Morality and Altruism: Evolutionary Foundations and Economic Implications

Ingela Alger (TSE, CNRS, IAST)

Bilkent University Economics Research Union Seminar October 12, 2021

What is economics about?



The Question: is resource allocation efficient?



The Question: is resource allocation efficient?

1st Theorem of Welfare Economics (1950's)

If: (1) if there is a market for every good,

(2) prices are known,

and (3) no individual or firm has power over prices,

then market equilibrium is efficient.

In particular, the 1st Theorem is true even if individuals care solely about own material well-being.

"It is not from the benevolence of the butcher, the brewer, or the baker that we expect our dinner, but from their regard to their own self-interest."

Adam Smith (The Wealth of Nations, 1776)



But... what happens in situations *not* covered by the 1st Theorem?

But... what happens in situations *not* covered by the 1st Theorem?

Such as situations with externalities...



But... what happens in situations *not* covered by the 1st Theorem?

... or interactions within smaller groups (firms, families, etc.)



But in reality *preferences* can help fix these problems.

But in reality *preferences* can help fix these problems.

Experiencing guilt when polluting the air reduces the propensity to pollute.

But in reality *preferences* can help fix these problems.

Caring about team-mates reduces the propensity to free-ride on them.

In situations not covered by the 1st Theorem, it is crucial to know people's preferences.

Economists need to take a closer look at human preferences!

Several paths to better understand human preferences:

- estimate preferences empirically
- build theory based on insights on human preferences in psychology and sociology
- build theory of the evolution of human preferences based on evolutionary logic (evolutionary biology, evolutionary anthropology)

- As living beings, we are the *product of evolution*
- For most of our evolutionary past we have had to adapt to local conditions to survive



- Evolutionary logic:
 - We all have trillions of ancestors
 - Our "recent" ancestors (500ky) lived in groups extending beyond the nuclear family
 - If we have inherited our ancestors' preferences, these should reflect the ability to survive and reproduce in such groups

 If *reproductive success* is the name of the game, shouldn't humans simply be expected to seek to maximize reproductive success?

- If *reproductive success* is the name of the game, shouldn't humans simply be expected to seek to maximize reproductive success?
- Economics literature on preference evolution [seminal papers: Frank (1987), Güth and Yaari (1992)]
- The main challenge of this research is to answer this question, and to understand why (or why not)

Roadmap

- I. A sketch of the evolutionary model in Alger and Weibull [Econometrica 2013, Games and Economic Behavior 2016]
- II. Main result: evolution favors *Homo moralis* preferences
- **III. Implications of** *Homo moralis* **preferences for economics**, with a comparison to altruistic preferences:
 - i. Voluntary contributions to public goods
 - ii. Tax compliance
 - iii. Incentives in firms
 - iv. Repeated interactions: an infinitely repeated prisoner's dilemma

A model of the evolution of preferences in social interactions

Imagine... a pre-industrial society



A model of the evolution of preferences in social interactions



Births, deaths, migration

Productive effort in team work

Material payoff (harvest, shelter) For example, suppose that:

- in each generation individuals work in teams of 2
- the material payoff to an individual making effort x when the other makes effort y is $\pi(x,y)$ For example: $\pi(x,y) = 5(x+y)^{1/2} - x$
- reproductive success increases with material payoff

For example, suppose that:

- in each generation individuals work in teams of 2
- the material payoff to an individual making effort x when the other makes effort y is $\pi(x,y)$ For example: $\pi(x,y) = 5(x+y)^{1/2} - x$
- reproductive success increases with material payoff

Suppose that each individual has a **goal function** *u*(*x*,*y*) that guides behavior

For example, suppose that:

- in each generation individuals work in teams of 2
- the material payoff to an individual making effort x when the other makes effort y is $\pi(x,y)$ For example: $\pi(x,y) = 5(x+y)^{1/2} - x$
- reproductive success increases with material payoff

Suppose that each individual has a **goal function** *u*(*x*,*y*) that guides behavior

Will evolution lead to **self-interested individuals**, with goal function $u(x,y) = \pi(x,y)$? Or some other goal function? • Extend the evolutionary game theory concept of evolutionarily stable strategy [Maynard Smith and Price (1973)]

- Extend the evolutionary game theory concept of evolutionarily stable strategy [Maynard Smith and Price (1973)]
- Consider an infinitely large population, and suppose that a share 1- ε carries the goal function $u_{\vartheta}(x,y)$ (residents) while a share ε carries some goal function $u_{\tau}(x,y)$ (mutants)

- Extend the evolutionary game theory concept of evolutionarily stable strategy [Maynard Smith and Price (1973)]
- Consider an infinitely large population, and suppose that a share 1- ε carries the goal function $u_{\vartheta}(x,y)$ (residents) while a share ε carries some goal function $u_{\tau}(x,y)$ (mutants)
- Individuals are matched at random to interact. Each matched group plays a game. They play some Nash equilibrium. Each individual obtains the associated material payoff.

- Extend the evolutionary game theory concept of evolutionarily stable strategy [Maynard Smith and Price (1973)]
- Consider an infinitely large population, and suppose that a share 1- ε carries the goal function $u_{\vartheta}(x,y)$ (residents) while a share ε carries some goal function $u_{\tau}(x,y)$ (mutants)
- Individuals are matched at random to interact. Each matched group plays a game. They play some Nash equilibrium. Each individual obtains the associated material payoff.
- If residents get a strictly higher average material payoff than rare mutants, then $u_{\vartheta}(x,y)$ is evolutionarily stable against $u_{\tau}(x,y)$

- Extend the evolutionary game theory concept of evolutionarily stable strategy [Maynard Smith and Price (1973)]
- Consider an infinitely large population, and suppose that a share 1- ε carries the goal function $u_{\vartheta}(x,y)$ (residents) while a share ε carries some goal function $u_{\tau}(x,y)$ (mutants)
- Individuals are matched at random to interact. Each matched group plays a game. They play some Nash equilibrium. Each individual obtains the associated material payoff.
- If residents get a strictly higher average material payoff than rare mutants, then $u_{\vartheta}(x,y)$ is evolutionarily stable against $u_{\tau}(x,y)$
- A goal function $u_{\vartheta}(x,y)$ is evolutionarily stable if it is stable against all possible mutant goal functions

In Alger and Weibull (2013, 2016):

- matching of individuals is assortative: there is assortativity if a rare mutant has a greater probability than residents of being matched with other rare mutants
 - such assortativity arises as soon as the population is structured into groups between which "migration" is limited (for example, (families, villages, cultural groups, firms, educational institutions, etc)
- minimalistic assumptions on the set of potential goal functions: any potential goal function u_∂ : Xⁿ → R is continuous and permutationinvariant in opponents' strategies

Result (n = 2)

$u_{\kappa}(x,y) = (1-\kappa) \cdot \pi(x,y) + \kappa \cdot \pi(x,x)$

Alger & Weibull (Econometrica, 2013)

These preferences are new to economics

$$u_{\kappa}(x,y) = (1-\kappa) \cdot \pi(x,y) + \kappa \cdot \pi(x,x)$$

They express a compromise between two goals: self-interest and "the right thing to do"

These preferences have a Kantian flavor

$$u_{\kappa}(x,y) = (1-\kappa) \cdot \pi(x,y) + \kappa \cdot \pi(x,x)$$



Immanuel Kant (1724-1804)

Implications of *Homo moralis* preferences for economics, with a comparison to altruistic preferences

Material payoff:
$$\pi\left(x_{i},x_{j}
ight)$$

$$\mathsf{Altruism:} \quad u_{\alpha}\left(x_{i}, x_{j}\right) = \pi\left(x_{i}, x_{j}\right) + \alpha \cdot \pi\left(x_{j}, x_{i}\right)$$

Homo moralis: $u_{\kappa}(x_i, x_j) = (1 - \kappa) \cdot \pi(x_i, x_j) + \kappa \cdot \pi(x_i, x_i)$

I. A public goods game

Material payoff:
$$\pi(x_i, \boldsymbol{x}_{-i}) = \left(x_i + \sum_{j \neq i} x_j\right)^{1/2} - x_i^2$$



II. Tax compliance

- Why do people pay taxes?
- *Homo oeconomicus*: tax compliance due to fear of being punished
- Altruistic preferences: if the taxpayer has a negligible impact on total tax revenues, it is still the fear of being punished that dominates the motivation to pay taxes
- *Homo moralis* preferences: tax compliance even if the taxpayer has a negligible impact on total tax revenues and there is no punishment
 - he/she evaluates what the tax revenues would be if a share κ of all the taxpayers chose the same level of tax compliance

II. Tax compliance

- Why do people pay taxes?
- *Homo oeconomicus*: tax compliance due to fear of being punished
- Altruistic preferences: if the taxpayer has a negligible impact on total tax revenues, it is still the fear of being punished that dominates the motivation to pay taxes
- *Homo moralis* preferences: tax compliance even if the taxpayer has a negligible impact on total tax revenues and there is no punishment
 - he/she evaluates what the tax revenues would be if a share κ of all the taxpayers chose the same level of tax compliance
- Bottomline: a positive degree of morality (κ > 0) induces tax compliance, even in the absence of monetary incentives

III. Incentives in firms

- A firm owner hires a team of workers
- She only observes total output \rightarrow salary based on total output
- Total output depends on all the employees' efforts
- If employees are altruistic towards each other: they internalize the effect of effort choice on the others
- If employees have *Homo moralis* preferences: each employee evaluates which total output would obtain if all chose the same effort as him/her

Sarkisian (Games, 2017)

III. Incentives in firms

- A firm owner hires a team of workers
- She only observes total output \rightarrow salary based on total output
- Total output depends on all the employees' efforts
- If employees are altruistic towards each other: they internalize the effect of effort choice on the others
- If employees have *Homo moralis* preferences: each employee evaluates which total output would obtain if all chose the same effort as him/her
- The firm owner prefers to hire either a team with altruistic employees or a team with *Homo moralis*

Sarkisian (Games, 2017)

IV. Repeated interactions: An infinitely repeated prisoner's dilemma

With Homo oeconomicus preferences:

	C	D
C	R	S
D	T	P

T>R>P>S

One condition is necessary for grim trigger (cooperate until someone defects, then defect forever) to sustain perpetual cooperation:

$$\delta \ge \frac{T-R}{T-P}$$

IV. Repeated interactions: An infinitely repeated prisoner's dilemma

With altruistic preferences:

	C	D
C	$(1+\alpha)R$	$S + \alpha T$
D	$T + \alpha S$	$(1 + \alpha) P$

Two conditions are necessary for grim trigger to sustain perpetual cooperation:

$$\delta \ge \frac{T - R - \alpha (R - S)}{T - P - \alpha (P - S)}$$
$$\alpha \le \frac{P - S}{T - P}$$

IV. Repeated interactions: An infinitely repeated prisoner's dilemma

With Homo oeconomicus preferences:

	C	D
C	R	$(1-\kappa)S+\kappa R$
D	$(1-\kappa)T+\kappa P$	Р

Two conditions are necessary for grim trigger to sustain perpetual cooperation:

$$\delta \ge \frac{T - R - \kappa (T - P)}{T - P - \kappa (T - P)}$$
$$\kappa \le \frac{P - S}{R - S}$$

Example I

T=10, R=8, P=4, S=0



Example II

T=10, R=6, P=2, S=0



Conclusion

- Economic theory + evolutionary logic: a theory of the long-term evolution of preferences
- Allows us to understand which forces in our evolutionary past have shaped our preferences
- May help us understand cultural differences
- May help us study how economic systems affect preferences. For example, do firms prefer to hire amoral or moral individuals ?

Merci !







INSTITUTE for A DVA N C E D S T U D Y in T O U L O U S E





