

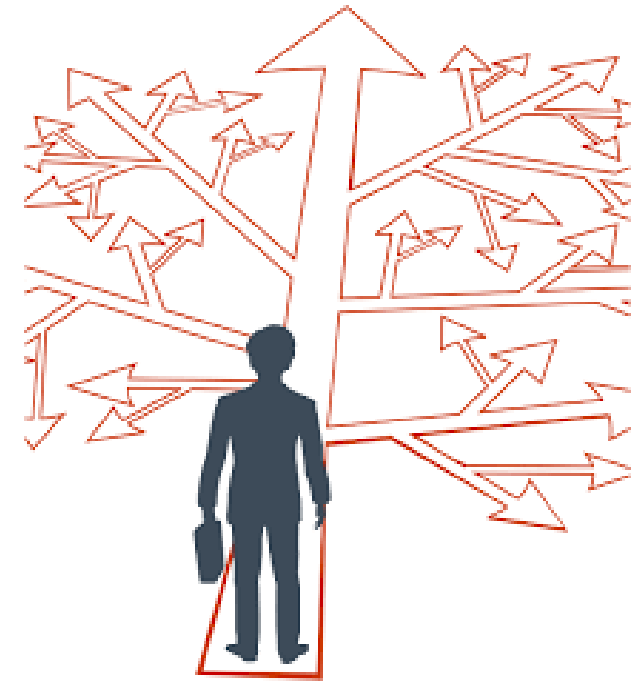
# Disentangling material, social, and cognitive determinants of human behavior and beliefs

Sergey Gavrilets

Center for the Dynamics of Social Complexity (DySoC),  
Department of Ecology and Evolutionary Biology,  
Department of Mathematics,  
University of Tennessee, Knoxville;  
School of Anthropology, University of Oxford, UK  
Complexity Science Hub Vienna, Austria

# Human decision-making

- Social instincts (biological)
- Material cost-benefit considerations
- Social influences (peers and different types of authorities)
- Personal/internalized norms and beliefs; heuristics
- Expectations about actions and beliefs of others
- Previous experience (learning)
- Errors



# Davis et al. (2014) “Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review”

Theory	First author theorist (date)	that ir	Theory	First author theorist (date)	that ir	Theory	First author theorist (date)
1 An Action Model of Consumption	Bagozzi (2000)		36 Integrative Theory of Health Behaviour Change	Ryan (2009)		77 Theory of Normative Social Behaviour	Rimal (2005)
2 Affective Events Theory	Weiss (1996)		37 Model of Pro-environmental Behaviour	Kolmuss (2002)		78 Theory of Planned Behaviour/Reasoned Action	Ajzen (1985)
3 AIDS Risk Reduction Model	Catania (1990)		38 Motivation Opportunity Abilities Model	Olander (1995)		79 Theory of Triadic Influence	Flay (1994)
4 Attitude-Social Influence – Efficacy Model and its successor I – Change	DeVries (1998)		39 Needs Opportunities Abilities (NOA) Model	Gatersleben (1998)		80 Transcontextual Model of Motivation	Hagger (2003)
5 Behavioural Ecological Model of AIDS Prevention	Hovell (1994)		40 Norm Activation Theory	Schwartz (1977)		81 Transtheoretical/Stages of Change Model	Prochaska (1983)
6 Change Theory	Lewin (1943)		41 Operant Learning Theory	Skinner (1954)		82 Value Belief Norm Theory	Stern (1999)
7 Classical Conditioning	Pavlov (1927)		42 Precaution Adoption Process Model	Weinstein (1988)			
8 COMB Model	Michie (2011)		43 Pressure System Model	Katz (2001)			
9 Consumption of Social Practices	Spaargaren (2000)		44 PRIME Theory	West (2006)			
10 Containment Theory	Reckless (1961)		45 Problem Behaviour Theory	Jessor (1977)			
11 Control Theory	Carver (1981/1982)		46 Prospect Theory	Kahneman (1979)			
12 Diffusion of Innovations	Rogers (1983)		47 Protection Motivation Theory	Rogers (1975)			
13 Differential Association Theory	Sutherland (1947)		48 Prototype Willingness Model	Gibbons (1995)			
14 Ecological Model of Diabetes Prevention	Burnet (2002)		49 Rational Addiction Model	Becker (1988)			
15 Extended Information Processing Model	Flay (1980)		50 Reflective Impulsive Model/Dual Process Theory	Strack (2004)			
16 Extended Parallel Process Model	Witte (1992)		51 Regulatory Fit Theory	Higgins (2000)			
17 Feedback Intervention Theory	Kluger (1996)		52 Relapse Prevention Theory	Marlatt (1980)			
18 General Theory of Crime	Goffredson (1990)		53 Risks as Feelings Model	Lowenstein (2001)			
19 General Theory of Deviant Behaviour	Kaplan (1972)		54 Self-determination Theory	Deci (2000)			
20 Goal Directed Theory	Bagozzi (1992)		55 Self-efficacy Theory	Bandura (1977)			
21 Goal Framing Theory	Lindenberg (2007)		56 Self-regulation Theory	Kanfer (1970)			
22 Goal Setting Theory	Locke (1968)		57 Six Staged Model of Communication Effects	Vaughan (2000)			
23 Health Action Process Approach	Schwarzer (1992)		58 Social Action Theory	Ewart (1991)			
24 Health Behaviour Goal Model	Gerbhardt (2001)		59 Social Action Theory	Weber (1991)			
25 Health Behaviour Internalisation Model	Bellg (2003)		60 Social Change Theory	Thompson (1990)			
26 Health Belief Model	Rosenstock (1966)		61 Social Cognitive Theory	Bandura (1986)			
27 Health Promotion Model	Pender (1982)		62 Social Consensus Model of Health Education	Romer (1992)			
28 Information-Motivation-Behavioural (IMB) Skills Model	Fisher (1992)		63 Social Development Model	Hawkins (1985)			
29 IMB Model of ART Adherence (extension of IMB)	Fisher (2008)		64 Social Identity Theory	Tajfel (1979)			
30 Integrative factors influencing smoking behaviour model	Flay (1983)		65 Social Influence Model of Virtual Community Participation	Dholakia (2004)			
31 Integrative model of health and attitude behaviour change	Flay (1983)		66 Social Ecological Model of Walking	Alfonzo (2005)			
32 Integrating the factors influencing smoking behaviour and the model of attitude and behaviour change	Flay (1983)		67 Social Ecological Model of Behaviour Change	Panter-Brick (2006)			
33 Integrative Model of Behavioural Prediction	Fishbein (2000)		68 Social Learning Theory	Miller (1941)			
34 Integrated Theory of Drinking and Behaviour	Wagenaar (1994)		69 Social Norms Theory	Perkins (1986)			
35 Integrated Theoretical Model for Alcohol and Drug Prevention	Gonzalez (1989)		70 Systems Model of Health Behaviour Change	Kershell (1985)			
			71 Technology Acceptance Models 1, 2 and 3	Venkatesh (1989, 2000, 2008)			
			72 Temporal Self-regulation Theory	Hall (2007)			
			73 Terror Management Health Model	Goldenberg (2008)			
			74 Terror Management Theory	Greenberg (1986)			
			75 Theory of Normative Conduct	Cialdini (1991)			
			76 Theory of Interpersonal Behaviour	Triandis (1977)			

# Mathematical models

- Social instincts (biological)
  - Population and quantitative genetics modeling; replicator equation; adaptive dynamics
- Material cost-benefit considerations
  - Different flavors of game theory: classical, **evolutionary**, mean field, and quantum
- Social influences
  - Models of conformity (descriptive norms), consensus formation (or fragmentation)
- Personal norms and beliefs; internalized norms
  - Some modeling attempts
- Expectations about actions and beliefs of others
  - Few models of injunctive social norms; level-k modeling, foresight
- Previous experience
  - Reinforcement learning, Bayesian learning
- Errors
  - Quantal Response Equilibrium

# Today's talk

## 1. Theory and mathematical models

Gavrilets (2021) *Evolutionary Human Sciences* 3: e44

## 2. Experiments

Tverskoi et al. 2021 (submitted)

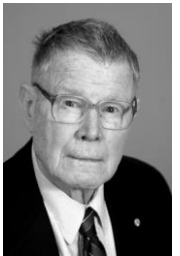
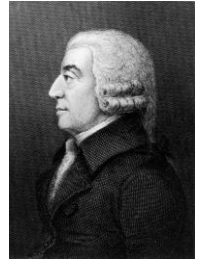
Tverskoi et al. 2022 (in preparation)

# Novel approach

- Starting point is the “Thomas theorem” of social psychology: “If men define situations as real, they are real in their consequences” (Thomas 1928)
- In my models, individual decisions in social situations are based on individual beliefs about the current situation as well as on their beliefs about others and their beliefs.
- Scaling up from individual behavior to group dynamics

# Beliefs in economics

- Jeremy Bentham (1789) "An Introduction to the Principles of Morals and Legislation": mental ingredients of utility
- Adam Smith (1759) "A Theory of Moral Sentiments": importance of beliefs as drivers of human behavior
- Thomas Schelling (1984) "The Mind as a Consuming Organ": "We also consume by thinking."
- Robert Abelson (1986) "Beliefs are Like Possessions"
- Roland Bénabou and Jean Tirole (2016) "Mindful Economics: The Production, Consumption, and Value of Beliefs"
- George Loewenstein and Andras Molnar (2018) "The renaissance of belief-based utility in economics"
- Andras Molnar and George Loewenstein (2022) "Thoughts And Players: An Introduction To Old And New Economic Perspectives On Beliefs"



## Recent calls:

- Mirta Galesic et al. (2018) "Integrating social and cognitive aspects of belief dynamics: towards a unifying framework"

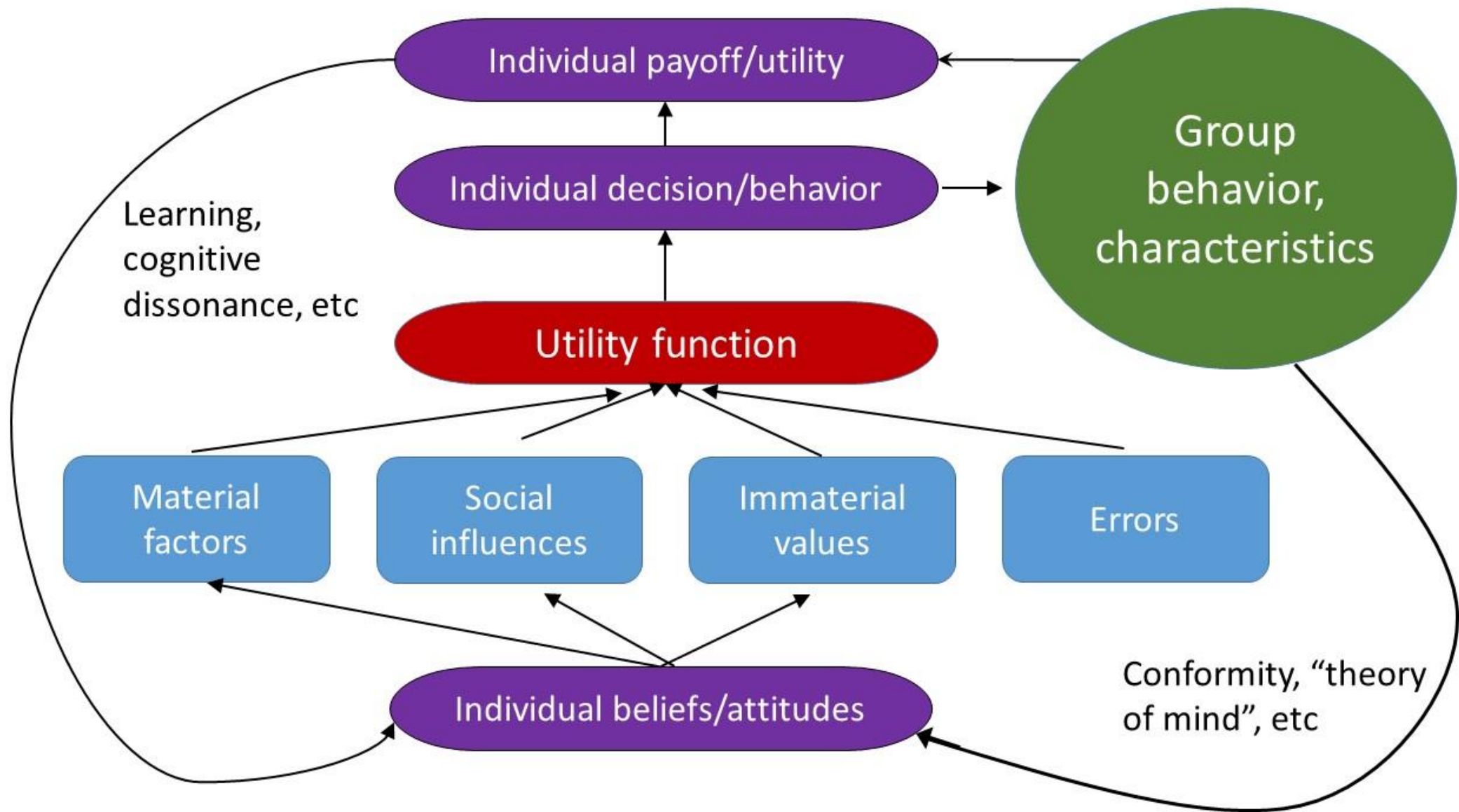


# Forces affecting human behavior in the model

- **Material forces:** (e.g., monetary) payoffs
- **Social forces:** conformity with peers' behavior (descriptive norms) and peers' expectations (injunctive norms), conformity with authority, social learning
- **Cognitive forces:**
  - Cognitive dissonance (a feeling of mental discomfort when the person's attitudes, behavior, or beliefs conflict; Festinger 1957)
  - "Theory of mind" and "social projection" (a tendency to assume that others are similar to oneself; Premack and Woodford 1979, Krueger 2007)
  - Logic constraints on different beliefs (Friedkin et al. 2016)

# A general setup

- Consider a group of people repeatedly engaged in a particular type of social interactions
  - Production or maintenance of a public good
  - Harvesting a common pool of resources
- Individuals:
  - care about their own material costs and benefits
  - want to do what they think is right
  - observe and are influenced by behavior of others
  - don't like to be disapproved by others
  - infer (but don't know exactly) their attitudes and beliefs
  - are influenced by an external authority (or propaganda)
- Questions
  - How do individuals find the “right” actions?
  - What happens to their attitudes, beliefs and behaviors as social interactions dynamically unfold?



# Standard modeling approaches

- Dynamics variables:

- Game theoretic models: strategies/actions  $x$  (von Neuman and Morgenstern 1944,...)
- Social influence models: opinions/beliefs  $y$  (Rashevsky 1949, DeGroot 1974, ...)

- Rules for changes

- Evolutionary game theory models: Choose strategy  $x$  to maximize payoff  $\pi(x | \dots)$
- Game theory with personal norms (Akerloff 1980, Rabin 1984,...): Choose strategy  $x$  to maximize utility, e.g.

$$u = \pi(x | \dots) - \alpha(x - y)^2$$

where  $\alpha$  measures the “weight” of personal norms (or strength of cognitive dissonance, Festinger, 1957)

- Social influence theories : Change opinion (personal norm)  $y$  in the direction of the average opinion  $Y$  of peers (which is assumed to be known):

$$y' = y + \beta(Y - y)$$

where  $\beta$  is a measure of conformity

# Some modeling details

- Four main dynamic variables (for each individual; continuous):
  - Action  $x$  chosen by a focal individual ,
  - First order belief:
    - Attitude  $y$  (personal belief about the most appropriate action in a given social situation the action you'll take in the absence of any social influences; “personal norm”; )
  - Second order beliefs:
    - Belief (an expectation)  $\tilde{x}$  about the average action of peers (“empirical expectation”)  
Question:
    - Belief (an expectation)  $\tilde{y}$  about the average attitude of peers (“normative expectation”)
  - Inspired by recent experimental work by D’Adda et al. (2020), Goges and Nosenzo (2020), Andreozzi et al. (2020), **Szekely et al. (2021)**, Gachter et al. (2021)
- External authority (cultural, religious, political, etc) promoting a particular action  $G$

# Some modeling details: actions

Subjective utility function for a focal individual

$$u = \underbrace{A_0 \pi(x, \tilde{x})}_{\text{material payoff}} - \underbrace{\frac{1}{2} A_1 (x - y)^2}_{\text{cognitive dissonance}} - \underbrace{\frac{1}{2} A_2 (x - \tilde{y})^2}_{\text{disapproval by peers}} - \underbrace{\frac{1}{2} A_3 (x - \tilde{x})^2}_{\text{conformity w/ peers}} - \underbrace{\frac{1}{2} A_4 (x - G)^2}_{\text{conformity w/ authority}}.$$

where nonnegative weights  $A_i$  sum up to 1.

If the payoff function  $\pi$  is quadratic, then the (best response) action maximizing  $u$  is

$$x_t = B_0 \theta_t + B_1 y_t + B_2 \tilde{y}_t + B_3 \tilde{x}_t + B_4 G,$$

where  $\theta_t$  is the action maximizing payoff  $\pi$ . Constant weights  $B_i$  are proportional to the corresponding values  $A_i$  ( $\sum B_i = 1$ )

# Some modeling details: cognitive and social effects

- Dynamics of beliefs for a focal individual

$$\begin{aligned} y_{t+1} = & y_t + \underbrace{\alpha_1(x_t - y_t)}_{\text{cognitive dissonance}} + \underbrace{\beta_1(X_t - y_t)}_{\text{conformity w/ peers}} + \underbrace{\gamma_1(G - y_t)}_{\text{conformity w/ authority}}, \\ \tilde{y}_{t+1} = & \tilde{y}_t + \underbrace{\alpha_2(y_t - \tilde{y}_t)}_{\text{social projection}} + \underbrace{\beta_2(X_t - \tilde{y}_t)}_{\text{learning about others}} + \underbrace{\gamma_2(G - \tilde{y}_t)}_{\text{conformity w/ authority}}, \\ \tilde{x}_{t+1} = & \tilde{x}_t + \underbrace{\alpha_3(\tilde{y}_t - \tilde{x}_t)}_{\text{logic constraints}} + \underbrace{\beta_3(X_t - \tilde{x}_t)}_{\text{learning about others}} + \underbrace{\gamma_3(G - \tilde{x}_t)}_{\text{conformity w/ authority}}, \end{aligned}$$

where  $X$  is an observed average action of groupmates (can be different for different people).

- I allow for variation between individuals in all parameters

# Analysis

- With  $n$  individuals,  $4n$  dynamic equations
- Analysis of equilibria and transient dynamics (for individuals and groups)
  - There are differences w/ game-theoretic predictions!
- What the model can predict/describe:
  - How individuals and groups behave
  - How their attitudes and beliefs change
  - Which norms are internalized
  - Which factors control beliefs about others
  - What the effects of external influences are



**Table S1:** Production function  $P$  (or  $p_i$ ) and expected payoffs  $\pi(x_i, \bar{x}_i)$  in different games. Games with quadratic payoff functions: Coordination, Public Goods Game (PGG) with quadratic costs, PGG with diminishing return, Common Pool Resource (CPR), Tragedy of the Commons (TC) with quadratic costs, and TC with diminishing return. Games with linear payoff functions: Dictator, Give-or-Take, Rule Following, and Linear PGG. Game with quasi-linear payoff function: continuous Prisoner's Dilemma. Nonlinear game: "us vs. nature" game. In all collective action games, the expected group effort is  $Z = x + (n - 1)\bar{x}_i$ . An empty entry in the table means that in the corresponding game the corresponding function or parameter is not defined. Parameters with subscript  $i$  (e.g.,  $b_i, c_i, d_i, r_i, v_i$ ) are specific for individuals. Parameters without subscripts (e.g.,  $b, d, R$ ) are the same for all individuals.  $R$  and  $r_i$  are the endowments. Note that variable  $P$  in collective action games is the production function while in the Prisoner's Dilemma game,  $P_i$  is the punishment payoff.

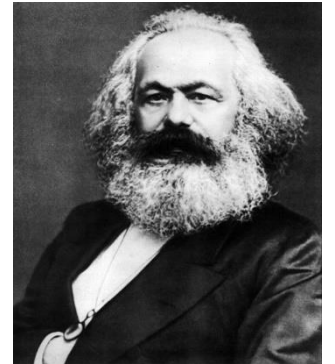
Game	Production $P/p_i$	Expected payoff $\pi_i$
Coordination		$b_i - 0.5c_i(x_i - \theta_i)^2 - 0.5d_i(x_i - \bar{x}_i)^2$
PGG w/ quadratic costs	$P = bZ$	$v_iP - 0.5c_ix_i^2$
PGG w/ diminishing return	$P = bZ - 0.5dZ^2$	$v_iP - c_ix_i$
CPR	$P = bZ - 0.5dZ^2$	$\frac{x_i}{Z}P - c_ix_i$
TC w/ quadratic costs	$p_i = b_ix_i$	$p_i - 0.5c_iZ^2$
TC w/ diminishing return	$p_i = b_ix_i - 0.5d_ix_i^2$	$p_i - c_iZ$
Public vs. private production	$P = bZ - 0.5dZ^2$	
	$p_i = b_iz_i - 0.5d_iz_i^2$	$v_iP + p_i$ , where $z_i + x_i = r_i$
Dictator		$R - x_i$
Give-or-Take		$R - x_i$
Rule Following		$R - x_i$
Linear PGG	$P = bZ$	$v_iP - c_ix_i$
Continuous PD		$x_i\bar{x}_iR_i + x_i(1 - \bar{x}_i)S_i \dots$ $+ (1 - x_i)\bar{x}_iT_i + (1 - x_i)(1 - \bar{x}_i)P_i$
"Us vs. nature"	$P = b\frac{Z}{Z+Z_0}$	$v_iP - c_ix_i$



# Some results

- When messaging of external authorities is weak (or absent), then
  - individual actions  $x$  converge to the Nash equilibrium  $x_{Nash}$
  - individual beliefs  $y$  converge to values “justifying” their actions.

"it is not the consciousness of men that determines their existence, but, on the contrary, their social existence that determines their consciousness." (Karl Marx)



# Some results

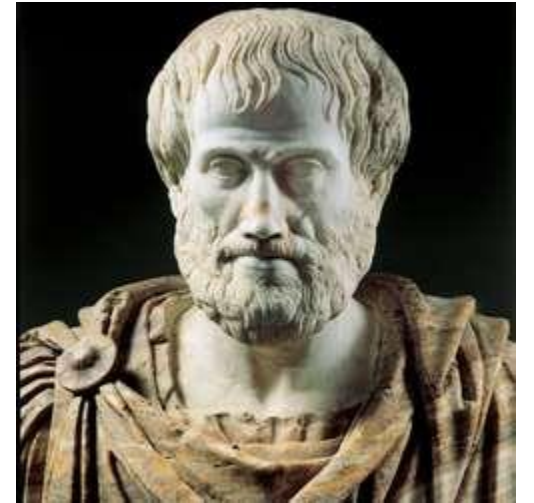
- But in general, long-term beliefs are determined by a balance between material payoffs and values promoted by external authority (propaganda/inculcation)



Bernays, E.  
Propaganda. (1928)

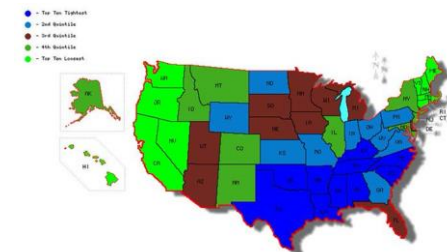
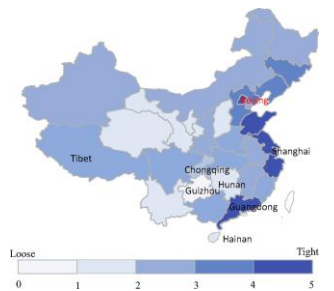
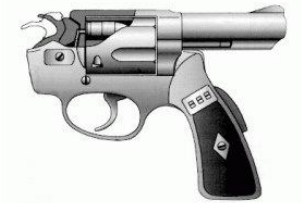
# Some results

- Differences with game theory predictions can be large due to social influences
  - “Man by nature is a social (political) animal” (Aristotle)



# Some results

- Within-society variation: largest in actions, followed by that in attitudes, followed by that in normative expectation, followed by that in empirical expectations
- People can internalize preferences for acts detrimental to them (potential application to the evolution of obedience; Milgram's and Zimbardo's experiments)
- Under some conditions the effort of an authority to promote certain actions can backfire and cause an opposite effect
- Explaining differences between “tight” and “loose” cultures of states, provinces, and countries
  - social heterogeneity, societal threat, propaganda effects, collectivism-individualism variation, population size, historical differences in subsistence style (e.g., rice vs. wheat)



## 2. Testing the theory

- Denis Tverskoi, postdoc at DySoC/NIMBioS
- Andrea Guido, postdoc, Laboratory of Agent-Based Social Simulation, Rome, Italy
- Giulia Andrighetto, Laboratory of Agent-Based Social Simulation, Rome, Italy
- Anxo Sanchez, Applied Mathematics, Universidad Carlos III de Madrid, Spain



# Common Pool Resources (CPR) game



- CPR are resources whose yield is subtractable and the exclusion from which is difficult (e.g., fish in bodies of water, pastures, and water used for irrigation)
- Ostrom et al. (1990): CPR game is “a much more realistic environment... than many of the [other] dilemma games.”



- Individual action/investment:  $x_i$
- Group investment  $Z = \sum x_i$
- Production function (w/ diminishing return):  $P = bZ - \frac{1}{2} dZ^2$ , where  $b$  and  $d$  are positive parameters
- Expected payoff



$$\pi_i = \pi_0 + v_i P - c x_i,$$

where  $\pi_0, c_i$  are the endowment and cost coefficient and the individual's share of the reward is

$$v_i = \frac{x_i}{Z}.$$

Social dilemma: Nash equilibrium > socially optimum effort



# Online behavioral economics experiment using the Common Pool Resources game

- 150 subjects w/out authority's messaging and 150 subjects w/ authority's messaging
- Spring 2021; mostly Spanish subjects
- CPR game in randomly formed groups of 6
- 35 rounds (5 weeks)
- Endowment: 30 units. Decision: how much to invest in the extraction of the CPR
- With parameters used,
  - Nash equilibrium is  $x = 24$
  - Socially optimum contribution is  $x = 14$ .
- Authority's message: "The group-beneficial contribution is 14 units"



# Beliefs elicitation



- Personal normative belief: How many points should a person in your group, including yourself, contribute?  $y$
- Normative expectation (incentivized): How many points will the other 5 people in your group think you should contribute?  $\tilde{y}$
- Empirical expectation (incentivized): How many points will the other 5 people contribute?  $\tilde{x}$

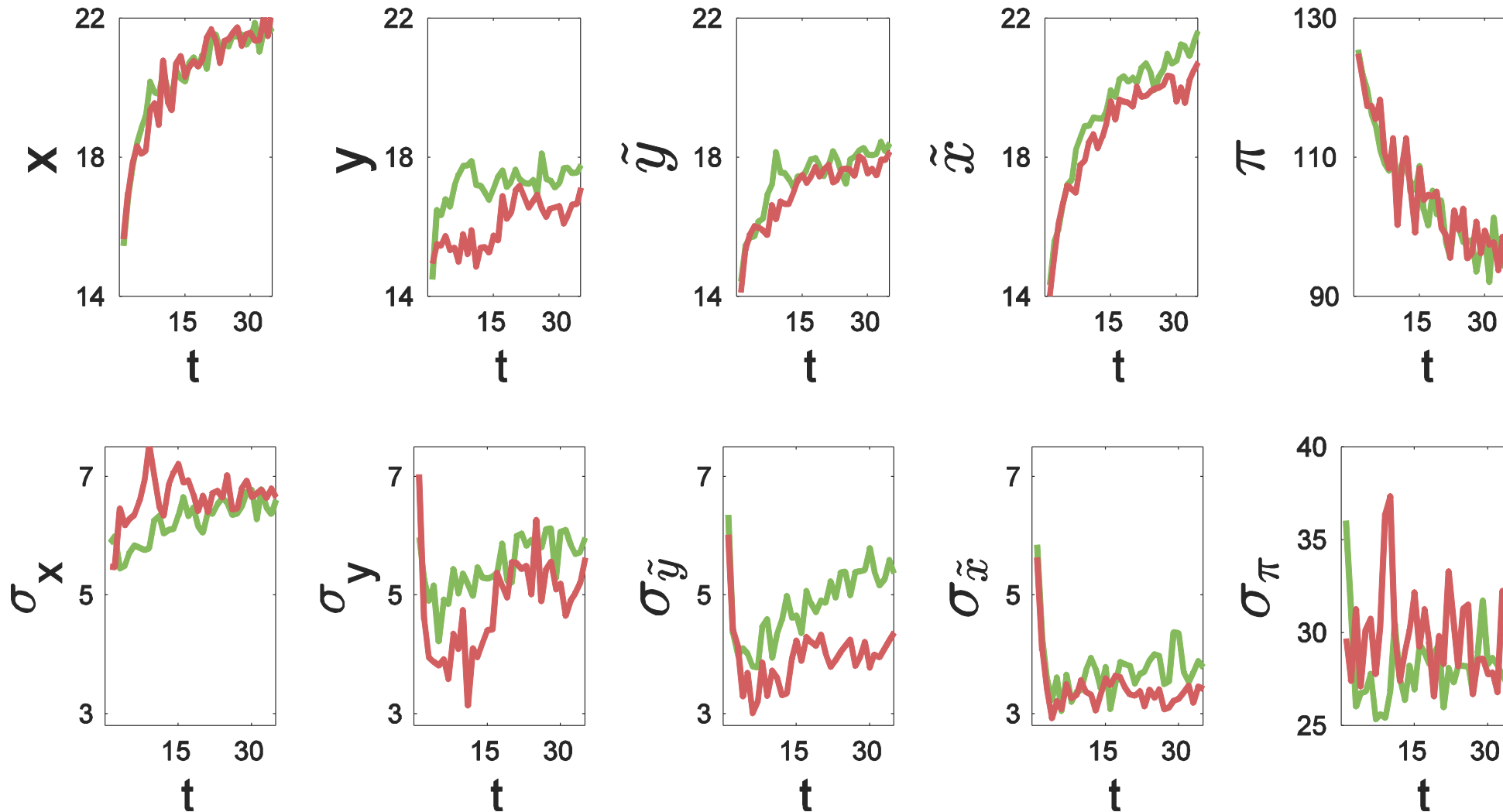
- 
- How many points will you contribute?  $x$ 
    - Participants are given a tool to estimate their expected payoff

- 
- Feedback: your payoff and the contributions of groupmates



Experimental protocol as in Szekely et al. (Nat. Comm., 2021)

# Observed dynamics of actions and beliefs **without** and **with** authority's messaging



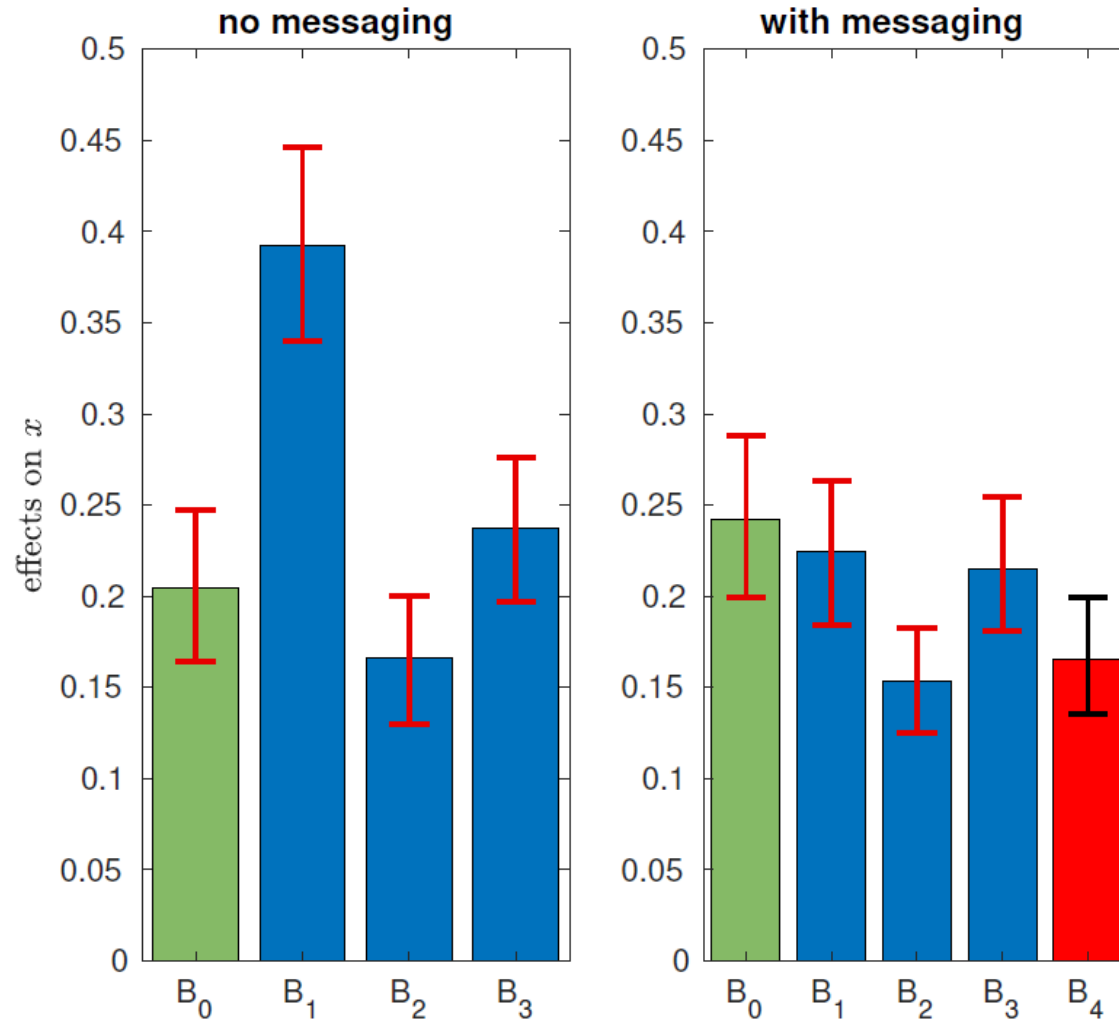
# Some observations

- The average individual effort appears to evolve to a value below Nash equilibrium
- Subjects make larger efforts (and expect others to make larger efforts) than what they believe is right. They also make larger efforts than what they think others will do.
- Personal norms equilibrate faster than other characteristics
- Variation between individuals is the highest in individual efforts, followed by that in personal norms, normative expectations, and empirical expectations, respectively.
- Authority's messaging does not affect the average contribution and payoff, but it increases their standard deviations.
- In contrast, the messaging decreases both the means and standard deviations of personal norms and both second-order beliefs.
- The average payoff continuously declines in time.

# Parameter estimation

- Dropped all individuals with less than 30 observations
- Multicollinearity:
  - Detected using condition index, variance decomposition proportion, and variance inflation factor. Also checked correlations between independent variables.
  - Treated by dropping or combining problematic columns in the design matrix
- There is a lot of variation between individuals but the multilevel model/random coefficients approach didn't work
- Estimated individual regressions for each individual separately and then took the averages across all individuals (i.e. used the mean group estimation method)
- Used seemingly unrelated regression to check the effects of correlations between errors
- Used numerical maximization of likelihood
- Model selection using the AIC, BIC and RMSE: compared with 10 alternative models for changes in behavior and 6 alternative models for changes in beliefs.
- Estimated confidence intervals using bootstrap

# Mean group estimates: parameters of utility function



B<sub>0</sub>: material payoffs

B<sub>1</sub>: cognitive dissonance

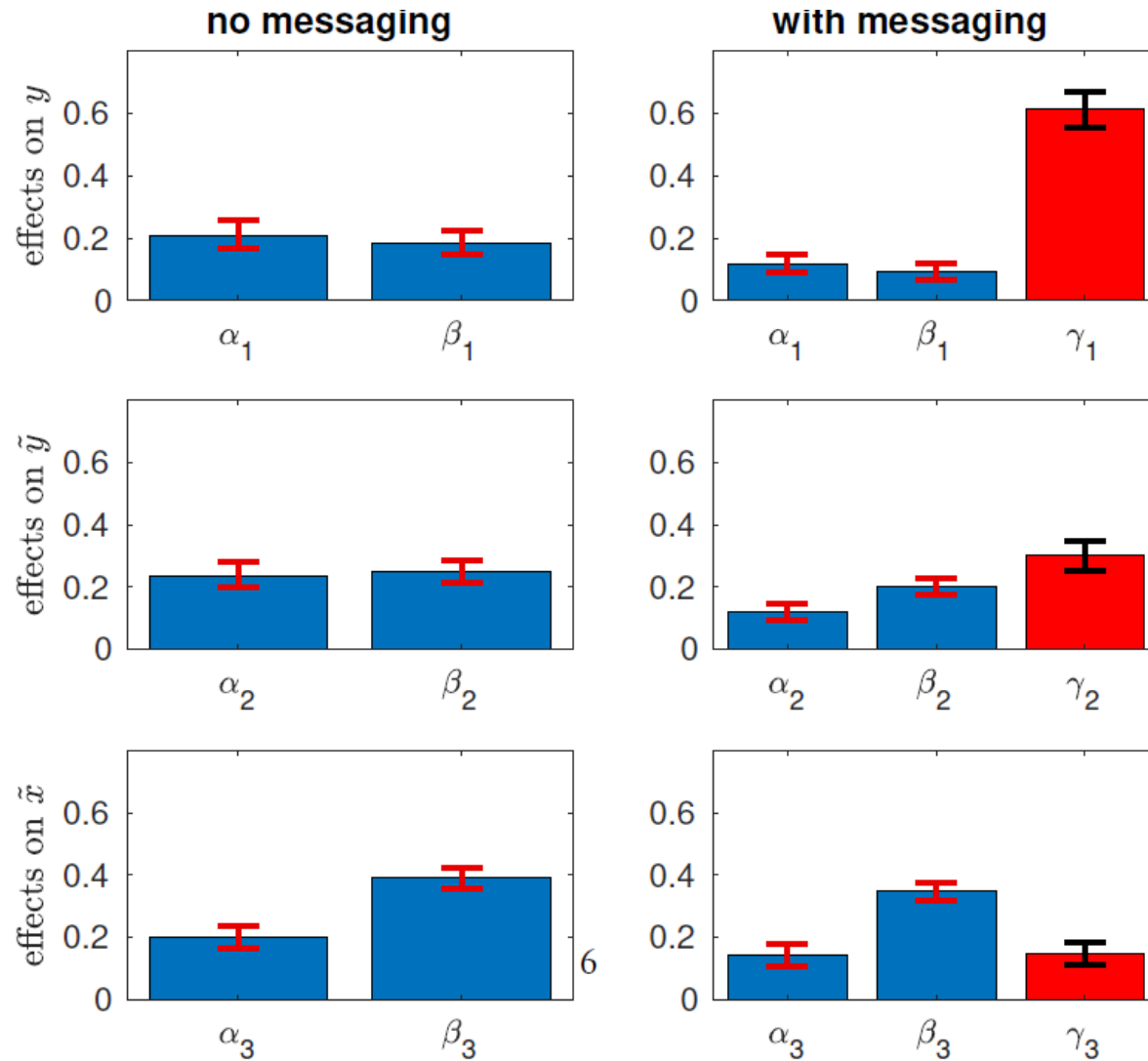
B<sub>2</sub>: disapproval by peers

B<sub>3</sub>: conformity with peers' actions

B<sub>4</sub>: conformity with authority's messaging (propaganda)

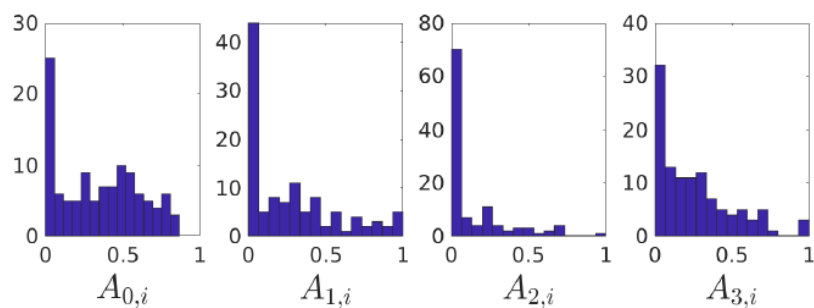
# Mean group estimates: parameters of beliefs dynamics

- $\alpha_1$ : Cognitive dissonance
- $\alpha_2$ : Social projection
- $\alpha_3$ : Logic constraints
  
- $\beta_1$ : Conformity with peers
- $\beta_2$ : Learning about peers
- $\beta_3$ : Learning about peers
  
- $\gamma_1, \gamma_2, \gamma_3$ : Conformity with authority's messaging (propaganda)

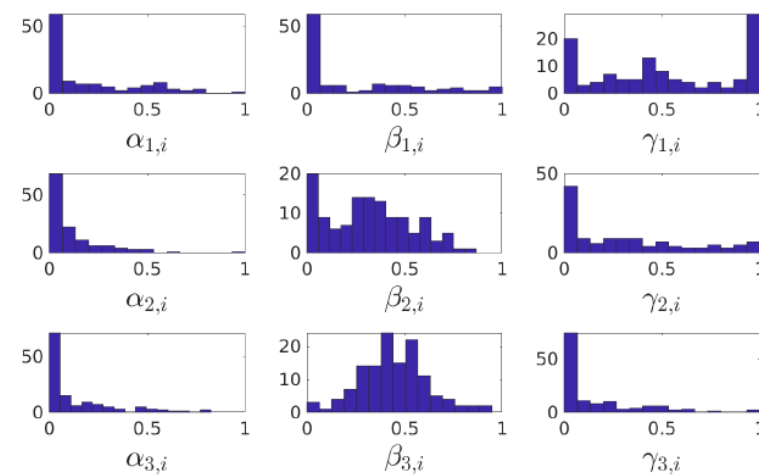
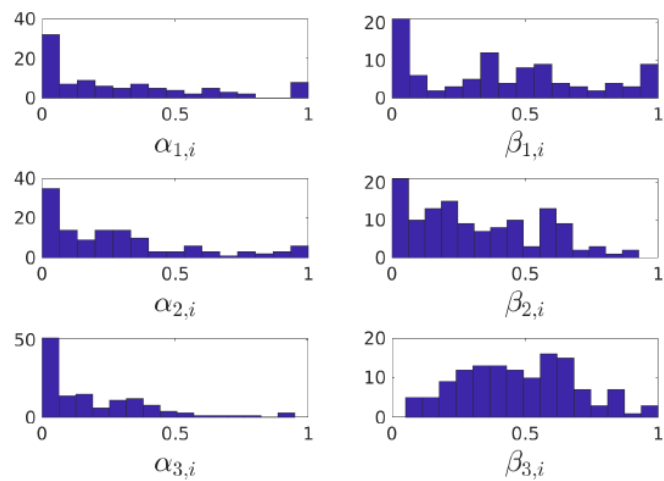
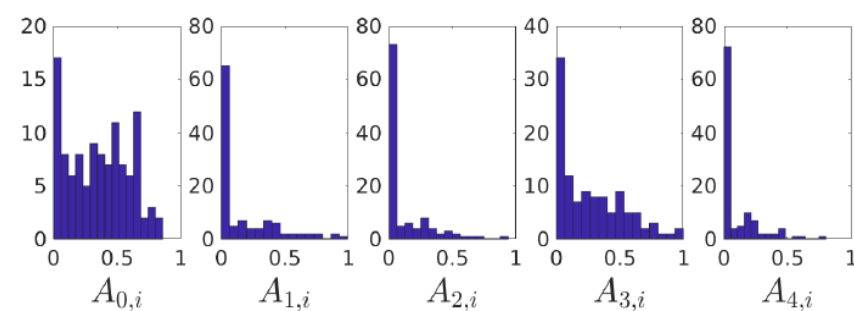


# Differences between individuals in parameters

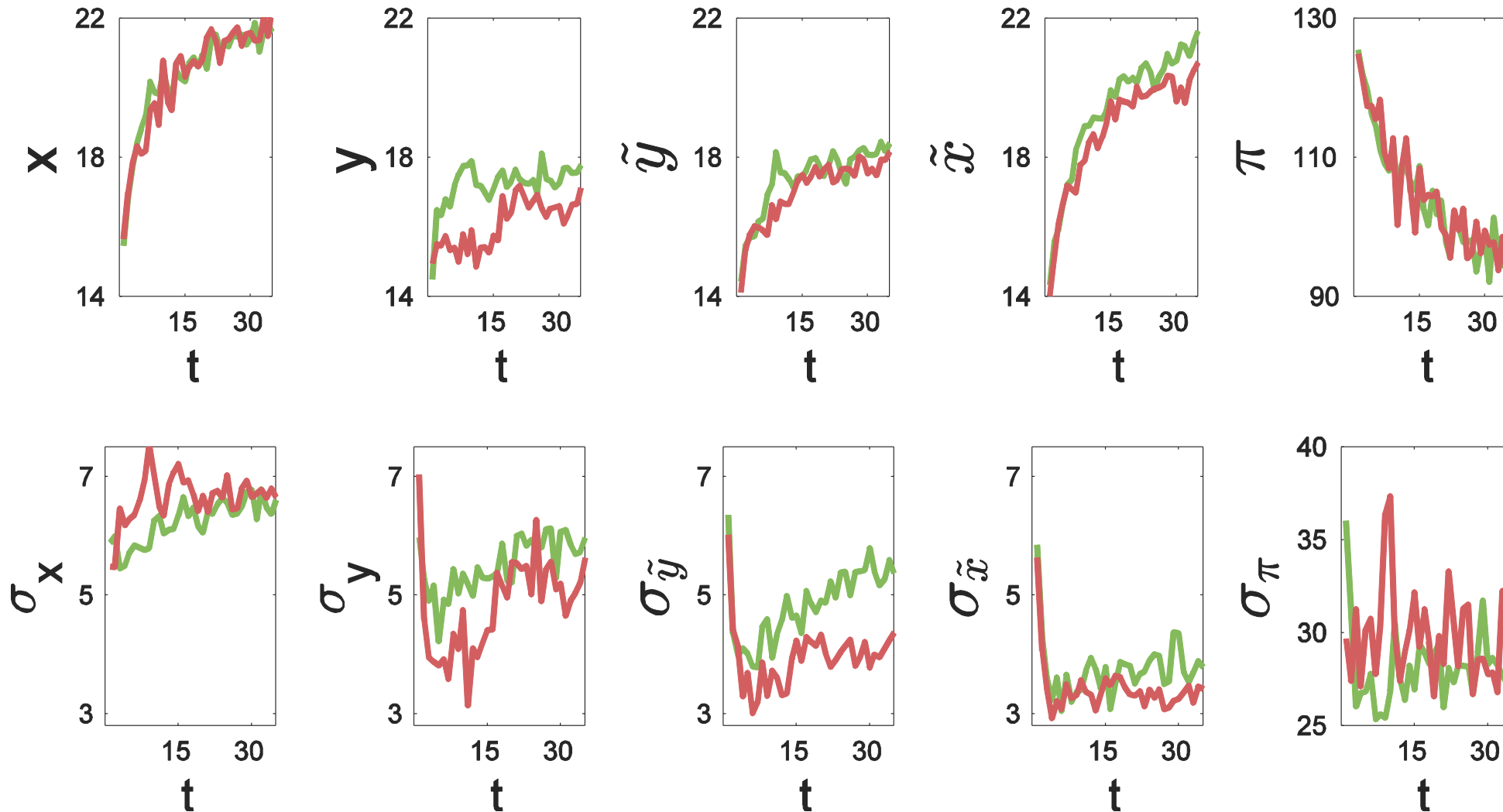
no messaging



with messaging

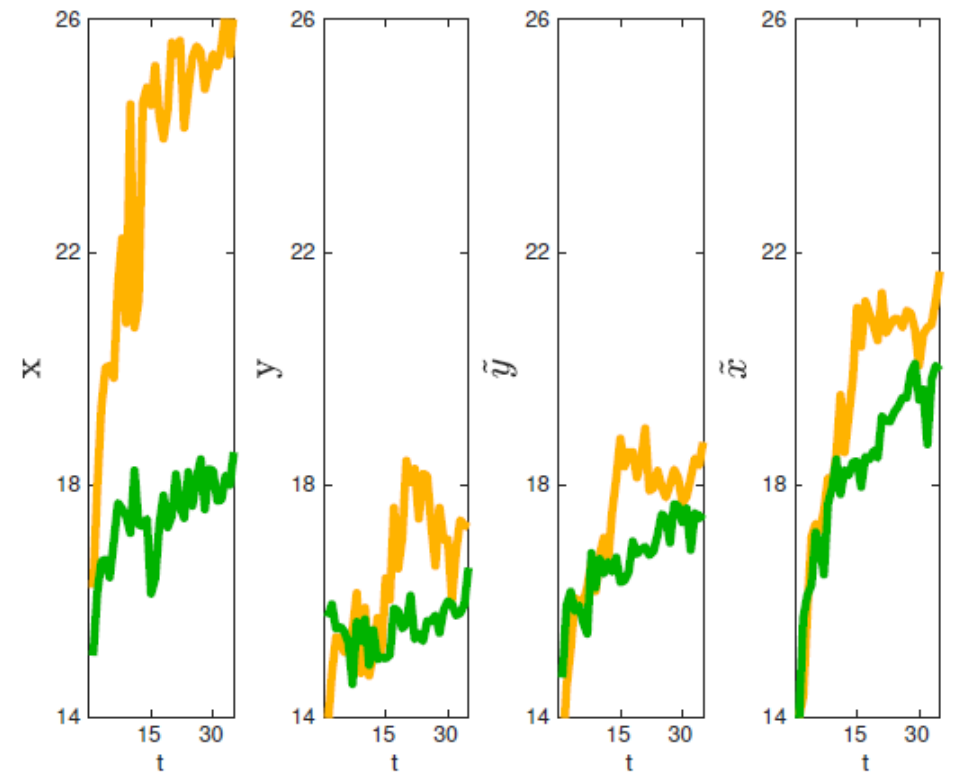
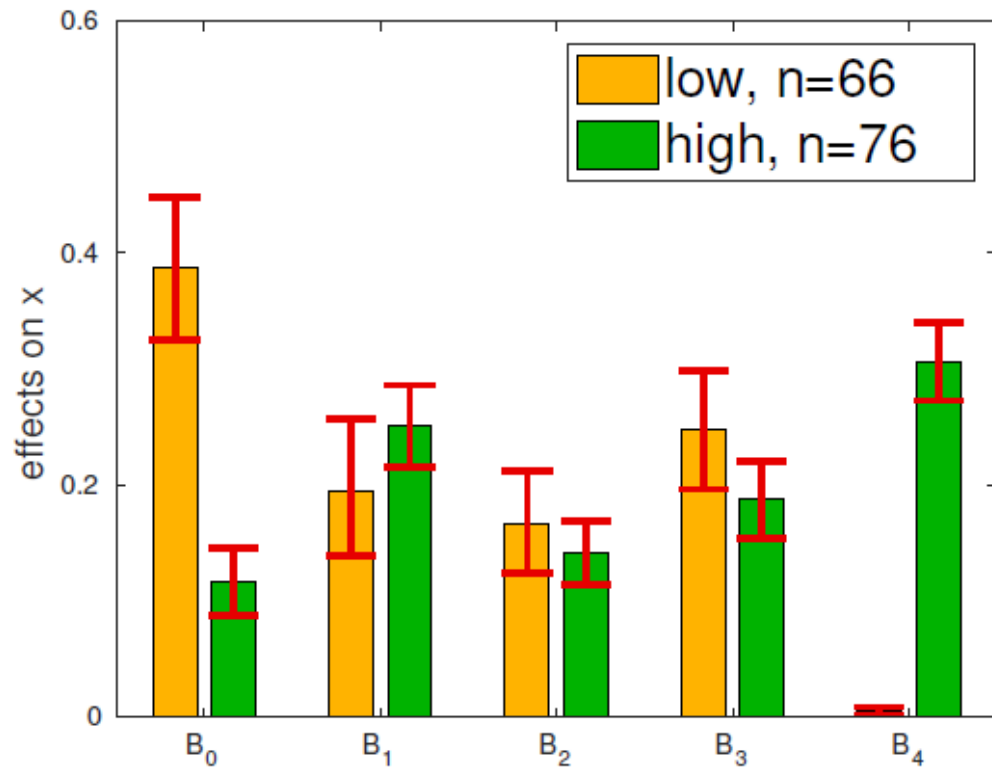


# Observed dynamics of actions and beliefs without and with authority's messaging

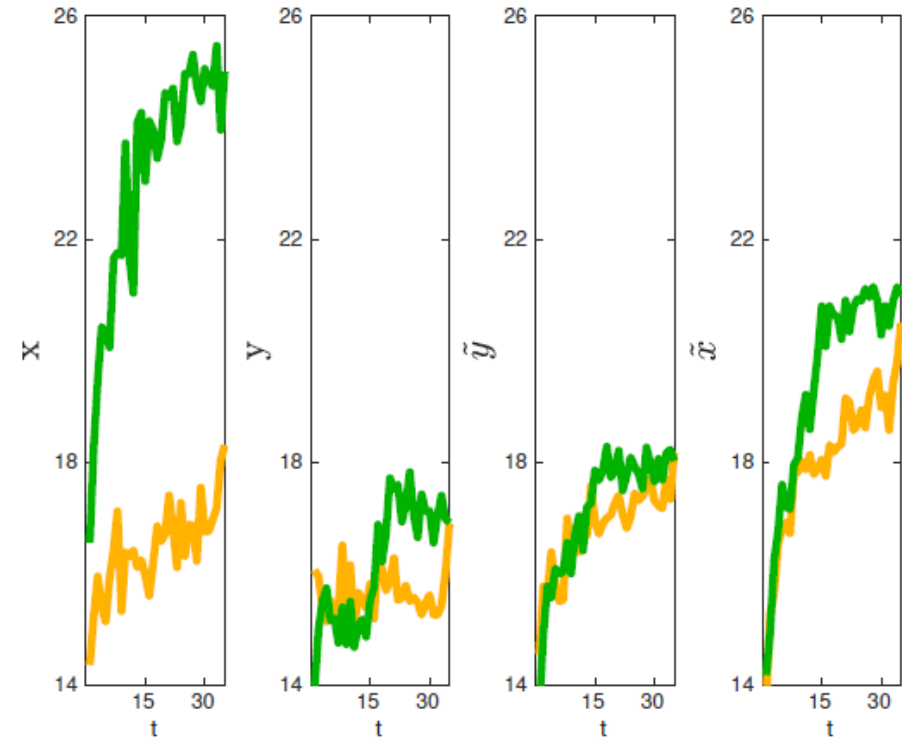
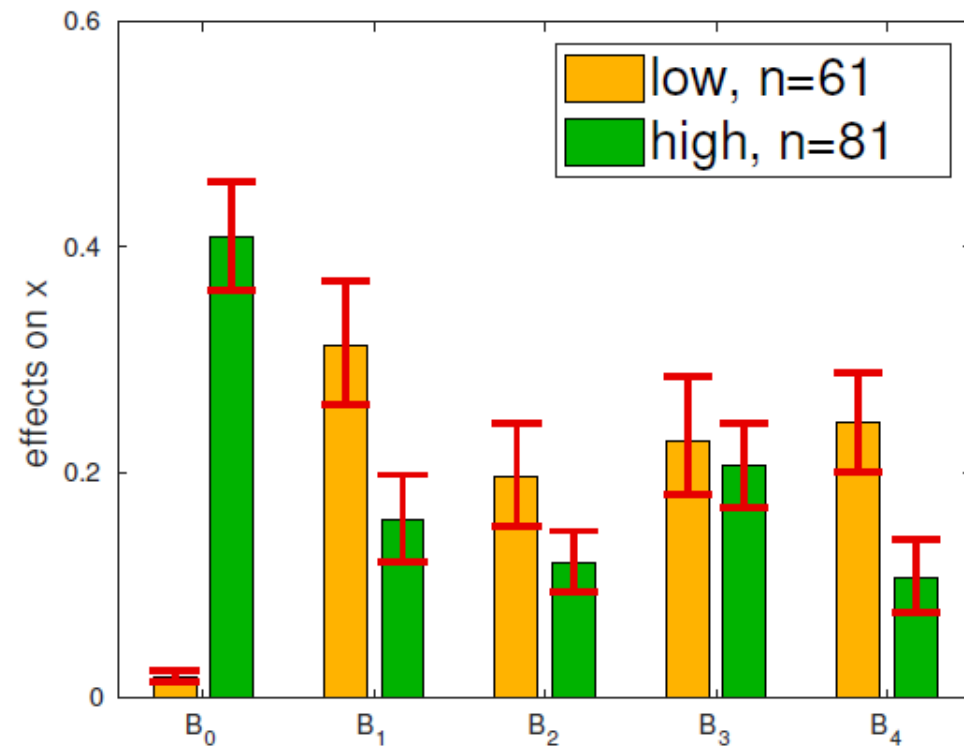




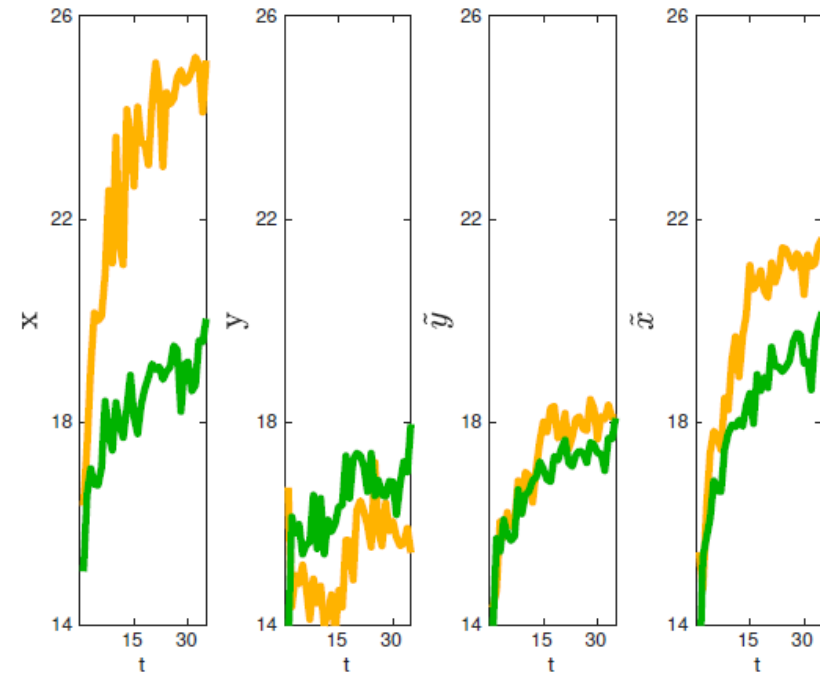
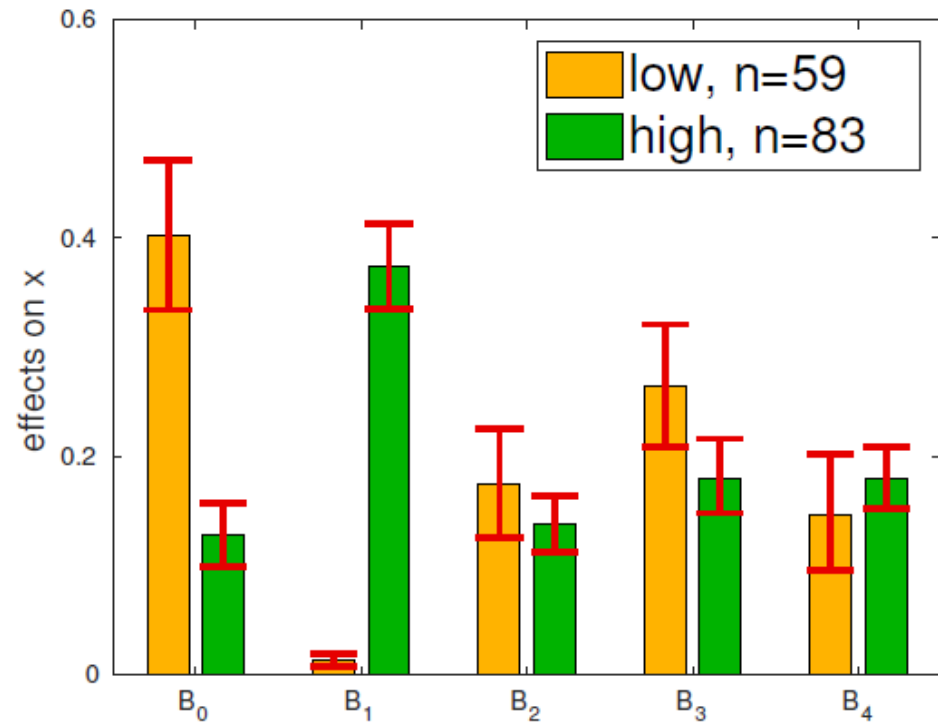
# Differences between individuals with high and low values of $B_4$ (conformity with authority)



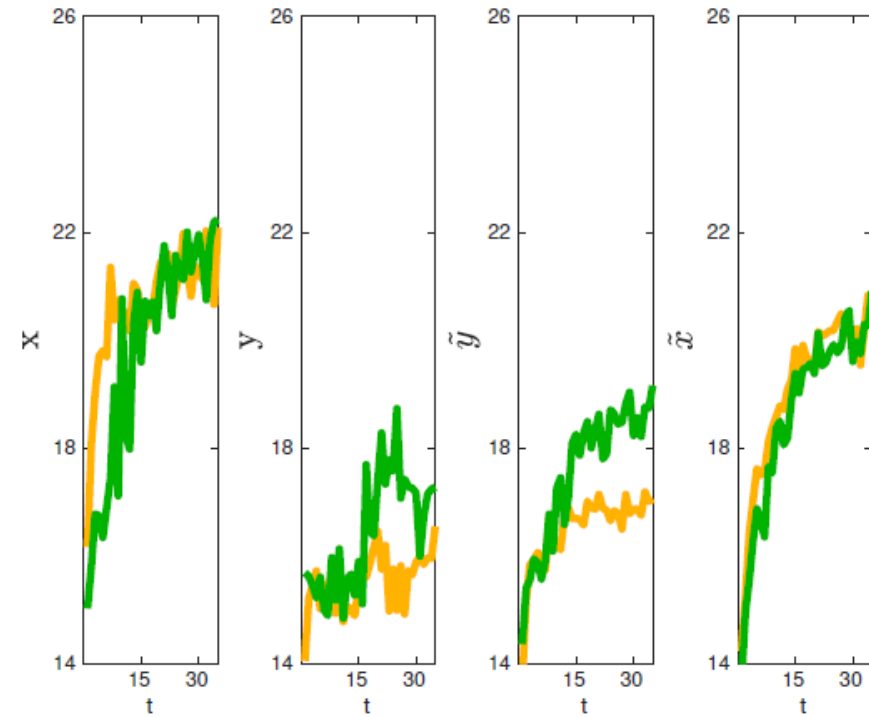
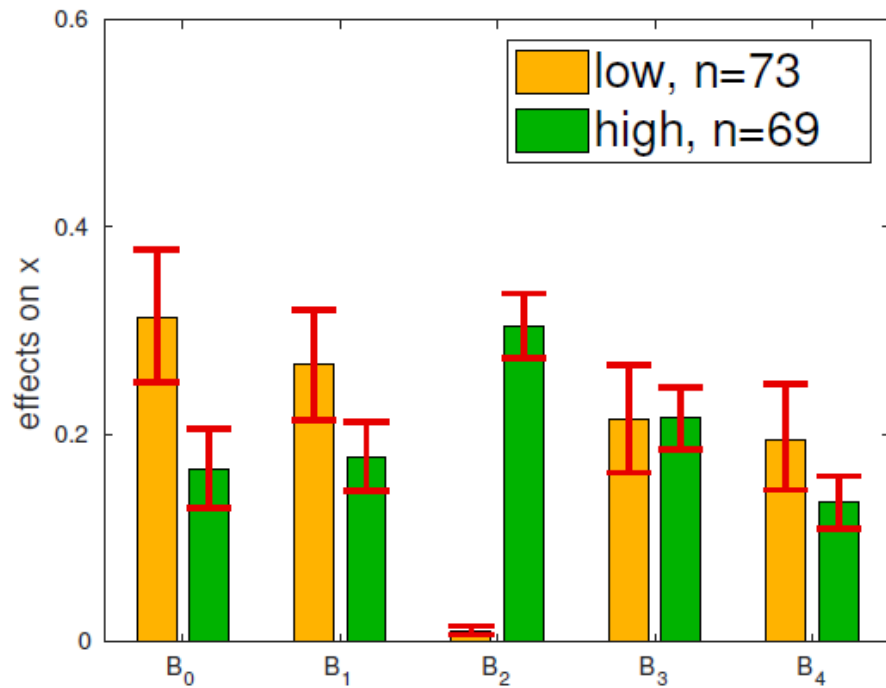
# Differences between individuals with high and low values of $B_0$ (material payoffs)



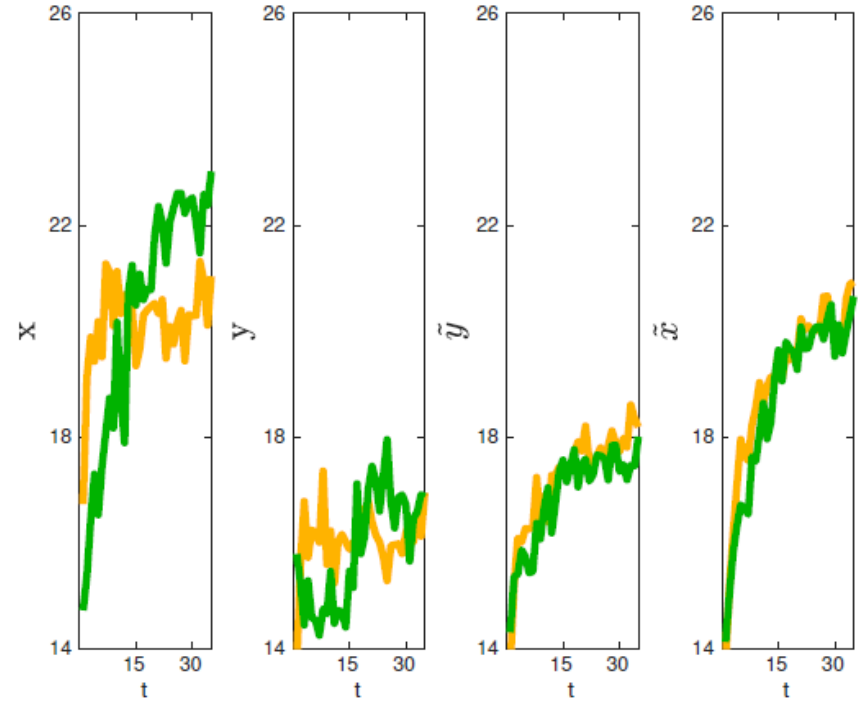
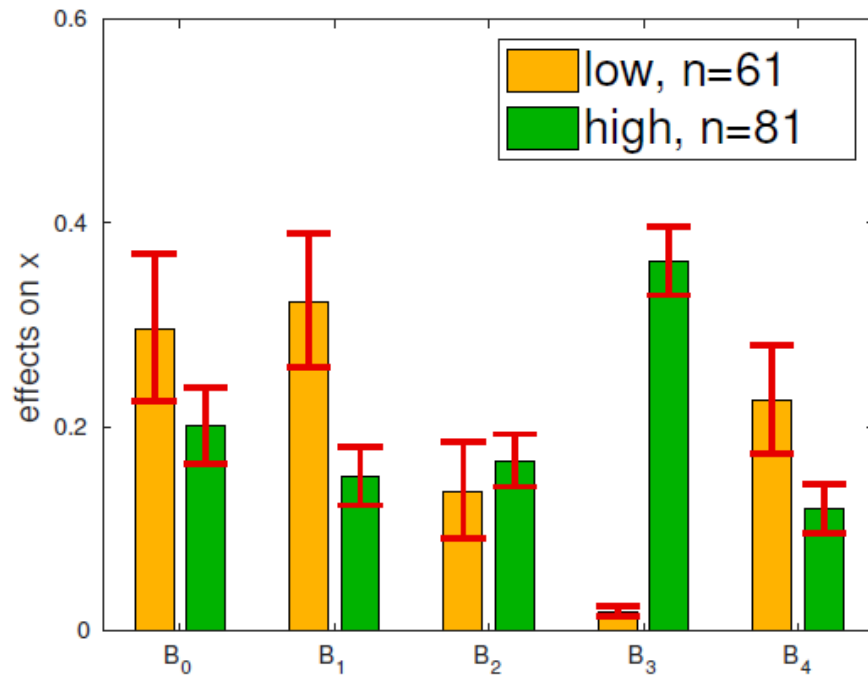
# Differences between individuals with high and low values of $B_1$ (personal norms)



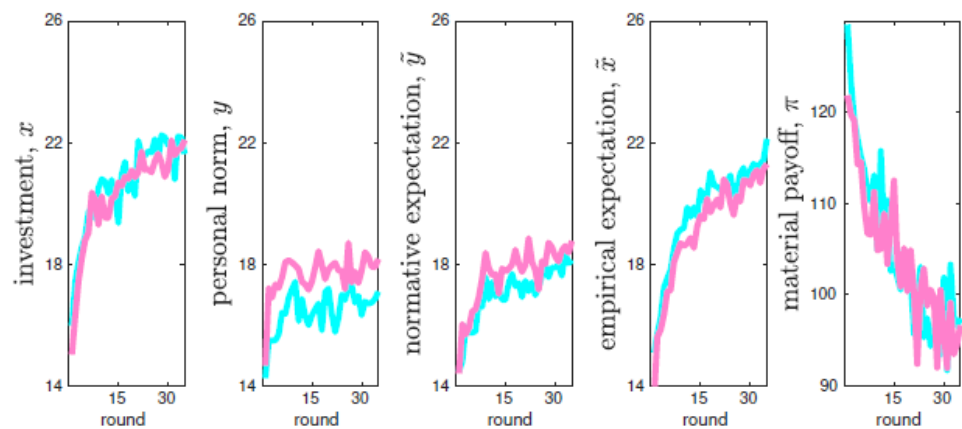
# Differences between individuals with high and low values of $B_2$ (expected disapproval)



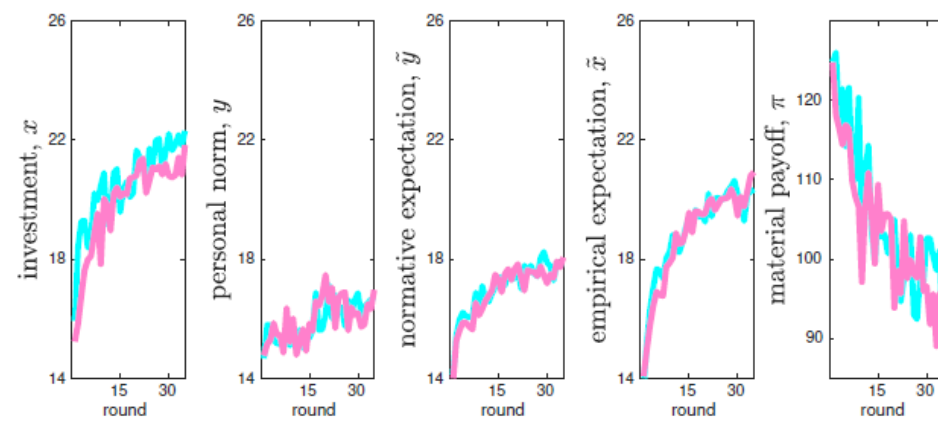
# Differences between individuals with high and low values of $B_3$ (conformity with peers)



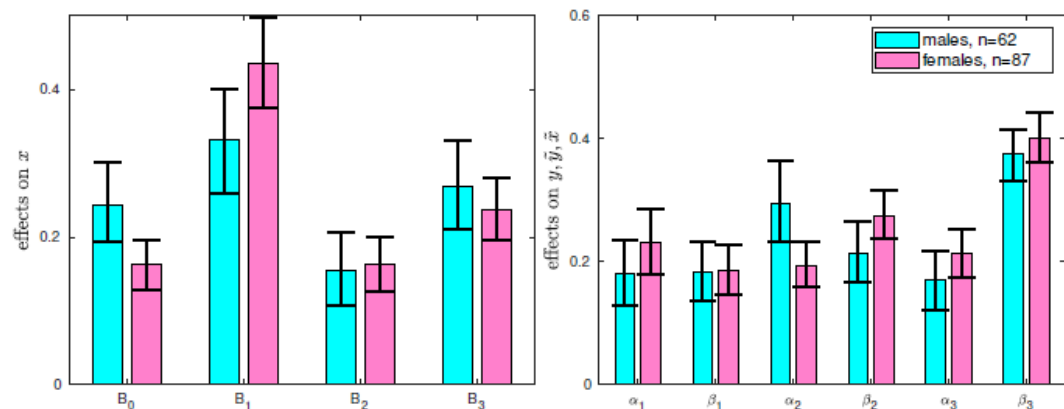
# Differences between sexes are small



(a)



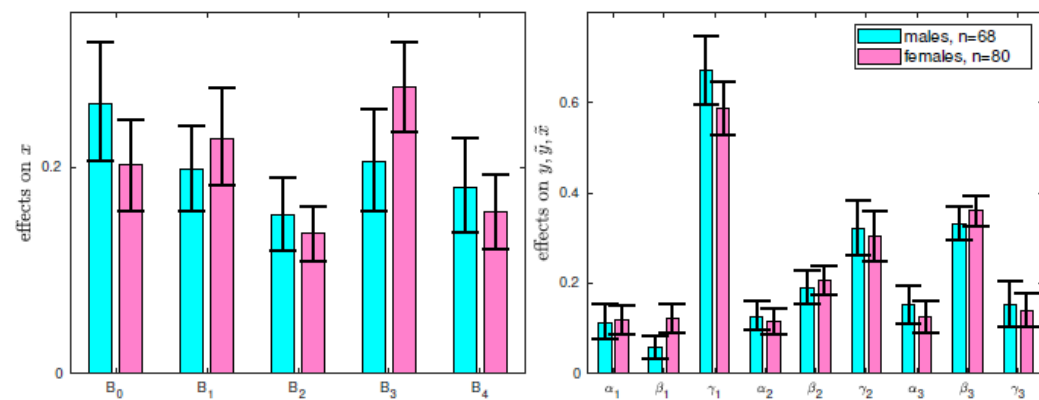
(a)



(b)

(c)

No messaging

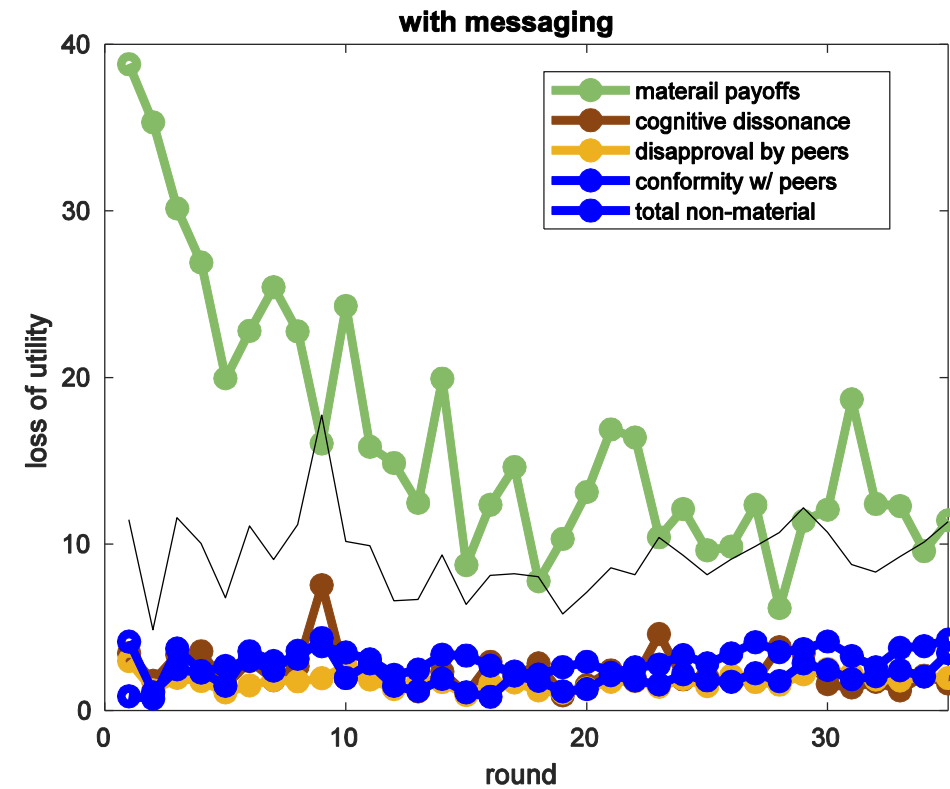
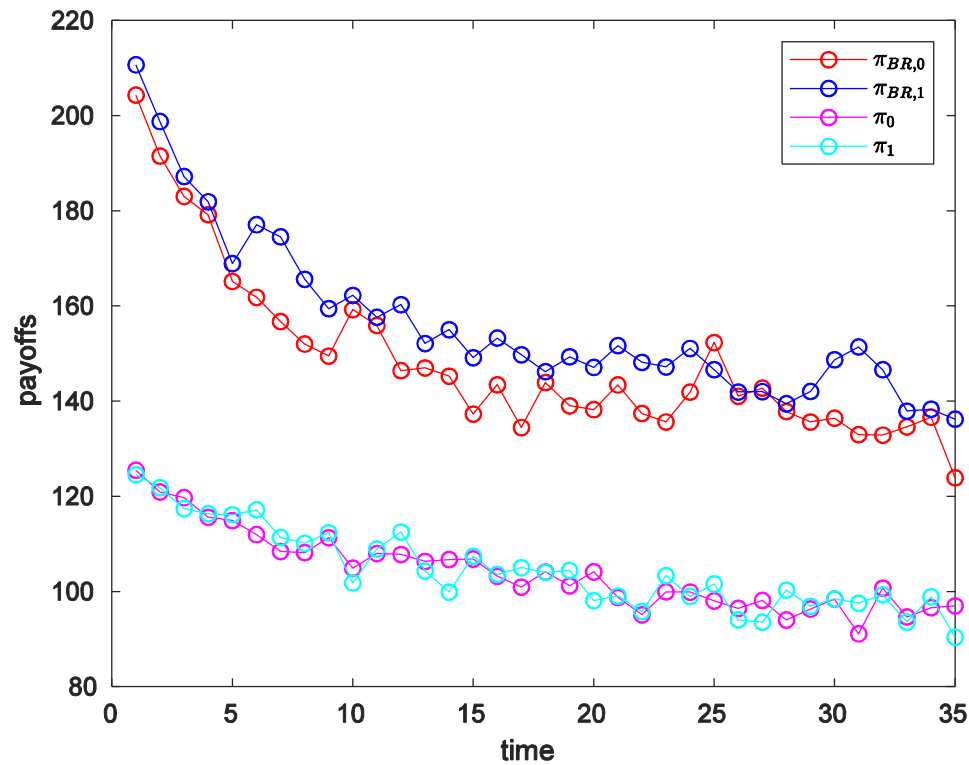


(b)

(c)

With messaging

# Economic significance of material and nonmaterial factors: loss of payoffs and utility



# Additional data/experiments

- CPR experiment
- Prosociality tests
- Rule Following experiments
- Collective Risk I (Szekely et al. 2021)
- Collective Risk II
  
- CPR experiment II (short)
- CPR experiment III (China)



# Additional independent tests

- Prosociality measure (Szekely et al. 2021)
  - Of the four types (competitive, individualistic, prosocial and altruistic), only individualistic and prosocial types are present in our data
    - Individualistic individuals have higher  $B_0$  (material), lower  $B_1$  (personal) and  $B_4$  (response to messaging) than prosocial ones
- Rule following (Kimbrough and Vostroknutov 2016)
  - Rule breakers have higher  $B_0$  (material) and lower  $B_4$  (response to messaging) than rule followers

## Evidence from a long-term experiment that collective risks change social norms and promote cooperation

Aron Szekely<sup>1,2,3,10</sup>✉, Francesca Lipari<sup>4,10</sup>, Alberto Antonioni<sup>4</sup>, Mario Paolucci<sup>1,5</sup>, Angel Sánchez<sup>4,6,7,8</sup>, Luca Tummolini<sup>1,2</sup> & Giulia Andrighetto<sup>1,2,9</sup>

- Groups of size 6; each individual has a budget of 100 units
- If the total group contribution is less than 300 units, the group risks losing everything with probability  $p=0.6$  (low risk treatment) or  $p=0.9$  (high risk treatment)

- Expected payoff  $\pi(x) = (100 - x)[1 - I(\Delta - x)]$ ,

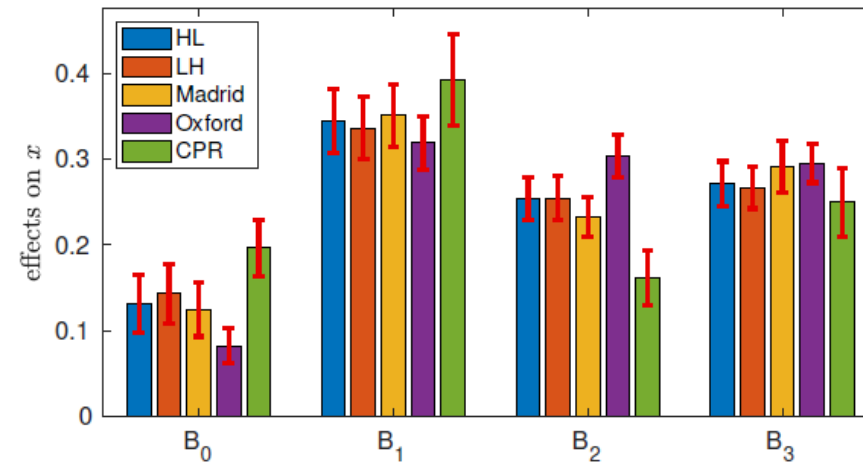
where  $I(\dots)$  is the step (indicator) function, and

$$\Delta = 300 - 5\tilde{x}$$

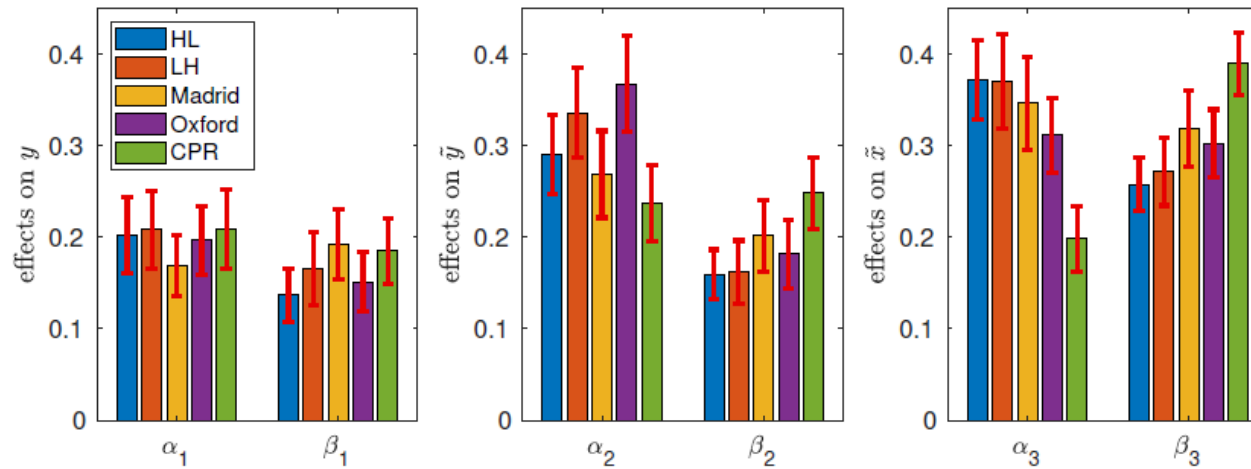
is the “missing” contribution as perceived by the subject

- Best response:  $\theta = \Delta$  if  $\Delta < 100p$ , or 0 otherwise.

# Collective Risk and CPR estimates



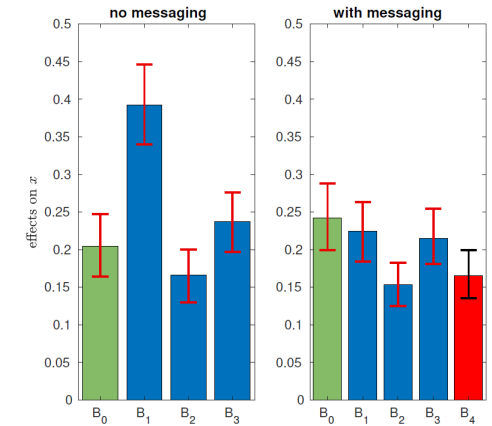
(a)



(b)

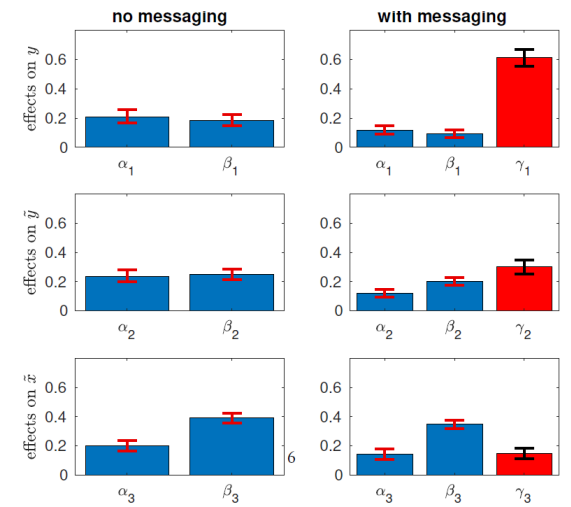
# Conclusions

- We all know that material, social and cognitive factors are important for our actions and beliefs. Research shows that their effects are statistically significant.
- We can go beyond this because we
  - use a model-based approach,
  - measure all variables on exactly the same scale.
- Our approach allows us to compare different factors quantitatively and within the same framework
- The weight of different factor in decision:
  - Material payoffs: about 20%
  - Conformity: about 25%
  - Cognitive dissonance and personal norms: about 40% (w/out authority) or 20% with authority messaging
  - (injunctive) Social norms: about 15%
  - Authority's messaging is about 15%
- Overall, individuals loose 30-40% of material payoffs due to non-material influences



# Beliefs dynamics

- Beliefs and values can change rapidly
- For personal norms and normative expectations, cognitive forces are comparable in strength with social influence by peers. For empirical expectations, learning about others is the stronger force.
- The effect of authority's messaging/propaganda is strong.
- The authority's messaging is most important for personal norms (when its weight is three times that of the joint effect of conformity and cognitive forces), then for normative expectations (when its weight is comparable with the joint effect of conformity and cognitive forces), and the least important for empirical expectations.



# Individual differences

- Are very important, can explain some puzzling outcomes, and need to be accounted for when developing practical policies
- Antisocial individuals (i.e. those who make large investments in CPR) are those who are greedy (high A0), without principles and shame (low A1), low sensitivity to disapproval by peers and authority (low A1 and A4), and conformists in behavior (high A3).
- The distributions of many parameters are highly skewed with a number of extreme individuals.
- Certain pairs of traits shows strong positive or negative correlations
- There are certain clusters (=“behavioral types of individuals”)
- But differences between sexes are rather small

- Our approach opens new directions for theoretical and empirical work as well as for practical applications
- Possible future extensions and applications
  - Different games, especially those stimulating the emergence of stronger social norms
  - Different types of messaging, especially those exploiting social identity
  - Repeating the experiment in different cultures and contrasting the results
  - Real-life analogs and applications

# Applications

- How to better understand and predict behavior and beliefs' change in a population as a result of policy interventions (e.g., aimed at environment protection or climate change mitigation) or certain shocks, such as an epidemic, a natural disaster, or social unrest
- Quantified how descriptive, injunctive and personal norms together with authority's messaging affect human behavior and how beliefs coevolve with actions.
- Different individuals respond differently to these forces
- Policy making must embrace heterogeneity and gathering information about the distribution of individuals' reactivity to different (material, cognitive, social) factors in the population, in order to design and implement suitable intervention strategies for different contexts
  - Messaging which targets different clusters of individuals
  - Highly personalized messaging



**Thank you!**