Moral Hazard

Bruno Salcedo

Reading assignments: Watson, Ch. 13

Cornell University · ECON4020 · Game Theory · Spring 2017

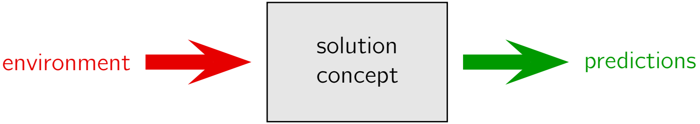
introduction to mechanism design

“Love and duty are not the cement of modern societies...

The mechanism is reciprocity. Seemingly altruistic behavior, based on versions of the I’ll-scratch-your-back-if-you-scratch-mine principle, require no nobility of spirit. Greed and fear will suffice as motivations: greed for the fruits of cooperation, and fear for the consequences of not reciprocating the cooperative overtures of others...

[A selfish rational agent] may not be an attractive individual, but he can cooperate very effectively with others like himself.”

— Ken Binmore (1994) Game Theory and the Social Contract Vol. I



- So far – given an environment, which outcomes are predicted by different solution concepts?

- Now – fixing a solution concept, how can we adjust the environment to achieve different outcomes?

- There is one principal (she) and one or more agents (him, they)
  - The principal chooses a mechanism (game, environment)
  - Agents decide whether to participate or not
  - Those who participate play the chosen game
- **Implementation:** which outcomes are implementable in some mechanism?
- **Efficient MD:** can socially desirable outcomes be implemented?
  - Principal = benevolent social planner
  - In this class, socially desirable = Pareto efficient
- **Optimal MD:** which mechanism maximizes the utility of the principal?
  - Examples: auctions, tax-revenue, contests, industrial organization

- Production in the USSR
- Braess' paradox
- Health insurance
- Prohibition
- Standardized exams
- Royal Economic Society
- Commissions in personal credit sales
- Point system in soccer

socially optimal?

- What does it means for an outcome to be socially optimal?
  - Different approaches: Mill vs. Rawls
  - Interpersonal utility comparisons are tricky
  - Social Choice Theory

- In this class, criterion proposed by Pareto (1927)

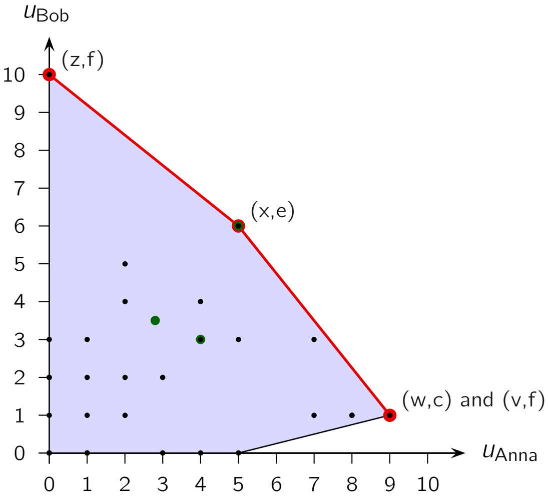
An outcome x Pareto dominates y if at least one agent is strictly better off, and no one is strictly worse off at x

An outcome is Pareto efficient if and only if it is not Pareto dominated

example – 5×6 game

	a	b	c	d	e	f
v	0 , 2	1 , 3	8 , 1	1 , 2	0 , 3	9 , 1
w	1 , 0	4 , 3	9 , 1	5 , 3	2 , 1	0 , 1
x	7 , 3	2 , 4	3 , 0	7 , 1	5 , 6	1 , 1
y	2 , 2	0 , 2	2 , 5	3 , 2	4 , 4	8 , 1
z	0 , 0	1 , 0	0 , 0	5 , 0	4 , 0	0 , 10

example – 5×6 game



example – 5×6 **game**

	a	b	c	d	e	f
v	0 , 2	1 , 3	8 , 1	1 , 2	0 , 3	9 , 1
w	1 , 0	4 , 3	9 , 1	5 , 3	2 , 1	0 , 1
x	7 , 3	2 , 4	3 , 0	7 , 1	5 , 6	1 , 1
y	2 , 2	0 , 2	2 , 5	3 , 2	4 , 4	8 , 1
z	0 , 0	1 , 0	0 , 0	5 , 0	4 , 0	0 , 10

- Everybody would agree on a Pareto improvement

- Pareto efficiency might not be attained for different reasons

  1. **Bounded rationality**: players might fail to fully understand the situation

  2. **Incomplete information**: required information not publicly available

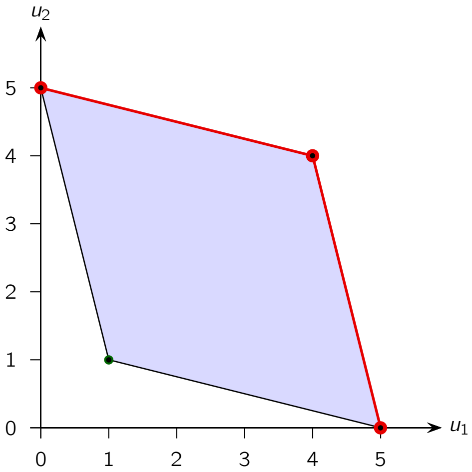
  3. **Moral hazard**: individual incentives for (socially) sub-optimal behavior

- The term “moral hazard” has different (but similar) definitions in different areas of Economics

- In this class, moral hazard means that all the equilibria lead to Pareto inefficient outcomes

prisoners’ dilemma

	C	D
C	4 , 4	0 , 5
D	5 , 0	1 , 1



- Each player i provides a voluntary contribution $c_i \in [0, 1]$ to a public good
- It takes 5$ to produce one unit of the public good

$$P = \frac{1}{5} \sum_i c_i$$

- Agents preferences are given by:

$$u_i(c) = P - c_i$$

- Each player gets a benefit of $1/5$ per dollar provided $\Rightarrow$ unique NE has $c_i^* = 0$ for all players
- If there are more than 5 players is greater than 5, this is Pareto dominated! (why?)
- Players only consider the private benefit of their contributions ignoring the "social benefit"

teamwork

- Teamwork is a form of public good
- Suppose that Anna and Bob are partners in a firm and provide levels of effort e_A and e_B
- They split the total revenue of the project given by

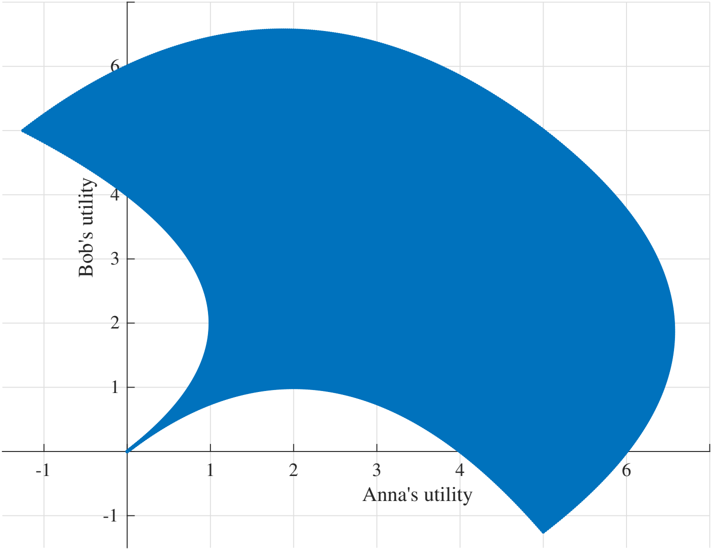
$$f(e) = 2e_A + 2e_B + \frac{1}{10} e_A e_B$$

- The cost of effort is:

$$c_i(e) = \frac{1}{4} e_i^2$$

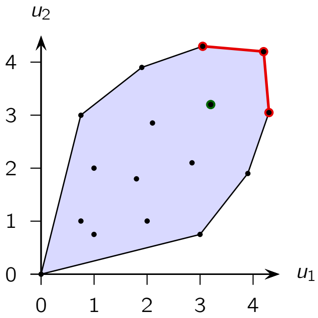
- Payoffs are then given by:

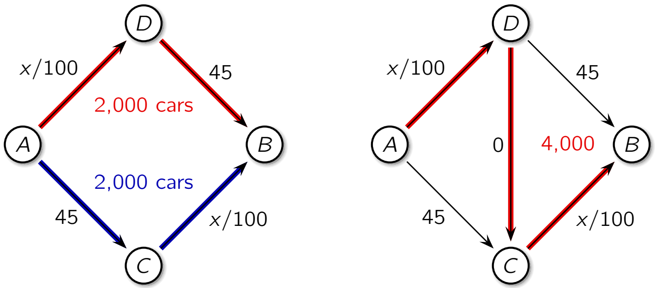
$$u_i(e) = \frac{1}{2} f(e) - c_i(e) = e_A + e_B + \frac{1}{20} e_A e_B - \frac{1}{4} e_i^2$$


6
4
3
2
1
-1
Bob's utility
-1
1
2
3
4
6
Anna's utility

teamwork

	1	2	3	4
1	0 , 0	1 , 0.75	2 , 1	3 , 0.75
2	0.75 , 1	1.8 , 1.8	2.85 , 2.1	3.9 , 1.9
3	1 , 2	2.1 , 2.85	3.2 , 3.2	4.3 , 3.05
4	0.75 , 3	1.9 , 3.9	3.05 , 4.3	4.2 , 4.2





- 4,000 drivers need to go from A to B
- Each driver chooses the fastest route taking traffic into account
- As a result, half the drivers take each route and takes 65 min
- A bridge connecting D to C is built
- Now, all cars take the route ADCB and take 80 min!

- Firms 1 and 2 produce the same good with constant marginal cost $c = 10$ and inverse demand function

$$P(q_1, q_2) = 100 - q_1 - q_2$$

- Each firm maximizes its profits

$$u_i(q_1, q_2) = (90 - q_1 - q_2)q_i$$

- Best response functions are given by:

$$BR_i(q_{-i}) = \frac{90 - q_{-1}}{2} = 45 - \frac{1}{2}q_{-i}$$

- The unique NE is $(q_1^C, q_2^C) = (30, 30)$

- Monopolist would maximize

$$u(q) = (90 - q)q$$

- And the optimal quantity would be $q^M = 45$
- An outcome of the Cournot environment is Pareto efficient if and only if the total supply equals q^M (why?)
- If each firm produced $q^M/2$ then profits for each firm would be:

$$u^* = u_1\left(q^M/2, q^M/2\right) = (90 - 45)\frac{45}{2} = 1012.5$$

- In contrast, in the Cournot equilibrium each firm makes:

$$u^C = u_1(q^C, q^C) = (90 - 60)30 = 900$$

- The Cournot equilibrium is inefficient from the perspective of the firms

games with contracts

- Efficient mechanism design
  - Look for mechanisms that induce Pareto efficient outcomes
  - Can we modify the incentives of the environment to eradicate Moral Hazard?
- Moral hazard can be solved when players can sign complete enforceable contracts at no cost and there are no bargaining failures
- Coase's conjecture:

$$\text{well defined property rights} + \text{no transaction costs} \Rightarrow \text{efficient outcomes}$$

- Suppose the prisoners have the option to sign the following (conditional) contract

> Those who sign this contract commit to keeping silent in case that everyone signs the contract, and confessing otherwise

- Signing the contract would be a Nash equilibrium of the resulting game and cooperation would be implemented

	C	D	Sign
C	4 , 4	0 , 5	0 , 5
D	5 , 0	1 , 1	1 , 1
Sign	5 , 0	1 , 1	4 , 4

- The minimax payoff for player i is the minimum payoff that it can guarantee even if everyone plays against him/her

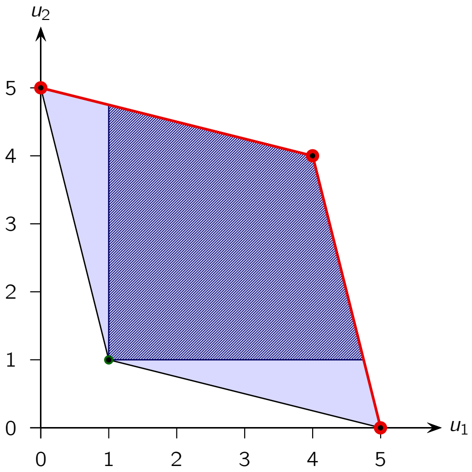
$$\bar{u}_i = \max_{s_i} \min_{s_{-i}} u_i(s_i, s_{-i})$$

- We say that an outcome is individually rational if every player gets at least his/her minimax payoff

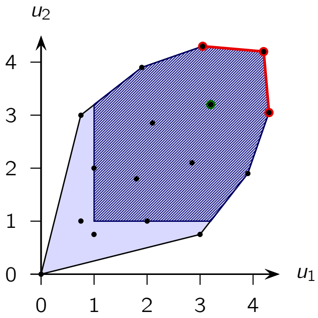
Theorem — There exists a game with contracts that implements an outcome as a Nash equilibrium if and only if the outcome is individually rational

prisoners' dilemma

	C	D
C	4 , 4	0 , 5
D	5 , 0	1 , 1



	1	2	3	4
1	0 , 0	1 , 0.75	2 , 1	3 , 0.75
2	0.75 , 1	1.8 , 1.8	2.85 , 2.1	3.9 , 1.9
3	1 , 2	2.1 , 2.85	3.2 , 3.2	4.3 , 3.05
4	0.75 , 3	1.9 , 3.9	3.05 , 4.3	4.2 , 4.2



teamwork with incomplete contracts

- Suppose not all contracts are possible because effort is not contractible
- An agent cannot prove in court that his/her partner did not work hard enough
- Agents can still write contracts contingent on the total output

If all partners sign this contract:

- If the total revenue is at least 17 it will be split evenly
- If the total revenue is less than 17 it will be thrown away

If at least one partner does not sign this contract then the total output will be split evenly

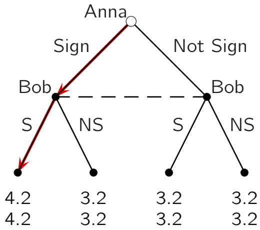
- The total outcome is greater than 17 if and only if both agents provide maximum effort

- If both agents sign the contract then the game becomes

	1	2	3	4
1	$-1, -1$	$-1, -4$	$-1, -9$	$-1, -16$
2	$-4, -1$	$-4, -4$	$-4, -9$	$-4, -16$
3	$-9, -1$	$-9, -4$	$-9, -9$	$-9, -16$
4	$-16, -1$	$-16, -4$	$-16, -9$	$4.2, 4.2$

- The Pareto efficient outcome now becomes a Nash equilibrium

- After one step of backward induction



- The Pareto efficient outcome is induced by a SPNE
- Throwing away excess profits might not be a credible threat unless there is a residual claimer
- Hölmstrom (1982) used this example as an argument for private ownership over the means of production