, emotions, commitments, social norms, psychology…

BELIEFS

- Our understanding of the outcomes that will result from actions *"I believe that action X would result in outcome Y"*
- Can depend on what you expect others to do.

CONSTRAINTS

- The feasible set of actions, meaning actions that are open to us.
- There are limits to what we can do. The set of actions that one can take is limited (constrained by feasibility)
- Limits are determined by law, physical capacities, social norms…

What does it mean to be *rational*?

- You are *rational* if your actions are consistent with your goals.
- You are *rational* if
  - *you can rank things* (outcomes or goods) from the one you like most to the one you like least in a consistent way.
  - *you act* in such a way as to try to get the things you like.
- Rationality has *nothing to do with the specific content* of your desires. It just has to do with having consistent desires and acting in accordance with them (whatever they are):
  - It can be perfectly rational to have unusual preferences (like preferring stale bread over ice-cream).

Operational definition of rationality:

A rational agent has *<u>complete</u>* *&* *<u>transitive</u>* *preferences.*

- Completeness:
  for any possible pair of outcomes, A and B, either $A \succ B$ or $A \prec B$ or $A \sim B$
  [rules out the case in which you throw up your hands and say "I just cannot compare these two outcomes"].

- Transitivity (or consistency):
  if $A \succ B$ and $B \succ C$, then $A \succ C$.
  [otherwise it is not a ranking!]
  ex: if you prefer pizza to falafel and falafel to macaroni & cheese, you cannot prefer macaroni & cheese to pizza (else you are not rational according to this definition).

Decision-making under risk

- Many decision-making processes involve *risk*.
- How are we to study decision-making under risk?
- Suppose you must decide whether to take a certain action or not. You don't know for sure what will be the outcome of the action.
- But assume that you do know:
  - all the different outcomes that could possibly result from the action;
  - the payoff that you would receive in each possible outcome;
  - how likely each possible outcome is (*probabilities*).
- You can then calculate the *expected payoff* (or *expected value* or *expected utility*) of the action. This depends on the payoff to each of the possible outcomes, and how likely each is.
- In particular, the expected payoff is *the weighted sum of the payoffs of all possible outcomes, with weights equal to probabilities.*

Decision-making under risk

- Risky action: betting \$5 on heads in a coin toss.
- Possible outcomes: heads or tails.
- Payoffs associated with the possible outcomes: +\$5 if heads comes out; -\$5 if tails comes out;
- Probability that heads comes out: 0.50 (50%);
- Probability that tails comes out: 0.50 (50%);
- $E(\pi_{heads})$ or $\hat{\pi}(heads)$ indicates the expected payoff from betting heads. It is equal to:

$$\underbrace{\hat{\pi}(heads)}_{\text{Expected payoff from betting on heads}} = \underbrace{0.5}_{\text{Probability of winning}}\underbrace{(\$5)}_{\text{Payoff if you win}} + \underbrace{0.5}_{\text{Probability of losing}}\underbrace{(-\$5)}_{\text{Payoff if you lose}} = \$0$$

Decision-making under risk

- 2 available actions: (1) Take an umbrella to class; (2) Don't take it.
- 2 possible outcomes: (1) Rain; (2) No Rain.
- Payoffs in each possible outcome:

		Rain	*No Rain*
Action	**Take**	15	8
	Don't take	3	20

- Probability of rain (from weather forecast): 0.6 (60%)
- Expected payoff of taking the umbrella:

$$\hat{\pi}(take) = 0.6(15) + 0.4(8) = 12.2$$

- Expected payoff of *not* taking the umbrella:

$$\hat{\pi}(don't\ take) = 0.6(3) + 0.4(20) = 9.8$$

- An expected payoff-maximizer would take the umbrella.

Decision-making under risk

A note about notation:

In your textbook (Sec 2.3), the expected payoff to a generic action *x* with 2 possible outcomes is indicated by

$$E(u_x, P\ ) = P\pi(x|1) + (1 - P)\pi(x|2)$$

- The same formula we have applied here, but for a generic action x;
- $E(u_x, P\ )$ is the expected payoff from action x, given that the probability of outcome 1 is P;
- P is the probability of outcome 1, and (1-P) is the probability of outcome 2;
- $\pi(x|1)$ is the payoff you get from action x in case outcome 1 occurs;
- $\pi(x|2)$ is the payoff you get from action x in case outcome 2 occurs;

Calculable risk vs. fundamental uncertainty

- This theory of decision-making under risk assumes that:
  - The agent has a list of all the outcomes that might possibly occur;
  - and can estimate the probability of each of them;
- This is called *calculable risk* (or just *risk*).
- A good approximation of reality? *Not always*
- Think about the onset of the Covid epidemics in 2019: did we have a full list of all possible ways it could develop? Could we calculate the probability of each possible outcome?
- When you don't have a full list of possible outcomes and/or you can't calculate the probability of each outcome, you face *fundamental uncertainty*.
- Economists' understanding of optimal decision-making under fundamental uncertainty is still limited/incomplete.

Expected payoffs in games

- Just like we calculate the expected payoff from a generic action, we can also calculate the expected payoff from a strategy in a game.
- Take the 'planting in Palanpur' game (an *Assurance Game*).
- Aram (A) wants to figure out the expected payoff to each of his possible strategies (Plant Early or Plant Late), and then choose the one with highest expected payoff;
- Aram believes that Bina (B) will plant early with probability P;

		B	
		Plant Early	Plant Late
A	Plant Early	4, 4	0, 3
	Plant Late	3, 0	2, 2

- **Expected payoff from planting early:**

$\hat{\pi}(Plant\ Early)$ = 4P + 0(1-P)= 4P

- **Expected payoff from planting late:**

$\hat{\pi}(Plant\ Late)$ = 3P + 2(1-P)= 2+P

- The expected payoff from a strategy is a function of P: the probability that the other player plants early.
- Which strategy gives highest expected payoff? This also depends on P.

The breakeven probability

- *Breakeven probability*: the value of P that makes the expected payoff of the two possible strategies equal;
- How to calculate the breakeven probability for player A in the 'Planting in Palanpur' game? Three steps

1. Write down the expected payoffs of the two possible strategies
   $\hat{\pi}(\boldsymbol{Plant\ Early})$ = 4P + 0(1-P)= 4P
   $\hat{\pi}(\boldsymbol{Plant\ Late})$ = 3P + 2(1-P)= 2+P
2. Write an equation setting them equal
   E(u(Plant Early))= E(u(Plant Late)) → 4P + 0(1-P) = 3P + 2(1-P)
3. Solve the equation for P
   P = 2/3 → this is the <u>breakeven probability</u>

- In words: when the probability of player B planting early is 2/3, player A gets the same expected payoff from planting early or late.
  This means that P=2/3 is the breakeven probability.
- The breakeven probability is the value of P that makes the player indifferent between her two available strategies;

Risk-dominance

- Suppose Aram has no information on the possible behavior of Bina;
- Should Aram plant early or late?
- *Principle of insufficient reason*: not knowing much, Aram assigns equal probability (½) to both possible actions of Bina (Plant Early or Plant Late).
- So Aram plays the strategy giving highest expected payoff when P=0.5;
- Expected payoffs with P=0.5:
  - $\hat{\pi}(\boldsymbol{Plant\ Early})$ = 4(0.5) + 0(1-0.5)= 2
  - $\hat{\pi}(\boldsymbol{Plant\ Late})$ = 3(0.5) + 2(1-0.5)= 2.5
- → Aram, who assumes P=0.5, will Plant Late;
- → Plant Late is the *risk-dominant* strategy for Aram;
- *Risk-dominant strategy:* that which maximizes expected payoff when P=0.5;
- Plant Late is the risk-dominant strategy also for Bina (check for yourself!)
- When both players have a risk-dominant strategy, as in this case, we have a *risk-dominant equilibrium*.
- The *risk-dominant equilibrium* is (Plant Late; Plant Late)

We can also see break-even probability and risk-dominance visually.

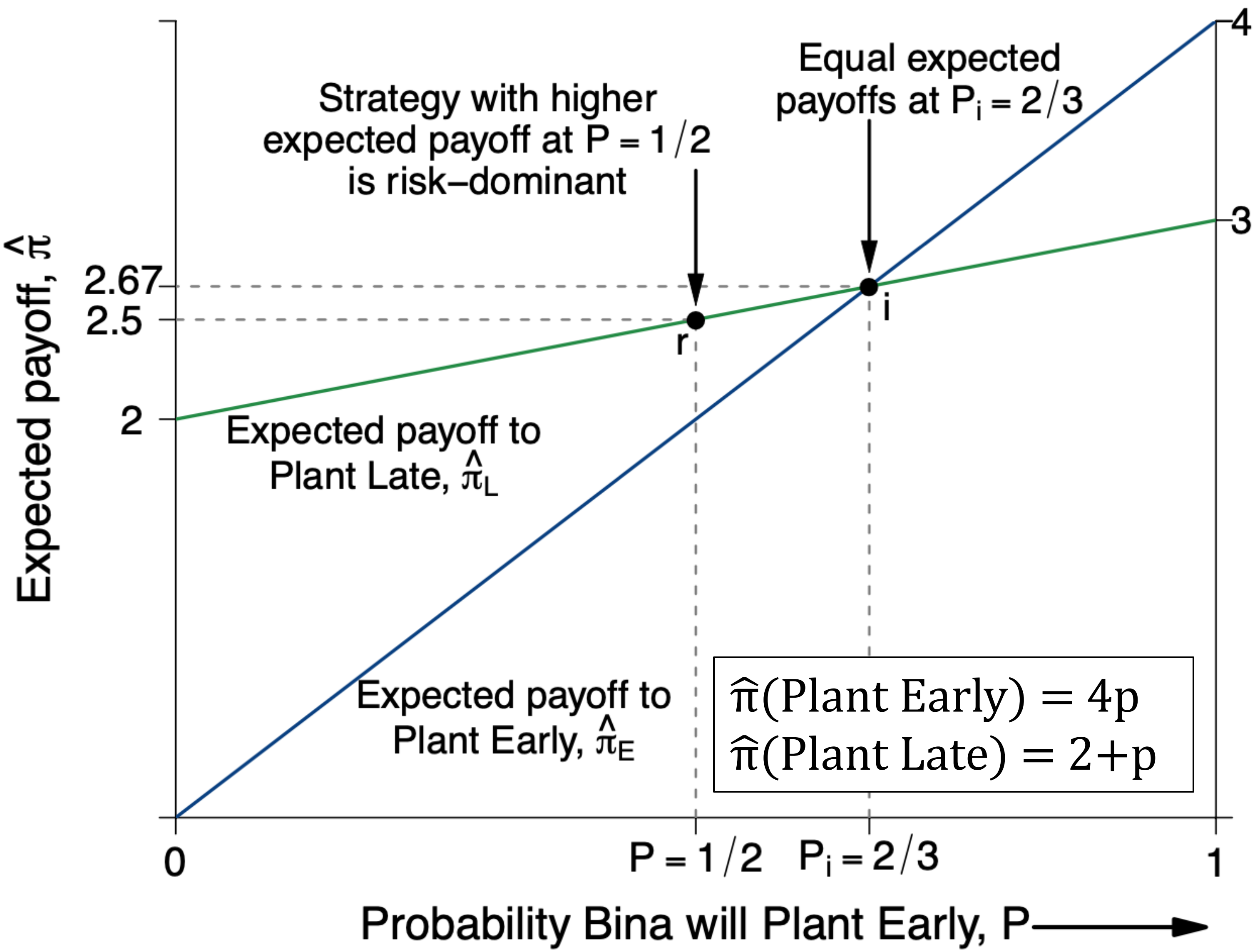
Make a graph with P on the horizontal axis and expected payoffs on the vertical axis

green line= $\hat{\pi}(\text{Plant Late})$ as a function of P;

blue line= $\hat{\pi}(\text{Plant Early})$ as a function of P.

Break-even probability is the P corresponding to the point where the two lines cross (P=2/3).

Risk-dominant strategy is the one that gives highest payoff with P=0.5;



Summing up:

- The *break-even probability* is the value of P that makes the player indifferent between her two available strategies, because it makes the expected payoffs equal;
- *Principle of insufficient reason*: in the absence of information, a player considers all possible actions by the other player equally likely;
  - It is an assumption we are making: we don't know if it's always true!
- The *risk-dominant strategy* for a player is the strategy that yields the highest expected payoff when she attributes equal probabilities to all possible actions of the other player (P=1/2 in 2x2 games).
- When both players have a risk-dominant strategy, the game has a *risk-dominant equilibrium* (= both players playing the risk-dominant strategy).

Breakeven probability & tipping points

- Imagine that in the Palanpur game, Aram was not facing one single farmer but many other farmers (more realistic in this case!).
- In this interpretation, player B is not one single farmer but *the rest of the village*.
- *P* is now re-interpreted as the proportion of farmers in the village who plant early.
- Each individual farmer (like Aram) will plant early as long as more than 2/3 of the population plants early.
- Each will want to plant late when less than 2/3 of the population plants late.
- 2/3 is the *tipping point*: whenever this threshold is passed, people will collectively shift one way or another.

Nash equilibrium vs. risk-dominant equilibrium

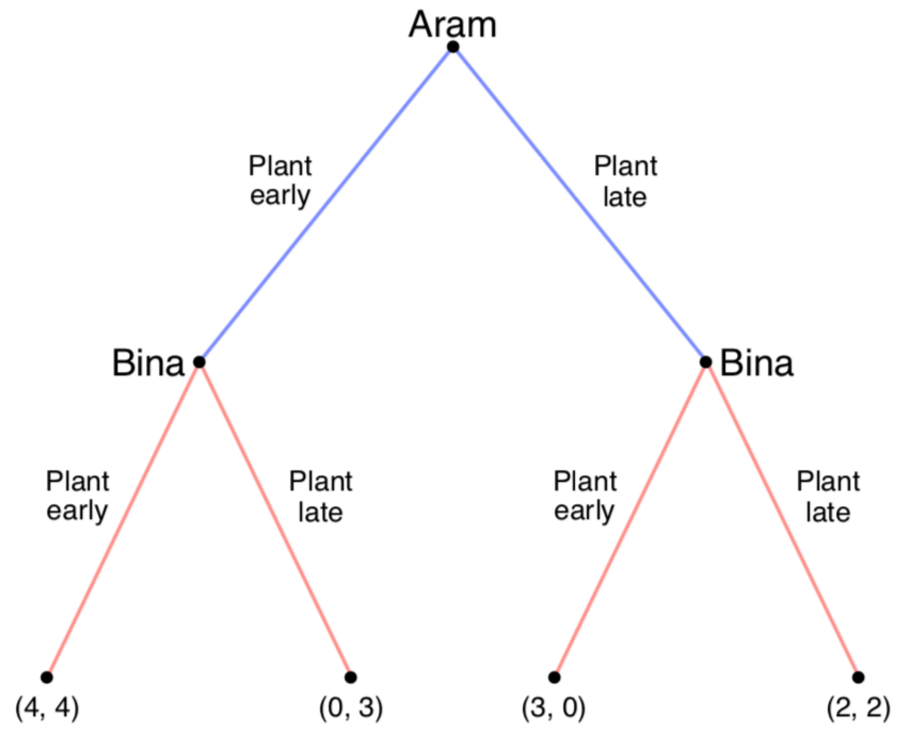
- *Nash Equilibrium*: a mutual best-response.
- Often games have multiple Nash Equilibria.
- Concepts we can use when there are multiple N.E.:
  - *Payoff-dominant equilibrium:* a N.E. equilibrium is payoff dominant if no other N.E. equilibrium exists that is Pareto-superior to it.
  - *Risk-dominant equilibrium:* the N.E. in which both players are playing a risk-dominant strategy.
- The payoff-dominant equilibrium is the one that would make players better with respect to the other equilibria.
- The risk-dominant equilibrium is the one that is likely to be realized when players are very uncertain about the actions of the other player(s).

		B	
		Plant Early	Plant Late
A	Plant Early	4, 4	0, 3
	Plant Late	3, 0	2, 2

- (Plant Early, Plant Early) is *payoff dominant*
- but (Plant Late, Plant Late) is *risk-dominant;*
- When each player is *uncertain* about the actions of the other players, they are likely to end up planting late (the risk-dominant strategy), although this equilibrium makes them worse off with respect to the plant-early equilibrium;
- This shows how *coordination failures* can happen also when there is a Pareto-efficient N.E.;

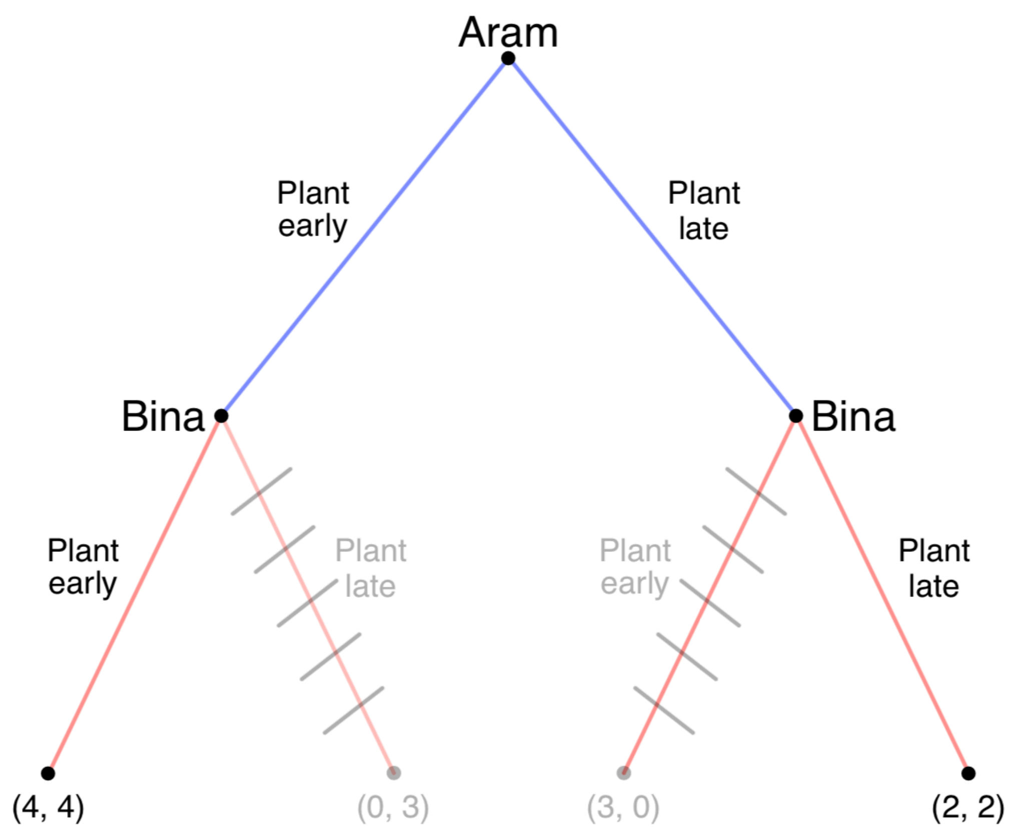
3 – Sequential games

- *Simultaneous games*: players make their choices simultaneously
  - all games we introduced so far were simultaneous;
- *Sequential games*: there is an order of play.
- *Game Tree* (or *extensive form*) is used to study sequential games.
- Suppose that the 'Planting in Palanpur' game was sequential instead of simultaneous: Aram moves first; Bina then responds.
- The corresponding Game Tree:



How to solve a sequential game? The method of *backward induction*

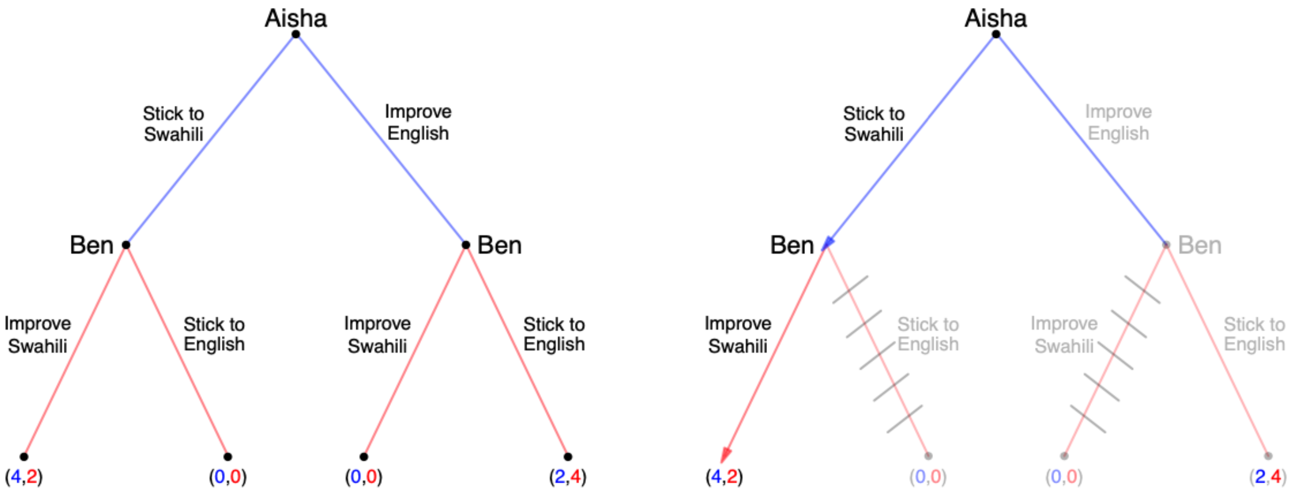
- *Aram (1st mover)* chooses when to plant using *backward induction:*
- He assumes that B wants to maximize her payoff
- If he plants early, B. will plant early (payoffs: 4,4);
- If he plants late, B. will plant late (payoffs: 2,2);
- Given B's expected response, Aram should plant early to maximize his payoff (4 > 2).



(b) Player 1's Reduced Choice of Actions

3 – Sequential games

- In some sequential games, 1st mover has an advantage over other players;
- Consider a sequential version of the language game (a *disagreement game*);
- Assume that Aisha moves first, and Ben then responds;
- When doing the simultaneous version, we found two equilibria: (English, English) and (Swahili, Swahili).
- Solving the sequential version by backward induction, we find only one equilibrium solution: (Swahili, Swahili)



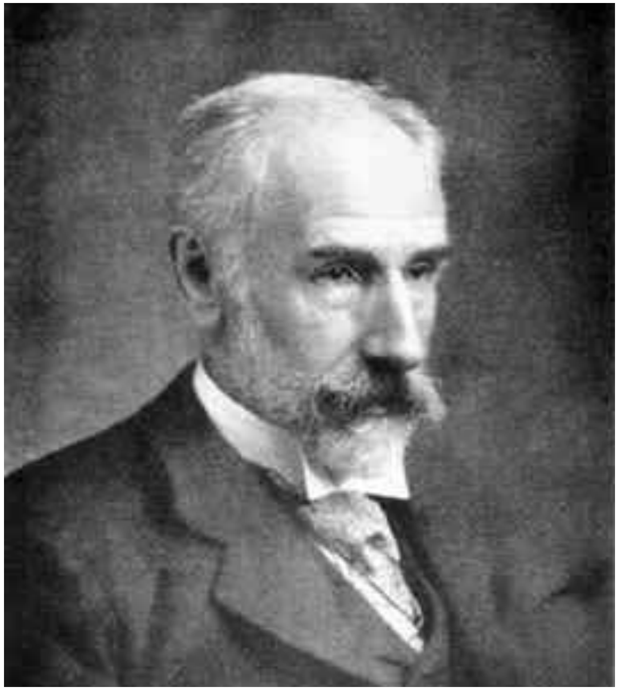
(a) Full game tree (b) Solved game tree

Takeaway: In the sequential version of a Disagreement Game, the first mover obtains their preferred outcome; an example of *first mover advantage*.

4 – Social preferences

- An implicit assumption we made so far: *self-regarding preferences*.
  - We assumed that a player cares only about her own payoff;
- In reality, do people care only about their own individual payoffs?
  - In many situations, no.
- Often people have *other-regarding* (or *social) preferences:* their evaluation of outcomes also depends on the results experienced by others.
- This can involve generosity or reciprocity, but also spite, retaliation, envy, …

The self-interest axiom



Francis Edgeworth

(influential late-19th Century economist):

"The first principle of economics is that every agent is actuated only by self-interest."

- *Self-interest axiom:* the assumption that people only have self-regarding preferences.
- Economics used to always adopt the self-interest axiom.
- Today we know that it can be a good approximation in some cases, but not others.

Social preferences

- People's behavior is often better explained by social preferences:
  - People care about what happens to others: altruism, but also spite and envy.
  - People like to reciprocate: meet generosity with generosity and punish those who behave 'badly'.
  - All this means that in choosing their actions people take into account not only the consequences for themselves, but also for others. This is what we mean by social (or other-regarding) preferences.

Self-interest and rationality

- Rationality and self-regarding behavior are two distinct concepts.
- Rationality $\not\Rightarrow$ self-regarding preferences.
- Rationality is about consistently pursuing your ends (transitivity & completeness), but it *does not* dictate what your ends should be.
- It is perfectly rational to act altruistically if you care about other people! But also acting on the basis of envy or spite can be perfectly rational.
- A person with other-regarding preferences (altruism, reciprocity, spite,…) can be perfectly rational.

Main types of preferences:

- *Self-regarding preferences ("homo economicus"):* you only care about your own payoff.
- *Other-regarding (or social) preferences:* You care about the payoffs received by others.
- 4 main types of other-regarding (or social) preferences:
  - *Altruistic preferences:* you place a positive value on the well-being of others. You are willing to personally incur a cost in order to benefit someone else.
  - *Reciprocal preferences*: you respond reciprocally, rewarding those who have been generous in the past, and seeking to penalize those who have acted badly.
  - *Inequality-adverse* (or *fairness-based*) *preferences*: you dislike unfairness and prefer more equal outcomes.
  - *Spite and 'us versus them' distinctions*: you place a negative value on outcomes experienced by some others. Often motivated by some type of group membership (ethnic, religious, racial, …).

5 - Experiments

- How do we study preferences?
- How do we tell whether people tend to have self-regarding or social preferences?
- If people tend to have social preferences, how do we know what types of social preferences are common?
- Social scientists have used experiments.
  - useful to test theories and highlight new facts;
  - they have been used to assess whether people's economic behavior is better described by the self-interest axiom or by some kind of social preferences.

Experiments in economics

- *Experiments* are used to empirically test hypotheses;
- *Field experiments:* you randomly assign a treatment to certain units of observation (individuals, villages, regions…) but not others;
  - *Example:* implement a micro-credit program in some randomly selected towns in a region. Then you compare outcomes of interest (start-ups, employment, average income…) between cities where the program was implemented (the treated group) and cities where it was not implemented (the control group). This allows you to estimate the effect of the program.
- *Lab experiments:* in a controlled environment (usually a computer lab), you make people play some "game", the rules of which have been established by the researcher;
  - *Example:* make people play a prisoners' dilemma, with monetary stakes as payoffs. Then see if people play their dominant strategy equilibrium or not, and if the prediction of the N.E. is borne out or not;
- Here we will mainly discuss lab experiments: generally most suited for studying preferences.

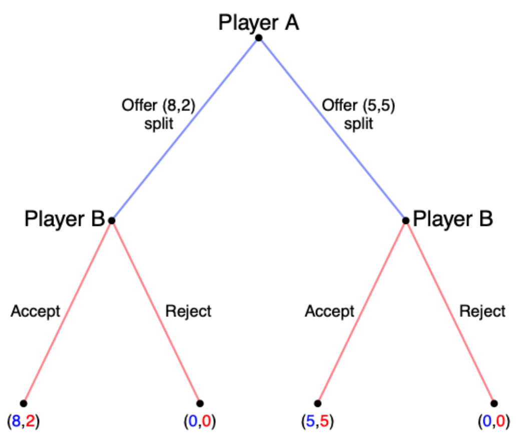
The Ultimatum Game (UG)

- Subjects are anonymously and randomly paired for a single interaction in a “take-it-or-leave-it” game.
- In each pair, one subject is randomly given the role of Proposer; the other will be the Responder
- Proposer is given a fixed endowment (say, 10$), to be divided with the Responder.
- Game sequence:
  1. The Proposer offers a share of the endowment of her choice to the Responder.
  2. The Responder can accept of reject.
- If Responder rejects the offer, they both get nothing.
- If Responder accepts, she gets the proposed share, while Proponent keeps the rest.

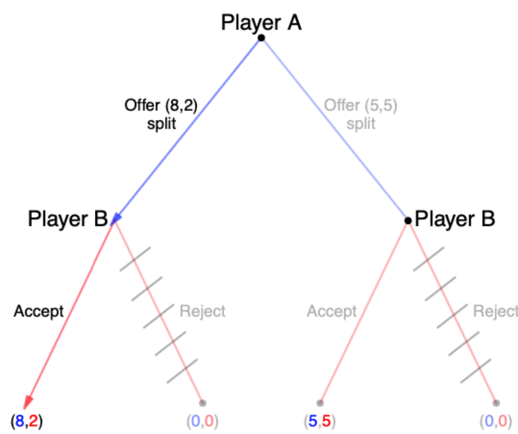
Prediction of the self-interest axiom:

- Proponent will offer a very small amount (ε), close to zero.
- Responder will accept, because $\varepsilon > 0$
- We obtain this prediction by solving the game under the assumption that players have self-regarding preferences.

Game tree for the UG:



(a) Full game tree



(b) Self-regarding players

- Assume for simplicity only 2 possible offers: (8,2) and (5,5)
- *Player A* (Proposer) uses backward induction:
  - in both cases Player 2 will accept, because both offers are better than 0.
  - so it is best for Player A to offer the (8,2) split.
- *Player B* accepts the (8,2) offer, because 2 > 0.

- When both players have self-regarding preferences, (8,2) is the outcome that should occur.
- In games where an offer lower than $2 is possible, the Proposer will offer the smallest possible amount greater than zero.

The Ultimatum Game (UG)

- In lab experiments, researchers made people play the UG all over the world, with real monetary stakes.
- The prediction of the self-interest axiom, according to which the proposer would make a very small offer and the responder would accept, invariably fails.
  - *Proposers* tend to offer over 40% of the prize.
  - *Respondents* usually reject offers smaller than 20% of the prize.
- Responders' behavior displays *social preferences*.
  - Specifically, they display *reciprocity*: they are willing to give up some money in order to punish a proposer who made a low offer.
- Proposers' behavior more difficult to interpret: compatible with both altruism and self-regarding preferences
  - (this is because even a self-regarding proposer would make a substantial offer, if she knew that the responder had reciprocal preferences)

5 - Experiments

- Variants of the UG have been tried, in order to better interpret results (see textbook for details)
  - these have provided clear evidence that responders' behavior is really motivated by reciprocity.
- Researchers have also performed *'lab in the field' experiments*: they went to small-scale societies around the world and performed the UG (as well as other games)
  - the prediction of the self-interest axiom fails everywhere: people across the world invariably tend to reject low offers;
  - but there is also differences in behavior across societies;
  - differences in behavior between groups is explained by their different socio-economic institutions: societies organized in a more egalitarian way tended to produce more equal splits, while more individualistic societies produced more unequal splits.

Public Goods Game (PGG)

- The Public Goods Game (PGG) is another important experiment.
- Goods can be of 4 types:

	Excludable	*Non-excludable*
Rival	Private good (clothing, food)	Common property (Pool) Resource (fishing stocks, potential buyers)
Non-rival	Club good (streaming music, online movies)	Public good (global climate, rules of calculus)

- 'Public goods' are non-excludable and non-rival.
  - *Non-excludable*: hard to stop people from accessing it.
  - *Non-rival*: one person's use of the good doesn't reduce the ability of others to use it as well.
- Classic example: a lighthouse.
- Public goods tend to be *underprovided*, because of externalities.

Public Goods Game (PGG)

- *n* players, each is given an endowment *z*.
- each must decide how much to keep for themselves and how much to contribute to a common pot.
  - Each starts with: z
  - Player i's contribution to common pot: e^i
  - Total in public pot: $E = M(e^1 + e^2 + e^3 + \cdots) = M(\sum_{j=1}^{n} e^j)$
  - M (rate of return) greater than 1/n but lower than 1.
  - Player i's payoff: $\pi^i = z - e^i + M\,(\sum_{j=1}^{n} e^j)$
- The common pot is a *public good*: no one is excluded from it (non-excludable), and each gets all of it (non-rival).
- *Free-riding*: you benefit even if you don't contribute

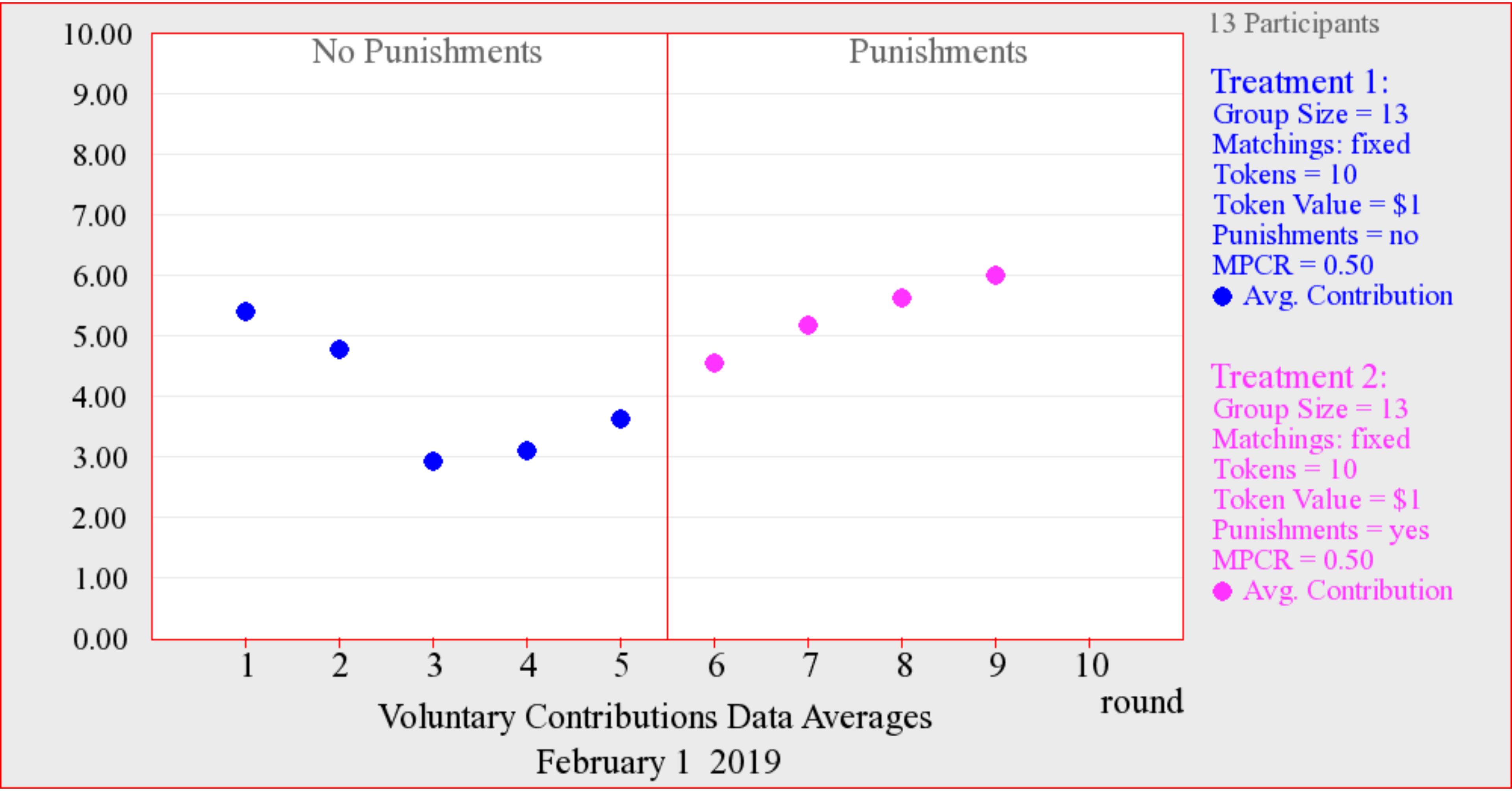
Public Goods Game (PGG)

- Players with self-regarding preferences would never contribute anything in the PGG (e^i=0 for all players);
- This is because for each dollar you contribute, you get M dollars back, and M<1. It is not individually profitable to contribute a positive amount.
- In experiments with the PGG:
  - *in one-shot games* contributions tend to be around ½ of the endowment. Higher if people can communicate before.
  - *in repeated games*, contributions start high, but often decrease towards zero in successive rounds.
- How to explain the decrease in contributions in successive rounds?
  - Self-regarding preferences? Perhaps through playing, people get to understand the game, and learn to play according to self-interest.
  - Sign of reciprocity? people decrease their contributions to punish those who are not contributing (the free-riders), or at least not be taken advantage of by them. This makes contributions decrease as the game goes on.

The Public Goods Game with punishment

- Same basic structure as in the PGG.
- BUT after each round:
  - contributions are made public;
  - each player has the possibility to *punish* some other players, reducing their payoff. But to do so, they must pay some cost;
- With punishment option, people tend to punish the non-cooperators, and contributions remain high throughout.
  - This proves that many people display *reciprocity*: they are willing to cooperate with cooperators, but they don't want to be exploited by selfish types.
  - Reciprocators are willing to personally pay a cost in order to punish the non-cooperators;

Last semester's Econ 203 students playing the public good game in a discussion section:



Social preferences are not irrational

- As we have seen, rationality does not imply self-interest. In principle, social preferences can be perfectly rational.
- But in practice, when people display non-selfish behavior, do they tend to act according to the basic axioms of rationality? (completeness & transitivity)
- Experimental evidence shows that altruistic and reciprocal behavior tend to be associated with rational behavior: *when people are being generous/reciprocal/inequality-averse, they are not less likely to respect completeness & transitivity.*
- Social preferences are not irrational!

Limits of experiments

- *Lab experiments* are very useful to study economic behavior and preferences…
- …but they do have some *shortcomings*:
  - being observed changes your behavior
  - in labs people usually interact anonymously through computers, rather than face-to-face as in real life
  - subjects that participate to experiments tend to be a particular subset of the population (WEIRD* samples)
  - usually make people play clearly strategic games, while real-life situations maybe not so neatly strategic.
- Notwithstanding these problems, experiments are among the best ways to assess preferences, and they have been found to generally predict well behavior in real economic situations.

* *usually from Western, Educated, Industrialized, Rich, and Democratic (WEIRD) societies*.

6 – Incentives & Social Preferences

- *Reciprocity* can change a Prisoner Dilemma into an Assurance Game: you cooperate if you think that the other player will cooperate too.
- In this case, by altering beliefs and expectations about what others will do, you can sustain the cooperative Pareto-efficient equilibrium.
- Combining this with material incentives that reward pro-social behavior may increase the likelihood of obtaining cooperative efficient outcomes.
- Economic incentives and social preferences may thus be *complementary*: reinforce each other.

6 – Incentives & Social Preferences

- …but it is also true that economic incentives can *crowd out* social preferences.
- This happens when introducing material incentives changes the context in which people think they are making decisions (*the "framing").*
- For example, material incentives may shift the framing of a decision from a moral/ethical realm to a mere business transaction.
- Examples:
  - Boston firemen
  - Haifa daycare centers case study

The case of Boston firefighters

- Before Dec 2001, firemen in Boston had unlimited paid sick days;
- *Boston Fire Commissioner* noticed a bunching of sick call-ins on Mondays and Fridays,
- Since Dec 1, 2001, the commissioner ended the Department's policy of unlimited paid sick days.
- He imposed a 15-day sick day limit.
- The pay of firemen exceeding that limit would be docked.
- What did the Boston firemen do?

The case of Boston firefighters

- On Christmas 2001, Boston firemen sick days increased *tenfold* relative to the previous Christmas;
- The commissioner retaliated by canceling their holiday bonus checks;
- Firemen increased sick days even more during 2002: 13,431 sick days, up from 6,432 in 2001; the opposite of what the commissioner wanted to achieve!
- Many firemen had previously *not* taken sick leave even when they could have;
- Now, with a limit and a cost to pay, they took the maximum they could;
- How do we explain that? Spite? Reciprocity?
- The policy backfired because it assumed firemen would behave like *homo economicus*, while they actually had social preferences!

The Haifa daycare centers experiment

- Field experiment in *Haifa daycare centers*
  - Half (randomly selected) introduced a system to fine parents for picking up their children late
  - Half used as control group: no fine introduced.
- What happened after the fine was implemented?
- Lateness actually increased! the fraction picking up their kids late more than doubled
- Fine was a contextual cue: it unintentionally provided information about the appropriate behavior.
- After introduction of the fine, being on time was no longer a *ethical obligation*, but an *economic transaction* with a price.

The Haifa daycare centers experiment

- What happened when the ‘treated’ daycare centers dropped the fine?
- Did the behavior return to the starting point?
- No!
- Parents continued to come late
- The re-framing of their decision had altered their perceptions of what was OK, and it was not sufficient to withdraw the fine in order to restore the initial sense of moral obligation.
- Takaway message: monetary incentives are powerful and often work; but sometimes they can also undermine (*crowd-out*) social preferences, thus producing unintended effects;

The Haifa daycare centers experiment

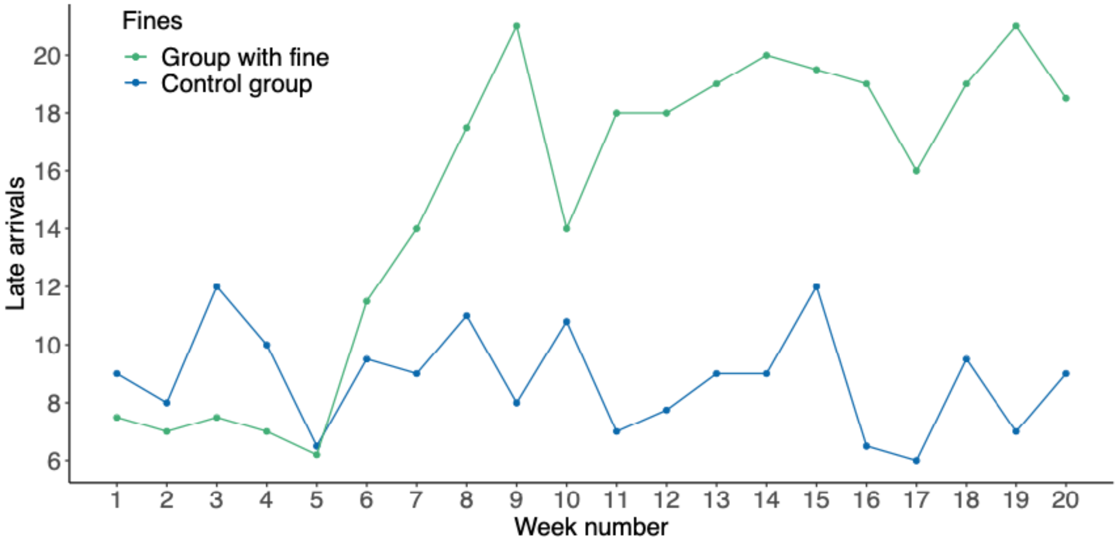


Figure 2.16: **The effect of a fine for lateness in Haifa's daycare centers.** Source: Gneezy and Rustichini (2000a). The fine was imposed in week 5 and retracted in week 17.