



**HAL**  
open science

# Community Enforcement of Pro-Social and Pro-Environmental Behavior in the Face of Inequalities : A Path to Climate Change Mitigation

Boris Wieczorek

► **To cite this version:**

Boris Wieczorek. Community Enforcement of Pro-Social and Pro-Environmental Behavior in the Face of Inequalities : A Path to Climate Change Mitigation. Economics and Finance. Université Grenoble Alpes [2020-..], 2023. English. NNT : 2023GRALE008 . tel-04613116

**HAL Id: tel-04613116**

**<https://theses.hal.science/tel-04613116>**

Submitted on 15 Jun 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de

**DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES**

École doctorale : SE - Sciences Economiques

Spécialité : Sciences économiques

Unité de recherche : GAEL - Laboratoire d'Economie Appliquée de Grenoble



## **Application par la Communauté des Comportements Pro-Sociaux et Pro-Environnementaux face aux Inégalités : Une Voie vers l'Atténuation du Changement Climatique**

### **Community Enforcement of Pro-Social and Pro-Environmental Behavior in the Face of Inequalities: A Path to Climate Change Mitigation**

Présentée par :

**Boris WIECZOREK**

#### Direction de thèse :

**Sabrina TEYSSIER**

CHARGE DE RECHERCHE, Université Grenoble Alpes

Directrice de thèse

#### Rapporteurs :

**Marie Claire Villeval**

DIRECTRICE DE RECHERCHE, National Center for Scientific Research (CNRS)

**Nikos Nikiforakis**

PROFESSEUR DES UNIVERSITES, New York University Abu Dhabi

#### Thèse soutenue publiquement le 4 décembre 2023, devant le jury composé de :

**Sabrina Teyssier**

CHARGE DE RECHERCHE HDR, National Research Institute for Agriculture, Food and the Environment (INRAE)

Directrice de thèse

**Marie Claire Villeval**

DIRECTRICE DE RECHERCHE, National Center for Scientific Research (CNRS)

Rapporteuse

**Nikos Nikiforakis**

PROFESSEUR DES UNIVERSITES, New York University Abu Dhabi

Rapporteur

**Nicolas Jacquemet**

PROFESSEUR DES UNIVERSITES, Paris School of Economics

Président

**Erte Xiao**

PROFESSEURE DES UNIVERSITES, Monash University

Examinatrice

**Laurent Muller**

DIRECTEUR DE RECHERCHE, National Research Institute for Agriculture, Food and the Environment (INRAE)

Examinateur

**Béatrice Roussillon**

MAITRESSE DE CONFERENCES, Grenoble Alpes University

Examinatrice

#### Invités :

**Régis Ferrière**

PROFESSEUR DES UNIVERSITES, Biology Institute of the Ecole Normale Supérieure (IBENS), Associate Professor at the University of Arizona (UoA), Deputy Director of France-Arizona Institute for Global Grand Challenges (FA), and Director of Interdisciplinary and Global Environmental Studies (IGLOBES)

# Abstract

Climate change is the most challenging issue we are facing today. It damages the environment, harms humankind, and undermines global growth and stability. Its resolution is paramount to ensure the perpetuation of our societies.

We identify two main concerns that limit cooperation regarding climate change solutions. The first concern is represented by the social dilemma, which highlights the preference of individual interests over the common good. Self-interested behavior leads to sub-optimal overall outcomes due to externalities. In the context of climate change action, self-interested individual behavior refers to the private benefits of carbon dioxide-emitting activities, while externalities refer to the impact of these polluting activities on the entire population. The inequality and heterogeneity of our world represent the second concern. Each region is subject to different levels of potential damage from climate change and is responsible to varying degrees for past and present polluting activities. Inequality exacerbates the social dilemma by bringing into play the concepts of fairness, justice, and heterogeneity.

Our commitment to contributing to climate change mitigation will lead us to explore social norms, belief-based mechanisms for strengthening cooperation to overcome social dilemmas in contexts of inequality. We investigate the role of social norm beliefs as an equilibrium selection mechanism and how they can reinforce pro-social and pro-environmental behaviors. Social norms are composed of descriptive and normative beliefs that influence behavior. However, they can be subject to misconceptions leading to impoverished cooperation. We will provide answers to the following research questions. First, how do social norms' beliefs respond to inequalities and communication mechanisms. Second, how do these mechanisms impact the relation between norms and pro-social and pro-environmental behavior. Third, can we represent social norms in individual decision-making. Fourth, how can social norms be used to promote pro-environmental behavior. To answer these questions, we implemented two complementary methodologies. We use three experiments based on behavioral and experimental economics to investigate the role and impact of social norms on human behavior. We use an ecological feedback model to model pro-environmental behaviors in a social-ecological system.

Our contributions to the understanding of how social norms shape behavior are the following: Firstly, we show how beliefs about social norms of cooperation vary across contexts and time. The mere presence of inequality in the environment causes an initial drop in normative and descriptive beliefs. Furthermore, while norma-

tive and descriptive beliefs tend to decline over time, simple on-shot-free communication supports them. Secondly, we have produced evidence on how social norms influence decision-making. We show that social norms' beliefs impact direct and strategic decision-making, although strategic decisions are less affected. More importantly, our investigation of unequal and heterogeneous frameworks showed that the importance of normative beliefs was similar to that of descriptive beliefs in the case of heterogeneity. Together, social norm beliefs can be seen as an equilibrium selection mechanism for achieving virtuous equilibria, even in the presence of inequality and heterogeneity. Finally, we demonstrate how social norms can induce behavioral change. We show three ways of using social norm information to improve pro-social and pro-environmental behavior. First, by revealing the actual value of social norms, leading to a revision of beliefs in the right direction. Second, by focusing attention on a particular population to highlight the right behaviors. Third, by deploying targeted information policies to influence marginal behaviors through static information and average behaviors through dynamic information.



# Résumé

Le changement climatique est le plus grand défi auquel nous sommes confrontés. Il nuit à l'environnement, damages la santé et compromet la croissance et la stabilité mondiales. Sa résolution est primordiale pour assurer la pérennité de nos sociétés.

Nous identifions deux difficultés principales qui limitent la coopération autour du changement climatique. La première difficulté relève du dilemme social, issue de la préférence des intérêts individuels par rapport au bien commun. Les comportements égoïstes conduisent à des résultats globaux sous-optimaux en raison d'externalités. Dans le contexte du changement climatique, les comportements égoïstes se réfèrent aux bénéfices privés des activités émettrices de dioxyde de carbone, tandis que les externalités se réfèrent à l'impact de ces activités polluantes sur l'ensemble de la population. La deuxième difficulté est représentée par l'inégalité de notre monde. Chaque région est soumise à différents niveaux de dommages potentiels et est responsable à différents degrés d'émission polluantes. Les inégalités exacerbent le dilemme social en ajoutant les concepts d'équité, de justice et d'hétérogénéité.

Notre volonté d'atténuer le changement climatique nous amène à explorer les normes sociales, un mécanisme fondé sur les croyances afin de surmonter les dilemmes sociaux dans des contextes d'inégalité. Nous étudions leurs rôles en tant que mécanisme de sélection d'équilibre et la manière de les utiliser pour améliorer les comportements. Les normes sociales sont composées de croyances descriptives et normatives qui influencent le comportement. Cependant, elles peuvent être erronées et nuire à la coopération. Nous répondrons aux questions suivantes. Comment les croyances sur les normes réagissent-elles aux inégalités et à la communication. Comment l'hétérogénéité et la communication influencent la relation entre les normes et les comportements. Comment représenter les normes dans la prise de décision individuelle. Comment les normes peuvent-elles être utilisées pour promouvoir les bons comportements. Pour répondre à ces questions, nous mettons en œuvre deux méthodologies complémentaires. À l'aide de trois expériences en économie comportementale, nous étudions le rôle et l'impact des normes sociales sur les comportements individuels. À l'aide d'un modèle de rétroaction écologique, nous modélisons les comportements pro-environnementaux dans un système socio-écologique.

Nos contributions sont les suivantes :

Premièrement, nous montrons comment les croyances concernant les normes sociales de coopération varient en

fonction du contexte et du temps. La présence d'inégalités entraîne une baisse initiale des croyances normatives et descriptives. En outre, alors que ces croyances ont tendance à diminuer avec le temps, une simple communication libre ponctuelle les supporte. Deuxièmement, nous montrons comment les normes sociales influencent la prise de décision. Nous montrons que les normes sociales influencent les décisions directe et stratégique. Les décisions stratégiques sont toutefois moins affectées. Plus important, nos études des cadres inégaux et hétérogènes ont montré que les croyances normatives sont aussi importantes que les croyances descriptives en cas d'hétérogénéité. Les normes sociales peuvent être considérées comme un mécanisme de sélection d'équilibre permettant d'atteindre des équilibres vertueux, même en présence d'inégalité et d'hétérogénéité. Enfin, nous montrons trois façons d'utiliser les normes sociales pour changer les comportements. D'abord, en révélant la véritable norme sociale, conduisant à une révision des croyances dans la bonne direction. Ensuite, en concentrant l'attention sur une population distincte afin de mettre en évidence les bons comportements. Finalement, en déployant des politiques d'information ciblées pour influencer les comportements marginaux au moyen d'informations statiques, et les comportements majoritaire au moyen d'informations dynamiques.

# Acknowledgements

First of all, I would like to thank the jury members for their time and valuable comments. Thanks to Marie Claire Villeval and Nikos Nikiforaki for accepting the role of rapporteur; and to Nicolas Jacquemet, Erte Xiao, Laurent Muller, and Béatrice Roussillon for accepting the role of reviewer. I look forward to developing collaboration and future research with you.

I am grateful to my supervisor, Sabrina Teyssier, firstly for giving me the opportunity to start my Ph.D., secondly for providing me with all the financial support I needed to carry out my experiments and participate in various international conferences, and thirdly for finding the patience to support my character and autonomy for three years. I am also grateful to Régis Ferrière for integrating me into a fantastic multidisciplinary project. Working alongside you has been a real pleasure and taught me much about science, methodology, and cooperation.

I thank the Grenoble Applied Economy Laboratory for caring for me for almost two years. I appreciated each member's benevolence, attention, and pleasant moments. Thank you for all the discussions and humor that brightened my working day during my countless breaks. To the best research support team Cédric Lanu, Erick Raberanto, Gwenaëlle Desury, Nadia Youkana, and Aline Fugerey-Scarbel. To the funniest researcher team Adrien Hervouet, Cédric Clastres, Paolo Crosetto, Stéphane Robin, Béatrice Roussillon, Olivier Bonroy, Frédéric Corolleur, Adélaïde Fadhuile, Laurent Muller, Alexis Garapin, Benjamin Ouvrard, Stéphane Lemarié, Daniel Llerena, Stephan Sémirat and Giorgio Fabbri. To my PhD Student partner, Blandin Lola.

I would also like to thank the Interdisciplinary and Global Environmental Studies and the Eller College of Management of the University of Arizona for their hospitality in hosting me for over a year. Thanks to Régis Ferrière, Sébastien Roux, and Ruth Gosset for your guidance and for discovering the American country and culture every day. Thanks to Charles Noussair for introducing me to the economic community at the University of Arizona. Thanks to Martin Dufwenberg for these philosophical reflections on human behavior. Thanks to Ashley Langer for the best course I have ever attended. Thanks to Jeong Yeol Kim for the pleasant moments that enabled me to break from work.

Special thanks to Martin's weekly meetings, which gave me a different perspective on my work and the relaxation required to complete a PhD. Thanks to their permanent members, Cédric Lanu, Adrien Hervouet, Erick Raberanto, Gwenaëlle Desury, as well as their visitors Nadia Youkana, Cédric Clastres, Olivier Bonroy, Paolo Crosetto, Stéphane Robin, Adélaïde Fadhuile, Alexis Garapin, Stéphane Lemarié, Aline Fugeray-Scarbel, and Laurent Muller.

Finally, I thank my parents for giving me the opportunity and support to pursue my university studies worldwide.

# Contents

<b>1</b>	<b>Introduction</b>	<b>9</b>
<b>2</b>	<b>Inequality, Social Norms and Cooperation: Strategy Choice in the Infinitely Repeated Prisoners' Dilemma</b>	<b>20</b>
2.1	Introduction . . . . .	20
2.2	The experiment . . . . .	23
2.2.1	The game . . . . .	23
2.2.2	Experimental treatments . . . . .	25
2.2.3	Procedures . . . . .	27
2.2.4	Theoretical predictions . . . . .	29
2.3	Results . . . . .	30
2.3.1	Social norms . . . . .	30
2.3.2	Decisions and strategies . . . . .	33
2.4	Discussion and conclusion . . . . .	37
<b>3</b>	<b>Inequality, Social Norms and Community Enforcement of Cooperation: Experimental Evidence</b>	<b>40</b>
3.1	Introduction . . . . .	40
3.2	The experiment . . . . .	42
3.2.1	Framework . . . . .	42
3.2.2	The basic game . . . . .	43
3.2.3	Experimental treatments . . . . .	45
3.2.4	Procedures . . . . .	46
3.3	Results . . . . .	47
3.3.1	Inequality . . . . .	47
3.3.2	Communication . . . . .	53
3.4	Discussion and conclusion . . . . .	59

<b>4</b>	<b>Dynamic Norms for Household Water Consumption</b>	<b>64</b>
4.1	Introduction . . . . .	64
4.2	Experimental design . . . . .	67
4.2.1	Procedure . . . . .	67
4.2.2	Experimental Design . . . . .	68
4.2.3	Survey . . . . .	70
4.3	Results . . . . .	71
4.3.1	Water evaluation and treatment effect . . . . .	71
4.3.2	Cognitive and Psychological processes . . . . .	74
4.4	Discussion . . . . .	76
<b>5</b>	<b>Disentangling Individual Attitudes, Expectations, and Experience in the Adoption of Proenvironmental Behavior</b>	<b>79</b>
5.1	Introduction . . . . .	79
5.2	Results . . . . .	81
5.2.1	Model overview . . . . .	81
5.2.2	Effect of attitudes and attitude strength . . . . .	84
5.2.3	Effect of normative expectations . . . . .	85
5.2.4	Effect of heterogeneity in attitudes' strength . . . . .	87
5.3	Discussion . . . . .	88
5.4	Methods . . . . .	91
<b>6</b>	<b>Conclusion</b>	<b>100</b>
<b>A</b>	<b>Inequality, Social Norms and Cooperation: Strategy Choice in the Infinitely Repeated Prisoners' Dilemma - Supplementary</b>	<b>105</b>
A.1	Equilibrium continuation probabilities . . . . .	105
A.1.1	Calculation of $\delta^{SPE}$ . . . . .	105
A.1.2	Calculation of $\delta^{RD}$ . . . . .	106
A.2	Instructions in the unequal-egalitarian treatment . . . . .	107
A.3	Classification of the menu of strategies . . . . .	117
A.4	Descriptive statistics . . . . .	118
A.4.1	Normative beliefs . . . . .	118
A.4.2	Descriptive beliefs . . . . .	119
A.4.3	Decisions in period 1 . . . . .	119

A.4.4	Conditional decisions in period 2 . . . . .	120
A.4.5	Memory-one strategies . . . . .	120
A.4.6	Menu of strategies . . . . .	121
<b>B</b>	<b>Inequality, Social Norms and Community Enforcement of Cooperation: Experimental Evidence - Supplementary</b>	<b>123</b>
B.1	Equilibrium continuation probabilities . . . . .	123
B.1.1	Calculation of $\delta^{SPE}$ . . . . .	123
B.1.2	Calculation of $\delta^{RD}$ . . . . .	125
B.2	Instructions . . . . .	126
B.3	Regressions explaining cooperation for the first and last three rounds . . . . .	144
B.4	Cooperation per matching in INEQ-EA and INEQ-EP . . . . .	146
B.5	Classification of messages in the communication treatments . . . . .	147
<b>C</b>	<b>Dynamic Norms for Household Water Consumption - Supplementary</b>	<b>149</b>
C.1	Household classification . . . . .	149
C.2	Long-term treatment effects . . . . .	151
C.3	Cognitive and Psychological processes . . . . .	153
C.3.1	Theory of planned behavior . . . . .	154
C.3.2	Normative evaluation . . . . .	154
C.4	Instruction . . . . .	155
C.4.1	Weekly-email . . . . .	155
C.4.2	Web page information . . . . .	159
C.4.3	Survey . . . . .	161
C.4.4	Preregistration . . . . .	177
<b>D</b>	<b>Disentangling Individual Attitudes, Expectations, and Experience in the Adoption of Proenvironmental Behavior - Supplementary</b>	<b>179</b>

# List of Figures

2.1	Distribution of second-order normative beliefs . . . . .	31
2.2	Distribution of descriptive beliefs . . . . .	31
3.1	Cumulative distribution of groups . . . . .	48
4.1	Different phases of the design . . . . .	68
4.2	Static information-based interventions . . . . .	69
4.3	Dynamic information-base policy . . . . .	70
4.4	Evolution of consumption over weeks, by category and treatment. <b>A</b> Represent the average weekly water consumption by category and treatment. <b>B</b> represent the average percentage of variation of water consumption compared to the first week by category and treatment. . . . .	72
5.1	Effect of attitudes and attitude strength on the adoption of the active behavior . . . . .	94
5.2	Exemplary trajectories in the model state space . . . . .	95
5.3	Effect of the strength of the descriptive and normative components of the social norm, $\delta_D$ and $\delta_N$ , on the adoption of the active behavior . . . . .	96
5.4	Effect of descriptive focus, $\rho_{aD}$ , and normative focus, $\rho_{aN}$ , on the adoption of the active behavior . . . . .	97
5.5	Effect of idealized public policy interventions . . . . .	98
C.1	Distribution of relative consumption for the first two weeks. Notes: Relative consumption refers to consumption per number of inhabitants. m-, m, and m+ represent the boundaries implemented to create categories. . . . .	149
C.2	Normative evaluation of water behavior . . . . .	155
D.1	Effect of attitudes and attitude strength on the adoption of the active behavior . . . . .	180
D.2	Effect of attitudes and attitude strength on the adoption of the active behavior . . . . .	181
D.3	Effect of the strength of the descriptive and normative components of the social norm, $\delta_D$ and $\delta_N$ , on the adoption of the active behavior . . . . .	182



D.4	Effect of descriptive focus, $\rho_{aD}$ , and normative focus, $\rho_{aN}$ , on the adoption of the active behavior . . .	183
D.5	Effect of descriptive focus, $\rho_{aD}$ , and normative focus, $\rho_{aN}$ , on the adoption of the active behavior . . .	184
D.6	Effect of descriptive focus, $\rho_{aD}$ , and normative focus, $\rho_{aN}$ , on the adoption of the active behavior . . .	185
D.7	Effect of descriptive focus, $\rho_{aD}$ , and normative focus, $\rho_{aN}$ , on the adoption of the active behavior . . .	186
D.8	Effect of descriptive focus, $\rho_{aD}$ , and normative focus, $\rho_{aN}$ , on the adoption of the active behavior . . .	187

# List of Tables

2.1	Egalitarian distribution of the benefits of cooperation . . . . .	24
2.2	Proportional distribution of the benefits of cooperation . . . . .	24
2.3	Payoffs in the Equal-D treatment . . . . .	25
2.4	Payoffs in the Equal-A treatment . . . . .	26
2.5	Payoffs in the Unequal-E treatment . . . . .	26
2.6	Payoffs in the Unequal-U treatment . . . . .	27
2.7	Beliefs, by treatment . . . . .	32
2.8	Decisions in periods 1 and 2 . . . . .	34
2.9	Memory-one strategies . . . . .	35
2.10	Menu of strategies . . . . .	36
3.1	Payoffs matrix . . . . .	44
3.2	Payoffs for two Low-type subjects . . . . .	45
3.3	Payoffs for two High-type subjects . . . . .	45
3.4	Payoffs for a Low-type subject (line) and a High-type subject (column) . . . . .	45
3.5	Distribution of participants by treatment . . . . .	46
3.6	Cooperation per subject type and treatment . . . . .	47
3.7	Cooperation per matching sort in INEQ-N treatment . . . . .	48
3.8	First-order normative beliefs per subject type and treatment . . . . .	49
3.9	Second-order normative beliefs per subject type and treatment . . . . .	50
3.10	Descriptive beliefs per subject type and treatment . . . . .	50
3.11	Decision to cooperate (Logit models - average marginal effects) . . . . .	52
3.12	Cooperation per participant type and treatment with communication . . . . .	54
3.13	First-order normative beliefs per subject type and treatment with communication . . . . .	55
3.14	Second-order normative beliefs per subject type and treatment with communication . . . . .	56
3.15	Descriptive beliefs per subject type and treatment with communication . . . . .	56

3.16 Decision to cooperate with communication (Logit models - average marginal effects) . . . . .	58
3.17 Estimation of average cooperation based on Equation (3.1) . . . . .	61
4.1 Water consumption (OLS models - average marginal effects). Cluster at the household level. Notes: Regression was conducted on weeks 1-5. Men and Women are over 13 years old, while Son and Daughter are under 13 years old. . . . .	73
4.2 Normative assessment of water-related behaviors. Notes: Score between -1 and 1, with -1 referring to low level and 1 to high level. <i>p-value</i> given by Wilcoxon statistical test. DI refers to Dynamic Information, and SI to Static Information. . . . .	74
4.3 Emotional evaluation of information-based policies. Notes: Score between -1 and 1, with -1 referring to low level and 1 to high level. <i>p-value</i> given by Wilcoxon statistical test. DI refers to Dynamic Information, and SI to Static Information. . . . .	75
5.1 Notations, description, and parameter default values. . . . .	93
A.1 Classification of the menu of strategies . . . . .	117
A.2 Share of participants who believe that the decision a participant should make is to cooperate (first- order normative beliefs) . . . . .	118
A.3 Share of participants who believe that disadvantaged participants believe that the decision a partici- pant should make is to cooperate (second-order normative beliefs for low-endowed participants) . . .	118
A.4 Share of participants who believe that advantaged participants believe that the decision a participant should make is to cooperate (second-order normative beliefs for high-endowed participants) . . . . .	119
A.5 Descriptive beliefs . . . . .	119
A.6 Decision to cooperate in period 1 . . . . .	119
A.7 Conditional decision to cooperate in period 2 . . . . .	120
A.8 Distribution of Memory-one strategies . . . . .	120
A.9 Distribution of strategies in the menu of strategies . . . . .	121
B.1 Decision to cooperate in the first three rounds (Logit models - average marginal effects) . . . . .	144
B.2 Decision to cooperate in the last three rounds (Logit models - average marginal effects) . . . . .	145
B.3 Cooperation per matching in INEQ-EA and INEQ-EP treatments . . . . .	146
B.4 Proportion of messages per category of message and treatment . . . . .	147
C.1 Average household characteristics by categories. Notes: Income score thresholds 1=less than 30 000€, 2=between 30 000 and 40 000€, 3=between 40 000 and 50 000€, 4=between 50 000 and 60 000€, 5=more than 60 000€. Water consumption in liters per week. . . . .	150

C.2 Water consumption (OLS models - average marginal effects). Cluster at the household level. Notes: Regression was conducted on weeks 1-6. Men and Women are over 13 years old, while Son and Daughter are under 13 years old. . . . . 151

C.3 Water consumption (OLS models - average marginal effects). Cluster at the household level. Notes: Regression was conducted on week 1-7. Men and Women are over 13 years old, while Son and Daughter are under 13 years old. . . . . 152

C.4 Reduction in per capita water consumption in phase 3 (OLS models - average marginal effects). Notes: All scores range between -1 and 1. . . . . 153

C.5 Direct and related components to Theory of planned behavior. Notes: Score between -1 and 1, with -1 referring to low and 1 to high levels. DI refers to Dynamic Information, and SI refers to Static information. . . . . 154

# Chapter 1

## Introduction

Climate change is one of the most challenging issues we are facing today. It damages the environment (Eslamian and Eslamian, 2017), presents direct harms to humankind (Bifulco and Ranieri, 2017), jeopardizes local economy (Tol, 2018), and undermines global growth and stability (World Bank, 2005; , 2013). Its resolution is paramount to ensure the perpetuation of our society as we know it.<sup>1</sup>

Climate change is a consequence of human activities such as burning carbon-based fuels, which is the primary factor of carbon dioxide emissions (Pörtner et al., 2022; Shukla, Skea and Slade, 2022). These gases trap heat, provoking an increase in land and ocean temperatures. As a consequence, six of the nine planetary boundaries defined by the Stockholm Resilience Center have already been exceeded (Persson et al., 2022; Rockström et al., 2009). These boundaries are the guardians of our hospitable world, preventing the planet's natural balance from triggering several feedback effects that deteriorate the environment (Steffen et al., 2015). A global change is thus expected with an increase in temperature about 3-5° by 2100 accompanied by extreme hit wave and precipitation, an increase in natural disasters such as tornadoes and hurricanes, and water scarcity problems (Intergovernmental Panel On Climate Change, 2023; Shukla, Skea and Slade, 2022).

The role of humankind in climate change is now widely accepted, both as an explanation of its origins and as the only way to mitigate it (Intergovernmental Panel On Climate Change, 2023; Persson et al., 2022; Shukla, Skea and Slade, 2022; Pörtner et al., 2022). However, identifying climate change solutions is insufficient to mobilize all stakeholders to resolve the issue. Behavioral sciences have identified several concerns that limit behavioral responses toward climate change solutions. The imperative of action at the individual level around an issue with collective consequences leads to a social dilemma. Social dilemmas emerge from the pursuit of personal interests despite the presence of externalities induced by everyone's actions. Self-interested individual behavior leads global outcomes to diverge from social surplus. In the context of climate change action, self-interested individual behavior refers to the private benefits of carbon dioxide-emitting activities, while externalities refer to the impact of these polluting

---

<sup>1</sup>Pessimistic predictions of late adaptation to climate change include rationing, war and collapse (Zhang et al., 2007).

activities on the entire population. To overcome social dilemmas, cooperation mechanisms are needed to connect individuals toward a common goal. In addition, the presence of inequality in individual decision-making exacerbates social dilemmas by bringing into play the questions of equity, justice, and heterogeneity. Inequality refers to a broad set of elements. Conventionally, it refers to the unequal distribution of consumption and expenditure, disposable and market income, wealth, etc. It can also be associated with differences in access to goods and services, production characteristics, or needs. In the context of climate change action, inequality has two important representations. On the one hand, individuals do not have the same sensitivity to climate change; on the other hand, they do not have the same capacity to adapt to climate change. Indeed, several regions are subject to different levels of potential damages, and others to potential benefits from climate change. While hot and arid regions are the most prone to damage, the poverty of many of them makes the costs of implementing measures to support or adapt to climate change unfeasible. On the contrary, countries located in the middle and higher latitudes are the least affected, although they are the main ones responsible for these changes through past industrialization and current over-consumption that have allowed them to prosper. Moreover, rising inequality has been a trend for decades, underlining its importance for numerous issues (Stiglitz, 2015; Piketty, Saez and Zucman, 2018). The World Inequality Report (Chancel et al., 2022) highlighted dramatic data on the increase in inequality from 1980 to the present day. They report that "the top 1% (richest) captured 38% of the global increment in wealth, while the bottom 50% (poorest) captured a frightening 2%".

The current thesis explores mechanisms to strengthen cooperation to overcome social dilemmas in contexts of inequality.

Our commitment to contributing to climate change mitigation at the individual level takes us away from traditional macro-level monetary policy tools and towards micro-level cooperation mechanisms that can be implemented in any context. Our investigation first led us to examine the mechanism of punishment. Fehr and Gächter (2000a) have developed a system of cheap punishments, i.e., the cost incurred by punishment does not exceed the gain an individual derives from non-cooperative behaviors. With this system, every individual can impose an ex-post punishment on any other member. This mechanism has proved effective in reducing non-cooperative behavior. Nonetheless, it is imperfect for several reasons, reducing overall outcome (Nikiforakis, 2008). Firstly, it does not entirely resolve cooperation problems. Secondly, it is associated with several implementation costs for the monitoring and enforcement procedure. Finally, it can lead to inefficient behaviors such as counter-punishment (Thöni, 2012). An interesting way of solving problems relating to costs and counter-punishments is to use non-monetary punishments. Masclet et al. (2003) have modified the system to replace the monetary aspect with information. In this way, individuals can express their approval or disapproval to others without incurring a monetary cost. Receiving disapproval implies that one's behavior is inappropriate for others, while receiving approvals implies that one's behavior is appropriate for others. Eventually, individuals may revise their beliefs and adapt their behavior accordingly. This

mechanism has proved almost as effective in increasing cooperation as monetary punishments. He also prevents counter-punishments and is associated with lower implementation costs.

This mechanism illustrates our ambition to improve behavior through low-cost information mechanisms that impose no restrictions on behavior. By focusing on individuals' beliefs and perceptions of their social environment, belief-based mechanisms represent easy-to-use and easy-to-implement tools for strengthening cooperation in favor of the commons. These mechanisms consider that individuals may have misconceptions about behaviors, environments, and societies. Correcting these misconceptions enables behaviors to be addressed at a low cost, with minimal monitoring over longer periods. Behavioral economists consider these beliefs as additional elements driving behavior, in addition to monetary elements. Representation of these non-monetary components, such as reciprocity concern (Fehr and Gächter, 1998), altruistic concerns (Eckel and Grossman, 1996), emotion concerns (Bentham, 1970; Loewenstein, 2000), among many others, have been examined and integrated to the comprehension of human behavior. Although they explain behavior, these components cannot be easily manipulated to lead to the revision of individuals' beliefs. Therefore, we turn our attention to social norms, beliefs-based cooperation mechanisms linked to misconceptions, from understanding how they work to how they can amplify cooperative behavior.

Social norms have been the subject of much academic debate, and their definition has evolved over time. The current study uses the terminology of Bicchieri (2005, 2016), which defines norms as implicit rules emerging from endogenous interactions among agents, differing across populations and cultures, that lead individuals to adopt a specific type of behavior. We consider that norms are divided into two categories: descriptive norms and social norms, depending on the expectations they raise. We distinguish two types of expectations: empirical expectation (also called descriptive beliefs) and normative expectation (also called normative beliefs or second-order normative beliefs). Empirical expectation refers to the beliefs in the behavior performed by the referential group. In contrast, normative expectation refers to the beliefs of the social appropriateness of behaviors by a referential group. Therefore, we will talk about descriptive norms when a behavior depends only on empirical expectations, and social norms when behavior depends on both empirical and normative expectations. Accordingly, descriptive norms apply to situations where individuals need to coordinate, such as recycling behavior, dress code at work, or the direction of traffic. Alternatively, social norms apply to situations where individuals have beliefs about the behavior of others, and beliefs about others' evaluation of the appropriateness of that behavior, such as littering, the use of a smartphone during a meeting, or eating practices in the company of others. Our definition differs from a part of the literature in its categorization. This literature distinguishes between descriptive norms and injunctive/prescriptive norms, which are constituted respectively by the descriptive beliefs and normative beliefs described above (Cialdini, Reno and Kallgren, 1990; Bénabou, Falk and Tirole, 2018). In this manuscript, we will focus on social norms since they encompass and extend the component of descriptive norms. In addition to having proven their effectiveness in a variety of contexts, they represent an informal institution that can play an essential role in policymaking, all the more so when formal institutions are unable to impose appropriate behavior, as is the case with climate change

policies (Nyborg et al., 2016).<sup>2</sup> Research into social norms has already led to recommendations for policy-makers (see Schultz et al. (2007); Allcott (2011); Ayres, Raseman and Shih (2013) on energy consumption, Otaki, Ueda and Sakura (2017); Landon et al. (2018) on water consumption, Perkins (2002); Foxcroft et al. (2015) on alcohol consumption, Corno, La Ferrara and Voena (2020); Gulesci et al. (2021) on genital cutting), and there is no doubt that further research will lead to new recommendations.

Considering the existence of a shared knowledge of the norms in place, compliance or non-compliance with norms results from a cost-benefit analysis. Compliance with norms is considered costly, as the absence of cost eliminates the need for norms. Likewise, its transgression represents a cost that can be direct in the form of punishment and ostracism, or indirect in the form of moral disapproval and peer pressure. Consequently, if a norm is strong enough, individuals may comply with it to avoid the cost of its transgression. On the contrary, if a norm is too weak or ambiguous, individuals may transgress it to benefit from the other behavior. In these respects, compliance with norms can be a sign of a norm's strength.

To develop research on the use of social norms to enhance cooperation, we explore the role of social norms as an equilibrium selection mechanism. We assumed that norms perception and compliance could be a method of equilibrium selection in this context, allowing to reach a focal equilibrium in multiple equilibrium games (Burke and Young, 2011; Young, 2015).

To this end, we have implemented two complementary methodologies over four chapters to answer the following research questions. (i) How do social norms' beliefs respond to inequalities and communication mechanisms? (ii) How do heterogeneity and communication mechanisms impact the relation between social norms and pro-social and pro-environmental behavior? (iii) How can we represent social norms in individual decision-making? (iv) How can social norms be used to promote pro-environmental behavior?

The following three chapters apply behavioral and experimental economics methodologies. Behavioral economics is a branch of economics that integrates various behavioral components, mainly from psychology, cognitive science, and sociology literature. Incorporating these components into theoretic representations of interactions has led to a new understanding of human behavior and more accurate representations of strategic interactions. Experimental economics represents a method for developing targeted experiments on specific issues to measure the effect of a particular characteristic. Experiments are used to measure the causal impact of a characteristic in a particular behavior, and to observe the correlation of the intermediate mechanism leading to behavioral change. We use behavioral and experimental economics to investigate the role of social norms on cooperative behavior. Our experiments enable us to quantify and represent how social norms vary and evolve according to different settings,

---

<sup>2</sup>Social norms have been very effective to explain and predict different phenomena by identifying the perception of appropriateness and the implemented behaviors that constitute norms, such as the variations of behavior linked to framing and culture (Krupka and Weber, 2013; Dimant, van Kleef and Shalvi, 2020; Kimbrough et al., 2022; Constantino et al., 2022), ethical conduct of financial advisors (Burks and Krupka, 2012), bribery (Banerjee, 2016b), discrimination (Barr, Lane and Nosenzo, 2018), and gendered occupational choice (Gangadharan et al., 2016).



how social norms can explain behavior, and how they can induce behavioral changes.

To achieve these objectives, we required a method for eliciting social norms and a framework to represent our social dilemma.

For a long time, social norms have been considered unobservable rules of society that deviate behavior from theories (Henrich et al., 2004). At times, norms were seen as the origin of "abnormal" behavior, such as in the case of involuntary unemployment (Akerlof, 1980), conformity behavior (Bernheim, 1994), or the use of costly punishment (Fehr and Gächter, 2000*b*). Formerly complex to formalize, the "inexpressible" concept of social norms has been thoroughly examined, and methods have been developed to capture them.<sup>3</sup> The first method is the one of (Bicchieri and Xiao, 2009) (BX), which elicits direct individual beliefs about social norms. For this purpose, the appropriateness of behaviors in a given situation is questioned. First, individually with first-order normative beliefs, which are equivalent to personal preferences. And secondly, with second-order normative beliefs, by asking what a referent group's answers are to the first question. Additionally, descriptive beliefs are addressed through questions about behavior performed by a referent group. The second method is the one of Krupka and Weber (2013) (KW), corresponding to a coordination game that enables the capture of second-order normative beliefs. Participants must find the modal response of a referential group on the social appropriateness of behavior in a four-sized Likert scale. These methods have proven effective in assessing social norms' beliefs across contexts.<sup>4</sup>

In spite of the existence of these methods, the study of norms remains complex due to its endogenous co-evolution with behavior, law, and culture (Lane, Nosenzo and Sonderegger, 2023; Eriksson et al., 2021). Similar situations may imply different norms according to the environment. Therefore, each norm must be considered local and involve the expectations of a specific reference group. A notable example concerns the compliance with traffic lights. Most of us assume that drivers ought (normative expectation) and do (empirical expectation) comply with traffic lights. However, some regions of Italy or Colombia have local norms against waiting at traffic lights when no other cars are on the adjacent road (Villaveces et al., 2012; Collotta et al., 2014; Campisi et al., 2020). Accordingly, waiting for a green light at these locations can be associated with inappropriate behavior that can and should be punished. This specificity leads us to investigate the existence of different social norms of cooperation during the interaction of heterogeneous populations by eliciting beliefs specific to each sub-population.

To better understand how norms emerge from endogenous interaction to be a driver of behavior, we are looking at indefinite iterated interaction, which we will represent in the laboratory by the indefinitely repeated prisoner dilemma game (IRPDG). IRPDG is an extension of the Prisoner's Dilemma Game (PDG). These games involve individuals having to decide between self-oriented choices, which benefit only themselves, and socially-oriented choices, which benefit all interacting individuals. To represent the social dilemma, earnings from self-oriented choices are higher

---

<sup>3</sup>Battigalli, Catonini and De Vito (n.d.) defined expressible concepts as "something that can be expressed in a clear, precise, and not self-referential language".

<sup>4</sup>These methodologies have been used to explain different phenomena involved in these games, such as reciprocity (Gächter, Nosenzo and Sefton, 2013; Nikiforakis, Oechssler and Shah, 2014), fair sharing (Gächter, Gerhards and Nosenzo, 2017), keeping promises (Krupka, Leider and Jiang, 2017), lying (d'Adda et al., 2017), among others.

than individual earnings from socially-oriented choices. Therefore, if all individuals commit to socially-oriented choices, the overall benefit is maximized. However, each individual has a personal interest in deviating from self-oriented choices to take advantage of socially-oriented choices made by others while maximizing personal earnings from his own choice. Yet, if all individuals adopt self-oriented choices, they all end up in an inefficient situation. IRPDG have been introduced by Roth and Murnighan (1978) with the implementation of a continuation probability to make the end-game of PDG indeterminate and simulate infinity. This led to the development of a new theoretical framework centered on continuation probability to represent behavior in indefinite games, the Folk theorem. The calculation of the continuation probability was initially designed to support Grim Trigger and Tit-For-That strategies against defection based on payoff rate efficiency. Afterward, Kandori (1992) defines the Folk theorem to justify the cooperation of individuals who aim to maximize their returns and efficiency. At the start of the game, no one has an interest in defecting as it would reduce earnings for the rest of the game. However, cooperation remains vulnerable to the first deviation, which triggers a contagious punishment that pushes all individuals into a suboptimal equilibrium of defection. The Folk theorem has been widely discussed.

First, Ellison (1994) demonstrated its robustness and ability to support some noise without excessive patience. Then Deb, González-Díaz and Renault (2016); Deb, Sugaya and Wolitzky (2020) further generalized the model by including a set of games with random and anonymous matching, introduced the notion of "blocks" representing periods of interaction between individuals, and probabilities of state transitions across blocks. In other words, they group anonymous and unidentifiable individuals into a community. Then, each individual will be paired in a "block" of indefinitely repeated interaction. After each block, individuals will be paired with another anonymous random individual from the community. In this way, every first interaction of a block is not subject to reputation effects but only to a perceived probability of cooperation attributable to the entire community. This probability of cooperation for each first interaction can thus be seen as a representation of the group norm. To focus on the effect of social norms, we represent only the first interactions of each block. Decision-making processes and strategy choices depend on individuals' beliefs about the community's cooperation, which we consider descriptive and normative beliefs about social norms. This representation makes cooperation one of the possible equilibriums among others, and social norms the main driver of behavior. For these reasons, we use IRPDG in our first two chapters.

In the first chapter, we investigate the effects of inequality on social norms of cooperation and how norm compliance, in turn, affects cooperation strategies. For this purpose, we developed an online experiment on Amazon Mechanical Turk on a pool of 500 participants from the USA in a between-subjects protocol. Our reference framework consists of an infinitely repeated prisoner's dilemma game, where participants are randomly matched with an anonymous member of their group at each period, with an initial beliefs elicitation of the group's social norms based on the BX methodology (Bicchieri and Xiao, 2009).<sup>5</sup> Treatments differ by the presence of inequalities within the group and

---

<sup>5</sup>Participants play only one period per match, to avoid reputation effects and target the influence of the group norm.

the type of matching participants may encounter during the game. We assign a type to each participant in the game, advantageous or disadvantageous, representing inequality in the population. Three scenarios of inequality are represented in the experiment. "Equal" when the group is homogeneous, "unequal-egalitarian" and "unequal-proportional" when the group is made up of half advantageous and half disadvantageous participants. In addition, while double cooperation in the "equal" and "unequal-egalitarian" PDG removes the initial inequality between the advantaged and disadvantaged participants, this is not the case in the "unequal-proportional" PDG, where inequalities are unavoidable. Comparing the two unequal scenarios allows us to assess the pure effect of inequality on beliefs and cooperation. The results of our experiment answer questions (i) and (ii), showing that inequalities reduce personal beliefs and social norms about cooperation, and lead to the selection of less cooperative strategies. The influence of social norms on strategies is low yet present.

The second chapter extends the first one on direct interaction to investigate the construction of social norms of cooperation and how norm compliance affects repeated direct cooperation. For this purpose, we developed a laboratory experiment on a pool of 600 students from the University of Grenoble in a between-subjects protocol. Our reference framework consists of an IRPDG where participants are randomly matched with an anonymous group member at each period, with a belief elicitation of the group's social norms at each period based on the BX methodology (Bicchieri and Xiao, 2009). Treatments differ by the presence of inequalities and communication mechanisms within the group. We assign a type to each participant in the game, advantageous or disadvantageous, representing inequality in the population. Two scenarios of inequality are represented in the experiment. Equal when the group is homogeneous, and Unequal when the group is heterogeneous. Both scenarios have the same PDG structure, where double cooperation removes the initial inequality between the advantaged and disadvantaged participants. In addition, three communication mechanisms have been introduced to observe their impact on social norms. Although communication is known to enhance cooperation, the processes leading to this increase are not fully explained (Cooper et al., 1992; Dvorak and Fehrler, 2018). Introducing communication allows us to understand better how it interacts with beliefs about social norms and test the robustness of the influence of social norms on cooperative behavior in the presence of noise. We have distinguished between "No communication" between matches, "ex-ante" one-shot-free communication between matches, and "ex-post" one-shot-free communication between matches to observe the influence of each on different beliefs about social norms.<sup>6</sup> The results of our experiment answer questions (i) and (ii), showing that inequalities initially reduce social norms' beliefs in cooperation at the beginning of the game. In addition, they decline over time without communication. Whereas only descriptive beliefs influence decision-making without communication, the integration of communication reveals the influence of normative beliefs in heterogeneous matching. The latest results have enabled us to model decision-making

---

<sup>6</sup>The dissociation between ex-post and ex-ante communication is explained in the literature by their different cooperation-enhancing mechanisms: ex-ante communication reduces strategic uncertainty, while ex-post communication reduces uncertainty about the history of the game (Dvorak and Fehrler, 2018).

using descriptive and normative beliefs about social norms with a structural model. In addition to providing a representation to answer the question (iii), the structural model has enabled us to simulate different levels of belief and their impact on equilibrium. These simulations show that manipulating beliefs has great potential for achieving cooperation between all individuals and answering the question (iv).

To further explore question (iv), we investigate the use of social norms by providing information on concrete behaviors. The term Norms-Nudge (NN) was coined to refer to the use of normative and/or descriptive information about the behavior of a referent group. The first term refers to information about the components of social norms, and the second to the fact that this information is intended to influence behavior without restricting it.<sup>7</sup> NN have proven their effectiveness in improving behavior in many fields, such as energy (Schultz et al., 2007; Allcott, 2011; Ayres, Raseman and Shih, 2013), the environment (Ahn, Kim and Kim, 2020; Borg, Curtis and Lindsay, 2020) and water (Ferraro and Price, 2013; Otaki, Ueda and Sakura, 2017; Landon et al., 2018). NN implementations raise the matter of information provision. Indeed, shared knowledge of beliefs is (in principle) necessary for the effectiveness of norms. This "contingency" criteria (Bicchieri, 2005) is an important characteristic that makes NN much more effective, as many situations do not involve clear information about the local norms.

Yet the exposure of common thoughts and behaviors must be used thoughtfully, as it does not necessarily lead to the desired behaviors. Indeed, Cialdini (2003); Cialdini et al. (2006) highlight the limitation of NN, which can cause boomerang effects through two mechanisms. First, by a decrease of expectation when an individual has a misconception of the social norms, overestimating the number of appropriate behaviors. In this case, the revelation of the actual level of the social norm shows that the behavior is not being followed or accepted as believed before the NN. Second, by increasing the salience of the non-compliance with a social norm. Highlighting norm-breaking behavior reduces the perceived peer pressure and cost of non-compliance, thereby reducing the weight of the norm in the decision-making process.

In chapter four, we investigate the effects of a new type of information-based policy aimed at influencing majority behavior. For this purpose, I developed a frame-filled experiment on a pool of 171 households from the Auvergne-Rhône Alpes region of France in a between-subjects protocol over seven weeks. In the experiment, I test the effect of a new type of information based on dynamic information on household domestic water consumption compared with traditional information based on static information. Whereas static information reveals the realized behaviors and/or the social appropriateness of behaviors of a reference group at a given time, dynamic information reveals the realized behavior and/or the social appropriateness of behaviors of a reference group from a spread of time. Static information have proven effective on high consumers with few to no effect on average consumers. For this reason, I have focused my study on average consumers. I measured household water consumption every week using a picture of the water meter in each household. The treatment group received dynamic information on the reduction in

---

<sup>7</sup>In other words, without reducing the set of possible actions or implementing punishments.

water consumption of related households, while the reference group received static information on the consumption of related households. The results of my experiment show that the dynamic information led to a slightly significant reduction in water consumption (about 12.79% per household on average, i.e., 133.7 liters per week) in the short term on below-average households. In contrast, we find no effect of dynamic information on above-average households in the short or long term.

Along these first chapters, we will see that social norms are an essential driver of both direct and strategic decision-making. However, social norms' beliefs vary across time and context. The mere presence of inequality in the environment causes an initial drop in them. Furthermore, while normative and descriptive beliefs tend to decline over time, simple on-shot-free communication supports them. In addition, we show that both descriptive and normative beliefs must be considered in heterogeneous contexts. Together, the manipulation of information on social norms makes it possible to enhance pro-social and pro-environmental behavior.

To go one step further and represent the interaction of a large number of agents in an infinite environment, we developed a social-ecological system using an ecological feedback model. Ecological models are a branch of mathematics specializing in representing complex models of organisms. Their modeling enables the study of a wide range of phenomena, from observing transitions towards equilibrium to identifying critical values of each component leading to different equilibria. They can easily represent "non-expressible" concepts, the interaction of an infinite number of agents, and observe the results of their interactions over infinite time horizons. In complement, feedback models are particularly useful for studying social norms. Feedback loops enable information about the system to be updated (such as beliefs) and the consequences of individual actions on the system to be represented (such as changes in environmental conditions). Our model allows us to observe transition paths, the role and interaction of each behavioral driving force, and to simulate the impact of policy on the system. It provides more general answers to questions (iii) and (iv).

More specifically, chapter five investigates the cross-influence of personal preferences, social norms, and environmental sensitivity on the adoption of pro-environmental behavior. For this purpose, we model a human-environment feedback system with two subpopulations having to choose between a costly active pro-environmental behavior or a baseline business-as-usual behavior. Behavioral change between active and baseline is shaped by personal preferences, social interactions leading to descriptive beliefs and normative beliefs about the social norms, and private environmental experience. Each sub-population is characterized by personal preferences that maintain it in an active or basic behavior. While social norms exert conformity pressure on the majority subpopulation and behavior, the environmental state exerts anti-conformity pressure against the majority behavior. The adoption of pro-environmental behavior depends on the cross-influence between personal preferences holding subpopulation

toward their preferred behavior, social norms exerting conformity pressure toward the majority subpopulation and behavior, and the environmental state exerting anti-conformity pressure against the majority behavior. The conclusion of the model shows that the adoption of a costly, actively pro-environmental behavior fundamentally depends on whether the pro-environmental subpopulation represents a minority or makes up the majority of the population. We show that a strong pro-environmental personal preference may not be sufficient to allow a minority of the population to express their preferences. The weight of social norms must be low but can counteracted by the weight of the environment. Furthermore, when the population is composed of a majority of pro-environmental subpopulations, this may not be sufficient to generalize its behavior to other subpopulations. A combination of descriptive and normative beliefs about social norms is needed to ensure the expression of active behavior in both subpopulations. Nevertheless, it is still possible to spread the active behavior without reaching a majority of the pro-environmental subpopulation by manipulating the descriptive and normative part of the social norms to focus on the pro-environmental subpopulation.

Together, our results underline the importance of social norms for pro-social and pro-environmental behavior choices. Their use in the context of climate policy can help to achieve tipping and transition points toward the generalization of ecological behavior and, thus, climate change mitigation. For this purpose, our model can be used to simulate the impact of different policies to select the most suitable according to the situation, as the variation of our parameters enables us to represent a wide range of situations and initial conditions.



## Chapter 2

# Inequality, Social Norms and Cooperation: Strategy Choice in the Infinitely Repeated Prisoners' Dilemma

*Co-authored with Sabrina Teyssier*

**Abstract** Our societies currently face important challenges of cooperation, which is needed among individuals who interact at a non-regular frequency. In this context, cooperation can be sustained if the social norm pushes in this direction. We design an online experiment where participants make strategic choices in an infinitely repeated prisoner's dilemma. We examine the effects of inequality on social norms of cooperation and how norm compliance, in turn, affects cooperation. Inequality exists when the two participants defect and cooperation gives equal payoffs in one treatment or keeps the unequal payoffs in the other. The results show that inequality weakens the social norm by limiting first- and second-order normative beliefs of cooperation, as well as descriptive beliefs about other participants' cooperation. Inequality reduces the likelihood of cooperation, mainly driven by changes in social norms. Overall, the mere existence of inequality causes these changes instead of specific behaviors depending on the participants' type.

### 2.1 Introduction

The role of inequality in power, status, income, or wealth on long-term economic efficiency has often been discussed (Stiglitz, 2015). Recent evidence supports that inequality leads to an inefficient allocation of resources and lower investment and innovation that undermine economic growth (Aghion and Williamson, 1998; World Bank,



2005; , 2013; Ostry, Berg and Tsangarides, 2014; Piketty, 2013). Furthermore, inequality is suggested as leading to the erosion of social cohesion, which undermines social norms of cooperation in the long run (Putnam, 2000; United Nations Development Programme, 2013). Cooperation in infinitely repeated games may depend on such social norms. Social norms are informal rules of behavior in groups and societies that individuals conform to if they believe that most people conform to it and also believe that most people believe that people ought to conform to it (Bicchieri, 2005). On the one hand, inequality may alter such beliefs and undermine the existence of the norm (Xiao and Bicchieri, 2010). On the other hand, a weakened norm may render sustaining cooperation more difficult (Ostrom, 2009; Bicchieri, 2016).

In daily life, people interact with many others at a non-regular frequency and nevertheless contribute to the same shared common good. This is, for instance, the case for inhabitants of a neighborhood who contribute to the well-being of the neighborhood or colleagues from the same firm or department who contribute to the global profit by contributing to various smaller projects, or more generally, the citizens of a country who contribute to the country revenues and well-being. In such context, cooperation can be sustained if there is a social norm that is to cooperate, pushing all actors to contribute to the shared common good (Fehr and Schurtenberger, 2018).<sup>1</sup> However, when actors are not all equal, the relevance of the social norm is challenged, and with it, the level of cooperation that can be expected.

The economic literature on finite cooperation games shows that inequality negatively impacts cooperation (Sheposh and Gallo, 1973; Zelmer, 2003; Cherry, Kroll and Shogren, 2005; Buckley and Croson, 2006; Ahn et al., 2007; Beckenkamp, Hennig-Schmidt and Maier-Rigaud, 2007; Anderson, Mellor and Milyo, 2008; Fischbacher, Schudy and Teyssier, 2014).<sup>2</sup> The efficiency of instruments such as punishment and communication is also reduced in the presence of inequality (Nikiforakis, Noussair and Wilkening, 2012; Reuben and Riedl, 2013; Gangadharan, Nikiforakis and Villeval, 2017; Koch, Nikiforakis and Noussair, 2021). The role of inequality in long-term cooperation has been rather under-investigated. Whereas it has been shown that long-term cooperation can be sustained in infinitely repeated games with stranger matching in case of repeated interactions of the same pair (Duffy and Ochs, 2009; Dal Bó and Fréchet, 2011, 2018) as well as in anonymous settings (Camera and Casari, 2009; Camera, Casari and Bigoni, 2012), only Camera, Deck and Porter (2020) study the role of inequality in this context. Camera, Deck and Porter (2020) find that inequality undermines efficient cooperation in donor–recipient pairs.

In this paper, we empirically examine the effects of inequality on social norms of cooperation and how norm compliance, in turn, affects cooperation in infinitely repeated games. We conducted an online experiment where

---

<sup>1</sup>Using the elicitation method of the social norms as appropriate behaviors introduced by (Krupka and Weber, 2013), it has been found that social norms explain an extensive series of phenomena such as reciprocity (Gächter, Nosenzo and Sefton, 2013; Nikiforakis, Oechssler and Shah, 2014), fair sharing (Gächter, Gerhards and Nosenzo, 2017), promise-keeping (Krupka, Leider and Jiang, 2017), lying (d’Adda et al., 2017), ethical conduct of financial advisers (Burks and Krupka, 2012), corruption (Banerjee, 2016a), discrimination (Barr, Lane and Nosenzo, 2018), and gendered occupational choices (Gangadharan et al., 2016). See also Fallucchi and Nosenzo (2022) for a discussion of the robustness of the Krupka-Weber elicitation method of social norms when other points are made salient for the coordination of the group. They find the method robust, particularly when beliefs about the appropriate behavior are clear. More recently, (d’Adda et al., 2020) explains all behaviors in a dictator game with personal values and social norms perception.

<sup>2</sup>Chan et al. (1996) and Visser and Burns (2015) are rare evidence showing that inequality increases cooperation.

participants made strategy choices in an infinitely repeated prisoner's dilemma game. A fairly large group of persons made choices for repeated interactions in pairs with a stranger-matching protocol, which reflects interactions of people in societies as a succession of interactions in small groups without individual reputation effects. Theoretically, based on the folk theorem, cooperation can be sustained over time with anonymous random matching (Kandori, 1992; Ellison, 1994; Deb, González-Díaz and Renault, 2016; Deb, Sugaya and Wolitzky, 2020). Two participants cooperating or two participants defecting are both subgame-perfect Nash equilibria. The norm perception could be a method of equilibrium selection in this context (Burke and Young, 2011).<sup>3</sup> We elicited norms by asking for the participants' first- and second-order normative expectations as well as their descriptive expectations (Bicchieri, 2005, 2016).

Achieving cooperation generates additional benefits to be shared between the participants. An equal share of these benefits is evident when the participants have the same amount available to invest in cooperation. When they are instead unequal, the distribution of the benefits is questionable: the benefits may either be equally shared or distributed proportionally according to their available investment. These two distribution rules imply different motives. Therefore, we compare settings that differ by inequality when the two participants defect or cooperate. In two treatments, the participants are equal; all are advantaged or disadvantaged. In two other treatments, the participants are unequal, with an equal share of advantaged and disadvantaged participants. In one treatment, the benefits from cooperation are equally shared between the participants, whereas in the other treatment, the benefits are proportionally distributed. The comparisons between treatments allow for the identification of the impact of inequalities on norms of cooperation and strategies in the infinitely repeated prisoner's dilemma game.

The results show that the large majority of the participants believe that the decision that should be chosen and that will be chosen is cooperation. Inequality weakens the social norm by decreasing these expectations of cooperation. For both unequal treatments, the mere presence of inequality changes first- and second-order normative beliefs, as well as descriptive beliefs, whatever the type of the participant and the type of his expected playmate. In turn, the social norm impacts the decision to cooperate with higher beliefs, leading to a higher likelihood of choosing cooperation. The strategy to always cooperate in the long run is less chosen, while the always defect strategy is more chosen in unequal treatments than equal treatments. Interestingly, the type of the participant (i.e., advantaged or disadvantaged) does not affect choices. The mere existence of inequality causes the changes instead of specific behaviors depending on the participants' type.

Our contribution to the literature is twofold. First, we study the impact of inequality on long-term cooperation in a controlled framework. On the one hand, in most previous studies, the introduction of inequality distorts the trade-off between equality and efficiency, introduces normative conflict, changes incentives to cooperate, and biases the evaluation of the role of inequalities (Gangadharan, Nikiforakis and Villeval, 2017). Our study aims to identify the

---

<sup>3</sup>See also Bicchieri (2005) and Cialdini, Reno and Kallgren (1990) for a discussion on the role of norms to select non-equilibrium behavior as part of Bayesian strategies.

pure effect of inequality on cooperation, prohibiting changes in trade-offs between equality and efficiency. On the other hand, inequalities and cooperation have mainly been studied in finite games, while long-term cooperation strategies can only be studied in infinite games. Our study aims to fill this gap. Second, we directly elicit social norms perception with normative and descriptive beliefs. Our study aims to explain whether inequality changes the perception of the social norm and whether strategies are affected by social norm perception.

The rest of the paper is organized as follows. Section 2.2 presents the experiment. Section 2.3 shows the results. Section 2.4 discusses the results and concludes.

## 2.2 The experiment

Participants in the experiment were in groups of 50 individuals. They had to make strategy choices in a repeated game where they are randomly matched in pairs at each period. The number of interactions is finite but uncertain, which makes the game similar to an infinitely repeated game.<sup>4</sup> At the end of each period, a random draw decides whether a new period starts with a continuation probability equal to 0.95. Participants could not identify the other player in their pair. In this section, we detail the game, the different treatments that the participants play in a between-subjects design, the elicitation of the social norm, the elicitation of decisions and strategies, as well as the experimental procedures.

### 2.2.1 The game

Participants play an infinitely repeated prisoner's dilemma game with two possible actions: cooperate (C) or defect (D).

**Payoffs** Gains of player  $i$  are calculated based on his voluntary contributions,  $g_i$ , and on his playmate  $j$ 's,  $g_j$ , to a public good with a return of  $a = 1.6$ . Individual contributions are supposed to be either 0 (D) or the player's entire endowment (C), i.e.,  $g_i \in \{0, E_i\}$  and  $g_j \in \{0, E_j\}$ . Cooperation between the two players generates benefits. The distribution of these benefits can be either egalitarian, i.e., player  $i$ 's gains are  $\Pi_i = E_i - g_i + 0.8(g_i + g_j)$ , or proportional to the players' endowments, i.e., player  $i$ 's gains are  $\Pi_i = E_i - g_i + 0.8(g_i + \frac{E_i}{E_j}g_j)$ . In these two settings, cooperation does not increase relative inequality, which avoids any normative conflict between efficiency and equality (Gangadharan, Nikiforakis and Villeval, 2017).

The payoffs depend on the two players' actions. When the distribution of the benefits of cooperation is egalitarian, payoffs are as follows.

---

<sup>4</sup>We use the random termination period, firstly introduced by Roth and Murnighan (1978). See Dal Bó and Fréchette (2018) for a survey of experiments using infinitely repeated games to study cooperation in this context and a discussion of the methods used to induce infinitely repeated games in the laboratory.

		Player $j$	
		Cooperate	Defect
Player $i$	Cooperate	$0.8(E_i + E_j) ; 0.8(E_i + E_j)$	$0.8E_i ; E_j + 0.8E_i$
	Defect	$E_i + 0.8E_j ; 0.8E_j$	$E_i ; E_j$

Table 2.1: Egalitarian distribution of the benefits of cooperation

When the distribution of the benefits of cooperation is proportional, payoffs are as follows.

		Player $j$	
		Cooperate	Defect
Player $i$	Cooperate	$1.6E_i ; 1.6E_j$	$0.8E_i ; 1.8E_j$
	Defect	$1.8E_i ; 0.8E_j$	$E_i ; E_j$

Table 2.2: Proportional distribution of the benefits of cooperation

When endowments are equal, the payoffs are the same for the egalitarian and proportional distribution. However, when endowments are unequal, the two types of distribution lead to different payoffs.

**Equilibria** Two subgame perfect equilibria exist for a sufficiently high probability of continuation,  $\delta$ , when the game is infinitely repeated. On the one hand, players defecting in all periods is an equilibrium because their individual interest drives them to best respond to defection by choosing defection as well. On the other hand, for sufficiently high  $\delta$ , if the player assumes that his playmate is adopting the grim trigger strategy, i.e., the cooperative strategy providing the strongest punishment when observing defection, he best responds by playing the grim trigger strategy as well (Kandori, 1992; Ellison, 1994; Camera and Casari, 2009; Camera, Casari and Bigoni, 2012; Dal Bó and Fréchette, 2018). The threshold of the probability of continuation that makes cooperation an equilibrium action,  $\delta^{SPE}$ , is identical for all players with equal endowments or unequal endowments with a proportional distribution of the benefits of cooperation. However,  $\delta^{SPE}$  differs depending on the relative endowments of the players in the egalitarian distribution:  $\delta^{SPE}$  is higher for advantaged players and lower for disadvantaged players.

The choice of cooperation depends not only on whether cooperation is an equilibrium action (Dal Bó and Fréchette, 2018). Indeed, a player may worry about his low payoff when cooperating while his playmate chooses to defect, which is not included in the  $\delta^{SPE}$  calculation. Assuming the always defect strategy and a cooperative strategy such as grim trigger in the infinitely repeated prisoner's dilemma game, Blonski and Spagnolo (2015) and Dal Bó and Fréchette (2018) define cooperation as risk dominant in the sense of Harsanyi and Selten (1988) if the grim trigger strategy is risk dominant, i.e. the best response to the strategy of the other player that is to randomize with equal probability between always defect and grim trigger. Cooperation is part of a risk-dominant equilibrium if the player's

payoff when he chooses the grim trigger strategy is higher than when he chooses the always defect strategy, which is the case for a sufficiently high continuation probability. The threshold for cooperation to be part of a risk-dominant equilibrium,  $\delta^{RD}$ , is identical for all players when endowments are equal or unequal with a proportional distribution of the benefits of cooperation. However, as for  $\delta^{SPE}$ ,  $\delta^{RD}$  differs depending on the relative endowments of the players in the egalitarian distribution:  $\delta^{RD}$  is higher for advantaged players and lower for disadvantaged players. Details of calculations are provided in Appendix A.1.

The parameters in the experiment verify  $\delta > \delta^{SPE}$  and  $\delta > \delta^{RD}$  to have both defects and cooperate as equilibrium actions.

## 2.2.2 Experimental treatments

We conducted four treatments: two equal treatments where all 50 players have the same endowment, either high or low, and two treatments where half of the 50 players have a high endowment, and the other half have a low endowment. The treatments are between subjects, meaning each participant takes part in only one treatment.

### Equal treatments

In the Equal treatments, we assume  $E_i = E_j$ . We consider two levels of endowment: the two players in the pair are either disadvantaged, i.e.,  $E_i = E_j = E_d = 10$  (*Equal-D treatment*), or advantaged, i.e.,  $E_i = E_j = E_a = 20$  (*Equal-A treatment*).

**Equal-D** The 50 participants have the low endowment  $E_d = 10$ . Applying the players' gains defined in the previous section, the payoff matrix is as follows.

		Player $j$	
		Cooperate	Defect
Player $i$	Cooperate	16 ; 16	8 ; 18
	Defect	18 ; 8	10 ; 10

Table 2.3: Payoffs in the Equal-D treatment

**Equal-A** The 50 participants have the high endowment  $E_a = 20$ . The payoff matrix is as follows.

		Player $j$	
		Cooperate	Defect
Player $i$	Cooperate	32 ; 32	16 ; 36
	Defect	36 ; 16	20 ; 20

Table 2.4: Payoffs in the Equal-A treatment

We calculate a range of  $\delta^{SPE}$  and  $\delta^{RD}$  with the lower bound corresponding to a punishment for defection directly at the next period and the upper bound corresponding to an equal probability to face cooperation at each period for 25 periods. <sup>5</sup>  $\delta^{SPE}$  ranges from 0.25 to 0.866 and  $\delta^{RD}$  ranges from 0.4 to 0.88. These two thresholds are identical in the two treatments and are lower than the continuation rate of 0.95 that applies in the experiment. To defect and to cooperate are thus two equilibrium actions in the equal treatments.

### Unequal treatments

Inequality is introduced assuming 25 players are disadvantaged with endowment  $E_d = 10$  and 25 other players are advantaged with endowment  $E_a = 20$ . The pair matching can be among players with the same or unequal endowments. If the endowments of the players in the pair are equal, the payoff matrices are the same as in the equal treatments. We present below the payoff matrices if endowments are unequal.

**Unequal-Egalitarian** When player  $i$  in the pair is advantaged, i.e. he receives  $E_a = 20$ , and player  $j$  is disadvantaged, i.e. he receives  $E_d = 10$ , the payoffs matrix is as follows.

		Player $j$	
		Cooperate	Defect
Player $i$	Cooperate	24 ; 24	16 ; 26
	Defect	28 ; 8	20 ; 10

Table 2.5: Payoffs in the Unequal-E treatment

<sup>5</sup>The upper bound is calculated under the assumption that gives important incentives to the player to defect.

**Unequal-Proportional** When player  $i$  in the pair is advantaged, i.e.  $E_a = 20$ , and player  $j$  is disadvantaged, i.e.  $E_d = 10$ .

		Player $j$	
		Cooperate	Defect
Player $i$	Cooperate	32 ; 16	16 ; 18
	Defect	36 ; 8	20 ; 10

Table 2.6: Payoffs in the Unequal-U treatment

The range of  $\delta^{SPE}$  and  $\delta^{RD}$  are identical in the Unequal-Proportional treatment, and the equal treatments:  $\delta^{SPE}$  ranges from 0.25 to 0.866 and  $\delta^{RD}$  ranges from 0.4 to 0.88. In the Unequal-Egalitarian treatment, these thresholds are changed when, in the pair, one player is advantaged and the other is disadvantaged.  $\delta^{SPE}$  ranges from 0.125 to 0.759, and  $\delta^{RD}$  ranges from 0.222 to 0.792 for disadvantaged players, whereas these thresholds are higher for advantaged players with  $\delta^{SPE}$  ranging from 0.5 to 0.942 and  $\delta^{RD}$  from 0.667 to 0.947. In all cases,  $\delta^{SPE}$  and  $\delta^{RD}$  are lower than the continuation rate of 0.95 used in the experiment. However, Dal Bó and Fréchette (2018) emphasize that the distance to these thresholds matters in the equilibrium selection. Advantaged players are, therefore, less likely to cooperate than disadvantaged players when the players in the pair have different endowments.

### 2.2.3 Procedures

The online sessions were conducted during 2021. In total, 500 US-located participants were recruited through Amazon Mechanical Turk to participate in the experiment. 35.6% were women, 21% were less than 30 years old, 56% between 30 and 45, 19.8% between 45 and 60 and 3.2% above 60, 19.4% had a degree lower than a bachelor, 61.4% had a bachelor's degree, and 19.2% had a master degree. In the three-item IQ test (Oechssler, Roider and Schmitz, 2009), 38.8% of the participants gave a correct answer to the three questions, 16.6% to two questions, 16.4% to one question and 28.2% gave no correct answer.<sup>6</sup> Participants also had to answer a question about trust toward other people ("Generally speaking, would you say that most people are trustworthy or that you can never be too careful with people?"), and 56.4% indicated that "most people are trustworthy". The distribution of these variables is not different between the treatments (ranksum Mann-Whitney tests:  $p > 0.1$ ). In the econometric analysis, we controlled for the individual characteristics aforementioned.

Each participant took part in a single treatment: 50 participants in Equal-D, 50 in Equal-A, 200 in Unequal-Egalitarian, and 200 in Unequal-Proportional. In each unequal treatment, 50 participants received endowment  $E_d$  and were matched with participants with endowment  $E_d$  as well, 50 participants received endowment  $E_a$  and were matched with participants with endowment  $E_a$  as well, 50 participants received endowment  $E_d$  and were

<sup>6</sup>The questions are given in the instructions in Appendix A.2.

matched with participants with endowment  $E_a$ , and 50 participants received endowment  $E_d$  and were matched with participants with endowment  $E_a$ .

Instructions were formulated neutrally (see Appendix A.2). The experiment was performed with oTree (Chen, Schonger and Wickens (2016)). The experiment lasted for about 5 minutes. We formed groups of 50 participants once all participants from treatment had completed the experiment to implement their strategic decisions and thus calculate the payoffs. We randomly drew the number of periods for each group using the continuation rate of 0.95. The participants earned additional earnings based on the elicitation of their beliefs. The ex-post implementation was done in Python with Jupyter. The average earnings were about \$5.

The participants started by answering questions about their beliefs regarding the decision that should or would be made and then reported their decisions and strategies for the game.

### **Social norm elicitation**

The social norm consists of three dimensions: first-order normative beliefs, second-order normative beliefs, and descriptive beliefs (Bicchieri, 2005). The first-order normative beliefs were evaluated asking the participants what decision, in their opinion, participants should make. The second-order normative beliefs were measured by asking the participants what share of the participants of their group they think would indicate that participants should choose to cooperate (paths of 10%). The descriptive beliefs were measured by asking the participants what share of the participants of their group they think would choose to cooperate in period 1.<sup>7</sup> In unequal treatments, the number of questions for each dimension is multiplied depending on the type of the participants. At this social norm elicitation stage, subjects only know the framework of the game and their own type.

### **Decisions and strategy elicitation**

The participants made their strategy choices in the prisoner's dilemma game, and the matching was done ex-post following Dal Bó and Fréchette (2019) who validated the method with dynamic experimental data.<sup>8</sup> This procedure allows to directly address strategy choices instead of simulating strategies from observed actions (Camera, Casari and Bigoni, 2012; Dal Bó and Fréchette, 2011; Engle-Warnick, McCausland and Miller, 2004; Engle-Warnick and Slonim, 2006).

The participants learned the type of their future playmates in the experiment after answering questions about their beliefs but before choosing their behavior and strategy. In period 1, the participant had to choose between cooperating and defecting. In contrast, in period 2, he had to choose conditionally on the decision of his previous playmate in period 1, i.e., cooperate or defect if the previous playmate had chosen to cooperate in period 1 and cooperate or defect if the playmate had chosen to defect in period 1.

---

<sup>7</sup>See Bicchieri and Xiao (2009) for the first use of this elicitation method.

<sup>8</sup>Another way to elicit strategies in the indefinitely repeated prisoner's dilemma game is provided in Romero and Rosokha (2018).



After period 2, strategies were elicited. First, we elicited the *memory-one strategy* that corresponds to choosing to cooperate or defect after the choices of the two playmates in the previous period, i.e., cooperate or defect if, in the previous period, the participant chose to cooperate or defect and his previous playmate chose to cooperate or defect. Second, we elicited more complex strategies that participants chose among a *menu of strategies*. From this menu, we identify the following main strategies: always cooperate, always defect, tit-for-tat, and grim-trigger (see Appendix A.3 that indicates how strategies are grouped). Always cooperate, tit-for-tat, and grim trigger are cooperation strategies, while always defect is a defection strategy. One of these two strategies elicitation method was randomly selected for implementation for the ex-post computation of payoffs.

## 2.2.4 Theoretical predictions

In any prisoner's dilemma game, the action to defect is an equilibrium. Additionally, the action to cooperate is an equilibrium and part of a risk-dominant equilibrium for both disadvantaged and advantaged players in all equal and unequal treatments. Strategies, then, should not be different between the treatments. Nevertheless, the perceived social norm,  $N$ , may help the players to select one of the two equilibria (Burke and Young, 2011). We suppose  $N \in \{0, 1\}$ , with  $N = 0$  if the social norm is to defect and  $N = 1$  if it is to cooperate. Following d'Adda et al. (2020), the social norm is expressed as  $N = r + \alpha(E(r) - r) + \beta(E(g) - r)$  with  $r \in \{0, 1\}$  the player's first order normative beliefs,  $E(r) \in \{0, 0.1, \dots, 1\}$  his second-order normative beliefs and,  $E(g) \in \{0, 0.1, \dots, 1\}$  his descriptive beliefs. The coefficients  $\alpha$  and  $\beta$  are the weights the player attributes to his second-order normative beliefs and descriptive beliefs, respectively, with  $\alpha, \beta > 0$  and  $\alpha + \beta < 1$ . Previous experimental work suggests that inequality decreases the social norm of cooperation (Reuben and Riedl, 2013; Fischbacher, Schudy and Teyssier, 2014; Xiao and Bicchieri, 2010). The first hypothesis we test is the following:

**Hypothesis 1:** The social norm, based on first-order and second-order normative beliefs and descriptive beliefs, is lower in the unequal treatments than in the equal treatments.

If we assume that the choice of the strategy depends on the social norm, the second hypothesis we test is:

**Hypothesis 2:** The likelihood of choosing a cooperative strategy increases with the social norm of cooperation.

If the social norm is lower in the unequal treatments than in the equal treatments, the cooperative strategy would be less likely to be chosen in the unequal treatments than in the equal treatments. Also, the weight the player attributes to the social norm may differ between the unequal and equal treatments. If it is lower in unequal treatments than in equal treatments, the cooperative strategy would be even less chosen in the former than in the latter.

Although the thresholds of the continuation rate are lower than 0.95 in any situation of the experiment, we may observe different decisions in the Unequal-Egalitarian treatment because the levels of  $\delta$  thresholds are different than in Equal-D, Equal-A, or Unequal-Proportional:  $\delta$  thresholds are lower (higher) for disadvantaged (advan-

tagged) players. The  $\delta$  thresholds are identical in Unequal-Proportional, Equal-D, and Equal-A. Differences between Unequal-Proportional and equal treatments would reflect the pure effect of inequality, keeping identical incentives for cooperation.

## 2.3 Results

In this section, we will answer two questions: (i) How does inequality change the participants' beliefs and the social norm? (ii) How do inequality and changes in beliefs influence the participant's strategy choices?

### 2.3.1 Social norms

First-order normative beliefs represent the participant's personal value regarding the action he thinks should be made in the game. This value is the individual reference of the social norm (d'Adda et al., 2020). In the equal treatments, 78% of the participants think that the decision that should be made is to cooperate, whereas this frequency decreases to about 63% in the Unequal-Egalitarian treatment and 58% in the Unequal-Proportional treatment (see Table A.2 in Appendix for detailed statistics by type).

Second-order normative beliefs represent the participant's beliefs about the other participants' opinions regarding the action that should be taken. These beliefs range between 0 when the participant believes that 0% of the other participants think that the decision to cooperate should be made and 1 when he believes 100% of the other participants think that the decision to cooperate should be made. Steps are of 0.1. On average, in the equal treatments, the participants believe that 74% of the participants think that the decision to cooperate should be made. This share decreases to 63% and 60% in the Unequal-Egalitarian and Unequal-Proportional treatments, respectively (see Tables A.3 and A.4 in appendix). The distribution of second-order beliefs is represented in Figure 2.1.

Descriptive beliefs provide the participant's beliefs about other participants' decisions. These beliefs range between 0 when the participant believes that 0% of the other participants will choose to cooperate and 1 when they believe 100% of the other participants will cooperate. Steps are of 0.1. In the equal treatments, participants believe that, on average, 66% of the participants will decide to cooperate instead of defect. In the unequal treatments, this share is about 63% in the Unequal-Egalitarian treatment and 57% in the Unequal-Proportional treatment (see Table A.5 in appendix). The distribution of descriptive beliefs is represented in figure 2.2.

We now test whether these beliefs differ between the treatments. We account for the type of the participant, as well as for the fact that the participant gives his beliefs regarding the same type as herself (homogeneous) or the other type (heterogeneous). Table 2.7 gives the results.

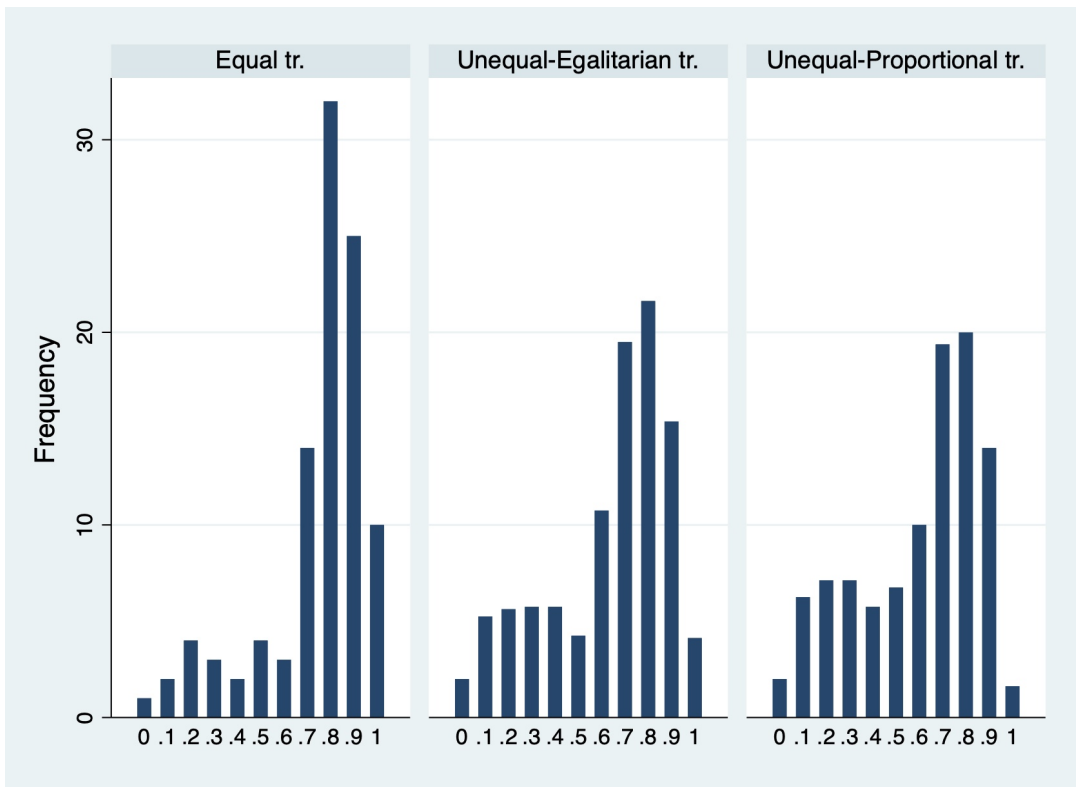


Figure 2.1: Distribution of second-order normative beliefs

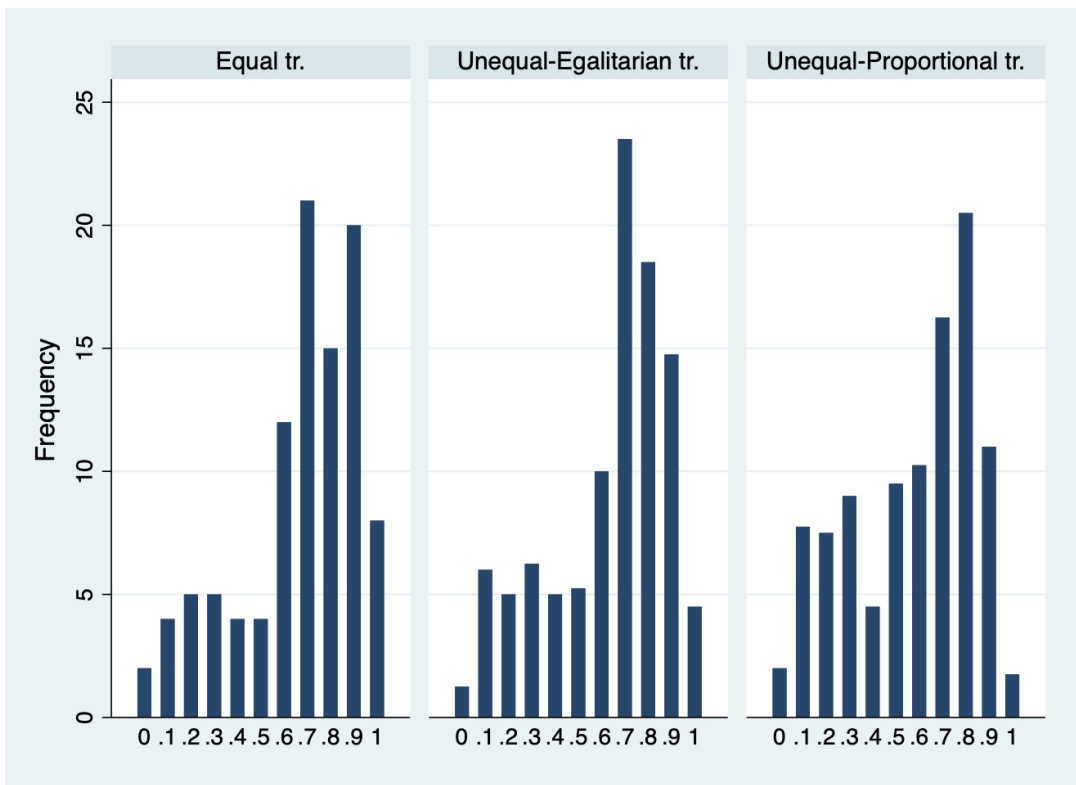


Figure 2.2: Distribution of descriptive beliefs

	First-order normative beliefs		Second-order normative beliefs		Descriptive beliefs	
	(1)	(2)	(3)	(4)	(5)	(6)
UE	-0.326*** (0.046)	-0.338*** (0.069)	-0.525*** (0.021)	-0.492*** (0.028)	-0.305*** (0.024)	-0.302*** (0.033)
UP	-0.371*** (0.044)	-0.285*** (0.070)	-0.563*** (0.025)	-0.501*** (0.035)	-0.368*** (0.023)	-0.351*** (0.033)
UE × Heterogeneous		-0.018 (0.042)		-0.010 (0.015)		-0.016 (0.018)
UP × Heterogeneous		-0.159*** (0.044)		-0.051*** (0.017)		-0.027 (0.024)
Advantaged	0.025 (0.028)	-0.002 (0.085)	-0.013 (0.018)	0.041 (0.024)	-0.003 (0.020)	-0.004 (0.036)
UE × Advantaged		0.055 (0.095)		-0.054 (0.039)		0.010 (0.049)
UP × Advantaged		0.006 (0.094)		-0.071* (0.036)		-0.007 (0.047)
Intercept			1.231*** (0.036)	1.203*** (0.036)	1.000*** (0.040)	1.000*** (0.040)
<i>N</i>	1000	1000	2000	2000	1000	1000
Clusters	500	500	500	500	500	500
pseudo <i>R</i> <sup>2</sup>	0.070	0.082	0.480	0.486	0.272	0.274

Standard errors in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Models (1) and (2) are logit models, reporting average marginal effects; Models (3) to (6) are Tobit model left-censored at 0 and right-censored at 1.

Table 2.7: Beliefs, by treatment

The social norm of cooperation is weakened in the two unequal treatments with a decrease of first- and second-order normative beliefs, as well as descriptive beliefs. The type of the participant has no significant impact on his beliefs, whatever the treatment. Interestingly, normative beliefs are significantly lower when the participants are asked about their beliefs about the action that should be made for other participants of the other type. This gives result 1.

**Result 1:** The existence of inequality weakens the social norm of cooperation by a decrease in first- and second-order normative beliefs and descriptive beliefs, whatever the type of the participants.

When the benefits of cooperation are proportionally shared between the participants, incentives to cooperate are kept exactly constant when inequality is introduced as compared to equality in endowments. Changes in beliefs in this game are then uniquely driven by inequality in endowments. The presence of inequality changes participants' beliefs that are not related to changes in incentives to cooperate. Compliance with the social norm of cooperation is lower in the presence of inequality.

## 2.3.2 Decisions and strategies

Inequalities weaken the social norm of cooperation. We now analyze whether inequalities or changes in the social norm affect decisions and strategies of cooperation.

### Decisions in periods 1 and 2

In period 1, 74% of the participants decide to cooperate in the equal treatments. This share is 67% and 65% in the unequal treatments, Unequal-Egalitarian treatment, and Unequal-Proportional treatment, respectively (see Table A.6 in Appendix for detailed statistics). In period 2, in equal treatments, 70% of the participants cooperated if their playmates cooperated in period 1, and 63% cooperated if their playmates defected. These rates are 62% and 56% in the Unequal-Egalitarian treatment and 59% and 50% in the Unequal-Proportional treatment (see Table A.7 in Appendix for detailed statistics).

Table 2.8 presents the impact of the treatments and the social norm on the decisions in periods 1 and 2. In period 2, the decision of the playmate in the previous period, as well as the participant's decision in period 1, are also estimated.<sup>9</sup>

In period 1, as well as in period 2, inequality decreases the likelihood of cooperation through a change in beliefs: first-order normative beliefs and descriptive beliefs positively impact the decision to cooperate. However, second-order normative beliefs have no significant impact. In period 2, the decision to cooperate also strongly depends on past behaviors from his previous playmate and himself. The existence of inequality decreases the likelihood of cooperation only through a change in beliefs. Besides, the intensity of the effect of the social norm on the decisions is not related to inequality.

---

<sup>9</sup>We also conducted regressions with crossed effects between heterogeneity and the unequal treatments and between the type of the participant and the treatments. We find no significant effect, so we do not report the coefficients in the table.

	Decision in period 1		Decision in period 2		
	(1)	(2)	(3)	(4)	(5)
UE	-0.078 (0.057)	-0.041 (0.055)	-0.079 (0.049)	-0.059 (0.049)	-0.048 (0.048)
UP	-0.090 (0.057)	-0.035 (0.054)	-0.120** (0.048)	-0.084* (0.048)	-0.076 (0.048)
Advantaged	0.022 (0.041)	0.023 (0.038)	0.014 (0.036)	0.012 (0.034)	0.006 (0.032)
$r$		0.617*** (0.081)		0.400*** (0.079)	0.204*** (0.075)
$E(r) - r$		0.138 (0.107)		-0.012 (0.106)	-0.055 (0.100)
$E(g) - r$		0.370*** (0.094)		0.314*** (0.093)	0.198** (0.088)
Playmate cooperated in period 1			0.072*** (0.025)	0.072*** (0.025)	0.072*** (0.025)
Participant cooperated in period 1					0.267*** (0.031)
$N$	500	500	1000	1000	1000
Clusters	500	500	500	500	500
pseudo $R^2$	0.045	0.143	0.024	0.061	0.113

Standard errors in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Logit models, reporting average marginal effects.

Table 2.8: Decisions in periods 1 and 2

## Strategies

Memory-one strategies imply four decisions from the participants based on choices in the previous period: whether the participant and his playmate chose to cooperate or to defect. In equal treatments, when the participants have chosen to cooperate in the previous period, 88% of the participants cooperate if the previous playmate cooperated, and 46% cooperate if the previous playmate defected. When the participant has chosen to defect in the previous period, 70% of the participants cooperate if the playmate also cooperated, and 50% cooperate if the playmate defected. These shares are 81%, 43%, 57%, and 43% in the unequal-egalitarian treatment and 79%, 49%, 50% and 43% in the unequal-proportional treatment (see Table A.8 in Appendix).

Table 2.9 presents the marginal effects of the estimation of the participant's likelihood to cooperate depending on the hypothetical decision of herself and his playmate in the previous period. Model (1) does not include normative and descriptive expectations, whereas Model (2) does.

	Model (1)	Model (2)
UE	-0.075** (0.032)	-0.053* (0.031)
UP	-0.080** (0.032)	-0.047 (0.031)
Advantaged	-0.001 (0.024)	-0.005 (0.023)
$r$		0.335*** (0.060)
$E(r) - r$		0.042 (0.063)
$E(g) - r$		0.196*** (0.055)
Participant cooperated in previous period	0.011 (0.024)	0.011 (0.024)
Playmate cooperated in previous period	0.111*** (0.026)	0.111*** (0.026)
Participant cooperated $\times$ Playmate cooperated in previous period	0.259*** (0.0364)	0.259*** (0.0364)
$N$	2000	2000
Clusters	500	500
pseudo $R^2$	0.079	0.104

Standard errors in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .  
Logit models, reporting average marginal effects.

Table 2.9: Memory-one strategies

We find that the likelihood of choosing to cooperate is significantly lower when there is inequality. The previous choice to cooperate with the playmate, whatever the participant did or did not cooperate, increases this likelihood. The increase is stronger when the previous choice of the participant was also to cooperate. Interestingly, additional regressions show that the impact of past cooperation by the previous playmate is significantly lower in the unequal proportional treatment, i.e., observing cooperation from others plays a lower role in the unequal proportional treatment compared to the equal treatments. When beliefs related to the social norm are introduced, dummies for unequal treatments are not significant anymore, while first-order normative beliefs and descriptive beliefs significantly explain the likelihood of cooperating. The effect of beliefs is not different between the treatments. Result 2 is as follows.

**Result 2.** The existence of inequality decreases the likelihood of the decision to cooperate because of a change in the social norm, whatever the type of the participant.

Four main strategies are elicited from the menu of strategy: always cooperate, always defect, Grim-trigger, and Tit-for-Tat. On average, in the equal treatments, always cooperate is chosen by 51% of the participants, always defect by 11%, grim-trigger by 23%, and tit-for-tat by 15%. In the unequal treatments, these shares are 34%, 24%, 23%, and 20% in the Unequal-egalitarian treatment and 33%, 24%, 21%, and 23% in the Unequal-proportional treatment, respectively.

The results show a lower share of participants who choose to always cooperate and a higher share who choose to always defect in the unequal treatments than in the equal treatments. Table 2.10 gives marginal effects of multinomial logit models that compare the effects of treatments, type of the participant, beliefs, and previous behavior on the participant's likelihood to adopt each strategy.

	Always defect			Always cooperate		
	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)
UE	0.16*** (0.029)	0.14*** (0.028)	0.13*** (0.027)	-0.16*** (0.028)	-0.13*** (0.028)	-0.12*** (0.027)
UP	0.14*** (0.029)	0.12*** (0.028)	0.11*** (0.027)	-0.17*** (0.028)	-0.12*** (0.028)	-0.12*** (0.027)
Advantaged	-0.029 (0.018)	-0.025 (0.017)	-0.023 (0.017)	0.036* (0.021)	0.027 (0.021)	0.023 (0.021)
$r$		-0.31*** (0.051)	-0.16** (0.053)		0.43*** (0.51)	0.31*** (0.53)
$E(r) - r$		0.049 (0.047)	0.086* (0.045)		0.14** (0.63)	0.12* (0.62)
$E(g) - r$		-0.24*** (0.055)	-0.15*** (0.055)		0.14** (0.43)	0.065 (0.42)
Participant coop. in period 1			-0.19*** (0.016)			0.16*** (0.024)
	Grim-trigger			Tit-for-Tat		
	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)
UE	0.022 (0.019)	0.014 (0.019)	0.014 (0.019)	-0.013 (0.029)	-0.026 (0.029)	-0.027 (0.029)
UP	-0.030 (0.020)	-0.031 (0.021)	-0.031 (0.020)	0.055** (0.028)	0.037 (0.028)	0.037 (0.028)
Advantaged	-0.046*** (0.015)	-0.044*** (0.015)	-0.043*** (0.014)	0.039* (0.020)	0.042** (0.020)	0.044** (0.020)
$r$		0.028 (0.033)	0.015 (0.035)		-0.15*** (0.047)	-0.16*** (0.049)
$E(r) - r$		-0.086** (0.043)	-0.091** (0.043)		-0.11* (0.058)	-0.11** (0.058)
$E(g) - r$		0.13*** (0.040)	0.12*** (0.040)		-0.030 (0.052)	-0.035 (0.053)
Participant coop. in period 1			0.011 (0.016)			0.014 (0.023)
$N$				500	500	500
pseudo $R^2$				0.0387	0.0740	0.0982

Standard errors in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Multinomial logit models, reporting average marginal effects.

Table 2.10: Menu of strategies



Constant strategies independent of other players' decisions, like choosing to always cooperate or always defect, depend on inequality: inequality decreases the likelihood of choosing to always cooperate but increases the likelihood of always defect. First-order normative beliefs and the participant's cooperation in period 1 negatively influence the choice to always defect and positively influence the choice to always cooperate. Descriptive beliefs also negatively influence the choice to always defect. The impact of the treatment is still significant once controlling for beliefs and the participant's past decision in period 1. This leads to result 3.

**Result 3.** The existence of inequality decreases the likelihood of choosing to always cooperate, and increases the likelihood of choosing to always defect because, partially, of a change in the social norm, whatever the type of the participants.

Strategies directly linked to the participant's playmate behavior do not depend on inequality and are not related to the participants' decisions in period 1. First-order normative beliefs negatively influence the choice of the tit-for-tat strategy, while descriptive beliefs positively influence the choice of the grim-trigger strategy.

Inequality changes normative and descriptive beliefs, weakening the social norm of cooperation. Such changes reduce, in turn, the decision to cooperate and the choice of the always cooperate strategy while raising the choice of the always defect strategy. The intensity of the impact of the social norm does not appear to impact behavior differently in the equal and unequal frameworks.

## 2.4 Discussion and conclusion

Cooperation in moving social interactions is essential to face today's main challenges, such as reducing greenhouse gas emissions, water conservation, and improving general well-being. Interactions with many others at a non-regular frequency make the situation complex to study and render the role of the social norm central. An essential stake in this context is the consideration of inequality. Indeed, many people benefit from the same common good but do not have the same possibilities to contribute to it. The experiment we conducted tackles these two dimensions: infinitely repeated interactions and the existence of inequality. We elicit the participants' beliefs about the social norm and their decision to cooperate or not in an infinitely repeated prisoner's dilemma that either gives equal or unequal payoffs to the participants for cooperation/defection according to the treatment.

The results show that the existence of inequality weakens the social norm of cooperation by decreasing first- and second-order normative beliefs as well as descriptive beliefs. Such changes in first-order normative beliefs and descriptive beliefs lower the likelihood of choosing to cooperate. The long-term strategy of always cooperating is also more likely for higher first-order normative expectations. Always defecting is more chosen for lower levels of first-order normative beliefs and descriptive beliefs. According to the results of the experiment, while the incentives for cooperation are not changed, the existence of inequality is detrimental to cooperation because of a weakened social norm. Interestingly, the participants' endowment level does not appear to influence either beliefs or behaviors.

We do not observe different influences of the social norm depending on whether inequality exists or not. The very existence of inequality leads to changes in the choice of cooperation.

An opposite effect could have been expected if assuming inequity aversion of the participants (Fehr and Schmidt, 1999). The payoffs in the unequal treatments have been determined to guarantee that cooperation would not lead to any conflicting norms (Gangadharan, Nikiforakis and Villeval, 2017). Therefore, some participants might be willing to cooperate to avoid the inequality. This is not what is observed in the experiment: the detrimental effect of inequality on the social norm of cooperation appears to be much stronger than the wish to reduce inequality by cooperating.

The results of the experiment emphasize the importance of transparency about the normative behavior to adopt and the adoption of this behavior by other people. Indeed, a main driver of lower cooperation in the presence of inequality is the change in beliefs. Transparency about the social norm should be disseminated to limit the negative effects of inequality on cooperation. The question of the sustainability of the common good in the presence of inequality needs to be further investigated (Baland, Bardhan and Bowles, 2007). Future research should address interactions between the participants in a dynamic setting and ask for the role of various instruments in this context.



## Chapter 3

# Inequality, Social Norms and Community Enforcement of Cooperation: Experimental Evidence

*Co-authored with Sabrina Teyssier*

**Abstract** This paper evaluates the impact of inequality on community enforcement of cooperation through the prospective implementation of a social norm. We explore players' perceptions of social norms and their cooperative tendencies by utilizing an infinitely repeated prisoner's dilemma game with anonymous random matching. Our study elicits and monitors the evolution of social norm beliefs, comprising first- and second-order normative beliefs and descriptive beliefs. The results show that inequality negatively affects cooperation and normative beliefs, whether or not the institution authorizes communication. When communication is allowed, heterogeneous matching modifies the weight of the impact of social norm beliefs on cooperation: descriptive beliefs influence the decision to cooperate less to the benefit of second-order normative beliefs. Our findings offer insights into the interplay of inequality, social norms, and communication in driving cooperative behavior.

### 3.1 Introduction

Cooperation among individuals and groups has become increasingly vital in tackling societal challenges, with climate change as a prominent example. Individuals' decisions and actions provide externalities on the environment, but individual interests lead to free-riding behavior. Community enforcement of cooperation has emerged as a potential equilibrium enabled by existing social norms – informal group and societal behavioral rules – (Kandori,

1992; Ellison, 1994; Bicchieri, 2005, 2016). However, the crucial questions of whether such social norms exist and whether they effectively encourage community enforcement of cooperation remain unanswered.

The current paper aims to address these questions with experimental evidence. We employ an infinitely repeated prisoner's dilemma game with anonymous random matching, allowing us to investigate players' perceptions of social norms and their cooperation tendencies when cooperating is an equilibrium strategy. Prior research has emphasized that although cooperation can be maintained if group size is sufficiently small, payoffs provide high returns on cooperation, and individual past cooperation is public information, it has also found a traditional impoverishment of cooperation despite the infinity of the game (Schwartz, Young and Zvinakis, 2000; Camera and Casari, 2009; Duffy and Ochs, 2009).<sup>1</sup> Social norm perception and individual beliefs have not been examined in these studies despite their significance in driving community enforcement. Our study aims to fill this gap by eliciting and monitoring the evolution of social norm beliefs, consisting in first- and second-order normative beliefs and descriptive beliefs (Bicchieri, 2005, 2016).<sup>2</sup> Since both cooperation and defection are Nash subgame perfect equilibria, compliance with the social norm can serve as a mechanism for selecting the equilibrium (Burke and Young, 2011; Young, 2015).<sup>3</sup>

Community enforcement plays a pivotal role in regulating collective actions and maintaining social order. The presence of a shared social norm among individuals within a group is fundamental to the effectiveness of community enforcement mechanisms (Granovetter, 2011; Bicchieri, 2005). Shared social norms foster a sense of collective identity and mutual understanding, encouraging individuals to internalize cooperative behavior and participate in monitoring and sanctioning norm deviations (Baland, Bardhan and Bowles, 2007; Ostrom, 2009). The alignment of individual behavior with the social norm enhances the likelihood of compliance, leading to more effective community enforcement in achieving desirable outcomes such as cooperation (Fehr and Gächter, 2000*b*). A social norm is, therefore, a critical determinant of the feasibility and success of community enforcement in economic contexts (Bicchieri, 2005; Young, 2015).

Inequality pervades society and has been exacerbated by globalization, social reproduction, and climate change (Banerjee and Duflo, 2005; Piketty, Saez and Zucman, 2018; Shukla, Skea and Slade, 2022). It becomes a crucial factor in our study because the presence of inequality poses challenges in establishing a single social norm, as different reference groups may be considered, leading to strengthened intra-group cooperation and reduced extra-group cooperation (Tajfel, 1970; Tajfel et al., 1971). The benefits of cooperation may be lower for some players, reducing adherence to the social norm (Andreoni, Brown and Vesterlund, 2002; Bicchieri and Mercier, 2013; Bicchieri, 2016). Consequently, inequality erodes cohesion and trust in society, weakening norms of cooperation and more generally pro-social behavior (Putnam, 2000; Ostrom, 2009; Bicchieri and Chavez, 2010; Xiao and Bicchieri,

---

<sup>1</sup>See Dal Bó and Fréchette (2018) for a survey on the main determinants of cooperation in infinitely repeated prisoners' dilemma games. The literature considering random matching in this context is still relatively scarce.

<sup>2</sup>To the best of our knowledge, our research is the first to measure the dynamics of social norm beliefs with experimental evidence, enhancing our understanding of cooperative behavior. Only Bandiera, Barankay and Rasul (2006) explain the change in cooperative behavior by an evolution of the cooperative norm in the field without directly measuring it.

<sup>3</sup>See also Bicchieri (2005) and Cialdini, Reno and Kallgren (1990) for a discussion on the role of norms to select non-equilibrium behavior as part of Bayesian strategies.

2010; Krupka and Weber, 2013; Bicchieri and Mercier, 2013; Bicchieri, 2016; Gangadharan, Nikiforakis and Villeval, 2017; Dimant and Gesche, 2023).<sup>4</sup>

Our study focuses on the impact of inequality on community enforcement of cooperation through the prospective implementation of a social norm. We explore institutions with varying degrees of communication: one with no communication allowed and two others permitting players to send free written messages to their playmates, either before or after playing the prisoner's dilemma game. Communication serves as a mechanism that aids coordination, reducing strategic uncertainty and promoting increased cooperation (Cooper et al., 1992; Charness, 2000; Camerer, 2011; Blume and Ortmann, 2007). By comparing institutions with varying levels of communication, we gain insights into the influence of communication on cooperative behavior.

Experimental results show that cooperation is difficult to enforce in the community. Inequality decreases cooperation, mainly due to the presence of different types of players according to the benefits they derive from cooperation rather than the specific behavior of one type of player or the sort of matching, homogeneous or heterogeneous. Inequality also reduces normative and descriptive beliefs about cooperation that positively affect the likelihood of cooperation. Communication increases the probability of cooperation and, essentially, prevents it from diminishing over the rounds. When communication is allowed, we observe that heterogeneous matching modifies the weight of the impact of social norm beliefs on cooperation: descriptive beliefs influence the decision to cooperate less to the benefit of second-order normative beliefs. Inequality modifies beliefs about social norms and the weight of these beliefs on cooperation as a function of matching. The matching between players is central in determining the beliefs responsible for decision-making.

The remainder of this paper is structured as follows. Section 3.2 outlines the design and methodology of the experiment. The results are presented in Section 3.3. In Section 3.4, we discuss the implications of our findings and conclude.

## **3.2 The experiment**

### **3.2.1 Framework**

In an indefinitely repeated game, cooperation is an equilibrium for a sufficiently high probability of continuation, meaning that cooperation may be sustained among individuals based on the folk theorem (Roth and Murnighan, 1978; Kandori, 1992; Ellison, 1994). Reputation effects, where two individuals interact repeatedly over an indefinite number of periods and are then randomly and anonymously matched with another playmate, facilitate cooperation (Deb, González-Díaz and Renault, 2016; Deb, Sugaya and Wolitzky, 2020). In the experiment, we eliminate any

---

<sup>4</sup>Many studies have reported a negative impact of inequality on cooperation in finite games such as the public good game (Anderson, Mellor and Milyo, 2008; Fischbacher, Schudy and Teyssier, 2014), the prisoner's dilemma game (Ahn et al., 2007; Beckenkamp, Hennig-Schmidt and Maier-Rigaud, 2007) or the common pool resources game (Gardner, Ostrom and Walker, 1990; Banerjee et al., 2001). See also Suchon and Theroude (2022) for a meta-analysis on the impact of inequality on cooperation finite public goods games.

effect of reputation and consider that individuals are randomly matched at every period of the game. We maintain a probability of continuation that allows for cooperation to be an equilibrium in this context, respecting the minimum probability of continuation as defined in Camera and Casari (2009); Camera, Casari and Bigoni (2012) for subgame perfect equilibrium and Blonski, Ockenfels and Spagnolo (2011); Blonski and Spagnolo (2015) for risk dominance.

One reason cooperation is maintained when individuals do not interact repeatedly with the same person but rather randomly with people in a larger group is that everyone in the group shares the same opinion about the social appropriateness of cooperation. In previous research, Duffy and Ochs (2009) have shown a decline in cooperation over time in this framework, prompting us to delve deeper into the factors that hinder community enforcement of cooperation in the absence of external enforcement mechanisms. We specifically aim to identify the individuals' beliefs that influence cooperation in a context where agents interact sporadically. We question whether perceptions of the social norm have a role in shaping cooperative behavior and how these beliefs evolve throughout the game. We define social norm beliefs as the combination of the personal value of the good thing to do, i.e., first-order normative beliefs, beliefs about other group members' personal values of the good thing to do, i.e., second-order normative beliefs, and beliefs about what other group members do between cooperating or defecting, i.e., descriptive beliefs (Bicchieri, 2005). We then ask how such beliefs impact the cooperation behavior and how it evolves as the game goes.

Societies are unequal, and social norms of cooperation are more challenging to achieve in this context. We investigate the impact of inequality on cooperation and individuals' perceptions of the social norm. We hypothesize that inequality and heterogeneous matching could hamper cooperation by hindering the establishment of a shared social norm. Additionally, we study the role of communication, implemented through free messages exchanged before or after the game, in facilitating community enforcement of cooperation. We aim to discern the distinct effects of ex-ante and ex-post communication on belief formation and cooperative behavior, anticipating that each form of communication influences specific types of beliefs. Ex-ante communication is believed to reduce strategic uncertainty, while ex-post communication reduces uncertainty about the history of the game (Dvorak and Fehrler, 2018). We then hypothesize that ex-ante communication would have a stronger effect on descriptive beliefs, while ex-post communication would have a stronger effect on normative beliefs. The experiment aims to gain empirical insights into the formation of beliefs about the social norm of cooperation and their influence on cooperative behavior under different contextual conditions, such as the existence of two types of individuals, homogeneous or heterogeneous matching, and message exchanges.

### **3.2.2 The basic game**

In all treatments, groups of eight subjects are randomly constituted and are fixed for the entire experiment. The game is played for an indefinite number of periods: at the end of each period, a random draw makes a new period

starting with a 95% probability. This continuation probability makes the duration of the game uncertain. Therefore, a time constraint was introduced, limiting the experiment to a maximum of 80 minutes. In order to prevent too few periods within an experiment, a new game starts in case the first game stops, with the same continuation probability.

Within each group, subjects are randomly matched in pairs at each period. They play a prisoner's dilemma game with two options: cooperating or defecting. Subject  $i$ 's payoffs,  $\Pi_i$ , are defined following a two-player public good game with contribution,  $g_i$ , that is either contributing nothing to the public good, i.e.,  $g_i = 0$ , or contributing her full endowment, i.e.,  $g_i = E_i$ , and a marginal per capita return of 0.8. In a pair of subjects, subject  $i$ 's and  $j$ 's payoffs are then calculated such as  $\Pi_i = E_i - g_i + 0.8(g_i + g_j)$  and  $\Pi_j = E_j - g_j + 0.8(g_i + g_j)$ , respectively. The matrix of payoffs for subjects  $i$  and  $j$  is as follows:

	Cooperate	Defect
Cooperate	$0.8(E_i + E_j) ; 0.8(E_i + E_j)$	$0.8E_i ; E_j + 0.8E_i$
Defect	$E_i + 0.8E_j ; 0.8E_j$	$E_i ; E_j$

Table 3.1: Payoffs matrix

To avoid framing effects in the experiment, cooperating is renamed "decision K", and defecting "decision P". In each period, before playing the game, the subjects had to provide their beliefs regarding their perception of the social norm in the sense of Bicchieri and Xiao (2009) including their first- and second-order normative expectations and descriptive expectations. The first-order normative beliefs (*NB1*) are elicited with the following question: "In your opinion, what is the most appropriate decision a participant should make?". The answer can either be "decision K" or "decision P". They represent the subject's personal values for cooperation. The second-order normative beliefs (*NB2*) are elicited with the question, "In your opinion, how many participants in your group (you excluded) will indicate K or P?" and the descriptive beliefs (*DB*) with "In your opinion, how many participants in your group (you excluded) will actually choose decision K or P?". To answer these two questions, the subjects had to move a slider to indicate their expectations between 0 participants who indicate "P" and 7 participants who indicate "K", and 7 participants who indicate "P" and 0 participants who indicate "K". Subjects also indicate whether they believe their playmate, i.e., the subject they are paired with, will choose "decision K" or "Decision P". Participants earn 3 additional units of experimental currency per correct answer to these questions with an acceptable margin of error of 1, except for NB1 which is not verifiable.

At the end of the period, the subjects are recalled their decision, "K" or "P", and receive feedback about their playmate's decision and their associated gain for the period. They do not learn answers from the questions about their expectations and beliefs but are informed of the number of correct answers at the end of the experiment.



### 3.2.3 Experimental treatments

The experiment consists of nine treatments using a  $2 \times 3$  design. The treatments differ in both the inequality of the subjects' endowments within the group, i.e., the existence of two types or only one type of subjects, and the opportunity to communicate with their playmate by sending them a message before or after playing the game.

**Inequality** Two types of subjects exist, either *Low* or *High*. Low-type subjects have an endowment equal to 10, whereas high-type subjects have an endowment equal to 20. In two treatments, groups of eight participants are equal, with all the subjects being of the same type (EQ treatments): all subjects have a low endowment, or all subjects have a high endowment. In a third treatment (INEQ treatment), half of the subjects are low-type, and the other half are high-type. At the beginning of the experiment, within each group of eight subjects, types are randomly attributed, with four subjects having a low endowment and four subjects having a high endowment. Subjects keep the same type for the whole duration of the experiment.<sup>5</sup> Pairs can be composed either of participants of the same type or participants with a different type. The group composition was common knowledge, while the playmates' type was privately known. The INEQ treatment asks the questions eliciting the subjects' beliefs, but duplicates them by asking the same questions about subjects of each type. The payoff matrices in experimental currency, depending on the type of the two playmates, are as follows:

	Cooperate	Defect
Cooperate	16 ; 16	8 ; 18
Defect	18 ; 8	10 ; 10

Table 3.2: Payoffs for two Low-type subjects

	Cooperate	Defect
Cooperate	32 ; 32	16 ; 36
Defect	36 ; 16	20 ; 20

Table 3.3: Payoffs for two High-type subjects

	Cooperate	Defect
Cooperate	24 ; 24	8 ; 28
Defect	26 ; 16	10 ; 20

Table 3.4: Payoffs for a Low-type subject (line) and a High-type subject (column)

At each period, the corresponding payoff matrix is presented to the subject before making her decision in the game. In the INEQ treatment, the subject also learns her playmate's type.

**Communication** Three treatments differ in the possibility of the participants to send a message to their playmates during the periods. In the "N" treatment, the subjects are not allowed to send any message, while in the "EA" treatment, each subject in the pair can send a free written message to her playmate before playing the game, and

<sup>5</sup>Only subjects in the INEQ treatment are aware of the concept of type.

in the "EP" treatment, each subject in the pair can send a free written message after having played the game. In the EA and EP treatments, subjects can send a message but choose not so. The messages were simultaneous and did not allow for a reply. Messages are free.

For all treatments, the minimum continuation probability supporting cooperation as a Subgame Perfect Equilibrium,  $\delta^{SPE}$ , is 0.574 for a homogeneous matching, and 0.773 (resp. 0.404) for high-type (resp. low-type) subjects in a heterogeneous matching. The minimum continuation probability ensuring cooperation as risk Dominant,  $\delta^{RD}$ , is 0.4 for a homogeneous matching and 0.533 for a heterogeneous matching.<sup>6</sup> The continuation probability in the experiment is 95% , which satisfies all of these conditions and makes cooperation a subgame perfect equilibrium that is risk dominant.

### 3.2.4 Procedures

The experiment was conducted at the GAEL laboratory in Grenoble, France. In total, 600 students from Grenoble Alpes University participated in the experiment. They were 58.5% female, with an average age of 20.4 years old. 45% had a high school diploma, 29% a bachelor's degree, 22% a master's degree, and 3% other qualifications. Recruitment was organized with Hroot software, and sessions were conducted using oTree software (Chen, Schonger and Wickens (2016)). Table 3.5 summarizes the treatments and the distribution of the subjects and independent groups in the treatments.

Treatment	Subjects type	Messages	Participants	Groups
EQ-N	Low	No	40	5
EQ-N	High	No	40	5
INEQ-N	Low and High	No	120	15
EQ-EA	Low	Ex-ante	40	5
EQ-EA	High	Ex-ante	40	5
INEQ-EA	Low and High	Ex-ante	120	15
EQ-EP	Low	Ex-post	40	5
EQ-EP	High	Ex-post	40	5
INEQ-EP	Low and High	Ex-post	120	15
Total	-	-	600	75

Table 3.5: Distribution of participants by treatment

For a comparison of heterogeneous and homogeneous groups, the analysis will analyze together treatments EQ-N with low-type subjects only and with high-type subjects only in order to count both types of subjects as in INEQ treatments.

Upon arrival, subjects were randomly assigned to a computer. The instructions appeared on the subjects' screen and were read aloud by the experimenter. Only neutral words were used in the instructions to ensure a neutral framing (instructions are given in Appendix B.2). The subjects had then to fill in a comprehension questionnaire, and

<sup>6</sup>Calculation is based on Camera and Casari (2009) giving conditions of a Subgame Perfect Equilibrium and on Blonski and Spagnolo (2015) for Risk Dominance. Details are provided in Appendix B.1.

the experiment started only once all subjects correctly understood the process of decision-making in the experiment. The subjects were randomly assigned a type, low or high, at the beginning of the experiment. At the end of the experiment, the subjects completed a short socio-demographic questionnaire including gender and their discipline of study. The average duration of the experiment was about 75 minutes, with average earnings of 15.45 euros.

### 3.3 Results

We first analyze the impact of inequality on cooperation and social norm beliefs when communication is not allowed. Then, we focus on institutions with communication that happens before or after playing the game. To report the role of repeated interactions within the community, we will study *a priori* beliefs and initial cooperation and observe cooperation and the social norm in the final periods.

#### 3.3.1 Inequality

##### Cooperation

Table 3.6 presents descriptive statistics of the frequency of cooperation, i.e., the average and the intra-group standard deviation into parentheses, when the group of eight subjects is equal or unequal. We separate the first three rounds and the last three rounds to analyze the evolution of behaviors. Because the game is indefinitely repeated, intermediate rounds vary depending on the total number of rounds in the experimental session, with an average of 24.

Rounds	First Three	Intermediate	Last Three	All
<b>EQ-N treatment</b>				
High-type	0.591 (0.483)	0.424 (0.462)	0.408 (0.417)	0.439 (0.463)
Low-type	0.558 (0.466)	0.309 (0.451)	0.316 (0.450)	0.327 (0.457)
<i>Average</i>	<i>0.575 (0.474)</i>	<i>0.367 (0.456)</i>	<i>0.362 (0.433)</i>	<i>0.383 (0.460)</i>
<b>INEQ-N treatment</b>				
High-type	0.455 (0.461)	0.332 (0.420)	0.216 (0.321)	0.332 (0.426)
Low-type	0.466 (0.484)	0.290 (0.406)	0.227 (0.344)	0.306 (0.423)
<i>Average</i>	<i>0.461 (0.482)</i>	<i>0.311 (0.434)</i>	<i>0.222 (0.372)</i>	<i>0.318 (0.440)</i>

Table 3.6: Cooperation per subject type and treatment

Cooperation decreases over time (Wilcoxon ranksum test between the first and last three rounds:  $z=3.329$ ,  $p=0.001$  in EQ-N and  $z=5.377$ ,  $p<0.001$  in INEQ-N). The absence of community enforcement of cooperation in our experiment confirms previous findings from Duffy and Ochs (2009). Inequality decreases cooperation both in the three first rounds (0.575 in EQ-N and 0.461 in INEQ-N, which are significantly different using a Wilcoxon ranksum test:  $z=2.133$ ,  $p=0.0330$ ) and the three last rounds (0.362 in EQ-N and 0.222 in INEQ-N, which are significantly

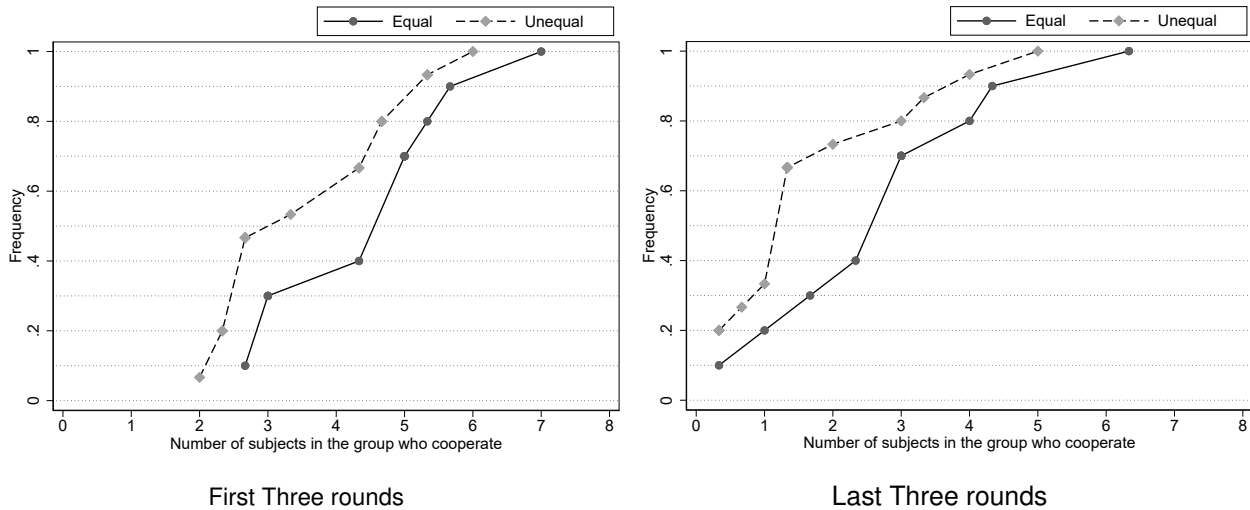


Figure 3.1: Cumulative distribution of groups

different using a Wilcoxon ranksum test:  $z=2.061$ ,  $p=0.0408$ ). Besides, we do not observe a significant change in the intra-group variance over time. The distribution of the number of cooperators per group shown in Figure 3.1 confirms this finding.

Inequality decreases the frequency of cooperation in groups. We ask whether such effect depends on the type of the decision maker, high or low, on the matching, homogeneous or heterogeneous, or on the existence of inequality in the group *per se*, i.e., the presence of the two types of subjects. From Table 3.6, we observe no significant difference in cooperation between high- and low-type subjects under inequality (Wilcoxon ranksum test:  $z=0.228$ ,  $p=0.8094$  for the first three rounds and  $z=0.400$ ,  $p=0.7041$  for the last three rounds). We study the role of the matching with numbers in Table 3.7 which reports, for the unequal treatment, cooperation depending on the matching for each subject type. In the last three rounds, cooperation is about 26% in homogeneous matching and 17% in heterogeneous ones but are not significantly different (Wilcoxon ranksum test:  $z=1.441$ ,  $p=0.1493$  for the first three rounds and  $z=1.414$ ,  $p=0.1578$  for the last three rounds).

Rounds	First Three	Intermediate	Last Three	All
<b>INEQ-N treatment</b>				
Homogeneous matching				
H-type	0.503 (0.476)	0.386 (0.449)	0.255 (0.337)	0.371 (0.455)
L-type	0.529 (0.483)	0.329 (0.449)	0.271 (0.337)	0.315 (0.455)
<i>Average</i>	<i>0.516 (0.480)</i>	<i>0.358 (0.449)</i>	<i>0.263 (0.337)</i>	<i>0.343 (0.455)</i>
Heterogeneous matching				
H-type	0.421 (0.465)	0.291 (0.417)	0.160 (0.349)	0.291 (0.426)
L-type	0.432 (0.462)	0.283 (0.417)	0.187 (0.352)	0.324 (0.426)
<i>Average</i>	<i>0.426 (0.463)</i>	<i>0.287 (0.417)</i>	<i>0.174 (0.350)</i>	<i>0.308 (0.426)</i>

Table 3.7: Cooperation per matching sort in INEQ-N treatment

The decision to cooperate, or not, does not appear to depend either on the subjects' type or on the matching.

The existence of inequality *per se*, i.e., the existence of two types of subjects, seems to explain the lower cooperation rates in unequal rather than equal treatments.<sup>7</sup> These observations lead to the following first result.

**Result 1.** Community enforcement of cooperation is weaker when groups are unequal rather than equal, mainly due to the presence of two types of subjects in the group.

Result 1 is valid both at the beginning and at the end of the game. The lower cooperation in the unequal treatment is present right from the start of the game and remains so until the end. Because the type of subjects and matching do not explain the difference in cooperation between unequal and equal groups, we turn to the analysis of the perception of the social norm.

## Social norm

Inequality may affect social norm beliefs that, in turn, may decrease cooperation in two ways: by a decrease in social norm beliefs or by a change in the weight of social norm beliefs in explaining cooperation. Social norm beliefs include first- and second-order normative beliefs as well as descriptive beliefs regarding the decision to cooperate. Means are presented in Tables 3.8, 3.9 and 3.10, respectively, and intra-group standard deviation are in parentheses.

Rounds	First Three	Intermediate	Last Three	All
<b>EQ-N treatment</b>				
High-type	0.7 (0.423)	0.656 (0.392)	0.65 (0.378)	0.659 (0.399)
Low-type	0.7 (0.391)	0.593 (0.478)	0.575 (0.490)	0.599 (0.476)
<i>Average</i>	<i>0.7 (0.407)</i>	<i>0.625 (0.435)</i>	<i>0.612 (0.434)</i>	<i>0.629 (0.437)</i>
<b>INEQ-N treatment</b>				
High-type	0.525 (0.472)	0.387 (0.456)	0.369 (0.429)	0.405 (0.460)
Low-type	0.533 (0.478)	0.420 (0.467)	0.388 (0.462)	0.432 (0.470)
<i>Average</i>	<i>0.529 (0.473)</i>	<i>0.403 (0.464)</i>	<i>0.379 (0.449)</i>	<i>0.418 (0.467)</i>

Table 3.8: First-order normative beliefs per subject type and treatment

<sup>7</sup>We observe that, in the last three rounds, cooperation in homogeneous matching in the unequal treatment is slightly lower than in the equal treatment (Wilcoxon ranksum test:  $z=1.707$ ,  $p=0.0895$ ). This result strengthens the role of the existence of inequality *per se*, suggesting that even when groups are homogeneous, the cooperation rate is lower in the unequal than equal treatments.

Rounds	First Three	Intermediate	Last Three	All
<b>EQ-N treatment</b>				
High-type	0.605 (0.304)	0.505 (0.272)	0.488 (0.225)	0.513 (0.278)
Low-type	0.653 (0.284)	0.413 (0.323)	0.417 (0.296)	0.429 (0.326)
<i>Average</i>	<i>0.629 (0.294)</i>	<i>0.459 (0.298)</i>	<i>0.452 (0.261)</i>	<i>0.471 (0.302)</i>
<b>INEQ-N treatment</b>				
High-type	0.529 (0.272)	0.388 (0.274)	0.333 (0.261)	0.399 (0.280)
Low-type	0.525 (0.262)	0.409 (0.272)	0.340 (0.255)	0.414 (0.275)
<i>Average</i>	<i>0.527 (0.266)</i>	<i>0.398 (0.273)</i>	<i>0.337 (0.259)</i>	<i>0.406 (0.278)</i>

Table 3.9: Second-order normative beliefs per subject type and treatment

Rounds	First Three	Intermediate	Last Three	All
<b>EQ-N treatment</b>				
High-type	0.551 (0.288)	0.380 (0.226)	0.359 (0.173)	0.395 (0.239)
Low-type	0.586 (0.270)	0.305 (0.252)	0.308 (0.203)	0.324 (0.263)
<i>Average</i>	<i>0.569 (0.279)</i>	<i>0.343 (0.239)</i>	<i>0.333 (0.188)</i>	<i>0.360 (0.251)</i>
<b>INEQ-N treatment</b>				
High-type	0.486 (0.319)	0.344 (0.288)	0.292 (0.265)	0.356 (0.297)
Low-type	0.476 (0.312)	0.366 (0.287)	0.315 (0.262)	0.372 (0.291)
<i>Average</i>	<i>0.481 (0.315)</i>	<i>0.355 (0.289)</i>	<i>0.303 (0.268)</i>	<i>0.364 (0.295)</i>

Table 3.10: Descriptive beliefs per subject type and treatment

First- and second-order normative beliefs are significantly lower when there is inequality between participants in the group, both in the first and last rounds (Wilcoxon ranksum tests between EQ-N and INEQ-N treatments: first-order normative beliefs,  $z=3.306$ ,  $p<0.001$  for the first three rounds and  $z=3.333$ ,  $p<0.001$  for the last three rounds; second-order normative beliefs,  $z=3.022$ ,  $p=0.0024$  for the first three rounds and  $z=2.718$ ,  $p=0.0064$  for the last three rounds). Descriptive beliefs are significantly lower in the unequal treatment than in the equal treatment in the first rounds (Wilcoxon ranksum tests:  $z=2.392$ ,  $p=0.0165$ ), but this difference is not observed in the last rounds (Wilcoxon ranksum tests:  $z=0.954$ ,  $p=0.3411$ ). The experience in the game erases differences in descriptive beliefs between equal and unequal treatments, whereas this is not the case for normative beliefs. Descriptive beliefs and second-order normative beliefs differ in their dynamics: descriptive beliefs evolve according to the experience in the game rather than the context of inequality, while normative beliefs conserve the difference due to the unequal context even after some experience. This leads to Result 2.

**Result 2.** Normative beliefs are lower when groups are unequal rather than equal, while descriptive beliefs are not different after some experience in the game.

We now measure the influence of social norm beliefs on individual decisions to cooperate. Table 3.11 presents the average marginal effects of logit models explaining each participant's decision to cooperate. In Model (1), we

measure the effect of inequality and the matching. In Model (2), we account for the type of the decision maker. Models (3) to (5) add the subjects' social norm beliefs: Model (3) accounts for first- and second-order normative beliefs and descriptive beliefs, Models (4) and (5) account for the impact of such beliefs depending on inequality and the matching and Model (5) additionally controls for the subject's expectation regarding her playmate's decision to cooperate in the current round. The latter variable comprises the subject's optimism or pessimism regarding her playmate's decision on top of her perception of the social norm. In all models, we included variables accounting for potential spillover effects in the group through experience. Following Galbiati, Henry and Jacquemet (2019), we included whether the subject experienced cooperation with the playmate at the previous round, the number of times he experienced cooperation with the playmates over the five previous rounds, and the number of times in a row cooperation has been experienced in the previous rounds. The total number of rounds is also controlled for in all models.

The presence of inequality or the heterogeneous matching do not significantly decrease the likelihood of cooperation. Instead, the history of the game, i.e., observing cooperation in the short term, either at the previous round or over the five last rounds, significantly increases the likelihood of cooperation. Including social norm beliefs in Models (3) to (5) removes explanatory power from these spillover effects. The impact of social norm beliefs is strong: both first-order normative beliefs and descriptive beliefs significantly increase the likelihood of cooperation throughout the experiment. This aligns with previous findings in the literature such as Bicchieri and Xiao (2009); Bašić and Verrina (2023). Accounting for the subject's expectation regarding her playmate's decision to cooperate in the current round keeps this result, which emphasizes the power of the social norm perception of the group on the decision. One may expect that if the decision maker thinks that her playmate does not follow the social norm, he would not be sensitive to the social norm. This result shows that this is untrue. Model (4) shows a lower impact of descriptive beliefs on the likelihood of cooperation in heterogeneous matching, but this is not observed with this control. This suggests that the influence of the descriptive beliefs of the social norm, i.e., concerning the total group as a whole, is lower under inequality. Still, this effect is offset by the influence of the subject's expectations regarding his playmate's decision. Besides, beliefs about the playmate's decision influence more the decision maker's likelihood of cooperation in the presence of inequality, which suggests that subjects' are more concerned by their beliefs regarding their direct playmate instead of the behavior of all members in the group. This leads to Result 3.

**Result 3.** First-order normative and descriptive beliefs increase the likelihood of cooperation with equivalent weights when groups are equal or unequal.

We separately run regressions for the first three rounds and the last three rounds to observe the evolution of the influence of social norm beliefs (see Tables B.1 and B.2 in Appendix B.3). At the beginning of the game, only first-order normative beliefs positively impact the decision to cooperate. The effect of descriptive beliefs is insignificant, possibly because subjects are not confident about their own beliefs because they have no information about the

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
INEQ	-0.073 (-1.46)	-0.079* (-1.65)	-0.060 (-1.44)	0.016 (0.24)	-0.088 (-1.20)
INEQ x Hetero.	-0.051 (-1.39)	-0.051 (-1.38)	-0.021 (-0.59)	0.025 (0.75)	0.052 (0.99)
High-type		-0.007 (-0.12)	-0.024 (-0.84)	-0.028 (-0.94)	-0.016 (-0.54)
High-type x INEQ		0.016 (0.17)	0.039 (0.77)	0.042 (0.87)	0.046 (1.00)
NB1			0.349*** (11.68)	0.404*** (7.37)	0.366*** (7.26)
NB1 x INEQ				-0.075 (-0.93)	-0.089 (-1.09)
NB1 x INEQ x Hetero.				-0.029 (-0.48)	0.017 (0.29)
NB2			-0.010 (-0.23)	-0.058 (-1.03)	-0.059 (-1.15)
NB2 x INEQ				0.050 (0.49)	-0.013 (-0.13)
NB2 x INEQ x Hetero.				0.107 (1.22)	0.044 (0.52)
DB			0.274*** (5.92)	0.386*** (11.69)	0.262*** (4.99)
DB x INEQ				-0.116 (-1.58)	-0.124 (-1.32)
DB x INEQ x Hetero.				-0.181** (-2.10)	-0.076 (-1.07)
Playmate's Cooperation Expected					0.134*** (3.48)
Playmate's Cooperation Expected x INEQ					0.203*** (3.49)
Playmate's Cooperation Expected x INEQ x Hetero.					-0.058 (-1.27)
Playmate's cooperation in the previous round	0.031** (2.31)	0.031** (2.30)	0.018* (1.66)	0.018 (1.61)	0.003 (0.30)
Number of times playmate's cooperation in the 5 previous rounds	0.066*** (4.05)	0.066*** (4.02)	0.014 (1.10)	0.011 (0.86)	-0.004 (-0.33)
Number of times playmate's cooperation in a row	-0.008 (-0.70)	-0.008 (-0.74)	-0.004 (-0.43)	-0.004 (-0.41)	-0.003 (-0.32)
Number of rounds	-0.004*** (-2.70)	-0.004*** (-2.50)	-0.004*** (-2.66)	-0.004** (-2.47)	-0.004** (-2.37)
Observations	4568	4568	4568	4568	4568
Clusters	25	25	25	25	25
Log-pseudolikelihood	-2677.04	-2676.35	-2051.89	-2032.7404	-1865.9501
Pseudo-R <sup>2</sup>	0.057	0.057	0.277	0.284	0.343

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Standard errors in parentheses.

Table 3.11: Decision to cooperate (Logit models - average marginal effects)

subjects they interact with. Interestingly, the effect of first-order normative beliefs is significantly lower when there is inequality in the group, which effect disappears at the end of the game. This suggests that the only presence



of inequality decreases the weight the subjects set on their normative beliefs only when they have no experience interacting with others in the group. Although community enforcement of cooperation is not observed, subjects in the group respect their first-order normative beliefs independently of the group's composition or playmate matching.

Inequality decreases normative beliefs that, in turn, decrease cooperation, but the impact of such beliefs on cooperation does not depend on inequality. These results apply to the considered structure that finally leads to low cooperation rates, i.e., 36% and 22% in equal and unequal treatments, respectively. Based on the literature, allowing communication among playmates should increase cooperation rates (Cooper et al., 1992; Charness, 2000; Camerer, 2011; Blume and Ortmann, 2007). This change in institution would provide a different setup as another case study of the impact of inequality on social norms' beliefs and cooperation.

### 3.3.2 Communication

When communication is allowed, we observe that only 4.38% of the participants never send a message in the equal treatments, while they are 14.58% in unequal treatments, resulting in less communication in unequal than equal treatments. The number of messages from ex-post communication increases between the first three and the last three rounds, from 44.2% to 57.9% in the equal treatment and 37.8% to 41.1% in the unequal treatment. Although the difference is not significant in the first three rounds, the number of messages that are sent in the last three rounds is higher in the equal than in the unequal treatment (Wilcoxon ranksum test:  $z=1.562$ ,  $p=0.140$  for the three first rounds and  $z=4.034$ ,  $p<0.001$  for the three last rounds). Ex-ante communication leads to a strong increase in the number of messages that are sent between the three first and the three last rounds, from 71.7% to 82.1% under equality and from 38.9% to 64.2% under inequality. The number of messages that are sent is higher in the equal than in the unequal treatment (Wilcoxon ranksum test:  $z=7.866$ ,  $p<0.001$  for the three first rounds and  $z=4.751$ ,  $p<0.001$  for the three last rounds). The number of messages sent is not different between homogeneous and heterogeneous matching in unequal treatments.

The proportion of messages sent by category of message is available in Table B.4 in Appendix B.5 according to the categorization of Dvorak and Fehrler (2018).<sup>8</sup>

### Cooperation

By the end of the experiment, cooperation rates equal 50% and 62% in the equal treatments and 43% and 44% in the unequal treatments, with ex-ante and ex-post communication, respectively. These rates are much higher than without the possibility to send any messages. Table 3.12 presents the mean and intra-group variance of cooperation

---

<sup>8</sup>Messages of the ex-ante communication are mostly in the "Coordination" category (60.6%), followed by the "Deliberation" category (19.2%), then the "Relationship" (11.1%) and "Trivia" (7.8%) categories, with a very small amount from "Information" category (1.1%). While the messages of the ex-post communication are mostly in the "Relationship" category (46.3%), followed by the "Coordination" category (22.7%), then the "Deliberation" (12.2%) and "Trivia" (11.2%) categories, with a small amount from "Information" category (7.3%).

in the first three rounds, intermediate rounds, last three rounds, and all rounds in the equal and unequal treatments with ex-ante and ex-post communication.

Rounds	First Three	Intermediate	Last Three	All
<b>EQ treatments</b>				
<b>Ex ante communication</b>				
High-type	0.6 (0.481)	0.410 (0.482)	0.408 (0.486)	0.436 (0.488)
Low-type	0.675 (0.419)	0.628 (0.386)	0.583 (0.405)	0.628 (0.419)
Average	0.637 (0.450)	0.519 (0.434)	0.495 (0.445)	0.532 (0.453)
<b>Ex post communication</b>				
High-type	0.583 (0.486)	0.507 (0.466)	0.558 (0.445)	0.523 (0.473)
Low-type	0.441 (0.476)	0.546 (0.432)	0.675 (0.369)	0.549 (0.443)
Average	0.512 (0.481)	0.527 (0.449)	0.616 (0.407)	0.536 (0.458)
<b>INEQ treatments</b>				
<b>Ex ante communication</b>				
High-type	0.488 (0.481)	0.457 (0.426)	0.361 (0.369)	0.445 (0.439)
Low-type	0.627 (0.478)	0.505 (0.454)	0.494 (0.390)	0.524 (0.459)
Average	0.558 (0.483)	0.481 (0.456)	0.427 (0.421)	0.485 (0.461)
<b>Ex post communication</b>				
High-type	0.561 (0.494)	0.439 (0.422)	0.411 (0.357)	0.455 (0.444)
Low-type	0.5 (0.464)	0.438 (0.435)	0.466 (0.403)	0.457 (0.453)
Average	0.530 (0.484)	0.438 (0.443)	0.438 (0.401)	0.456 (0.457)

Table 3.12: Cooperation per participant type and treatment with communication

Although cooperation decreases over rounds, communication tends to lower this decreasing trend, especially when communication is ex-post (Wilcoxon ranksum test between treatments with and without communication: for equal groups,  $z=1.181$ ,  $p=0.239$  when communication is ex-ante  $z=3.851$ ,  $p<0.001$  when communication is ex-post, and for unequal groups  $z=1.950$ ,  $p=0.051$  when communication is ex-ante and  $z=2.380$ ,  $p=0.017$  when communication is ex-post). Cooperation rates even increase for equal groups with ex-post communication. Community enforcement of cooperation appears to be easier when communication is allowed but still without reaching full cooperation in the group.

In the first three rounds, there is no significant difference in cooperation between equal and unequal treatments, whatever the subject's type and matching. The possibility to communicate thus helps to eliminate the *a priori* feeling of difference between the subjects in the same group in the unequal treatments. Community enforcement is less observed in the presence of inequality only with ex-post communication (Wilcoxon ranksum test between equal and unequal treatments for the three last rounds:  $z=1.166$ ,  $p=0.245$  with ex-ante communication,  $z=2.887$ ,  $p=0.004$  with ex-post communication). Ex-ante communication removes the negative impact of inequality on cooperation, which is valid for both high- and low-type subjects. With ex-post communication, both high- and low-type subjects cooperate less under inequality (Wilcoxon ranksum test between equal and unequal treatments for the three last rounds:  $z=1.752$ ,  $p=0.082$  for high-type participants,  $z=2.460$ ,  $p=0.013$  for low-type participants). The matching

does not impact cooperation.<sup>9</sup> Result 4 is as follows.

**Result 4.** Community enforcement of cooperation is facilitated by communication whether it is ex-ante or ex-post; whereas ex-ante communication removes the negative effect of inequality on cooperation, cooperation remains weaker when groups are unequal with ex-post communication due to the existence of two types of subjects in the group.

## Social norms

Descriptive statistics of social norm beliefs are presented in Tables 3.13, 3.14 and 3.15.

Rounds	First Three	Intermediate	Last Three	All
<b>EQ treatments</b>				
<b><i>Ex ante communication</i></b>				
High-type	0.575 (0.497)	0.529 (0.496)	0.566 (0.502)	0.540 (0.497)
Low-type	0.725 (0.382)	0.775 (0.370)	0.8 (0.335)	0.774 (0.372)
<i>Average</i>	<i>0.65 (0.440)</i>	<i>0.652 (0.433)</i>	<i>0.683 (0.418)</i>	<i>0.657 (0.434)</i>
<b><i>Ex post communication</i></b>				
High-type	0.7 (0.423)	0.656 (0.392)	0.65 (0.378)	0.659 (0.399)
Low-type	0.466 (0.485)	0.613 (0.447)	0.675 (0.345)	0.605 (0.451)
<i>Average</i>	<i>0.604 (0.455)</i>	<i>0.631 (0.427)</i>	<i>0.658 (0.352)</i>	<i>0.635 (0.432)</i>
<b>INEQ treatments</b>				
<b><i>Ex ante communication</i></b>				
High-type	0.525 (0.495)	0.448 (0.480)	0.411 (0.468)	0.453 (0.482)
Low-type	0.591 (0.486)	0.506 (0.479)	0.497 (0.471)	0.516 (0.481)
<i>Average</i>	<i>0.558 (0.491)</i>	<i>0.477 (0.481)</i>	<i>0.454 (0.470)</i>	<i>0.485 (0.484)</i>
<b><i>Ex post communication</i></b>				
High-type	0.505 (0.485)	0.464 (0.480)	0.483 (0.475)	0.475 (0.481)
Low-type	0.527 (0.493)	0.486 (0.484)	0.483 (0.484)	0.493 (0.485)
<i>Average</i>	<i>0.516 (0.487)</i>	<i>0.475 (0.482)</i>	<i>0.483 (0.478)</i>	<i>0.484 (0.483)</i>

Table 3.13: First-order normative beliefs per subject type and treatment with communication

<sup>9</sup>For exact numbers, see Table B.3 in Appendix B.4 that presents the mean cooperation and intra-group standard errors for the unequal treatments.

Rounds	First Three	Intermediate	Last Three	All
<b>EQ treatment</b>				
<b>Ex ante communication</b>				
High-type	0.658 (0.309)	0.575 (0.326)	0.517 (0.353)	0.579 (0.330)
Low-type	0.709 (0.257)	0.735 (0.236)	0.763 (0.260)	0.737 (0.250)
<i>Average</i>	<i>0.683 (0.283)</i>	<i>0.655 (0.281)</i>	<i>0.640 (0.307)</i>	<i>0.658 (0.290)</i>
<b>Ex post communication</b>				
High-type	0.635 (0.285)	0.598 (0.268)	0.639 (0.246)	0.611 (0.277)
Low-type	0.503 (0.297)	0.561 (0.276)	0.689 (0.196)	0.567 (0.282)
<i>Average</i>	<i>0.569 (0.291)</i>	<i>0.579 (0.272)</i>	<i>0.664 (0.221)</i>	<i>0.589 (0.279)</i>
<b>INEQ treatment</b>				
<b>Ex ante communication</b>				
High-type	0.536 (0.282)	0.496 (0.279)	0.456 (0.273)	0.494 (0.285)
Low-type	0.532 (0.292)	0.518 (0.281)	0.520 (0.267)	0.519 (0.287)
<i>Average</i>	<i>0.534 (0.286)</i>	<i>0.507 (0.281)</i>	<i>0.488 (0.272)</i>	<i>0.507 (0.287)</i>
<b>Ex post communication</b>				
High-type	0.503 (0.293)	0.504 (0.295)	0.493 (0.262)	0.502 (0.299)
Low-type	0.522 (0.280)	0.508 (0.281)	0.491 (0.257)	0.507 (0.285)
<i>Average</i>	<i>0.513 (0.286)</i>	<i>0.506 (0.288)</i>	<i>0.492 (0.259)</i>	<i>0.505 (0.292)</i>

Table 3.14: Second-order normative beliefs per subject type and treatment with communication

Rounds	First Three	Intermediate	Last Three	All
<b>EQ treatment</b>				
<b>Ex ante communication</b>				
High-type	0.621 (0.284)	0.481 (0.273)	0.422 (0.227)	0.493 (0.278)
Low-type	0.628 (0.268)	0.554 (0.228)	0.55 (0.212)	0.563 (0.245)
<i>Average</i>	<i>0.625 (0.276)</i>	<i>0.517 (0.250)</i>	<i>0.486 (0.219)</i>	<i>0.528 (0.261)</i>
<b>Ex post communication</b>				
High-type	0.575 (0.273)	0.546 (0.256)	0.571 (0.224)	0.556 (0.260)
Low-type	0.446 (0.244)	0.506 (0.252)	0.636 (0.175)	0.512 (0.261)
<i>Average</i>	<i>0.510 (0.259)</i>	<i>0.526 (0.254)</i>	<i>0.604 (0.199)</i>	<i>0.534 (0.260)</i>
<b>INEQ treatment</b>				
<b>Ex ante communication</b>				
High-type	0.5 (0.315)	0.444 (0.290)	0.412 (0.266)	0.447 (0.299)
Low-type	0.547 (0.314)	0.509 (0.302)	0.497 (0.257)	0.512 (0.302)
<i>Average</i>	<i>0.524 (0.315)</i>	<i>0.477 (0.301)</i>	<i>0.455 (0.269)</i>	<i>0.479 (0.305)</i>
<b>Ex post communication</b>				
High-type	0.482 (0.325)	0.485 (0.304)	0.483 (0.266)	0.484 (0.311)
Low-type	0.524 (0.314)	0.493 (0.301)	0.473 (0.261)	0.495 (0.305)
<i>Average</i>	<i>0.503 (0.320)</i>	<i>0.489 (0.303)</i>	<i>0.478 (0.263)</i>	<i>0.490 (0.309)</i>

Table 3.15: Descriptive beliefs per subject type and treatment with communication

First- and second-order normative beliefs are significantly lower in the unequal rather than equal treatments whether the message is sent before or after playing the game. When communication is ex ante, descriptive beliefs are significantly lower in the unequal than equal treatments at the beginning of the game but not in the end (Wilcoxon ranksum tests between unequal and equal treatments with ex ante communication: respectively for first-order nor-

mative, second-order normative and descriptive beliefs,  $z=2.128$  ,  $p=0.033$ ,  $z=5.234$  ,  $p<0.001$ ,  $z= 3.388$ ,  $p<0.001$  in the first three rounds and  $z=3.929$ ,  $p<0.001$ ,  $z=3.612$ ,  $p<0.001$ ,  $z=0.828$ ,  $p=0.409$  in the last three rounds) while they are significantly lower in the end of the game but not at the beginning when communication is ex-post (Wilcoxon ranksum tests between unequal and equal treatments: respectively for first-order normative, second-order normative and descriptive beliefs,  $z=1.770$  ,  $p=0.076$ ,  $z=1.667$  ,  $p=0.097$ ,  $z= 0.259$ ,  $p=0.797$  in the first three rounds and  $z=3.064$  ,  $p=0.002$ ,  $z=3.955$ ,  $p<0.001$ ,  $z=3.059$  ,  $p=0.002$  in the last three rounds). The latter difference is due to higher descriptive beliefs when communication is ex-post with equality than in the other treatments with ex-ante communication of inequality. Descriptive beliefs are then not decreased with inequality. Thus, Communication does not remove the negative effect of the presence of inequality on the perception of the social norm: inequality especially reduces normative beliefs about cooperation as the appropriate behavior to adopt. Besides, we observe no apparent impact of the timing of the communication on the subjects' beliefs, which suggests that such timing does not target some social norm beliefs more than others. When groups are equal, we only observe that descriptive beliefs are lower when communication is ex-ante instead of ex-post. Result 5 is as follows.

**Result 5.** Communication prevents descriptive beliefs from being lower when groups are unequal rather than equal, while normative beliefs are still lower.

We now investigate whether inequality impacts the weight of social norm beliefs on the decision to cooperate when communication is allowed. Table 3.16 presents logit estimations of the decision to cooperate. The five models include the same explaining variables as in Table 3.11 but add whether the communication is ex-ante and controls for whether subjects have received messages. We also cross this variable with unequal treatment and whether the matching in the current period is heterogeneous or not.

The results show that a subject is more likely to cooperate when she receives a message. When messages are sent before playing the game, the likelihood of cooperation increases in the current round, and when messages are sent after playing the game, the likelihood of cooperation increases in the following round. The impact of the message is stronger in the presence of inequality, suggesting that there is a lower trust in the message when the playmate is of a different type than the decision maker. The impact of received messages is lower in heterogeneous matching. This may indicate a reluctance to cooperate with playmates of a different type. This effect is consistent with the minimal group paradigm established by Akerlof (1980): unequal treatments assign types to subjects, leading to in-type favoritism and out-type defection.

Such as without communication, we observe that first-order normative beliefs and descriptive beliefs increase the likelihood of cooperation. A new and interesting result is observed: beliefs are more or less influential depending on the matching. Indeed, we find that second-order normative beliefs increase the probability of cooperation in the case of heterogeneous matching only, and the impact of descriptive beliefs is significantly reduced in this sort of matching. When the matching is heterogeneous, the loss of influence of descriptive beliefs on cooperation is offset

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
INEQ	-0.006 (-0.15)	0.002 (0.03)	-0.020 (-0.38)	-0.065 (-1.11)	-0.160*** (-2.67)
INEQ x Hetero.	-0.039 (-1.60)	-0.040 (-1.61)	-0.000 (-0.00)	0.031 (0.76)	0.059 (1.43)
Ex-Ante	-0.089*** (-3.13)	-0.092*** (-3.17)	-0.093*** (-3.80)	-0.090*** (-3.68)	-0.054** (-2.52)
Message received	0.103*** (4.25)	0.103*** (4.18)	0.057*** (2.67)	0.065*** (2.84)	0.048** (2.55)
Ex-Ante x Message received	0.068*** (2.81)	0.069*** (2.83)	0.076*** (3.05)	0.0727*** (2.81)	0.017 (0.73)
Message received x INEQ	0.080** (2.43)	0.080** (2.37)	0.114*** (3.16)	0.104*** (2.85)	0.084*** (2.71)
Message received x INEQ x Hetero.	-0.083*** (-3.34)	-0.080*** (-3.14)	-0.092*** (-3.82)	-0.100*** (-3.75)	-0.084*** (-3.13)
High-type		-0.020 (-0.48)	-0.043 (-1.19)	-0.036 (-0.92)	-0.045 (-1.26)
High-type x INEQ		-0.022 (-0.40)	-0.008 (-0.19)	-0.014 (-0.29)	-0.001 (-0.03)
NB1			0.236*** (8.63)	0.293*** (6.49)	0.241*** (5.13)
NB1 x INEQ				-0.087 (-1.50)	-0.081 (-1.42)
NB1 x INEQ x Hetero.				-0.053* (-1.78)	-0.009 (-0.32)
NB2			0.004 (0.11)	-0.060 (-0.84)	-0.055 (-0.80)
NB2 x INEQ				0.005 (0.05)	-0.021 (-0.25)
NB2 x INEQ x Hetero.				0.211*** (3.07)	0.150** (2.40)
DB			0.306*** (6.62)	0.274*** (3.19)	0.129 (1.60)
DB x INEQ				0.202** (2.08)	0.210** (2.34)
DB x INEQ x Hetero.				-0.204*** (-3.48)	-0.127** (-2.15)
Playmate's Cooperation Expected					0.210*** (7.81)
Playmate's Cooperation Expected x INEQ					0.103*** (2.84)
Playmate's Cooperation Expected x INEQ x Hetero.					-0.044* (-1.69)
Playmate's cooperation in the previous round	0.047*** (5.08)	0.047*** (5.05)	0.046*** (4.71)	0.046*** (4.56)	0.034*** (3.64)
Number of times playmate's cooperation in the 5 previous rounds	0.069*** (7.78)	0.069*** (7.90)	0.018** (2.01)	0.017** (1.85)	-0.004 (-0.78)
Number of times playmate's cooperation in a row	0.021*** (4.16)	0.021*** (4.18)	0.016*** (4.81)	0.016*** (4.52)	0.015*** (4.75)
Number of rounds	0.003* (1.69)	0.002 (0.76)	0.000 (0.09)	0.001 (0.28)	0.001 (0.66)
Observations	6768	6768	6768	6768	6768
Clusters	50	50	50	50	50
Log-pseudolikelihood	-3856.598	-3852.15	-3194.16	-3161.68	-2850.38
Pseudo-R <sup>2</sup>	0.178	0.179	0.319	0.326	0.392

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Standard errors in parentheses.

Table 3.16: Decision to cooperate with communication (Logit models - average marginal effects)

by a stronger influence of second-order normative beliefs. This leads to results 6 and 7 that follow.

**Result 6.** Communication is more beneficial when groups are unequal rather than equal, but only for homogeneous matching.

**Result 7.** First-order normative beliefs increase the likelihood of cooperation with equivalent weights when the groups are equal or unequal, while the effect of descriptive beliefs is reduced when the matching is heterogeneous and replaced by a positive effect of second-order normative beliefs.

This last result confirms that social norm beliefs play a different role depending on the type of interactions individuals are subject to. Depending on the context, the subject's decision to cooperate depends on different types of social norm beliefs. Suppose that public policies need to change behavior using information about the social norm. The type of information given depends on the type of interactions: second-order normative beliefs positively influence behavior in heterogeneous matching, while descriptive beliefs are lower in this case. First-order normative beliefs always influence behavior, whatever the context.

### 3.4 Discussion and conclusion

Inequality decreases cooperation. We find no difference in behavior specific to high- or low-type subjects but observe instead lower cooperation due to the only presence of inequality, i.e., the existence of two types of subjects. Community enforcement of cooperation is positively influenced by first-order normative beliefs and descriptive beliefs, additionally to direct expectations regarding the playmate. Inequality both decreases first- and second-order normative beliefs. The former, as well as descriptive beliefs, strongly influence the decision to cooperate. A new result is that, when communication is allowed, the influence of beliefs on cooperation is modified: under heterogeneous matching, descriptive beliefs have a lower impact on the likelihood to cooperate to the profit of second-order normative beliefs that have no impact in homogeneous matching. Social norm beliefs affect cooperative behavior differently depending on the matching. Our results give insights into the impact of inequality on social norm beliefs and the impact of the latter on cooperation when subjects interact in pairs within the same group — the type of interaction matters. While interactions between similar agents involve only the descriptive component of social norms, interactions between different agents give significance to the second-order normative component of social norms at the expense of the descriptive component.

Based on the observed behaviors in the experiment, we are able to estimate the weight of each part of social norm beliefs with a structural model accounting for the difference between homogeneous and heterogeneous matching. We expand the model of Bašić and Verrina (2023) defining individual utility as a combination of monetary concerns,  $\pi(x)$ , for choosing action  $x$ , that is to cooperate or to defect, and social norm concerns includ-

ing first-order normative beliefs,  $r$ , second-order normative beliefs,  $E(r)$ , and descriptive beliefs,  $E(g)$ .<sup>10</sup> We permit a different influence of second-order and descriptive beliefs depending on the sort of matching,  $M$ , that can be homogeneous,  $M^{hom}$ , or heterogeneous,  $M^{het}$ . We then estimate the following utility function:  $U(x) := \pi(x) - \beta_1 \times r - [\beta_2^m \times E(r) + \beta_3^m \times E(g)]\mathbb{1}(M)$ . We control for the decision maker's expected cooperation of the playmate. Second-order normative beliefs being insignificant for homogeneous matching, equation (3.1) presents the model estimation results without this variable.

$$U(x) := \pi(x) - 8.19 \times r^* - 7.20 \times E(g)^* \times M^{Hom} - (2.57 \times E(r)^* + 3.80 \times E(g)^*) \times M^{Het} \quad (3.1)$$

$r^*$  takes value 1 if the first-order beliefs of the subject are different from her actual behavior.  $E(g)^*$  and  $E(r)^*$  represent the proportion of other group members respectively not taking action  $x$  and thinking action  $x$  is not the socially appropriate behavior.<sup>11</sup> This model correctly predicts 75.8% of the decision to cooperate in the experiment, with a balanced estimation error by subject's type, round, expected cooperation from the playmate, and treatments.

From estimated coefficients, we simulate the implementation of social information campaigns, i.e., *norm-nudges* informing on either what action other subjects perform or what action other subjects think is the most appropriate, on community enforcement of cooperation.<sup>12</sup> We use the model described in Equation (3.1) to estimate the rate of cooperation in the last round played for each experimental session. We assume that the subjects correctly change their beliefs following the social information they receive. Table 3.17 presents the average of cooperation in the last round of the experimental sessions resulting from the estimations. The column "experiment" presents the average of cooperation observed in the experiment while column "Estimation" gives the average of cooperation estimated by the model defined by equation (3.1). Columns "NB2=DB=1", "NB2", and "DB" give the average of cooperation in the last round when some social information campaigns are implemented: giving the false information that all subjects in the group both think that the appropriate behavior is to cooperate and do cooperate in "NB2=DB=1", informing of the observed share of subjects in the group who think that cooperating is the appropriate behavior in "NB2", and informing of the observed share of subjects in the group who do cooperate in "DB".

<sup>10</sup>To represent monetary concerns in the game, we reduce the payoff matrix according to the individual's beliefs regarding the choice of cooperation of her playmate. Consequently, the actions have deterministic consequences on the monetary part of utility.

<sup>11</sup>Values of  $E(g)^*$  and  $E(r)^*$  are between -1 which means that everyone else in the group thinks or acts like the decision maker and 1 which means that everyone else in the group thinks or acts differently than the decision maker. Value 0 is attributed when there is no majority of action or thought in the group.

<sup>12</sup>See Cialdini et al. (2006); Schultz et al. (2007); Goldstein, Cialdini and Griskevicius (2008) for different applications of norm-nudges and (Dimant, van Kleef and Shalvi, 2020; Bicchieri and Dimant, 2022) for their backfire effects.



Treatments	Experiment	Estimation	NB2 = DB = 1	NB2	DB	NB2 and DB
EQ-N	41.25	41.25	100	41.25	65	65
EQ-EA	48.75	55	100	55	46.25	46.25
EQ-EP	65	61.25	100	61.25	60	60
INEQ-N	24.16	24.16	100	27.5	18.33	10.83
INEQ-EA	43.33	40.83	100	41.66	30.83	19.16
INEQ-EP	47.5	43.33	100	40	27.50	18.33

Table 3.17: Estimation of average cooperation based on Equation (3.1)

Simulating a perfect social norm by changing normative and descriptive beliefs toward the perception that all group members cooperate leads all subjects to cooperate, regardless of their first-order normative beliefs. Under such excessive information, the model predicts total community enforcement of cooperation. This is not the case under realistic social information. We observe that norm-nudges giving social information to the subjects do not help to improve cooperation in unequal groups. Cooperation rates are indeed lower with the implementation of norm-nudges when groups are unequal. In equal groups, the cooperation rate only increases with social information when no communication is available. These results contribute to the debate on the efficiency of using social information to change behavior (Bolton, Dimant and Schmidt, 2021; Dimant, van Kleef and Shalvi, 2020; Dimant and Shalvi, 2022). Inequality is then not favorable to the use of social information. When communication between the subjects is available, even if the group is equal, social information does not increase cooperation either. By contrast, social information focusing on the actions of other group members can be highly effective in increasing cooperation rates when the group is equal and communication is not available. In this case, social information leads each group to reach extreme distributions of cooperation in the group, with either all or almost all participants in the group who cooperate or all or almost all participants in the group who defect. Social information reduces pluralistic ignorance but cannot eliminate the conformity trap, as this type of intervention has no coordinating mechanism.<sup>13</sup>

Our findings have important implications for policies that aim to change behavior using social information and may explain their failure in some situations. In particular, we find that in an environment with similar agents, only information about others' actions can increase pro-social behavior, while in an environment with heterogeneity between agents, information about others' normative beliefs has a stronger role. Whatever the type of matching, personal values, i.e., first-order normative beliefs about the appropriate behavior to follow, affect the decision to cooperate. Our results highlight the need to rethink the design of norm-nudges more sophisticatedly, depending on the context of the decision. In situations where cooperation has already deteriorated, simply revealing actual behaviors may not be sufficient to improve cooperation. Overall, our findings highlight the need for more nuanced and complex approaches to norm-nudges to address complex situations (See Nyborg et al. (2016); Carlsson et al.

<sup>13</sup>Pluralistic ignorance represents situations in which individuals refrain from engaging in a behavior because of the belief that this behavior is not socially approved by other members of the reference group. See (Halbesleben, Wheeler and Buckley, 2005; Geiger and Swim, 2016) for some applications. Conformity traps represent situations where people do not find a way to coordinate on a new behavior, due to the fact that first movers incur a high cost of non-compliance. See (Andreoni, Nikiforakis and Siegenthaler, 2021) for experimental evidence.

(2019); Cialdini and Jacobson (2021); Dimant and Shalvi (2022); Howley and Ocean (2022) for examples in this direction). It may also be necessary to examine how to influence personal values, which are often regarded as stable but can respond to exogenous dispositions (Binmore, 2005; Bicchieri, 2005). Such values are difficult to change, but implementing education programs could progressively lead to some changes (Thaler and Sunstein, 2008). A change of personal values for only a part of the population may be sufficient to lead to the transformation of the Society. Further research on the analysis and conditions of tipping points is necessary.



## Chapter 4

# Dynamic Norms for Household Water Consumption

**Abstract** Evidence of the effectiveness of social norms information in increasing pro-social and pro-environmental behavior has been widely reported in the literature and by policy-makers. Static information on social norms has proved effective in influencing marginalized individuals in the areas of water, energy, and many others. I developed a field experiment on household water consumption by implementing a new information policy based on the literature on dynamic norms to influence average individuals. Dynamic information exploits the propagation of pro-social or pro-environmental behaviors in order to promote them. Results show a significant effect of dynamic information on below-average individuals. Yet their effectiveness disappears over time, suggesting their potential is fully exploited on special occasions such as a temporary drought. Finally, I explore explanations for the effectiveness or ineffectiveness of the different information provided by cognitive and psychological processes.

### 4.1 Introduction

Although 71% of the earth's surface is covered by water, only 3% is fresh water, of which an even smaller percentage is available for human use (Feldman, 2013). The restricted water supply is threatened by climate change and increasing water pollution (Konapala et al., 2020). In parallel, water demand is increasing every year due to population growth, rising consumption of goods due to the water needed for the construction process, and improved access to water and goods in developing countries (Jorgensen, Graymore and O'Toole, 2009; Klemeš, 2012). The small amount of water remaining is then shared by sectors with divergent interests, such as agriculture, industry, and domestic consumption. Since it is not possible to increase the amount of water available, it is necessary to

reduce demand.<sup>1</sup>

Water-related concerns were, for now, mainly focused on developing countries since the population from these countries didn't have access to enough drinking water (Gadgil, 1998; World Health Organization, 2015). However, climate change alters this trend, making water concern a worldwide priority due to its adverse effects on water supply and demand. On the one hand, climate change is increasing water demand for both domestic and agricultural sectors (Wada et al., 2013; Goodchild, 2003). On the other hand, climate change decreases water supply with rainfall reduction, water pollution, and acidity (Amarasinghe et al., 2016; Singh and Agrawal, 2008). As a consequence, serious water scarcity problems have spread in developed countries, becoming a global concern (UNEP, 2023; Shukla, Skea and Slade, 2022).<sup>2</sup> The worldwide threats of drought to humankind are threefold; first, it harms the environment by damaging ecosystems and bio-diversity (Eslamian and Eslamian, 2017); secondly, it causes direct harm to humanity by increasing mortality and disease (Bifulco and Ranieri, 2017); Finally, it jeopardizes local economies by reducing human performance and destroying agriculture (Smith, 2022).

The current study aims to propose a new tool for reducing water demand for domestic consumption. Although domestic demand accounts for only 10% of total demand (Wada, van Beek and Bierkens, 2011), it has increased by 600% between 1960 and 2014 (Otto and Schleifer, 2020) and is forecast to rise by another 30% by 2050 (Kitamori et al., 2012). This increase can be explained by the historically low price of water, which has accustomed households to wasting water (Nauges and Thomas, 2003). Given the low price elasticity for water (Howe and Linaweaver Jr., 1967; Scheierling, Loomis and Young, 2006), the limited short-term reaction of households (Nauges and Thomas, 2003), and the vital character of water (Agthe and Billings, 1987), it is essential to explore policy options that are not price-based.

The provision of information, in contrast, makes it possible to influence behavior through targeted, low-cost policies that do not constrain behavior (Haaland, Roth and Wohlfart, 2023). Information policies involving only best practices and their consequences have little effect on individual behaviors (Han, Hsu and Sheu, 2010; Owens and Driffill, 2008; Kollmuss and Agyeman, 2002). Recent literature has introduced a promising new type of information provision through the components of social norms (Allcott, 2011; Ahn, Kim and Kim, 2020; Borg, Curtis and Lindsay, 2020). Social norms are implicit rules emerging from endogenous interactions among agents and differing across populations and cultures that lead individuals to adopt a specific type of behavior (Bicchieri, 2005, 2016). They are constituted of descriptive beliefs, which refer to beliefs about the behaviors adopted by a reference group, and normative beliefs, which refer to beliefs about the social appropriateness of behaviors by a reference group. Several studies have focused on the use of so-called static information on social norms, revealing the realized behaviors and/or the social appropriateness of behaviors of a reference group at a given point in time, to influence pro-environmental behaviors (Cialdini, 2003; Bolton, Dimant and Schmidt, 2021; Bicchieri and Dimant,

---

<sup>1</sup>Although technologies such as desalination can increase water supply, they are not yet suitable for large-scale low-cost development (S. Mauter and S. Fiske, 2020).

<sup>2</sup>Spencer and Altman (2010) evaluate that almost one-third of the U.S. will encounter water scarcity issues by 2050.

2023). Regarding the water sector, Han, Hsu and Sheu (2010) and Goldstein, Cialdini and Griskevicius (2008) have succeeded in increasing towel reuse in hotels by providing information on the average reuse by former hotel members. Similarly, Landon et al. (2018) and Otaki, Ueda and Sakura (2017) have succeeded in reducing overall domestic water consumption by providing information on efficient and average household consumption.<sup>3</sup> Despite their effectiveness in reducing overall water consumption, these policies have only significantly affected high water users, with little or no effect on average consumers (Bhanot, 2021; Landon et al., 2018). As average consumers represent the majority of actors, an improvement in their behavior would substantially affect the average. In addition to having a major impact, this will influence high-water users all the more.

In order to target average consumers, I have developed a policy based on the use of so-called dynamic information on social norms, revealing the spread of pro-social or pro-environmental behavior. The underlying theory behind this mechanism lies in the fact that people tend to assume that a change in one direction will continue in the future (Hubbard, 2015; Markman and Guenther, 2007). Therefore, in order to conform to new trends, individuals may be willing to adopt these pro-social and pro-environmental behaviors. The development of mechanisms to influence the majority could ultimately contribute to reaching the tipping point toward the generalization of green behavior.

Sparkman and Walton (2017) and Mortensen et al. (2019) are the only studies, to my knowledge, that have proven the effectiveness of dynamic information through experiments in the psychological literature. Sparkman and Walton (2017) conducted information provision experiments, changing the semantics of the information from static to dynamic. Results show that dynamic information about social norms has a greater impact than static messages on the intention to consume less meat, the choice of a meatless meal in a cafe, and the use of full-load washing machines in a university residence. Likewise, Mortensen et al. (2019) show the greater impact of dynamic information over static information on the water used in a toothbrushing task in the laboratory, and on time-donation to help an environmental organization by completing an additional survey.

In this study, I test the implementation of information provision using dynamic social norms information with a framed field experiment (Harrison and List, 2004) to reduce the domestic water consumption of average households. The advantages of our application of dynamic information on domestic water consumption are twofold. First, application to real behavior in its natural environment over the long term limits the experimenter's demand effect and increases the robustness and external validity of the results. Second, the private nature of domestic water consumption reduces the observability of realized behaviors, which attenuates the effect and pressure of social norms (Nyborg et al., 2016), turning our results into a lower bound on the potential effect of our mechanisms.

We recruited 171 similar households of the Auvergne-Rhône Alpes region for a 7-week experiment, eliciting water consumption through weekly measures of their water meter. After categorizing households into four categories of water consumption, the baseline group received a static information-based policy, and the treatment group received

---

<sup>3</sup>These studies are based on the numerous implementations carried out in the energy consumption sector (Ayres, Raseman and Shih, 2013; Allcott, 2011; Schultz et al., 2007).

a dynamic information-based policy. We focus on Medium households divided into two categories, below and above average, by a discontinuity relative to average consumption.

The results show that the dynamic information-based policy induces a slightly significant reduction in water consumption (12.91% per household on average, i.e., 133.8 liters per week) in the short term for Medium households below the average. By contrast, Medium households above the average did not react significantly in the short or long term. In addition to our main results, I have also explored explanations for the effectiveness of information provision through an emotional assessment of the information policy ("PANAS - Positive and Negative Affect Schedule" Watson, Clark and Carey (1988)), the theory of planned behavior assessment (Ajzen, 1991), and a normative assessment of water-related behavior (Krupka and Weber, 2013). On one side, results show that effective dynamic information-based policy involved more positive emotional reactions and induced an appropriate belief revision of elements motivating behavior related to the theory of planned behavior. On the other side, information-based policies do not alter normative perceptions of water-related behavior.

The remainder of this paper is structured as follows. Section 4.2 outlines the design and methodology of the experiment. The results are presented in Section 4.3. In Section 4.4, we discuss the implications of our findings and conclude.

## 4.2 Experimental design

### 4.2.1 Procedure

The framed field experiment took place in the Auvergne-Rhône Alpes region of France between May 30 and July 18, 2023. We recruited 201 households for the study, of which 171 completed the experiment, from the laboratory panel of true consumers.<sup>4</sup> 56.67% of the attrition happened during the Phase 1 (before any intervention), 26.67% during the Phase 2 (after the first static intervention), 13.33% during the Phase 3 (after the treatment intervention), 3.33% at the end-of-experiment survey. Households who completed the survey were all living in an apartment with a washing machine and 74.27% with a dishwasher. Households were divided into 4.68% single-person households, 40.35% 2-person households, 23.98% 3-person households, 25.73% 4-person households and 5.26% 5-8-person households. Participants earned 20€ for correctly sending their water meter every week and completing the end-of-experiment survey. Moreover, they could earn an additional 5€ in the end-of-experiment survey. The timeline of the study, as well as the incentive, were common knowledge.

---

<sup>4</sup>This study was preregistered on AsPredicted #131119 (See Appendix C.4.4). The lower-than-expected number of participants did not alter the focus on average consumers.

## 4.2.2 Experimental Design

**The task.** The water consumption of participating households was measured using weekly photos of their water meters for 7 weeks (8 photos). Every Tuesday morning, participants received an email telling them to add their water meter to their personal cloud storage space. If participants forgot to send a photo of their water meter, they received a first reminder at noon and a second one in the evening (see Appendix Section C.4.1 for details). Participants who did not upload their photos were progressively excluded from the study each week.

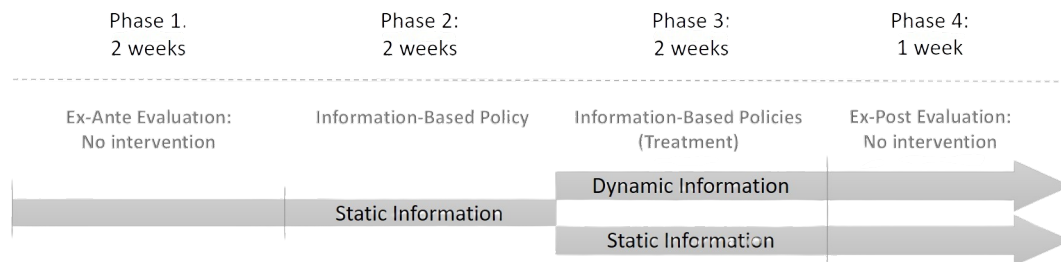


Figure 4.1: Different phases of the design

The study is divided into 4 phases over 7 weeks in a between-subject design. During the first four weeks, all participants followed the same script, composed of two phases: phase 1, with no policy, and phase 2, with a static policy based on water consumption relative to household size (See Figure 4.2). At the end of week 4 (phase 2), stratified random sampling divided participants into treatments according to the information provision of phase 2, the reference household size, the availability of a dishwasher, and whether or not they had reduced their consumption by at least 5% during the second phase. During the last 3 weeks (phases 3 and 4), 47.37% (81) households were in the baseline group, and 52.63% (90) households were in the treatment group.

**Phase 1.** During the first two weeks, water consumption is evaluated without any intervention to segregate households into 4 consumption categories: Efficient households, Medium-Low households, Medium-High households, and High households. The dissociation between the two medium categories was carried out to observe the impact of dynamic information on the discontinuity regarding the overall mean. The thresholds dividing households into 4 categories are 200, 448, and 770 liters per person per household in phase 1.<sup>5</sup> Accordingly, 8.19% (14) households were classified as Efficient, 38.60% (66) households as Medium-Low, 38.01% (65) households as Medium-High, and 15.20% (26) households as High (average households characteristics are available in Figure C.1 in Appendix).

<sup>5</sup>Household segregation thresholds between categories were chosen to exclude outliers from Medium categories while maintaining the highest possible proportion of households in the Medium categories (see Figure C.1 in Appendix).



**Phase 2.** On Wednesday of the first week of this phase, static information-based policies were delivered by email to all participants according to their consumption level (See Figure 4.2). These policies have been introduced to induce an initial reduction in consumption to build the dynamic information-based policy for phase 3. In addition, participants received a personal link to a web page containing their information-based policy as well as information on water conservation practices (See Figure C.4.2 in Appendix). During the second week, participants received emails with reminders to visit their web page. After the second phase, participants were assigned to treatment groups.

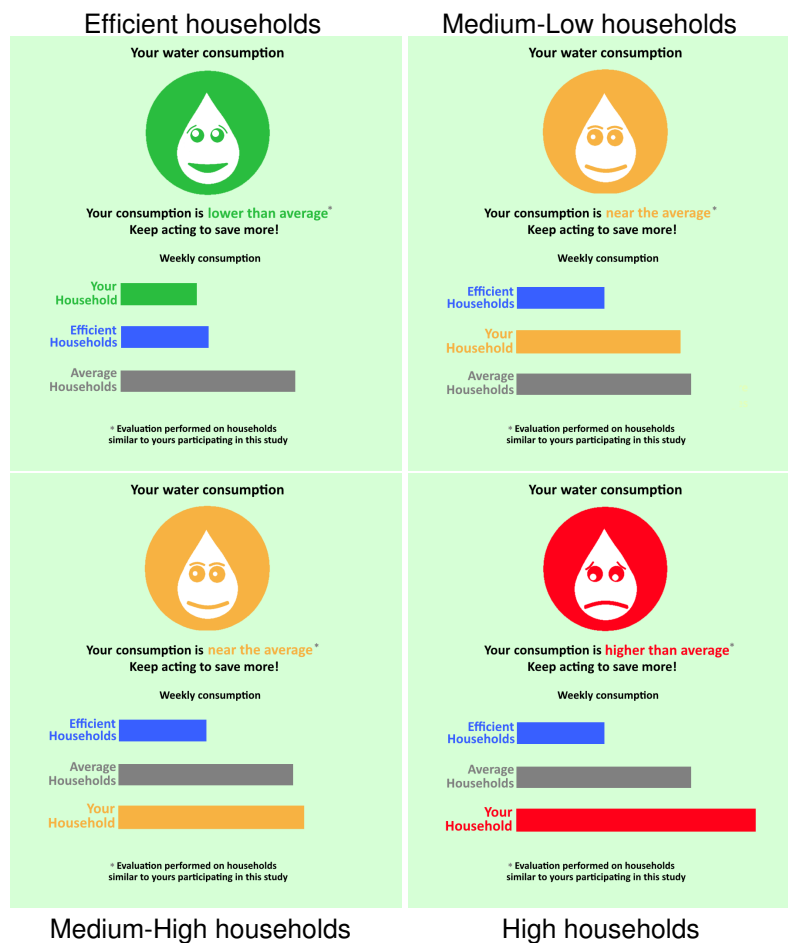


Figure 4.2: Static information-based interventions

**Phase 3.** On Wednesday of the first week of this phase, participants in the treatment group received a dynamic information-based policy (See Figure 4.3), while participants in the baseline group received the same static information-based policy as in phase 2. In addition, participants received a personal link to a web page identical to the one of phase 2, with the updated information policy. During the second week, participants received emails with reminders to visit their web page.



Figure 4.3: Dynamic information-base policy

The dynamic information-base policy announced that at least 30% of the participants in the experiment had reduced their water consumption during the second phase. Between the second and the first phase, 45.61% of households effectively reduced their consumption by at least 10% (64.28%, 33.33%, 49.29%, and 57.69% respectively, for Efficient, Medium-Low, Medium-High, and High categories).

**Phase 4.** For the last week, water consumption has been evaluated without any intervention.

### 4.2.3 Survey

At the end of the seventh week, I controlled various cognitive and psychological processes to assess better how information-based policies lead to behavior change. These assessments respond to the gap in the literature highlighted by Bicchieri (2023) (See Appendix C.4.3 for details).

First, I assessed the appropriateness of water-related behavior to observe if information-based policies induce a different perception of water behavior. To this end, I implemented the Krupka and Weber (2013) methodology framed on water conservation behavior as well as water wasting behavior. The Krupka and Weber (2013) methodology consists of a coordination game where participants must determine the level of social appropriation of water-related behavior provided by the largest number of participants. Correct answers earn participants an additional 5€.

Second, I assessed the emotional response to each information-based policy to observe the link between emotion and behavior change. To this end, I implemented a "PANAS - Positive and Negative Affect Schedule" (Watson, Clark and Carey (1988); French version by Caci and Baylé (2007)) on each information-based policy to measure the affect they generate. The Watson, Clark and Carey (1988) methodology consists of asking participants about their emotional response to the provided policies, using 10 positive and 10 negative descriptors on a 4-point Likert scale ranging from "strongly disagree" to "strongly agree".

Third, I assessed components of the Theory of Planned Behavior (Ajzen, 1991) to see which components are impacted by effective information-based policy. To this end, I implemented several questions assessing Perceived Behavioral control, Attitude toward conservation, Personal/Moral norms. As well as Expected results, Past/Current water conservation Behavior/Equipment.<sup>6</sup>

Lastly, I collected household socio-demographic information, including residential characteristics and the number of inhabitants in the household for each week of the study.

## 4.3 Results

In this study, I focus on the water consumption of the Medium-Low and Medium-High categories, which are the targeted population of the intervention. Household characteristics across categories are presented in Figure C.1 in the Appendix. I present the aggregated data for all categories, including Efficient and High categories. However, unlike the preregistration plan, I do not perform statistical analysis on these latter categories due to the small number of participants in these latter categories.

### 4.3.1 Water evaluation and treatment effect

Despite a reduction in water consumption of at least 10% for 33.33% of households in the Medium-Low and 49.29% of households in the Medium-High categories between phases 1 and 2, I find no statistical effect of static information on their consumption (Wilcoxon ranksum test comparing water consumption between phase 1 and phase 2,  $z=1.10$ ,  $p=0.272$  for Medium-Low category,  $z=0.94$ ,  $p=0.350$  for Medium-High category).

Nevertheless, a direction seems to be emerging between the categories, with an average increase of 16.08% (i.e., 123.71 liters per week) for the Medium-Low category and an average reduction of 6.98% (i.e., 92.43 liters per week) for the Medium-High category. The variation in water consumption between phases 1 and 2 are significantly different across categories (Wilcoxon ranksum test comparing variation of water consumption of Medium-Low with Medium-High categories between phase 1 and phase 2,  $z=2.49$ ,  $p=0.012$ ).

These results are consistent with Bhanot (2021) and Landon et al. (2018), which find little to no effect of static information on average consumers, with a stronger effect for the above-average consumers than under-average consumers.

---

<sup>6</sup>Chaudhary et al. (2017); Clark and Finley (2007); Lam (2006) found a significant correlation between these components and the component of the theory of planned behavior with water conservation behavior.

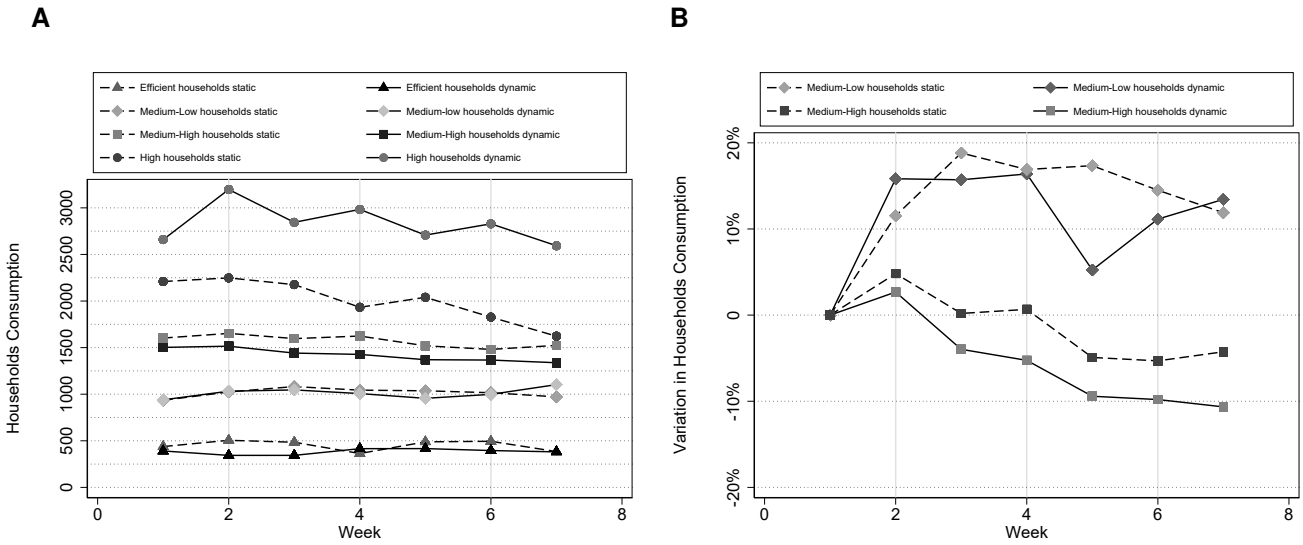


Figure 4.4: Evolution of consumption over weeks, by category and treatment. **A** Represent the average weekly water consumption by category and treatment. **B** represent the average percentage of variation of water consumption compared to the first week by category and treatment.

The effect of our treatment compares the impact of dynamic information on social norms with that of repeated static information on social norms (phase 3). Figures 4.4A-B represent respectively household consumption by category and treatment, and household variation in water consumption compared to week 1 by category and treatment.

7

As illustrated in figure 4.4B, the Medium-Low category has significantly more reduced consumption with the dynamic information (average reduction of 10.76%) than with the static information (average increase of 11.43%) in phase 3 (Wilcoxon ranksum test comparing variation in consumption of dynamic and static information in the Medium-Low category between phase 3 and phase 2,  $z=2.61$ ,  $p=0.009$ ). However, the Medium-High category responded in the same way to both information (average reduction of 8.81% and 6.30%, respectively, for dynamic and static information, Wilcoxon ranksum test comparing variation in consumption of dynamic and static information in the Medium-High category between phase 3 and phase 2,  $z=1.36$ ,  $p=0.17$ ). Statistical tests indicate that dynamic information leads to a significant and large reduction in water consumption in the medium-low category.

<sup>7</sup>The fragmentation of the high category is induced by its low number of households combined with a high standard deviation (1251) due to the presence of outliers. In comparison, the medium-low and medium-high categories have similar standard deviations of 518 and 585, respectively.

	Medium-Low (1)	Medium-High (2)
Treatment x Post	-133.8* (-1.86)	8.80 (0.09)
Treatment	-66.05 (-1.03)	18.42 (0.19)
Post	21.41 (0.41)	-117.1 (-1.40)
Living area size ( $m^2$ )	0.56 (0.25)	4.53 (1.34)
Dishwasher	68.93 (0.64)	46.34 (0.46)
Paying for water	-56.92 (-0.54)	-327** (-2.15)
Nb. Men	148.8*** (2.78)	311.1*** (3.69)
Nb. Women	235.5*** (5.78)	146.7 (1.53)
Nb. Son	293.6*** (6.42)	173.4 (1.55)
Nb. Daughter	232.9*** (3.81)	256.5*** (2.98)
Income	52.62* (1.92)	-4.44 (-0.11)
Owner	-118.5 (-1.14)	-47.71 (-0.32)
Observations	330	325
Cluster	66	65
$R^2$	0.41	0.39

*t* statistics in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table 4.1: Water consumption (OLS models - average marginal effects). Cluster at the household level. Notes: Regression was conducted on weeks 1-5. Men and Women are over 13 years old, while Son and Daughter are under 13 years old.

I used ordinary least squares regression clustered at the household level to model the difference in difference of water consumption with control for household characteristics. The results confirm the significant decrease in water consumption following the dynamic information on the Medium-Low category. Modeling until week 5 to measure the short-term effect of our dynamic information, I find a slightly significant average treatment effect with a reduction in water consumption of 133.7 liters for the medium-low category. However, the effect of our dynamic information does not persist in the long term (See Table C.2 and C.3 in Appendix). Regarding the Medium-High category, I do not find short or long-term effects of the dynamic information. Our results emphasize the large and slightly significant impact of dynamic information on Medium-Low households in the short term, with no long-term difference between dynamic information and the repetition of static information. Meanwhile, Medium-High households did not respond differently to dynamic information in the short or long term compared to the repetition of static information.

### 4.3.2 Cognitive and Psychological processes

In this part, I investigate the role of cognitive and psychological processes on the impact of information-based policies on water consumption. I will now refer to participants instead of households, as only one member of each household took part in the survey.

**Normative assessment** Table 4.2 shows the average normative appropriateness of water conservation and water waste behavior for each treatment and category.<sup>8</sup> Results show no significant differences between categories or treatments that might have been induced by the different information presented. Whereas Figure C.2 in the Supplementary shows the relatively shared knowledge of the social appropriateness of water conservation behaviors (with a concentration of 87% of participants with a positive score above 0.5), it also shows a broader disparity of the social inappropriateness of water wasting behavior (with only 61% of participants with a negative score under -0.5). These results suggest that it could be more effective to reinforce the social inappropriateness of wasting behaviors, and hence to revise beliefs, rather than highlight the social appropriateness of conservation behaviors, which may not alter the value attached to norms. However, OLS regression of water reduction per capita in phase 3 by Cognitive and Psychological processes (Table C.4 in the Supplementary) only finds an effect of social appropriateness of water conservation behavior on Medium-High households in the wrong direction. The normative assessment of water conservation and wasteful behavior is neither influenced by the provision of information, nor is it responsible for the reduction in water consumption in phase 3.

	Efficient participants	Medium-Low participants	<i>p-value</i> of Mediums	Medium-High participants	High participants
Water conservation behavior with DI	0.551	0.677	0.524	0.707	0.722
<i>p-value</i> of DI vs SI		0.344		0.551	
Water conservation behavior with SI	0.725	0.733	0.710	0.788	0.712
Water wasting behavior with DI	-0.352	-0.509	0.952	-0.521	-0.375
<i>p-value</i> of DI vs SI		0.594		0.558	
Water wasting behavior with SI	-0.625	-0.473	0.309	-0.551	-0.640

Table 4.2: Normative assessment of water-related behaviors. Notes: Score between -1 and 1, with -1 referring to low level and 1 to high level. *p-value* given by Wilcoxon statistical test. DI refers to Dynamic Information, and SI to Static Information.

**Emotional assessment** Table 4.3 shows the average arousal for positive and negative valences in response to each policy implemented in phase 3.<sup>9</sup> Results show that the dynamic information induced significantly more positive emotion for both Medium categories, as well as less negative emotion in the Medium-High category (Wilcoxon ranksum test of positive emotion  $z=2.16$ ,  $p=0.030$  for Medium-Low, and  $z=1.95$ ,  $p=0.051$  for Medium-High; Wilcoxon ranksum test of negative emotion  $z=4.35$ ,  $p<0.001$  for Medium-High). However, Table C.4 in the Supplementary finds

<sup>8</sup>Average normative appropriateness score represents the average of 12 items (for more detail see section ?? and C.4.3 in Supplementary).

<sup>9</sup>Average arousal score represents the average of 10 items (for more detail see section ?? and C.4.3 in Supplementary).

no link between emotional valence induced by our information provision and the reduction of consumption in phase 3. Dynamic information induces stronger positive emotions than static information in Medium participants and less negative emotion in Medium-High participants, but emotions do not seem to directly contribute to the reduction in water consumption in phase 3.

	Efficient participants	Medium-Low participants	<i>p-value</i> of Mediums	Medium-High participants	High participants
Positive emotions with DI	0.144	0.132	0.356	0.107	0.13
<i>p-value</i> of DI vs SI		0.030**		0.051*	
Positive emotions with SI	0.453	0.004	0.852	-0.009	-0.15
Negative emotions with DI	-0.639	-0.654	0.864	-0.714	-0.497
<i>p-value</i> of DI vs SI		0.169		0.000***	
Negative emotions with SI	-0.7	-0.488	0.001***	-0.271	-0.075

Table 4.3: Emotional evaluation of information-based policies. Notes: Score between -1 and 1, with -1 referring to low level and 1 to high level. *p-value* given by Wilcoxon statistical test. DI refers to Dynamic Information, and SI to Static Information.

**Theory of planned behavior assessment** Table C.5 in the Supplementary shows the average score for components related to the theory of planned behavior by category and treatment.<sup>10</sup> Results show that participants provided with the dynamic information exhibited significantly more Moral/Personal norms in the Medium-Low category, as well as significantly more Perceived control and significantly less Expected results in the Medium-High category (Wilcoxon ranksum test in the Medium-Low category  $z=2.04$ ,  $p=0.041$  for moral/personal norms; Wilcoxon ranksum in the Medium-High category  $z=3.05$ ,  $p=0.002$  for perceived control, and  $z=2.06$ ,  $p=0.040$  for expected result). In addition, Table C.4 in the Supplementary shows that Moral/Personal norms significantly explain a part of the reduction in per capita water consumption during phase 3 for the Medium-Low category, while Expected results significantly explain a part of the reduction in per capita water consumption during phase 3 for the Medium-High category. Whereas the increase in Moral/Personal norms exhibited with dynamic information in the Medium-Low category favors the reduction of water consumption, the decrease in Expected results exhibited with the dynamic information in the Medium-High category disfavors the reduction of water consumption.

**Cognitive and psychological processes** In this part, we have seen that dynamic information is correlated with several elements linked to cognitive and psychological processes, suggesting an impact of information on these components. In the Medium-Low category, individuals who have received the dynamic information feature greater arousal of positive emotions and Moral/Personal norms. Moral/Personal norms had a significant impact on reducing water consumption in phase 3, which may explain the effectiveness of the information policy in this category. In the Medium-High category, individuals who have received the dynamic information feature lower Expected results. However, as this element is linked to the reduction of water consumption, this may explain the ineffectiveness of the information policy in this category.

<sup>10</sup>Information on the construction of each component are available in section ?? and C.4.3 of Supplementary).

## 4.4 Discussion

This article examines the efficacy of information provision in the field of water consumption to influence average consumers, who represent the majority of the population. I implemented a new type of information, using dynamic information about the social norms, in face to traditional static information about the social norms.

By creating a discontinuity in the information provided between below- and above-average households, the dynamic information revealed a slightly significant average treatment effect only on below-average consumers that have not responded to the first static information (average decrease of 205 liters per household per week) in the short term. This reduction can be partly explained by the increased Moral/Personal norms which is correlated with the provision of dynamic information. Meanwhile, above-average households have not reacted to our policy, either in the short or long term. The policy's ineffectiveness on above-average households can be partly explained by a reduction in Expected Results which is correlated with the provision of dynamic information.

The effect of providing information about social norms on below-average consumers is promising. Firstly, it shows that it is possible to use information on social norms to influence average behavior. Secondly, since it concerns a large proportion of individuals, this change could have a considerable impact on the evolution of the average. It could, therefore, contribute to the creation of new norms that will eventually impact personal preferences. Thirdly, it reminds us of the effect of information provision, which can be used to increase the pro-social and pro-environmental behavior of most individuals at low cost without changing the incentives.

Our results have significant implications for public policy on water management. I show how low-cost information provision can induce an average reduction of over 12.91% of domestic water consumption for below-average households per week, which represents a considerable quantity on a city scale. For example, in a city of 500,000 inhabitants with around 30-40% below-average households, the water saved in one week is equivalent to the amount of water needed to produce 45,000-60,000 kilograms of potatoes (Pimentel et al., 1997). However, as I have only observed a short-term effect from our specific policy, it should be used on special occasions, such as during a period of drought. Our results are especially important as domestic water-related behaviors are private behaviors that are less subject to social pressure (Nyborg et al., 2016). Consequently, the use of dynamic information on visible behaviors could have greater and long-term effects.

Nevertheless, further studies are needed to consolidate our results and increase their robustness. First, it would enable us to test the provision of dynamic information on different samples. On the one hand, this will allow us to apply our policy to a larger number of households and eliminate any link with potential field specificity. On the other hand, it will enable us to test cross-cultural implications. Cultural differences could imply different attitudes to social norms, as well as different appreciation towards water. Secondly, it would allow for the implementation of dynamic information provision alone, as opposed to our implementation, which follows a first static information provision. Lastly, further research is needed to clarify the relationship between information policy, behavior, and



these components for the development of more effective policies. The increase in pro-social and pro-environmental behaviors induced by these policies could contribute to reaching tipping points towards the generalization of green behaviors.



## Chapter 5

# Disentangling Individual Attitudes, Expectations, and Experience in the Adoption of Proenvironmental Behavior

*Co-authored with Régis Ferrière*

**Abstract** How do attitudes, descriptive and normative expectations, and environmental experience influence the adoption of a costly, actively proenvironmental behavior? The answer fundamentally depends on whether proenvironmental individuals represent a minority or majority population. We show that a strong proenvironmental attitude may not be sufficient to allow a minority of the population to express their preferences. Furthermore, even if the population is composed of a majority of proenvironmental attitudes among which the active behavior spreads, this may not be sufficient for the active behavior's adoption among individuals with business-as-usual attitude. Regarding the impact of the environment, our analysis generalizes the previous finding that a small but not negligible dose of environmental feedback is necessary for the propagation of the active behavior. Altogether, fixed gains influenced by non-material payoffs and attitude strength, together with the degree of individual focus in forming their descriptive and normative expectations, are key to the widespread adoption of active behavior.

### 5.1 Introduction

An effective response to global environmental challenges requires that a large fraction of the populations that are most directly responsible for the environmental impact adopts virtuous behaviors. Focusing on climate warming driven by anthropogenic greenhouse gas emissions, the wealthiest countries are overwhelmingly responsible, and

some regions of their geographic range begin to suffer severe consequences, such as mega wildfires in the western US and Canada, more frequent, intense, and wide-ranging tornados in the eastern US, and unprecedented regimes of summer heat throughout. For as long as climate experts and national and international environmental organizations have been alerted to the looming climate crisis, the required large-scale transitions in individual behavior have not happened.

From governmental and industrial leaders with their far-reaching decisions to all individual citizens through their local actions, fundamentally, the same kind of behavioral change is required – from business-as-usual to actively seeking reduced environmental impacts, in spite of potential direct costs to the individual. Given the risk of collective losses that are at stake, game-theoretic models have repeatedly demonstrated the long-run superiority of cooperative solutions. But individuals are not rational *Homo economicus*, and the long-term risk of collective losses makes a cognitively weak force of behavioral change.

While costly changes driven by long-term benefits are elusive, behavior may shift rapidly in response to short-timescale changes in the social and environmental context. Behavioral change is shaped by attitudes (personal norms), social interactions leading to first-order expectations about the behaviors of others (descriptive beliefs) and second-order expectations about others' personal norms (normative beliefs), and private environmental experience. This is according to the theory of planned behavior as extended by (Schill et al., 2019) to reflect that both social and environmental contexts can influence behavior. Importantly, as individuals' behavior changes, social and environmental contexts may change. For example, environmental action by some individuals may result in individuals updating their descriptive beliefs and perceiving the environment as less vulnerable to further degradation – leading to adjusting their actions or intentions accordingly. Thus, individuals and their social and environmental contexts are entangled in a feedback loop.

The perceived environmental feedback may add to the perceived cost of behavior to limit the spread of pro-environmental intention or action. This is the case when individuals tend to act when their private perception of the environment indicates a high level of environmental degradation or vulnerability, whereas they tend to relax when external cues signal improvement in the environmental situation. In a population where individuals do not differ in their attitude towards environmental action, (Ecotièrre et al., 2023) developed a simple model of the individual-social-environmental feedback loop to identify conditions under which individuals may consistently renounce business-as-usual and engage in pro-environment (costly) behavior. They found that the consistent adoption of costly active behavior is possible but requires a very strong effect of the descriptive expectations and a small but non-negligible dose of environmental feedback. The dose of environmental feedback is critical to allow the pro-environmental behavior to rise from low frequency, and past a frequency tipping point, the strong social effect drives it to adoption

across the population.

The dynamics of proenvironmental behavior may be quite different for more realistic populations composed of individuals differing in their attitude towards environmental action. Here we address the most straightforward case involving two attitudes, one favoring proenvironment action (the  $a$  attitude) and the other, business-as-usual ( $b$  attitude). (Ecoti ere et al., 2023) model can be extended to investigate how the presence of the  $b$  population influences the adoption of the active behavior in the  $a$  population as well as overall. We know from (Ecoti ere et al., 2023) that individual sensitivity to the perceived environment and peer social pressure are key parameters. Factors that also come into play in the extended model include the strength of the attitude - how behaviorally rigid or flexible individuals of a given attitude are; the weights of the descriptive and normative expectations in the social context; and the focus of individuals on one type or the other when forming descriptive and normative expectations. The model presented here is designed to evaluate the relative influence of these factors and their interaction on the dynamics of the active behavior in the  $a$  and  $b$  subpopulations.

## 5.2 Results

### 5.2.1 Model overview

We consider a large population of agents that interact among themselves and perceive the state of the environment (more or less degraded) through a single environmental variable,  $e$ . Individual behavior and their environmental impact are modeled on a short enough timescale so that the population, as a set of individuals, does not change (i.e., no birth or death). The  $e$  variable measures the perceived environmental state on a continuous scale, with larger  $e$  corresponding to an environment perceived as more degraded. The  $e$  variable can be seen as an indicator or summary statistics of the perceived level of environmental degradation, whose variation is driven by the population level of environmental action, intention, or awareness, such as the spread of renewable energy, the adoption of plant-based diets, the reduced consumption of non-essential goods, or the prominence of pro-environmental demonstrations and other public calls for proenvironmental action.

Two attitudes are represented in the population: attitude  $a$  is actively proenvironment, and attitude  $b$  prefers 'business as usual'. Each agent is characterized by its attitude; agents do not change attitude. At any time  $t$ , an individual of attitude  $a$  or  $b$  can express one of two behaviors with respect to the environment: active (denoted by  $A$ ), which aligns with attitude  $a$ , or baseline (denoted by  $B$ ) aligned with attitude  $b$ . When expressing behavior  $A$ , an agent actively seeks to reduce their negative environmental impact compared to the baseline impact of behavior  $B$ .

This is modeled by assuming that an agent in state  $A$  (irrespective of its attitude) increases the perceived environmental impact of the population by an amount  $l_A$ , which is less than the environmental impact,  $l_B$ , of behavior  $B$ . We also assume that  $l_A$  is positive, meaning that the population influences the environment by its own existence. In a population where all agents express behavior  $A$  ( $B$ , respectively), the rate of change of the environment perceived as minimally (maximally) degraded is proportional to  $l_A$  ( $l_B$ ), and the stationary value of the perceived environmental state is  $l_A$  ( $l_B$ ). In a population where both behaviors are expressed, the perceived environmental state varies between  $l_A$  and  $l_B$ .

Agents do not change attitude, but they can switch between behaviors  $A$  and  $B$  as a result of social encounters and their individual perception of the environmental state. We define  $\kappa$  as a scaling parameter controlling the rate of social encounters. The rate at which an agent changes its behavior upon encountering another agent depends on the attractiveness of the alternate behavior, which is determined by four factors: the relative material payoff (perceived payoff differential between the two behaviors); the cognitive pressure to align the individual's attitude and behavior, or 'attitude pull'; behavioral conformity driven by descriptive expectations; and social disapproval driven by normative expectations.

We use parameters  $\gamma_i$ ,  $i = A$  or  $B$ , the payoff from adopting behavior  $i$ , and denote the payoff difference between behaviors  $A$  and  $B$ , or payoff differential, by  $\beta = \gamma_A - \gamma_B$ . We say that the active behavior  $A$  is costly when the payoff differential,  $\beta$ , is negative. The payoff differential may be positive if, for example, the active behavior  $A$  is incentivized through public policy. The 'pull' of attitude  $a$  ( $b$ , respectively) is measured by parameter  $\alpha_a$  (resp.  $\alpha_b$ ). When an individual forms first-order (descriptive) and second-order (normative) expectations,  $a$  individuals receive a fraction  $\rho_{aD}$  (for descriptive expectations formation) or  $\rho_{aN}$  (for normative expectations formation) of the attention, or focus, of the expectations-forming individual. The weight of descriptive and normative expectations in the individual decision is measured by parameters  $\delta_D$  and  $\delta_N$ , respectively.

The individual behavioral decision based on social encounters is then determined by a fixed gain (frequency-independent) and a variable gain (frequency-dependent). For an  $a$ -individual engaged in behavior  $B$ , the chance of meeting an individual ( $a$  or  $b$ ) engaged in behavior  $A$  is  $(x_a P_a + x_b(1 - P_a))$ . Then, the rate of switching to behavior  $A$  is given by the sum of the fixed gain

$$G_A = \underbrace{(\gamma_A - \gamma_B)}_{\text{material payoff}} + \underbrace{\alpha_a}_{\text{attitude}} + \underbrace{\delta_N \theta_N \rho_{aN} P_a - \delta_N \theta_N (1 - \rho_{aN})(1 - P_a)}_{\text{social disapproval}} \quad (5.1)$$

and variable gain (frequency-dependent descriptive expectation)

$$H_A = \delta_D \theta_D (\rho_{aD} x_a P_a + (1 - \rho_{aD}) x_b (1 - P_a)) - \delta_D \theta_D (\rho_{aD} (1 - x_a) P_a + (1 - \rho_{aD}) (1 - x_b) (1 - P_a)). \quad (5.2)$$

Likewise, for a  $b$  individual, the rate of switching from behavior  $A$  to behavior  $B$  (in the context of social encounters) is given by the sum of the fixed gain of behavior  $B$

$$G_B = (\gamma_B - \gamma_A) + \alpha_b + \delta_N \theta_N (1 - \rho_{aN})(1 - P_a) - \delta_N \theta_N \rho_{aN} P_a \quad (5.3)$$

and variable gain

$$H_B = \delta_D \theta_D (\rho_{aD} (1 - x_a) P_a + (1 - \rho_{aD}) (1 - x_b) (1 - P_a)) - \delta_D \theta_D (\rho_{aD} x_a P_a + (1 - \rho_{aD}) x_b (1 - P_a)). \quad (5.4)$$

The fixed gain of a behavior in a given subpopulation thus depends on the attitude pull,  $\alpha_a$  or  $\alpha_b$ , and captures the effect of the normative expectations; this effect itself is controlled by the product of the weight of normative expectations,  $\delta_N$ , attitude frequency,  $P_a$  or  $1 - P_a$ , and attitude focus,  $\rho_{aN}$  or  $1 - \rho_{aN}$ . Thus, varying these parameters (especially the  $\alpha$ s and  $\delta_N$ , which are positive and not bounded by one) can turn the fixed gain of the costly active behavior  $A$  from negative to positive. It is also to be noted that  $G_B = -G_A + \alpha_a + \alpha_b$ .

Any agent may also switch behavior at any time based on their assessment of the environmental state. Agents tend to adopt the alternate behavior when they perceive the environmental impact of their current behavior as relatively large compared to the alternate behavior. Accordingly, we use parameter  $\tau$  to set the timescale of the behavioral switch based on individual assessment and model the rates of switching from  $B$  to  $A$  and from  $A$  to  $B$  as  $\tau(e - l_A)$  and  $\tau(l_B - e)$ , respectively. Finally, we define parameter  $\ell$  as representing the timescale at which individuals' behavior affects the perceived environmental state: the higher  $\ell$ , the faster the perceived environmental state changes due to individuals' behavior.

In the Methods, we integrate the model components described above into a system of differential equations for the joint dynamics of the frequency of behavior  $A$  in subpopulations  $a$  and  $b$  and the perceived environmental degradation. Notations are listed in Table 5.1. Note that throughout the paper, the observabilities are assumed to be maximum and equal to 1, i.e.,  $\theta_N = \theta_D = 1$ . The model enables us numerically to experiment on the effect of different components and parameters of the behavior-environment system on its dynamics and attractor(s). The structure of the following analysis reflects our particular interest in better understanding the effect of normative expectations as influenced by the attitudes' pull, the weight of normative vs. descriptive expectations, and the difference between the attitudes' pull – in the social and environmental context set by the attitudes' frequencies and the strength of the population's perception of and influence on the environmental state.

## 5.2.2 Effect of attitudes and attitude strength

In this subsection, we make the simplifying assumption that the pull, descriptive focus, and normative focus are the same for both attitudes  $a$  and  $b$ , that is,  $\alpha_a = \alpha_b$ ,  $\rho_{aD} = 0.5$ ,  $\rho_{aN} = 0.5$ . The weight of the normative expectations is set to  $\delta_N = 2$ .

In a pure population of proenvironmental individuals ( $a$  attitude,  $P_a = 1$ ), the adoption of the costly active behavior  $A$  strongly depends on the strength of the attitude and an interaction between descriptive expectations and the environmental feedback (Fig. D.1). If the attitude pull is weak, the fixed gain of behavior  $A$  is positive but close to zero. In this case, adoption of  $A$  is still possible if the environmental feedback is weak enough and for sufficiently strong descriptive expectations (Fig. D.1A-C). Descriptive expectations that are even stronger, however, cause bistability (Fig. D.1A-C). In this case, interestingly, increasing the environmental feedback helps with the adoption of  $A$ : with an initially high frequency of behavior  $B$ , the environmental state is perceived as degraded, which promotes switching to behavior  $A$ ; this can drive the frequency of  $A$  above the bistability tipping point, leading eventually to a steady state with high  $A$  frequency.

The effect of attitude pull is fully mediated by the fixed gain of behavior  $A$ . In a mixed population, with attitudes  $a$  and  $b$ , the attitudes' pull is mediated by both the fixed gain  $G_A$  of  $A$  among  $a$  individuals and the fixed gain  $G_B$  of  $B$  among  $b$  individuals (equations (1) and (3)). These fixed gains also include the normative expectations, and thus depend on the attitude frequency  $P_a$ . In general, stronger attitudes (of both  $a$  and  $b$  individuals), by increasing the fixed gains  $G_A$  and  $G_B$ , favor the adoption of  $A$  in the  $a$  population and  $B$  in the  $b$  population (Fig. 5.1, compare Fig. 5.1A-C, 5.1D-F, 5.1G-I). On the other hand, increasing the  $a$  attitude frequency,  $P_a$ , also facilitates the adoption of behavior  $A$  in the  $a$  population (through the fixed gain  $G_A$ ) (Fig. 5.1, compare Fig. 5.1A,D,G, 5.1B,E,H, 5.1C,F,I). But  $P_a$  also influences descriptive expectations (variable gain), which interacts with the individuals' environmental sensitivity.



When a low frequency of  $a$  individuals is sufficient to make the net gain of  $A$  positive, behavior  $A$  will spread in the  $a$  population provided descriptive expectations are not too strong – otherwise, descriptive expectations overcome the (positive) fixed gain and impose the spread of behavior  $B$  among  $a$  individuals, due to the widespread expression of  $B$  in a population largely dominated by  $b$  individuals (Fig. 5.1G). We note that the effect of stronger descriptive expectations (larger  $\delta_D$ ) can be attenuated by stronger environmental sensitivity (larger  $\tau$ ) (Fig. 5.1G).

In populations with  $a$  individuals at high frequency, a high frequency of  $A$  in the  $a$  population may be reached at a steady state or along a stable cycle (Fig. 5.1C, F, I, Fig. 5.2A-B). Even if behavior  $A$  is adopted in the dominant  $a$  population,  $A$  does not necessarily spread in the  $b$  population. However, the rise of the strength of descriptive expectations can drive the system through a turning point where the frequency of behavior  $A$  rises steeply in the  $b$  subpopulation. In this case, a high frequency of  $A$  behavior can be established in the whole population, either permanently at a steady state, or periodically along a stable cycle (Fig. 5.1C, Fig. 5.2B). As in the case of a pure population ( $P_a = 1$ ) (Fig. D.1), the occurrence of stable cycles is favored by slowing down environmental reactivity (lower  $l$ ) relative to the timescale of environmental sensitivity set by  $\tau$  (given the timescale of social interactions, set by parameter  $\kappa$ ) (Fig. D.2).

To sum up, if proenvironmental individuals (attitude  $a$ ) only form a minority (low  $P_a$ ), descriptive expectations strongly oppose the adoption of the active behavior ( $A$ ), even in the  $a$  subpopulation. For  $A$  to spread among  $a$  individuals, the effect of descriptive expectations must be countered by strong attitudes (high  $\alpha_a = \alpha_b$ ) because stronger attitudes increase the fixed gain. The adoption of  $A$  can also be facilitated by environmental sensitivity. When proenvironmental individuals make the majority, strong descriptive expectations can drive the adoption of the  $A$  behavior in both  $a$  and  $b$  subpopulations. In this case, slow environmental reactivity, relative to the timescale of environmental sensitivity, can result in fluctuations of  $A$  and  $B$  frequencies along a stable cycle alternating high and low frequencies.

### 5.2.3 Effect of normative expectations

The effect of normative expectations is entirely mediated by the fixed gains of behavior  $A$  in the  $a$  subpopulation and  $B$  in the  $b$  subpopulation. In the case where normative expectations-forming individuals focus equally on attitudes  $a$  and  $b$  ( $\rho_{aN} = 0.5$ ), we have

$$G_A = \beta + \alpha_a + \delta_N(P_a - 0.5) \quad (5.5)$$

and

$$G_B = -\beta + \alpha_b - \delta_N(P_a - 0.5). \quad (5.6)$$

Hereafter, we contrast two cases:  $a$  individuals either occur at low frequency in a population of individuals with strong attitudes, or at high frequency in a population of individuals with weak attitudes (Fig. 5.3). The first case corresponds to a situation where business-as-usual is the dominant attitude, opposed by a small group of individuals with a proenvironmental attitude, in a population where all individuals have strong attitudes towards the environment. We illustrate this case numerically with  $P_a = 0.25$  and  $\alpha_a = \alpha_b = 2.5$  (Fig. 5.3A). The second case addresses the situation where the active behavior may be dominant but potentially undermined by the weakness of attitudes. We illustrate this case numerically with  $P_a = 0.9$  and  $\alpha_a = \alpha_b = 0.5$  (Fig. 5.3B).

In the first case ( $a$  individuals make a minority, attitudes are strong), the fixed gain of behavior  $B$  in the  $b$  type,  $G_B$ , is always a positive, increasing function of the strength of normative expectations,  $\delta_N$ . In contrast, the fixed gain of behavior  $A$  in the  $a$  type,  $G_A$ , is a decreasing function of  $\delta_N$ , turning from positive to negative as  $\delta_N$  crosses a threshold (equal to 6 in our numerical example, Fig. 5.3A). As a consequence, for a given strength of the descriptive expectations,  $\delta_D$ , the  $A$  behavior may become established in the  $a$  population provided  $\delta_N$  remains below a maximum value, which decreases as  $\delta_D$  increases, due to the dominance of  $b$  individuals in the population (Fig. 5.3A). Greater environmental sensitivity, measured by  $\tau$ , attenuates the effect of descriptive expectations and results in a wider range of  $\delta_N$  for which the  $A$  behavior becomes adopted by the  $a$  type, although its equilibrium frequency decreases as  $\tau$  increases (Fig. D.3A). In the  $b$  type, the steady state is always one of very low  $A$  frequency (Fig. 5.3A, Fig. D.3A).

In the second case, ( $a$  individuals make a majority, attitudes are weak, Fig. 5.3B), the strength of normative expectations,  $\delta_N$ , drives variation in the fixed gains,  $G_A$  and  $G_B$ , in the opposite direction: as  $\delta_N$  increases,  $G_A$  increases from negative to positive, while  $G_B$  decreases from positive to negative. As a consequence, the population can be in one of three regimes depending on  $\delta_N$  (Fig. 5.3B). Below a lower threshold on  $\delta_N$ ,  $G_A$  is negative (but increases with  $\delta_N$ ) while  $G_B$  is positive (but decreases with  $\delta_N$ ). Then behavior  $A$  is disfavored in both  $a$  and  $b$  types. At best for the adoption of  $A$ , bistability occurs if descriptive expectations are strong enough. Above a higher threshold on  $\delta_N$ ,  $G_A$  is positive (and increases with  $\delta_N$ ) while  $G_B$  is negative (and decreases with  $\delta_N$ ). Then, behavior  $A$  is adopted at high frequency among both  $a$  and  $b$  individuals. For  $\delta_N$  between these two thresholds, the adoption of behavior  $A$  in the  $b$  population is possible but critically dependent on descriptive expectations being strong enough, with larger  $\delta_N$  facilitating the spread of  $A$ . When this happens,  $A$  is also expected to reach high frequency in the  $a$  type. When  $A$  does not spread among  $b$  individuals, as a result of both  $\delta_N$  and  $\delta_D$  being too low,  $A$  reaches intermediate frequencies among  $a$  individuals, or fluctuates in frequency along a stable cycle (and  $A$  in  $b$  as well, results not shown).

To sum up, the effect of normative expectations strength, together with attitudes pull, is entirely mediated by the fixed gains of behavior  $A$  in the  $a$  type and behavior  $B$  in the  $b$  type. The effect depends on the  $a$  type frequency,  $P_a$ , which is also a factor of descriptive expectations. When  $a$  is a minority and attitudes are strong, stronger normative expectations make it more difficult for  $A$  to reach high frequency, and all the more so as descriptive expectations are stronger. In fact, there is a turning point on  $\delta_N$  above which the frequency of  $A$  in the  $a$  type shifts from high to low. When  $a$  makes a majority and attitudes are weak, the effect of  $\delta_N$  is turned around, whereby increasing  $\delta_N$  tends to promote the adoption of  $A$  in both  $a$  and  $b$  types. Both descriptive and normative expectations are essential determinants of the adoption of  $A$  in the  $b$  type, whereas the equilibrium frequency of  $A$  in the  $a$  type is more sensitive to the strength of normative expectations. Furthermore, varying the latter can drive transitions between equilibria of low vs. high frequency of  $A$  through regimes of bistability or stable cycles.

## 5.2.4 Effect of heterogeneity in attitudes' strength

Here, we ask how variation in attitude strength between the two types can influence the adoption of the costly active behavior. Specifically, we examine whether a weaker attitude in the  $b$  type can facilitate the spread of  $A$ . We hypothesize that the difference in attitude strength interacts with variation in focus on  $a$  vs.  $b$  attitudes in the descriptive and normative expectations, measured by  $\rho_{aD}$  and  $\rho_{aN}$ ; and that the strength of descriptive and normative expectations further modulates this effect.

To evaluate this hypothesis, we proceed with the following numerical experiment (Fig. 5.4, Fig. D.4-D.8). We compare populations differing in the frequency of the  $a$  type ( $P_a$ ) with either equal, strong attitudes of  $a$  and  $b$  ( $\alpha_a = \alpha_b = 2.5$ ), or a strong attitude in  $a$  and a weak attitude in  $b$  ( $\alpha_a = 2.5$ ,  $\alpha_b = 0.5$ ). For each scenario, we sample combinations of strength for the descriptive expectations ( $\delta_D$ ) and normative expectations ( $\delta_N$ ). Finally, for given  $P_a$ ,  $\alpha_a$  and  $\alpha_b$ , and  $\delta_D$  and  $\delta_N$ , we determine the population dynamics and compute the equilibrium frequency of behavior  $A$  in types  $a$  and  $b$  as a function of focus on type  $a$  in the descriptive expectations ( $\rho_{aD}$ ) and normative expectations ( $\rho_{aN}$ ).

When the  $a$  frequency is very low (see  $P_a = 0.1$  in Fig. 5.4A-B and Fig. D.5-D.6), behavior  $A$  generally fails to be adopted in the  $a$  subpopulation (as well, of course, as in the  $b$  subpopulation) and a weaker attitude in the  $b$  type does not change that. This outcome, however, depends on the attitude focus,  $\rho_{aD}$  and  $\rho_{aN}$ , in the descriptive and normative expectations. When  $a$  individuals are very focused on their own type to form both their descriptive and normative expectations ( $\rho_{aD}$  and  $\rho_{aN}$  close to 1), then  $A$  can be adopted in the  $a$  subpopulation – an effect that, again, dominates that of the reduction in  $b$  attitude strength (Fig. 5.4B, Fig. D.6).

Increasing the frequency of  $a$  individuals (see  $P_a = 0.5$  in Fig. 5.4C-D and Fig. D.7-D.8) lowers the thresholds on  $\rho_{aD}$  and  $\rho_{aN}$  above which behavior  $A$  is adopted among  $a$  individuals. In addition, at the largest values of  $\rho_{aD}$  and  $\rho_{aN}$ , the adoption of  $A$  in the  $b$  subpopulation becomes possible (Fig. 5.4C-D). The effect is especially pronounced when the fixed gain of  $B$  in the  $b$  type is reduced by a weaker attitude among  $b$  individuals (Fig. 5.4D, Fig. D.8).

The strength of descriptive and normative expectations,  $\delta_D$  and  $\delta_N$ , scale the threshold values of  $\rho_{aD}$  and  $\rho_{aN}$  above which the  $a$  and  $b$  subpopulations shift from behavior  $B$  to behavior  $A$ . They do so in opposite direction in the  $a$  and  $b$  types, in the sense that the larger  $\delta_D$  and/or  $\delta_N$ , the lower the thresholds on  $\rho_{aD}$  and/or  $\rho_{aN}$  for  $A$  to spread in  $b$ , but the higher the thresholds on  $\rho_{aD}$  and/or  $\rho_{aN}$  for  $A$  to spread in  $a$  (Fig. D.5-D.8).

To sum up, the attitude focus, both in the descriptive and normative expectations, turns out to be a critical determinant of the adoption of proenvironmental behavior  $A$ . Increasing type- $a$  focus promotes the spread of  $A$  in the  $a$  subpopulation;  $b$  individuals having a strong or weak attitude makes little difference. In populations where the  $a$  type is more frequent, if descriptive and normative expectations are strong, large  $a$ -type focus can even promote the adoption of  $A$  in both  $a$  and  $b$  types, an effect that is facilitated in populations where the attitude of  $b$  individuals is weaker.

## 5.3 Discussion

How can a costly, actively proenvironmental behavior spread in a population where some individuals favor proenvironmental action while others prefer business as usual? We answer this question by constructing and analyzing a simple model of behavioral choice, in which social encounters and private experiences of the environmental state shape individual decisions. As they interact with others, individuals decide to switch behavior or not based on the material costs of the behaviors, descriptive expectations, and normative expectations. As they experience the environment, individuals favor action (at a cost) when the environment is perceived as degrading, and otherwise relax from the active behavior.

Our model extends a previous study ((Ecoti re et al., 2023)) by assuming variation among individuals in their attitude towards the environment. The importance of such personal components is paramount, as they represent a significant part of the decision-making process (B nabou and Tirole, 2006; Ba i c and Verrina, 2023). Non-compliance with them imposes a cost on the individual that can be related to cognitive dissonance (Festinger, 1962). Furthermore, having two subpopulations with different environmental attitudes goes with the social approval/disapproval of

a subpopulation, which is represented in the model with a normative component ((Cialdini, Reno and Kallgren, 1990; Bicchieri, 2005, 2016)). As a result, the mere existence of a subpopulation exerts conformist pressures through the behaviors' fixed gains. To keep our model simple, focus on behavioral decisions, and also in accordance with the literature (Binmore, 2005; Bicchieri, 2005), we consider attitudes to be fixed over the model timescale.

The adoption of a costly, actively proenvironmental behavior thus depends, in the model, on attitudes, descriptive and normative expectations, and environmental experience. The outcome fundamentally depends on whether proenvironment individuals represent a minority or the majority population. We show that even if the proenvironment attitude is strong, the active behavior may not spread among the proenvironment individuals when they only form a minority. And even if the proenvironment group makes the majority, this may not be sufficient for the spread of the active behavior to the rest of the population.

When proenvironment individuals represent a minority, the adoption of the active behavior may only occur within the proenvironment subpopulation. Even therein, a strong proenvironment attitude may not be enough; the weight of descriptive and normative expectations must also be low enough. Then, the anti-conformism effect of the environmental feedback facilitates the active behavior adoption by further countering the conformism effect of descriptive expectations.

When proenvironment individuals make the majority, these individuals tend to adopt the active behavior under the effect of descriptive expectations combined with the environmental feedback. However, this is not always the case, and normative expectations turn out to play a critical role. Indeed, strong descriptive expectations promote the *A* adoption if normative expectations also weigh heavily in the social context, whereas strong descriptive expectations oppose the *A* adoption if the social weight of normative expectations is too low. Furthermore, the active behavior can also spread among individuals with 'business-as-usual' attitude. This occurs if descriptive and normative expectations have enough weight in the decision process, with the additional condition that individuals are strongly focused on their own type when forming their expectations. In this case, a weaker business-as-usual attitude will have little influence on the adoption of *A* by proenvironmental individuals, but will facilitate the adoption of *A* by the business-as-usual type. The conjunction of expectations weighing heavily in the decision process, strong same-type focus, and weak attitude in the business-as-usual type make the adoption of active behavior possible in both types, even if the proenvironment subpopulation makes a large minority instead of a majority.

Our results resonate with several lines of previous research. The positive and negative effects of descriptive beliefs have been highlighted and discussed on many issues Cialdini (2003); Cialdini et al. (2006). Moreover, the fact that it creates bi-stability when it takes on high values, making initial conditions decisive, can be associated with

conformity traps. (The trap of conformity keeps individuals in equilibria that differ from their attitude, even though they represent a majority. Because of the cost of non-compliance with majority behavior, individuals maintain a behavior inconsistent with their attitude. This phenomenon, as well as a coordinating mechanism enabling reaching an equilibrium respecting the attitudes of the majority, can be found in Andreoni, Nikiforakis and Siegenthaler (2021).) Similarly, the importance of the force of attitudes that can overcome social and environmental components can be associated with the concept of personal norms. (Personal norms represent the internalization of a norm. They are predominant and difficult to influence, as the expectations of others are no longer relevant (Bicchieri, 2005, 2016).) Furthermore, variation in  $\rho$  could be the consequence of implementing norm-nudge. Norm-nudge represents the use of normative and/or descriptive information about the behavior and attitude of individuals to influence behavior without restriction. Although the implementation of norm-nudge on a heterogeneous population has been studied (Dimant, 2024), we have not found a representation similar to our  $\rho$  that interacts with the proportion of a subpopulation. Finally, our results highlight the importance of the normative component. In contrast with many previous experiments showing little effect of normative expectations in homogeneous populations ((Bicchieri and Xiao, 2009; Bicchieri and Mercier, 2013; Bicchieri, Dimant and Sonderegger, 2023)), we find here, as in Teyssier & Wiczorek 2023, that the normative component of the behavioral decision is an essential determinant of behavior in heterogeneous interactions.

Our model can help guide environmental decision-making by simulating the implementation of idealized public policies. By manipulating specific variables, one can mimic the effect of public policies from transition paths to steady state. Three policy types can be represented in our model and are described below. First, the implementation of taxes and subsidies to adjust the relative costs of behavior can be simulated by varying  $\gamma_A$  and  $\gamma_B$ . Their variation impacts fixed gains, which has a critical effect on the system dynamics. Similarly, information and education policies can be simulated by varying  $\alpha_a$  and  $\alpha_b$ , and have similar effects via changes in fixed gains. Secondly, beliefs can be influenced through observability levels and access to behavioral information. Setting up specific infrastructures or programs to make behavior or attitude more visible can alter the effect of the descriptive or normative components. The effectiveness of such strategies can be assessed by varying  $\theta$ . The same evaluation can be carried out on information policies by varying  $\rho$ , which enables the social component to be focused on a specific population, as we demonstrated earlier. Thirdly, the relationship between individuals and their environment can be altered. Public decision-makers can design and implement laws and infrastructures to modify the perceived environmental state. This may be reflected in the model by changing the environmental sensitivity,  $\tau$ .

Considering the previous examples of equilibrium trajectories (Fig. 5.2), we show how little intervention could be used to overcome transition or tipping points. A tax policy, an information policy, and an education policy can impact behavior differently but lead to a similarly improved equilibrium (compare Fig. 5.5A, B, and D). However, unlike the

others, the information policy does not suppress bi-stability (Fig. 5.5C). Without a proper assessment of the current state of the environment, the information policy (in this specific case) may or may not advance proenvironmental action. We note that even in scenarios where there is no bi-stability or limit cycles (Fig. 5.2A-C), convergence to a globally stable equilibrium (Fig. 5.2 does not prevent subpopulation trajectories from following complex transients.

Modeling social norms with feedback effects was called for by Nyborg et al. (2016). Furthermore, their discussion highlights the need for considering other behavioral drivers for a more predictive understanding of environment-related behavior. Here, we briefly discuss some possible extensions of our model. This model assumes perfect and complete information as in (Ecoti re et al., 2023). The assumption of complete information could be relaxed to allow for misinterpretations of the model's component (on  $x_a, x_b, P_a$ ) and develop a private history and belief based on individuals' different encounters. In addition, people's attitudes ( $\alpha$ ) could be more flexible, allowing them to anchor their beliefs or move from one subpopulation to another. Furthermore, individuals with different attitudes may express different environmental sensitivities ( $\tau$ ). Correlation between attitude ( $\alpha$ ) and individual sensitivity to the environment ( $\tau$ ) may have interesting consequences for the behavioral inertia of individuals. With a broader perspective, the literature on network theory and innovation theory could also be employed to develop the encounter mechanism and define the construction of individuals' beliefs Centola (2010); Guilbeault and Centola (2021); Centola (2021); Horsevad et al. (2022). Tipping points could also be set up to represent the effect that the law or authorities have on social norms, as assumed in Weitz et al. (2016) and Gavrillets (2021).

## 5.4 Methods

We denote the proenvironmental attitude by  $a$  and the business-as-usual attitude by  $b$ . As in (Ecoti re et al., 2023), we denote the active (proenvironmental) behavior by  $A$  and the baseline (business-as-usual) behavior by  $B$ . At any time  $t$ , the frequency of behavior  $A$  is denoted by  $x_a$  in the  $a$  subpopulation and  $x_b$  in the  $b$  subpopulation. Using the parameters defined in Table 1 and the rationale presented in the main text (section Results, Model overview), we obtain the following system of differential equations for the frequencies  $x_a$  and  $x_b$  and the environment state variable,  $e$ . In the explanatory terms,  $a_A$  ( $a_B$ , respectively) or  $b_A$  ( $b_B$ , respectively) stands for individuals with attitude  $a$  or  $b$  expressing behavior  $A$  ( $B$ , respectively). See (Ecoti re et al., 2023) for a detailed discussion of the model timescales.

$$\begin{aligned}
\frac{dx_a}{dt} = & \overbrace{\kappa(1-x_a)(x_a P_a + x_b(1-P_a))}^{a_B \text{ encounters with } a_A \text{ and } b_A} \times \text{Max}[0, \\
& \overbrace{\underbrace{\underbrace{\gamma_A - \gamma_B}_{\text{material payoff}} + \underbrace{\alpha_a}_{\text{attitude}} + \underbrace{\delta_N \theta_N \rho_{aN} P_a - \delta_N \theta_N (1 - \rho_{aN})(1 - P_a)}_{\text{social disapproval}}}_{\text{Fixed gain}}}_{\text{material payoff}} \\
& + \underbrace{\delta_D \theta_D (\rho_{aD} x_a P_a + (1 - \rho_{aD}) x_b (1 - P_a)) - \delta_D \theta_D (\rho_{aD} (1 - x_a) P_a + (1 - \rho_{aD}) (1 - x_b) (1 - P_a))}_{\text{behavioral conformity}}] \\
& + \overbrace{\kappa x_a ((1 - x_a) P_a + (1 - x_b) (1 - P_a))}^{a_A \text{ encounters with } a_B \text{ and } b_B} \times \text{Max}[0, \\
& \overbrace{\underbrace{\underbrace{\gamma_B - \gamma_A}_{\text{material payoff}} - \underbrace{\alpha_a}_{\text{attitude}} + \underbrace{\delta_N \theta_N (1 - \rho_{aN})(1 - P_a) - \delta_N \theta_N \rho_{aN} P_a}_{\text{social disapproval}}}_{\text{Fixed gain}}}_{\text{material payoff}} \\
& + \underbrace{\delta_D \theta_D (\rho_{aD} (1 - x_a) P_a + (1 - \rho_{aD}) (1 - x_b) (1 - P_a)) - \delta_D \theta_D (\rho_{aD} x_a P_a + (1 - \rho_{aD}) x_b (1 - P_a))}_{\text{behavioral conformity}}] \\
& + \underbrace{\tau(e - l_A)(1 - x_a) - \tau(l_B - e)x_a}_{\text{environmental feedback}}
\end{aligned} \tag{5.7}$$

$$\begin{aligned}
\frac{dx_b}{dt} = & \overbrace{\kappa(1-x_b)(x_a P_a + x_b(1-P_a))}^{b_B \text{ encounters with } a_A \text{ and } b_A} \times \text{Max}[0, \\
& \overbrace{\underbrace{\underbrace{\gamma_A - \gamma_B}_{\text{material payoff}} - \underbrace{\alpha_b}_{\text{attitude}} + \underbrace{\delta_N \theta_N \rho_{aN} P_a - \delta_N \theta_N (1 - \rho_{aN})(1 - P_a)}_{\text{social disapproval}}}_{\text{Fixed gain}}}_{\text{material payoff}} \\
& + \underbrace{\delta_D \theta_D (\rho_{aD} x_a P_a + (1 - \rho_{aD}) x_b (1 - P_a)) - \delta_D \theta_D (\rho_{aD} (1 - x_a) P_a + (1 - \rho_{aD}) (1 - x_b) (1 - P_a))}_{\text{behavioral conformity}}] \\
& + \overbrace{\kappa x_b ((1 - x_a) P_a + (1 - x_b) (1 - P_a))}^{b_A \text{ encounters with } a_B \text{ and } b_B} \times \text{Max}[0, \\
& \overbrace{\underbrace{\underbrace{\gamma_B - \gamma_A}_{\text{material payoff}} + \underbrace{\alpha_b}_{\text{attitude}} + \underbrace{\delta_N \theta_N (1 - \rho_{aN})(1 - P_a) - \delta_N \theta_N \rho_{aN} P_a}_{\text{social disapproval}}}_{\text{Fixed gain}}}_{\text{material payoff}} \\
& + \underbrace{\delta_D \theta_D (\rho_{aD} (1 - x_a) P_a + (1 - \rho_{aD}) (1 - x_b) (1 - P_a)) - \delta_D \theta_D (\rho_{aD} x_a P_a + (1 - \rho_{aD}) x_b (1 - P_a))}_{\text{behavioral conformity}}] \\
& + \underbrace{\tau(e - l_A)(1 - x_a) - \tau(l_B - e)x_a}_{\text{environmental feedback}}
\end{aligned} \tag{5.8}$$

$$\frac{de}{dt} = \ell e(l_A(x_a P_a + x_b(1 - P_a)) + l_B((1 - x_a) P_a + (1 - x_b)(1 - P_a)) - e) \tag{5.9}$$



We explored the model dynamics over large ranges of parameter values. In the analysis reported here, the parameter ranges were chosen to reflect the diversity and patterns of dynamics observed in our initial numerical exploration.

Using Mathematica(c), for each parameter combination, the number of equilibria was evaluated, and for each equilibrium, its local stability was assessed by computing the eigenvalues of the associated Jacobian matrix.

Table 5.1: **Notations, description, and parameter default values.**

<b>Notation</b>	<b>Parameter description</b>	<b>Default value</b>
$A, B$	Active behavior, baseline behavior	
$a, b$	Proenvironment attitude, business-as-usual attitude	
$x_a, x_b$	Frequencies of active behavior in $a$ and $b$ subpopulations	
$e$	Perceived environmental state	
$\kappa$	Encounter rate	1
$P_a$	Frequency of subpopulation A	
$\gamma_A, \gamma_B$	Material payoff of behavior A, B	$\gamma_A = 1, \gamma_B = 2$
$\beta$	Payoff differential	-1
$\alpha_a, \alpha_b$	Pull of attitudes $a, b$	
$\delta_D, \delta_N$	Weight of descriptive, normative expectations	
$\theta_D, \theta_N$	Observability of behavior	1
$\rho_{aD}, \rho_{aN}$	Focus on $a$ subpopulation	0.5
$G_A, G_B$	Fixed gain of behavior $A$ in type $a$ , fixed gain of behavior $B$ in type $b$	
$H_A, H_B$	Variable gain of behavior $A$ in type $a$ , variable gain of behavior $B$ in type $b$	
$\tau$	Individual sensitivity to the environment	0.5
$l_A, l_B$	Individual environmental impact of behavior $i$	$l_A = 0.1, l_B = 1$
$l$	Environmental reactivity	1

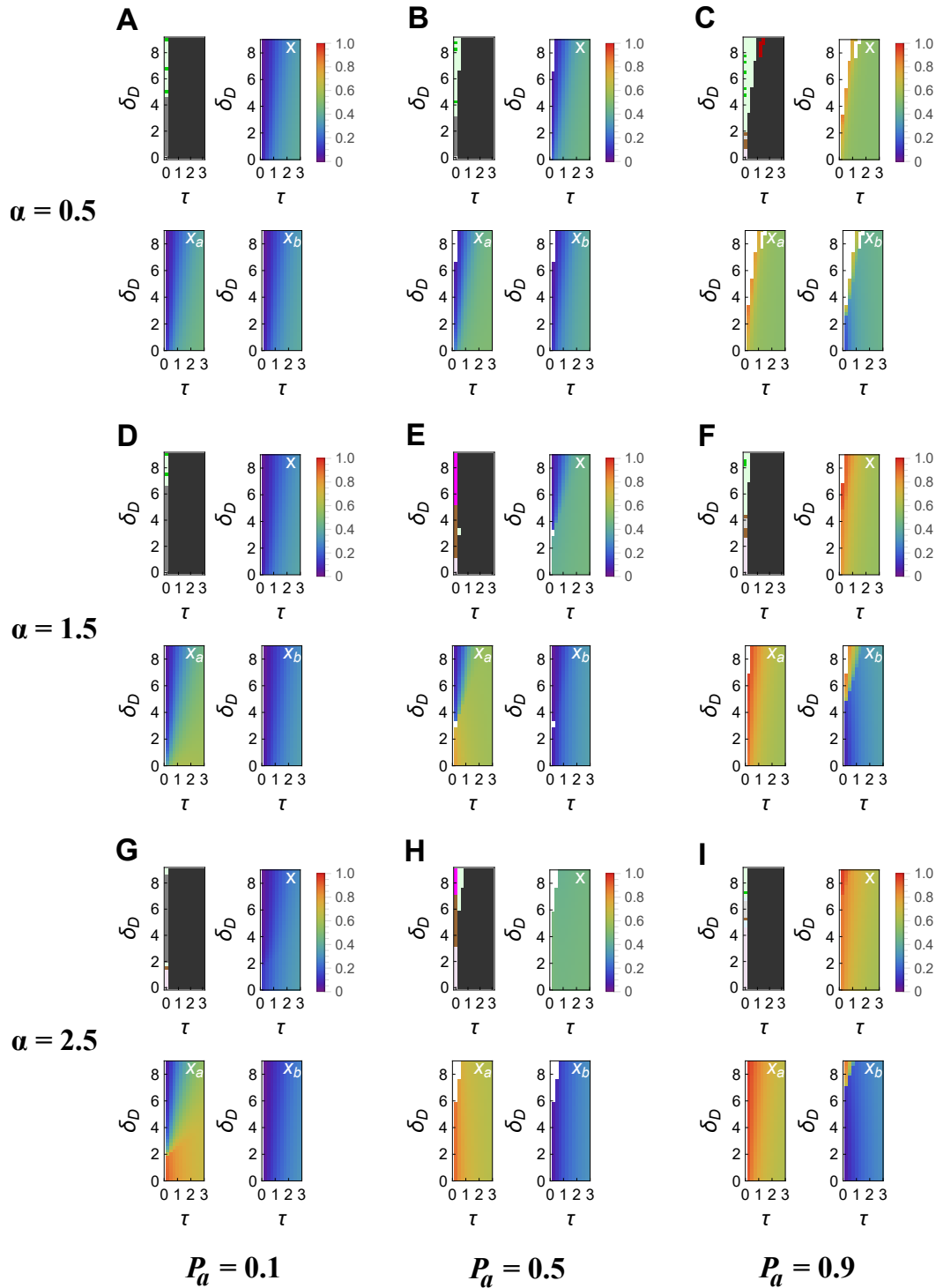


Figure 5.1: **Effect of attitudes and attitude strength on the adoption of the active behavior.** For each value of proenvironmental attitude frequency,  $P_a$ , and attitude strength,  $\alpha_a = \alpha_b = \alpha$ , the equilibrium frequency of the active behavior is shown as a function of the environmental sensitivity,  $\tau$ , and strength of the descriptive component of the social norm,  $\delta_D$ . For each panel, **A** through **I**, the upper left graph displays the nature of positive equilibria (black: one stable equilibrium, light green: bistability with two stable and one unstable equilibrium, red: one unstable equilibrium and one stable limit cycle); the upper right graph shows the overall frequency of the active behavior when there is a single stable equilibrium (white area indicates otherwise); bottom left and right graphs show the frequency of the active behavior at equilibrium (when it exists, otherwise the area is left white) in the  $a$  and  $b$  subpopulations, respectively. **A**,  $P_a = 0.1$ ,  $\alpha = 0.5$ . **B**,  $P_a = 0.5$ ,  $\alpha = 0.5$ . **C**,  $P_a = 0.9$ ,  $\alpha = 0.5$ . **D**,  $P_a = 0.1$ ,  $\alpha = 1.5$ . **E**,  $P_a = 0.5$ ,  $\alpha = 1.5$ . **F**,  $P_a = 0.9$ ,  $\alpha = 1.5$ . **G**,  $P_a = 0.1$ ,  $\alpha = 2.5$ . **H**,  $P_a = 0.5$ ,  $\alpha = 2.5$ . **I**,  $P_a = 0.9$ ,  $\alpha = 2.5$ . Parameters are set to their default values (Table 5.1). The strength of the normative component of the social norm is  $\delta_N = 2$ . Environmental reactivity is  $l = 1$ .

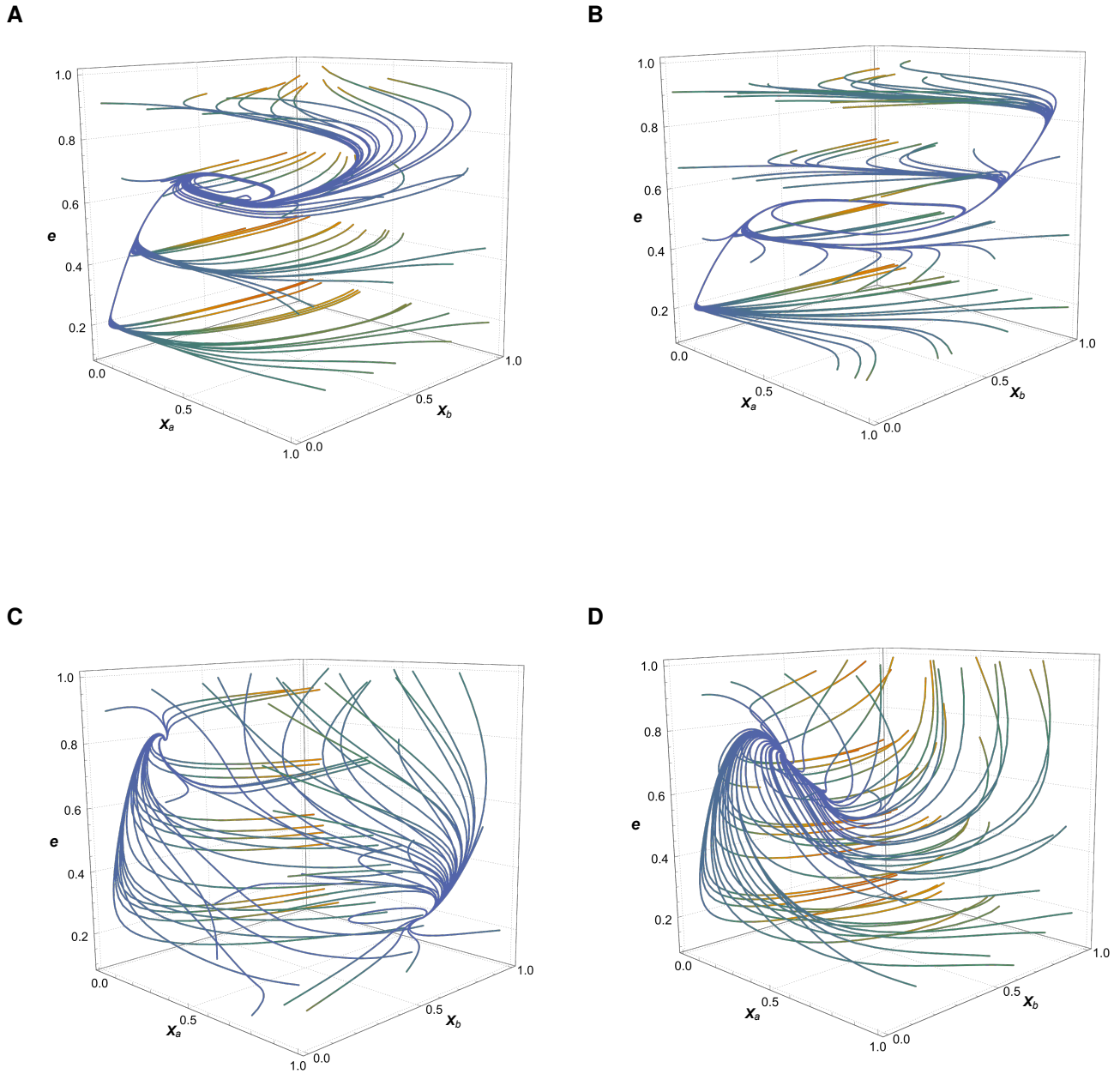


Figure 5.2: **Exemplary trajectories in the model state space** ( $A$  frequency in the  $a$  subpopulation,  $A$  frequency in the  $b$  subpopulation,  $e$  environmental state) from multiple initial conditions, in four dynamical regimes: **A**, small limit cycle. **B**, large limit cycle. **C**, bistability. **D**, globally stable equilibrium with nonmonotonic transients. **A**,  $\kappa=1$ ,  $\theta_i=1$ ,  $\rho_{ai}=0.5$ ,  $l=0.1$ ,  $l_A=0.1$ ,  $l_B=1$ ,  $\gamma_A=1$ ,  $\gamma_B=2$ ,  $\alpha_i=1.5$ ,  $P_a=0.5$ ,  $\tau=0.8$ ,  $\delta_D=5$ ,  $\delta_N=2$ . **B**,  $\kappa=1$ ,  $\theta_i=1$ ,  $\rho_{ai}=0.5$ ,  $l=0.1$ ,  $l_A=0.1$ ,  $l_B=1$ ,  $\gamma_A=1$ ,  $\gamma_B=2$ ,  $\alpha_i=0.5$ ,  $P_a=0.9$ ,  $\tau=1.25$ ,  $\delta_D=6$ ,  $\delta_N=2$ . **C**,  $\kappa=1$ ,  $\theta_i=1$ ,  $\rho_{ai}=0.5$ ,  $l=1$ ,  $l_A=0.1$ ,  $l_B=1$ ,  $\gamma_A=1$ ,  $\gamma_B=2$ ,  $\alpha_i=0.5$ ,  $P_a=0.9$ ,  $\tau=0.5$ ,  $\delta_D=6$ ,  $\delta_N=1.75$ . **D**,  $\kappa=1$ ,  $\theta_i=1$ ,  $\rho_{ai}=0.5$ ,  $l=1$ ,  $l_A=0.1$ ,  $l_B=1$ ,  $\gamma_A=1$ ,  $\gamma_B=2$ ,  $\alpha_i=1.5$ ,  $P_a=0.5$ ,  $\tau=0.5$ ,  $\delta_D=4$ ,  $\delta_N=4$ .

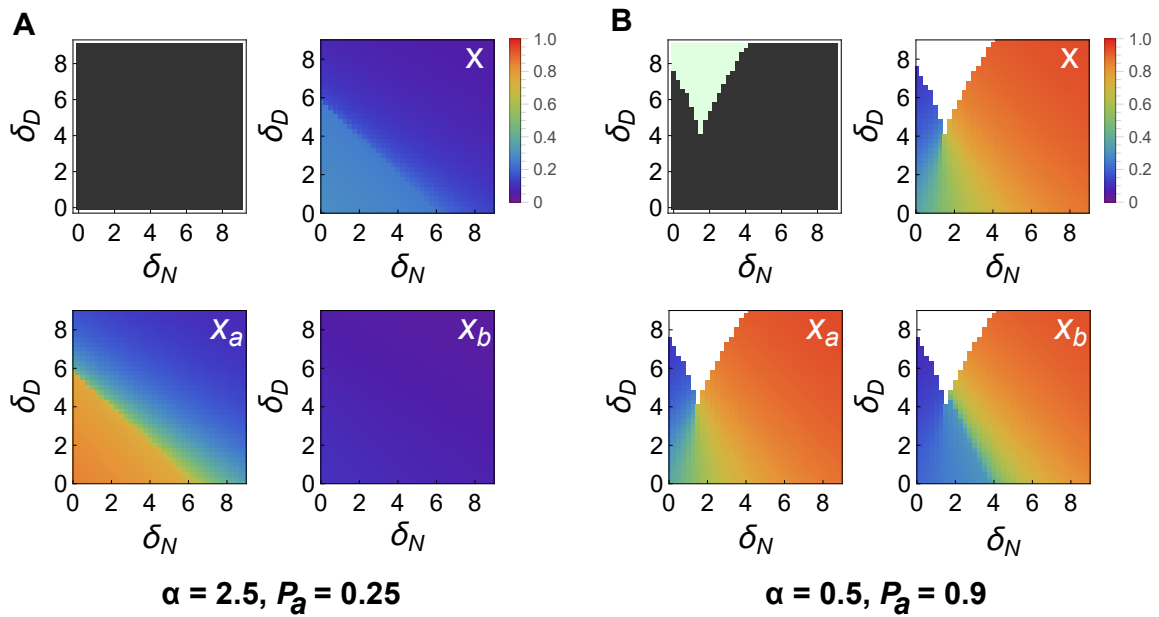


Figure 5.3: **Effect of the strength of the descriptive and normative components of the social norm,  $\delta_D$  and  $\delta_N$ , on the adoption of the active behavior.** **A**, attitude strength is  $\alpha_a = \alpha_b = \alpha = 2.5$  and the proenvironmental attitude frequency is  $P_a = 0.1$ . **B**, attitude strength is  $\alpha_a = \alpha_b = \alpha = 0.5$  and the proenvironmental attitude frequency is  $P_a = 0.9$ . See Fig. 5.1 for details. Parameters are set to their default values (Table 5.1). Other parameters: environmental reactivity is  $l = 1$  and environmental sensitivity is  $\tau = 0.5$ .

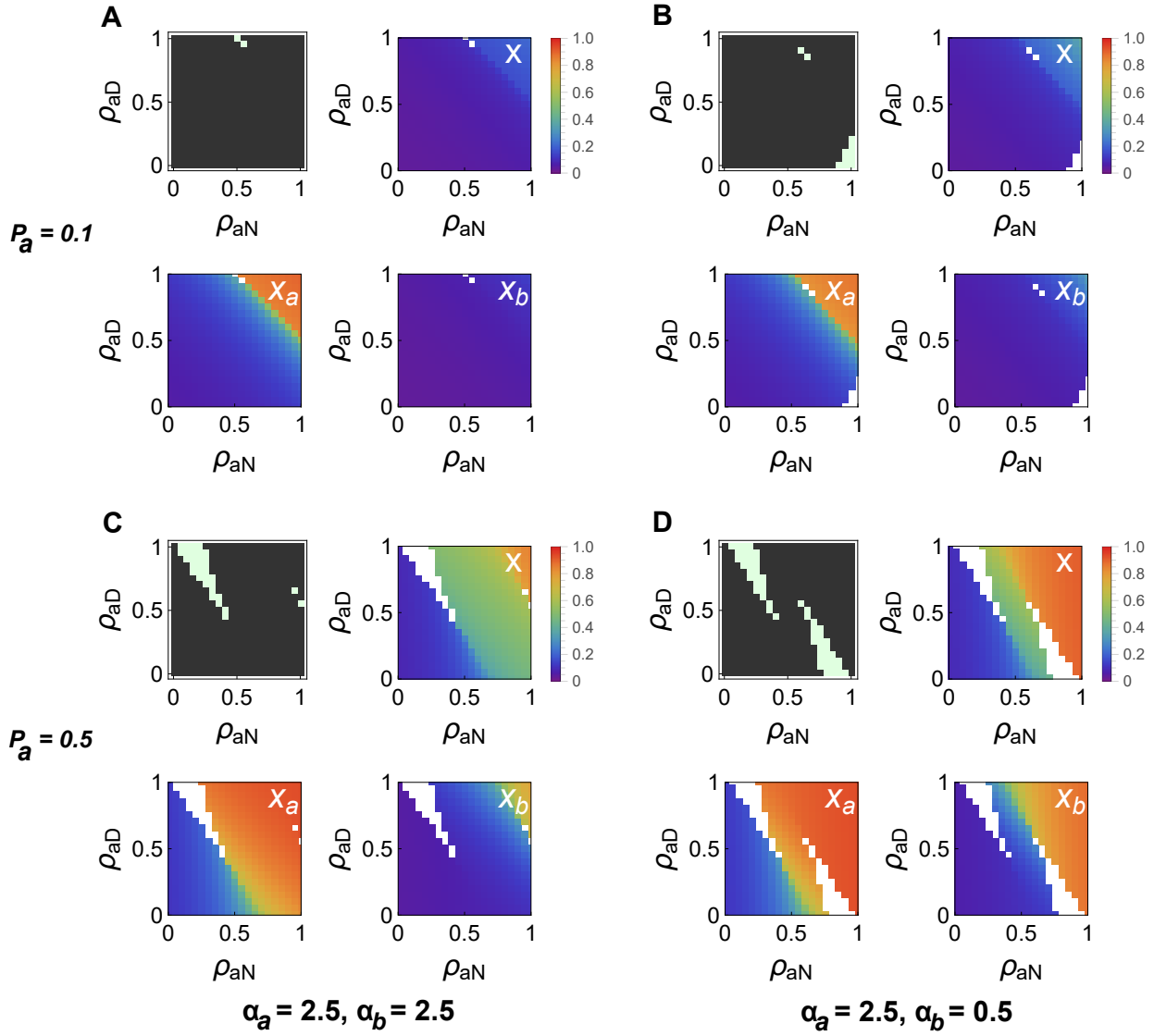


Figure 5.4: **Effect of descriptive focus,  $\rho_{aD}$ , and normative focus,  $\rho_{aN}$ , on the adoption of the active behavior.** **A**,  $P_a = 0.1$ ,  $\alpha_a = 2.5$ ,  $\alpha_b = 2.5$ . **B**,  $P_a = 0.1$ ,  $\alpha_a = 2.5$ ,  $\alpha_b = 0.5$ . **C**,  $P_a = 0.5$ ,  $\alpha_a = 2.5$ ,  $\alpha_b = 2.5$ . **D**,  $P_a = 0.5$ ,  $\alpha_a = 2.5$ ,  $\alpha_b = 0.5$ . See Fig. 5.1 for details. Parameters are set to their default values (Table 5.1). Other parameters: strength of descriptive expectations is  $\delta_D = 6$ , strength of normative expectations is  $\delta_N = 4.75$ , environmental reactivity is  $l = 1$ , environmental sensitivity is  $\tau = 0.5$ .

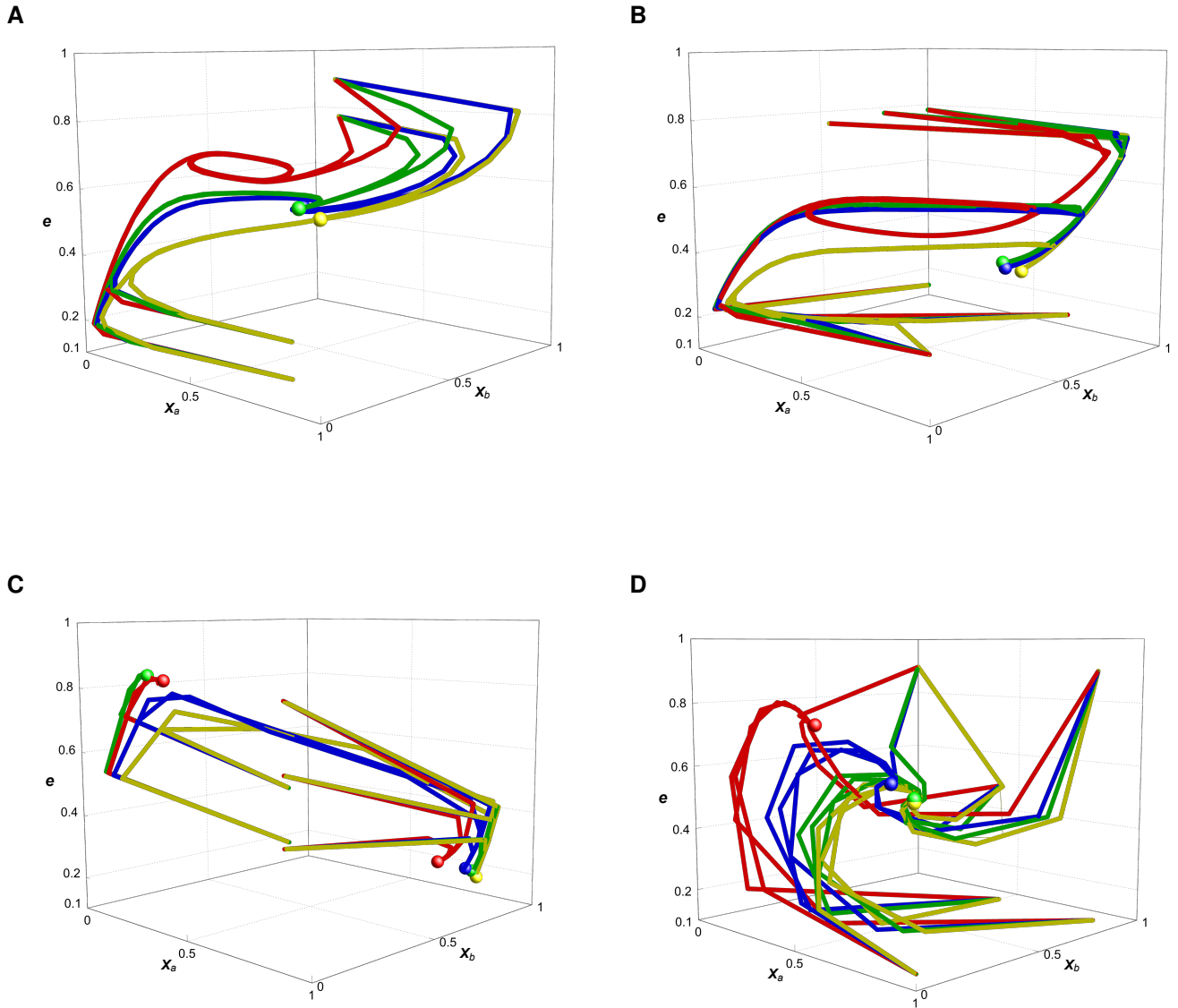


Figure 5.5: **Effect of idealized public policy interventions:** exemplary trajectories in the model state space ( $x_a$  frequency in the  $a$  subpopulation,  $x_b$  frequency in the  $b$  subpopulation,  $e$  environmental state) from multiple initial conditions, in the four dynamical regimes shown in Fig.5.2: **A**, small limit cycle. **B**, large limit cycle. **C**, bistability. **D**, globally stable equilibrium with nonmonotonic transients. Red: baseline parameters values, as in Fig. 5.2. Yellow: impact of tax policies and subsidies on material payoff (new  $\gamma_A=\gamma_B=1.5$ ). Green: impact of an information policy on  $a$ -type focus (new  $\rho_{aD} = \rho_{aN} = 0.65$ ). Blue: impact of an education policy on attitude strength (new  $\alpha_a = \alpha_a + 0.5$ , new  $\alpha_b = \alpha_b - 0.5$ ).



## Chapter 6

# Conclusion

Our research has made a significant contribution to the understanding of how social norms shape behavior. Firstly, our findings have shown how beliefs about social norms of cooperation vary across contexts and time: The mere presence of inequality in the environment causes an initial drop in normative and descriptive beliefs; while normative and descriptive beliefs tend to decline over time, a simple one-shot-free communication supports them. Secondly, we have produced evidence on how social norms influence decision-making: Social norms' beliefs impact both direct and strategic decision-making, although strategic decisions are less affected. More importantly, our investigations into unequal and heterogeneous frameworks have shown the importance of normative beliefs, which complement descriptive beliefs in the case of heterogeneity. Together, social norm beliefs can be seen as an equilibrium selection mechanism for achieving virtuous equilibria, even in the presence of inequality and heterogeneity. Finally, we demonstrate how social norms can induce behavioral change. We have shown three ways of using social norm information to improve pro-social and pro-environmental behavior. First, by revealing the actual value of social norms, leading to a revision of beliefs in the right direction. Second, by focusing attention on a particular population to highlight the right behaviors. Third, by deploying targeted information policies to influence marginal behavior through static information, and average behavior through dynamic information.

The complexity of implementing global regulations through existing formal institutions reinforces the need to work through informal institutions such as social norms to change behavior and consumption patterns (Nyborg et al., 2016). Our results reinforce the relevance of social norms in climate change mitigation. On the one hand, we demonstrate the ability of social norms to increase pro-social and pro-environmental behavior of both marginal and average individuals through information-based policies. The effectiveness of these policies is even stronger as many situations do not involve clear information about the local norms. Individuals hold beliefs about the social appropriateness of behavior and the behavior implemented in society. Information acquisition necessary to correct misconceptions can be difficult to access, and experience acquired in different situations requires time to correct



them accurately. During this process, people behave according to their beliefs, making these misconceptions the actual norm in their eyes.<sup>1</sup> This characteristic could explain the significance of the influence of social norms on behavior, even when there is no clear norm shared between individuals. This could also explain a well-known phenomenon in the field of social norms: pluralistic ignorance. Pluralistic ignorance occurs when everyone believes that others believe that one behavior is or is not socially appropriate when they individually do not have this belief. As a result, everyone refrains from expressing the behavior they perceive as appropriate to avoid the cost of transgressing this fictitious norm.<sup>2</sup> Therefore, information about social norms constitutes essential tools for climate change mitigation.

On the other hand, we have shown the specificities to be considered when using social norms to obtain positive effects and avoid boomerang effects. Whereas several research underline the limited influence of normative expectations on behavior (Bicchieri and Xiao, 2009; Bicchieri and Mercier, 2013; Bicchieri, Dimant and Sonderegger, 2023), we show that they are as important as descriptive expectations in the case of heterogeneous matching. Therefore, information provision to heterogeneous populations must reconcile the two components of social norms. In this respect, the development of social norms in a homogeneous neighborhood will require coordination of behavior to reach the tipping point and build a convention. However, the development of social norms across countries will require a shared vision of their appropriateness as well as coordination to reach the tipping point where enough people will behave and think the same way.

Nevertheless, information should not be seen as an omnipotent tool, as a simple characteristic such as a discontinuity in the average can be enough to counteract the positive effects of information policies or even induce boomerang effects. Conformity traps represent that limit where belief revision is no longer an issue. Even though everyone knows that everyone prefers one behavior over another, they may not be able to coordinate on the preferred behavior. This difficulty stems from the fact that the first deviation from the behavior of the majority may be associated with a transgression of norms and, therefore, incur a cost.<sup>3</sup> Further implementations are needed to test how best to use the information to improve pro-environmental behavior, depending on the population, culture, and position of individuals relative to the information provided. Progressive development is recommended until further research results are available.

In addition to social norms, personal preferences should not be forgotten. More than being an additional force driving individual behavior, they constitute, at the aggregate level, the normative component of social norms. Education and information policies must ensure the prevalence of positive values in society to avoid the emergence of detrimental norms.

---

<sup>1</sup> Similar phenomena are described by the Thomas theorem, "if men define situations as real, they are real in their consequences" (Thomson and Knopf, 1928).

<sup>2</sup> Bursztyn, González and Yanagizawa-Drott (2020) studied this phenomenon through a field experiment on non-working women in Saudi Arabia and contributed to its resolution through simple information disclosure.

<sup>3</sup> Andreoni, Nikiforakis and Siegenthaler (2021) have represented this effect in the laboratory, showing that providing information is insufficient to guarantee release from the conformity trap.

Social norms have become a fast-growing and increasingly studied area of the economic literature. Still, further studies are required to support and extend our findings toward better understanding and use. Regarding the first two chapters, complementary studies should dissociate the effect of heterogeneity from that of inequality. In this way, it will be possible to observe whether the interaction of these two characteristics has an over-additive or under-additive effect. In addition, we have found that one-shot communication supports beliefs over time but is not enough for individuals to align their expectations with simple games. It would be interesting to test different communication methods to observe which channels, structure, and content would converge expectations on the same beliefs.<sup>4</sup> This would provide evidence for the contingency criterion defined by Bicchieri (2005, 2016), which conflicts with our evidence of the influence of social norms despite the absence of shared knowledge about norms between group members. Regarding the fourth chapter, large-scale experiments should be implemented to consolidate the effect of dynamic norms on several behaviors with different levels of observability and cultures with varying levels of social conformity. Combining observability and social conformity levels could help build new representations of social norms on decision-making. In addition, complementary studies have to be implemented to investigate the cognitive and psychological processes representing the black box between the provision of information and the behavioral changes (Bicchieri and Dimant, 2023). A better understanding of all the steps leading to behavior change can contribute to developing better information-based policies. Regarding the fifth chapter, several extensions can be implemented to represent our society better; we describe here the most relevant. The assumption of perfect and complete information could be removed to allow misperception of social norms, each individual having a private history that shapes their beliefs based on different encounters. The literature on network and innovation theory could hence be employed to develop the encounter mechanism and define the construction processes of individuals' beliefs Centola (2010); Guilbeault and Centola (2021); Centola (2021); Horsevad et al. (2022).<sup>5</sup> The assumption of fixed preferences anchoring individuals in a sub-population could also be removed to allow for variations in heterogeneity over time; each individual could switch from one sub-population to another and be more or less anchored in personal preferences. At the same time, it would allow trendsetters to be represented.<sup>6</sup> Representing a model that integrates all these features is a promising and engaging interdisciplinary project that represents the essence of the COOCLIMATH program directed by Régis Ferrière.<sup>7</sup>

---

<sup>4</sup>The channels refer to the tools used for communication, the structure encompasses parameters/rules that govern the communication system, and the content of the communications represents the messages themselves (Schram and Ule, 2019).

<sup>5</sup>Recent work in network theory addresses "complex contagion", which represents the propagation of behavior analogous to theories of social norms (see more with Guilbeault and Centola (2021); Centola (2021); Horsevad et al. (2022)).

<sup>6</sup>Bicchieri (2005); Bicchieri and Mercier (2013) defines trendsetters as individuals who are less affected by norms and have maximizing behaviors that can trigger new behaviors in a group.

<sup>7</sup>The Emergence and Dynamics of Cultures of Cooperation: Mathematical Modeling in the Context of Climate Change and Climate Extremes (COOCLIMATH) is an interdisciplinary program for developing new mathematics to analyze cooperative joint adaptive dynamics models.

Social norms are a promising literature that can have far-reaching consequences for human interactions. For this reason, the number of articles focusing on or related to social norms has increased significantly in recent years. Social norms are at the core of human interactions, and many aspects remain to be discovered. Much research is still needed in economics and the surrounding human sciences to perfect our knowledge. Integrating different domains and representations of our world will prove crucial as we progress our understanding. Together, progress in social norms could enable tipping points to be reached toward generalizing green behavior and, therefore, mobilize all stakeholders towards climate change mitigation. The spread of this perspective could be the first transition point to achieve this ambitious objective. The European Research Council has joined this transition with funding of 1.8 million euros for GREEN TIPPING - From Niches to Norms: Drivers and Diffusion of Green Social Tipping, directed by Alessandro Tavoni.



## Appendix A

# Inequality, Social Norms and Cooperation: Strategy Choice in the Infinitely Repeated Prisoners' Dilemma - Supplementary

### A.1 Equilibrium continuation probabilities

#### A.1.1 Calculation of $\delta^{SPE}$

We calculate the minimum continuation rate that allows cooperation to be an equilibrium.

**Egalitarian distribution of the benefits of cooperation** In each round, if the playmate is adopting the grim trigger strategy, the player would receive  $\frac{1}{1-\delta}0.8(E_i + E_j)$  as payoff if she chooses any cooperative strategy and  $\frac{1}{1-\delta}E_i + 0.8E_j$  if she chooses to defect. In a stranger matching protocol, the player who defects is not necessarily directly punished in the next round by the other player because the latter may not have already experienced defection, and then adopting the grim trigger strategy leads him to choose to cooperate. Accounting for this stranger matching, the player would receive  $\sum_{t=1}^n \rho_t \left[ (E_i + 0.8E_j) \frac{1-\delta^t}{1-\delta} + E_i \frac{\delta^t}{1-\delta} \right]$  as expected payoff if she chooses to defect, with  $\rho_t$  being the probability of being paired with a player choosing to cooperate in period  $t$ ,  $\sum_{t=1}^n \rho_t = 1$ ,  $\rho_1 \geq \rho_2 \dots \geq \rho_n$ , and  $n$  being the total number of periods played. Indeed, the diffusion of defection may take several rounds to spread in the group; then the player would benefit from defecting for more than only one round but with a decreasing probability (see Camera and Casari (2009) and Camera, Casari and Bigoni (2012) who have the same reasoning assuming a specific matching with four players in the group). Players choosing to cooperate in each round is then an equilibrium if  $\delta > \delta^{SPE}$  with  $\delta^{SPE}$  such that  $\frac{1}{1-\delta}0.8(E_i + E_j) = \sum_{t=1}^n \rho_t \left[ (E_i + 0.8E_j) \frac{1-\delta^t}{1-\delta} + E_i \frac{\delta^t}{1-\delta} \right]$ . The

grim trigger strategy is an equilibrium strategy if  $\delta > \delta^{SPE}$ . As an example, if the player who defects is matched only once with a playmate choosing to cooperate ( $n = 1$  and  $\rho_1 = 1$ ), the player chooses to cooperate only if  $\frac{1}{1-\delta}0.8(E_i + E_j) > E_i + 0.8E_j + \frac{\delta}{1-\delta}E_i \Leftrightarrow \delta > \frac{E_i}{4E_j}$ , i.e.  $\delta^{SPE} = \frac{E_i}{4E_j}$ . Cooperation is an equilibrium action more easily (lower  $\delta^{SPE}$ ) when the relative weight of one player's endowment is low compared to her playmate's endowment.

**Proportional distribution of the benefits of cooperation** In each round, if the playmate is adopting the grim trigger strategy, the player would receive  $\frac{1}{1-\delta}1.6E_i$  as payoff if she chooses any cooperative strategy and  $\frac{1}{1-\delta}(1.8 - 0.8\delta)E_i$  if she chooses to defect. Because of the stranger matching, the player would receive  $\sum_{t=1}^n \rho_t \left[ 1.8E_i \frac{1-\delta^t}{1-\delta} + E_i \frac{\delta^t}{1-\delta} \right]$  as expected payoff if she chooses to defect. Players choosing to cooperate in each round is an equilibrium if  $\delta > \delta^{SPE}$  with  $\delta^{SPE}$  such that  $\frac{1}{1-\delta}1.6E_i = \sum_{t=1}^n \rho_t \left[ 1.8E_i \frac{1-\delta^t}{1-\delta} + E_i \frac{\delta^t}{1-\delta} \right]$ . As an example, if the player who defects is matched only once with a playmate choosing to cooperate ( $n = 1$  and  $\rho_1 = 1$ ), the player chooses to cooperate only if  $\frac{1}{1-\delta}1.6E_i > 1.8E_i + \frac{\delta}{1-\delta}E_i \Leftrightarrow \delta > \frac{1}{4}$ , i.e.  $\delta^{SPE} = \frac{1}{4}$ . The  $\delta^{SPE}$  level does not depend on the players' endowment.

## A.1.2 Calculation of $\delta^{RD}$

We calculate the minimum continuation rate that allows cooperation to be part of a risk-dominant equilibrium.

**Egalitarian distribution of the benefits of cooperation** The stranger matching protocol implies the following expected payoff for the player who chooses the always defect strategy:  $\frac{1}{2} \frac{E_i}{1-\delta} + \frac{1}{2} \sum_{t=1}^n \rho_t \left[ (E_i + 0.8E_j) \frac{1-\delta^t}{1-\delta} + E_i \frac{\delta^t}{1-\delta} \right]$ . If the player chooses the grim trigger strategy, her expected payoff is:  $\frac{1}{2} \frac{1}{1-\delta} \left[ (1.6 + 0.2\delta)E_i + 0.8E_j \right]$ . Cooperation is then part of a risk dominant equilibrium if  $\delta > \delta^{RD}$  with  $\delta^{RD}$  such that  $\frac{1}{1-\delta} \left[ (1.6 + 0.2\delta)E_i + 0.8E_j \right] = \frac{E_i}{1-\delta} + \sum_{t=1}^n \rho_t \left[ (E_i + 0.8E_j) \frac{1-\delta^t}{1-\delta} + E_i \frac{\delta^t}{1-\delta} \right]$ . As an example, if the player who defects is matched only once with a playmate choosing to cooperate ( $n = 1$  and  $\rho_1 = 1$ ), the player chooses to cooperate only if  $\frac{1}{1-\delta} \left[ (1.6 + 0.2\delta)E_i + 0.8E_j \right] > \frac{1}{1-\delta}2E_i + 0.8E_j \Leftrightarrow \delta > \frac{2E_i}{E_i + 4E_j}$ .

**Proportional distribution of the benefits of cooperation** If the player chooses the always defect strategy, her expected payoff is:  $\frac{1}{2} \frac{E_i}{1-\delta} + \frac{1}{2} \sum_{t=1}^n \rho_t \left[ 1.8E_i \frac{1-\delta^t}{1-\delta} + E_i \frac{\delta^t}{1-\delta} \right]$ . If she chooses the grim trigger strategy, her expected payoff is:  $\frac{1}{2} \frac{1}{1-\delta} \left[ (2.4 + 0.2\delta)E_i \right]$ . Cooperation is part of a risk dominant equilibrium if  $\delta > \delta^{RD}$  with  $\delta^{RD}$  such that  $\frac{1}{1-\delta} \left[ (2.4 + 0.2\delta)E_i \right] = \frac{E_i}{1-\delta} + \sum_{t=1}^n \rho_t \left[ 1.8E_i \frac{1-\delta^t}{1-\delta} + E_i \frac{\delta^t}{1-\delta} \right]$ . As an example, if the player who defects is matched only once with a playmate choosing to cooperate ( $n = 1$  and  $\rho_1 = 1$ ), the player chooses to cooperate only if  $\frac{1}{1-\delta} \left[ (2.4 + 0.2\delta)E_i \right] > \frac{1}{1-\delta}2E_i + 0.8E_i \Leftrightarrow \delta > \frac{2}{5}$ .

## A.2 Instructions in the unequal-egalitarian treatment

# Welcome to the Grenoble Applied Economics Laboratory

The duration of the experiment is about 20 minutes. **In this experiment, you will be matched with other participants who are real persons and whose earnings may depend on your decisions as it is explained in the next pages.**

**You will receive a flat fee of \$2** for your participation to the experiment. **You will also earn a bonus up to about \$8.** During the experiment, you will make decisions. Your decisions will be matched with other participants' decisions to determine your bonus as well as the other participants' bonus. You will receive your payment within the next two weeks.

The experiment consists in two parts that we will detail next and a short questionnaire. You must complete both parts of the experiment in one sitting to receive your payment (flat-fee and bonus); if one part is not complete, you will not receive your payment.

Remember to **NEVER CLOSE THE WINDOW**: if you close the window, you will be disconnected from the experiment, you will not be able to reconnect to participate to the experiment and you will be excluded from payment.

Thank you for your participation!

Instructions read

## Part 1

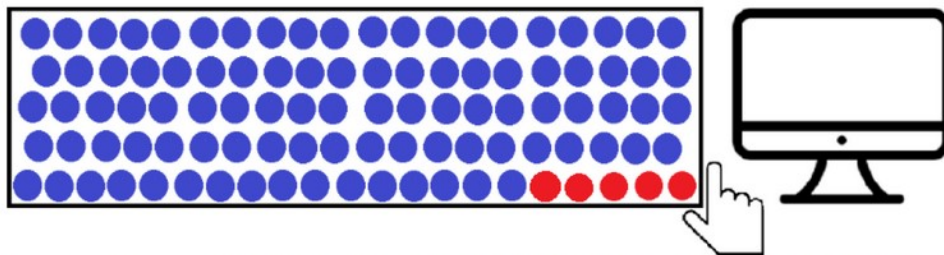
At the beginning of the experiment, you will be randomly assigned to a group of 50 participants. 25 participants have role Lambda and 25 participants have role Delta. You will learn your role at the beginning of the experiment and keep this role for the entire experiment. The experiment consists in several periods. In each period, you will be randomly paired with a participant of your group. This means that you will not be paired with the same participant for all the periods.

In this part **you will be asked to specify a plan of action for the periods of the experiment.** You must make your choice in period 1, period 2, and the periods after period 2.

**You will also have to answer questions about others' choices.**

The number of periods is not fixed in advance and is instead randomly selected: after each period, there is a 95% probability that another period will start. This probability is the same at any period of the experiment. In other words: After each period, the computer randomly draws one ball among 95 blue balls and 5 red balls, as you can see in the picture below:

- If the computer draws a blue ball, you will be randomly paired with a participant of your group and you will start another period paired with him.
- If the computer draws a red ball, this part will stop and you will be automatically directed to the second part of the experiment.



In each period, **your choice and the choice of the individual with whom you will be paired give the earnings that are presented in the following tables**, depending on your role and on the role of the participant you will be paired with. The first entry in each cell represents your earnings, while the second entry represents the earnings of the person you are paired with. At the end of the period, the two individuals in the pair learn their earnings as well as the choice and earnings of the other individual in the pair.

- If you have role **Lambda** and the other participant you are paired with has role **Lambda**, the table of earnings is as follows:

		The other's decision (Role Lambda)	
		K	P
Your decision (Role Lambda)	K	32 ; 32	16 ; 36
	P	36 ; 16	20 ; 20

With this table, the earnings are as follows:

- If you and the other choose K, you will earn **32** ECUs and the other will earn 32 ECUs.
- If you choose K and the other chooses P, you will earn **16** ECUs and the other will earn 36 ECUs.
- If you choose P and the other chooses K, you will earn **36** ECUs and the other will earn 16 ECUs.
- If you and the other choose P, you will earn **20** ECUs and the other will earn 20 ECUs.

- If you have role **Lambda** and the other participant you are paired with has role **Delta**, the table of earnings is as follows:

		The other's decision (Role Delta)	
		K	P
Your decision (Role Lambda)	K	24 ; 24	16 ; 26
	P	28 ; 8	20 ; 10

With this table, the earnings are as follows:

- If you and the other choose K, you will earn **24** ECUs and the other will earn 24 ECUs.
- If you choose K and the other chooses P, you will earn **16** ECUs and the other will earn 26 ECUs.
- If you choose P and the other chooses K, you will earn **28** ECUs and the other will earn 8 ECUs.
- If you and the other choose P, you will earn **20** ECUs and the other will earn 10 ECUs.



- If you have role **Delta** and the other participant you are paired with has role **Lambda**, the table of earnings is as follows:

		The other's decision (Role Lambda)	
		K	P
Your decision (Role Delta)	K	24 ; 24	8 ; 28
	P	26 ; 16	10 ; 20

With this table, the earnings are as follows:

- If you and the other choose K, you will earn **24** ECUs and the other will earn 24 ECUs.
- If you choose K and the other chooses P, you will earn **8** ECUs and the other will earn 28 ECUs.
- If you choose P and the other chooses K, you will earn **26** ECUs and the other will earn 16 ECUs.
- If you and the other choose P, you will earn **10** ECUs and the other will earn 20 ECUs.

- If you have role **Delta** and the other participant you are paired with has role **Delta**, the table of earnings is as follows:

		The other's decision (Role Delta)	
		K	P
Your decision (Role Delta)	K	16 ; 16	8 ; 18
	P	18 ; 8	10 ; 10

With this table, the earnings are as follows:

- If you and the other choose K, you will earn **16** ECUs and the other will earn 16 ECUs.
- If you choose K and the other chooses P, you will earn **8** ECUs and the other will earn 18 ECUs.
- If you choose P and the other chooses K, you will earn **18** ECUs and the other will earn 8 ECUs.
- If you and the other choose P, you will earn **10** ECUs and the other will earn 10 ECUs.

Instructions read

# Part 1

**The plan of action you will specify will take all decisions for you.** At the end of the experiment, **we will play the game based on your plan of action and the plan of action of the other persons in your group.** The pairs between individuals and the number of periods will be determined based on the procedures detailed above: in each period, you will be randomly paired with a participant of your group, and the number of periods is not fixed in advance and is instead randomly selected (after each period, there is a 95% probability that another period will start).

You will receive the sum of the earnings you make in every period of the game (based on your plan of action and the plan of action of the other persons in your group). **You will also have to answer questions, which will give you the opportunity to earn 50 additional ECUs for each correct answer** (only 2 questions with an asterisk (\*) do not give you additional ECUs). The exchange rate is 200 ECUs = \$1.

Instructions read

## Questions

Let us recall that you will be randomly assigned to a group of 50 participants. 25 participants have role Lambda and 25 participants have role Delta. Each participant keeps the same role for the entire experiment. The experiment consists in several periods. In each period, you will be randomly paired with a participant of your group. This means that you will not be paired with the same participant for all the periods. The number of periods is not fixed in advance and is instead randomly selected (after each period, there is a 95% probability that another period will start).

### You have role Lambda.

The tables of earnings are as follows: [Hide all tables](#)

		The other's decision (Role Lambda)	
		K	P
Your decision (Role Lambda)	K	32 ; 32	16 ; 36
	P	36 ; 16	20 ; 20

		The other's decision (Role Delta)	
		K	P
Your decision (Role Lambda)	K	24 ; 24	16 ; 26
	P	28 ; 8	20 ; 10

		The other's decision (Role Delta)	
		K	P
Your decision (Role Delta)	K	16 ; 16	8 ; 18
	P	18 ; 8	10 ; 10

		The other's decision (Role Lambda)	
		K	P
Your decision (Role Delta)	K	24 ; 24	8 ; 28
	P	26 ; 16	10 ; 20

In your opinion, what decision participants in role Lambda should make?(\*)

- K    - P

In your opinion, what decision participants in role Delta should make?(\*)

- K    - P

In your opinion, what share of the participants in role Lambda of your group will indicate that participants in role Lambda (your role) should make decision K or P? *(you will earn 50 additional ECUs if your guess is correct, we accept a 10% error rate)*

-- % will indicate P  -- % will indicate K

In your opinion, what share of the participants in role Lambda of your group will indicate that participants in role Delta should make decision K or P? *(you will earn 50 additional ECUs if your guess is correct, we accept a 10% error rate)*

-- % will indicate P  -- % will indicate K

In your opinion, what share of the participants in role Delta of your group will indicate that participants in role Delta should make decision K or P? *(you will earn 50 additional ECUs if your guess is correct, we accept a 10% error rate)*

-- % will indicate P  -- % will indicate K

In your opinion, what share of the participants in role Delta of your group will indicate that participants in role Lambda (your role) should make decision K or P? *(you will earn 50 additional ECUs if your guess is correct, we accept a 10% error rate)*

-- % will indicate P  -- % will indicate K

In your opinion, what share of the participants in role Lambda of your group will make decision K or P in period 1? *(you will earn 50 additional ECUs if your guess is correct, we accept a 10% error rate)*

-- % will choose P  -- % will choose K

In your opinion, what share of the participants in role Delta of your group will make decision K or P in period 1? *(you will earn 50 additional ECUs if your guess is correct, we accept a 10% error rate)*

-- % will choose P  -- % will choose K

Submit

# Decision - Period 1

Recall that you have role Lambda.

In each round, you will be randomly paired with a participant who has role Delta.

In each round, **because you have role Lambda and you will be randomly paired with a participant of your group who has role Delta**, your table of earnings is as follows:

		The other's decision (Role Delta)	
		K	P
Your decision (Role Lambda)	K	24 ; 24	16 ; 26
	P	28 ; 8	20 ; 10

In your opinion, what decision the other participant you will be paired with in period 1 will make? (you will earn 50 additional ECUs if your guess is correct)

- K    - P

What is your choice in period 1?

- K    - P

Submit

# Decision - Period 2

Recall that you have role Lambda.

In each round, **because you have role Lambda and you will be randomly paired with a participant of your group who has role Delta**, your table of earnings is as follows:

		The other's decision (Role Delta)	
		K	P
Your decision (Role Lambda)	K	24 ; 24	16 ; 26
	P	28 ; 8	20 ; 10

In case the choice of the other participant you were paired with in period 1 was : K

Recall, your choice in period 1 was : P

In your opinion, what decision the other participant you will be paired with in period 2 will make? (you will earn 50 additional ECUs if your guess is correct)

- K  - P

What is your choice in period 2?

- K  - P

In case the choice of the other participant you were paired with in period 1 was : P

Recall, your choice in period 1 was : P

In your opinion, what decision the other participant you will be paired with in period 2 will make? (you will earn 50 additional ECUs if your guess is correct)

- K  - P

What is your choice in period 2?

- K  - P

Submit

## Plan of action - period 3 and after

**Recall that you have role Lambda.**

In each round, **because you have role Lambda and you will be randomly paired with a participant of your group who has role Delta**, your table of earnings is as follows:

		The other's decision (Role Delta)	
		K	P
Your decision (Role Lambda)	K	24 ; 24	16 ; 26
	P	28 ; 8	20 ; 10

What is your plan of action for period 3 and after (period 3, period 4... until the game ends)?

We propose two different lists of plans of action. You will choose a plan of action in each list. At the end of the experiment, we will randomly select one of the two lists (**each list has a 50% chances of being selected**) to apply your choices to determine the earnings of individuals in the group.



---

List 1 (4 choices)

- If in the previous period you chose K and the other chose K, then choose  for the new matching
  - If in the previous period you chose K and the other chose P, then choose  for the new matching
  - If in the previous period you chose P and the other chose K, then choose  for the new matching
  - If in the previous period you chose P and the other chose P, then choose  for the new matching
- 

List 2 (1 choice)

- Choose K in every period
- Choose P in every period
- Choose K for X periods, then choose P until the end
- Choose K X% of the time and P 1-X% of the time
- Choose K for the new matching if both always chose K in the previous periods ; otherwise choose P
- Choose K for the new matching if the other chose K in the previous period; Choose P for the new matching if the other chose P in the previous period
- Choose K for the new matching if both made the same choice (both chose K or both chose P) in the previous period; otherwise choose P
- Choose P for the new matching if in X consecutive periods either the others or myself chose P; otherwise Choose K
- Choose P for the new matching if the others chose P in all of the previous X periods; otherwise Choose K
- Start by choosing K and do so until one of the others or myself chooses P, in that case Choose P for X rounds. After that go back to the start

Submit

# Survey

Please answer the following questions.

How confident are you about your answers in the first part regarding the questions that can give you 50 additional ECUs?

When you chose which decision a participant should make in the first part, what was the most important criterion for your answers?

- Maximization of the group's profit
- Equality of individual profits
- Justice
- Maximization of individual profit
- Ethics
- Equity

Next

# Survey

Please answer the following questions.

A bat and a ball cost \$1.10 in total. The bat costs \$1.00 more than the ball. How much does the ball cost? (in cents)

If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? (in minutes)

In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? (in days)

Next

# Survey

Please answer the following questions.

What is your age?

What is your gender?

- Male
- Female

Do you have any brothers or sisters?

- Yes
- No

Please write a word of more than 7 letters?

Are you a student?

- Yes
- No

Discipline studied (currently or when you were a student)?

- Economics-Management
- Law
- Political Science
- Psychology
- Other social sciences
- Mathematics
- Sciences
- Others



What is your level of education or equivalent?

- Before High School
- High School
- Bachelor
- Master
- PhD
- Others

Generally speaking, would you say that most people are trustworthy or that you can never be too careful with people?

- I do not know
- You can never be too careful
- Most people are trustworthy

Next

### A.3 Classification of the menu of strategies

Menu of strategies	Final strategies
Choose K in every period	Always cooperate
Choose P in every period	Always defect
Choose K for X periods, then choose P until the end	Always defect if $X < 10$ Always cooperate if $X \geq 10$
Choose K X% of the time and P 1-X% of the time	Always defect if $X < 50$ Always cooperate if $X \geq 50$
Choose K for the new matching if both always chose K in the previous periods; otherwise choose P	Grim-trigger
Choose K for the new matching if the other chose K in the previous period; Choose P for the new matching if the other chose P in the previous period	Tit-for-Tat
Choose K for the new matching if both made the same choice (both chose K or both chose P) in the previous period; otherwise choose P	Tit-for-Tat
Choose P for the new matching if in X consecutive periods either the others or myself chose P; otherwise choose K	Tit-for-Tat
Choose P for the new matching if the others chose P in all the previous X periods; otherwise choose K	Tit-for-Tat
Start by choosing K and do so until one of the others or myself chooses P, in that case choose P for X rounds. After that go back to the start	Tit-for-Tat if $X < 5$ Grim-trigger if $X \geq 5$

Table A.1: Classification of the menu of strategies

## A.4 Descriptive statistics

### A.4.1 Normative beliefs

	Beliefs about a participant of type:		
	D	A	Average
<b>Equal tr.</b>			
Equal-D	0.78	-	0.78
Equal-A	-	0.78	0.78
<i>Average Equal</i>	<i>0.78</i>	<i>0.78</i>	<b>0.78</b>
<b>Unequal-E tr.</b>			
Unequal-E Participant D	0.56	0.62	0.59
Unequal-E Participant A	0.62	0.72	0.67
<i>Average Unequal-E</i>	<i>0.59</i>	<i>0.67</i>	<b>0.63</b>
<b>Unequal-U tr.</b>			
Unequal-U Participant D	0.68	0.48	0.58
Unequal-U Participant A	0.49	0.66	0.58
<i>Average Unequal-U</i>	<i>0.59</i>	<i>0.57</i>	<b>0.58</b>

Table A.2: Share of participants who believe that the decision a participant should make is to cooperate (first-order normative beliefs)

	Beliefs about a participant of type:		
	D	A	Average
<b>Equal tr.</b>			
Equal-D	0.72	-	<b>0.72</b>
<b>Unequal-E tr.</b>			
Unequal-E Type D	0.65	0.65	0.65
Unequal-E Type A	0.61	0.60	0.61
<i>Average Unequal-E</i>	<i>0.63</i>	<i>0.62</i>	<b>0.63</b>
<b>Unequal-U tr.</b>			
Unequal-U Type D	0.65	0.59	0.62
Unequal-U Type A	0.56	0.57	0.57
<i>Average Unequal-U</i>	<i>0.61</i>	<i>0.58</i>	<b>0.59</b>

Table A.3: Share of participants who believe that disadvantaged participants believe that the decision a participant should make is to cooperate (second-order normative beliefs for low-endowed participants)

	Beliefs about a participant of type:		
	D	A	Average
<b>Equal tr.</b>			
Equal-A	-	0.77	<b>0.77</b>
<b>Unequal-E tr.</b>			
Unequal-E Type D	0.63	0.62	0.63
Unequal-E Type A	0.62	0.67	0.64
<i>Average Unequal-E</i>	<i>0.63</i>	<i>0.64</i>	<b>0.63</b>
<b>Unequal-U tr.</b>			
Unequal-U Type D	0.60	0.61	0.60
Unequal-U Type A	0.53	0.65	0.59
<i>Average Unequal-U</i>	<i>0.57</i>	<i>0.63</i>	<b>0.60</b>

Table A.4: Share of participants who believe that advantaged participants believe that the decision a participant should make is to cooperate (second-order normative beliefs for high-endowed participants)

#### A.4.2 Descriptive beliefs

Type of other participant:	D	A	All
Equal-D	0.68	-	0.68
Equal-A	-	0.65	0.65
<i>Average Equal</i>	<i>0.68</i>	<i>0.65</i>	<b>0.66</b>
Unequal-E Type D	0.64	0.62	0.63
Unequal-E Type A	0.63	0.64	0.63
<i>Average Unequal-E</i>	<i>0.63</i>	<i>0.63</i>	<b>0.63</b>
Unequal-U Type D	0.57	0.59	0.58
Unequal-U Type A	0.53	0.60	0.56
<i>Average Unequal-U</i>	<i>0.55</i>	<i>0.60</i>	<b>0.57</b>

Table A.5: Descriptive beliefs

#### A.4.3 Decisions in period 1

Type of the other participant:	D	A	Average
Equal-D	0.7	-	0.7
Equal-A	-	0.78	0.78
<i>Average Equal</i>	<i>0.7</i>	<i>0.78</i>	<b>0.74</b>
Unequal-E Type D	0.66	0.7	0.68
Unequal-E Type A	0.68	0.64	0.66
<i>Average Unequal-E</i>	<i>0.67</i>	<i>0.67</i>	<b>0.67</b>
Unequal-U Type D	0.62	0.62	0.62
Unequal-U Type A	0.72	0.62	0.67
<i>Average Unequal-U</i>	<i>0.67</i>	<i>0.62</i>	<b>0.65</b>

Table A.6: Decision to cooperate in period 1

#### A.4.4 Conditional decisions in period 2

In period 1 if the other participant: Type of the other participant:	Cooperates			Defects		
	D	A	Average	D	A	Average
<b>If cooperates in period 1</b>						
Equal-D	0.8	-	0.8	0.63	-	0.63
Equal-A	-	0.74	0.74	-	0.72	0.72
<i>Average Equal</i>	<i>0.8</i>	<i>0.74</i>	<b>0.77</b>	<i>0.63</i>	<i>0.72</i>	<b>0.68</b>
Unequal-E Type D	0.70	0.83	0.76	0.55	0.66	0.60
Unequal-E Type A	0.79	0.72	0.76	0.85	0.72	0.79
<i>Average Unequal-E</i>	<i>0.75</i>	<i>0.78</i>	<b>0.76</b>	<i>0.70</i>	<i>0.69</i>	<b>0.69</b>
Unequal-U Type D	0.84	0.55	0.69	0.65	0.65	0.65
Unequal-U Type A	0.69	0.71	0.70	0.61	0.65	0.63
<i>Average Unequal-U</i>	<i>0.76</i>	<i>0.63</i>	<b>0.70</b>	<i>0.63</i>	<i>0.65</i>	<b>0.64</b>
<b>If defects in period 1</b>						
Equal-D	0.53	-	0.53	0.33	-	0.33
Equal-A	-	0.45	0.45	-	0.73	0.73
<i>Average Equal</i>	<i>0.53</i>	<i>0.45</i>	<b>0.5</b>	<i>0.33</i>	<i>0.73</i>	<b>0.5</b>
Unequal-E Type D	0.41	0.33	0.38	0.47	0.13	0.31
Unequal-E Type A	0.31	0.22	0.26	0.44	0.11	0.26
<i>Average Unequal-E</i>	<i>0.36</i>	<i>0.27</i>	<b>0.32</b>	<i>0.45</i>	<i>0.12</i>	<b>0.29</b>
Unequal-U Type D	0.53	0.37	0.45	0.21	0.26	0.24
Unequal-U Type A	0.14	0.42	0.30	0.14	0.32	0.24
<i>Average Unequal-U</i>	<i>0.36</i>	<i>0.39</i>	<b>0.38</b>	<i>0.18</i>	<i>0.0.29</i>	<b>0.24</b>

Table A.7: Conditional decision to cooperate in period 2

#### A.4.5 Memory-one strategies

Strategy:	Participant past choice		Playmate past choice	
	Cooperate	Cooperate	Defect	Defect
	Cooperate	Defect	Cooperate	Defect
Equal-D tr.	88%	48%	68%	48%
Equal-A tr.	88%	44%	72%	52%
Unequal-Egalitarian tr.	81%	42.5%	57%	43%
Unequal-Proportional tr.	79%	48.5%	50%	43%
<i>Average</i>	<i>81.6%</i>	<i>45.6%</i>	<i>56.8%</i>	<i>44.4%</i>

Table A.8: Distribution of Memory-one strategies

#### A.4.6 Menu of strategies

Strategy:	AC	AD	Grim	TFT
Equal-D tr.	48%	18%	14%	20%
Equal-A tr.	54%	4%	8%	34%
Unequal-Egalitarian tr.	34%	23.5%	14.5%	28%
Unequal-Proportional tr.	33%	23.5%	9%	34.5%
<i>Average</i>	<i>37%</i>	<i>21%</i>	<i>11.6%</i>	<i>30.4%</i>

Table A.9: Distribution of strategies in the menu of strategies



## Appendix B

# Inequality, Social Norms and Community Enforcement of Cooperation: Experimental Evidence - Supplementary

### B.1 Equilibrium continuation probabilities

#### B.1.1 Calculation of $\delta^{SPE}$

We calculated the minimum continuation rate,  $\delta^{SPE}$ , that allows cooperation to be a Subgame Perfect equilibrium following Camera and Casari (2009). They defined  $\delta^{SPE}$  as the minimum continuation rate allowing cooperation as an equilibrium, with the idea that a Defection that increases the payoff at a round  $t$  will trigger a contagious punishment reducing the expected payoff at a round  $t+1$ . Contagious punishment is represented by assigning each participant a grim trigger strategy. Thus, the adaptation of their application to a group of 8 participants is as follows: We defined  $\rho_d=(\rho_1, \rho_2, \dots, \rho_8)$  the vector probability for a player who select to defect, to meet a cooperators given  $d=1, 2, \dots, 8$  the number of defectors in the group. With a random matching protocol, we obtain  $\rho_d=(1, \frac{6}{7}, \frac{5}{7}, \frac{4}{7}, \frac{3}{7}, \frac{2}{7}, \frac{1}{7}, 0)$ .

In addition, we have  $A$ , the transition matrix where the bold numbers in the rows (columns) indicate the number  $D_t$  ( $D_{t+1}$ ) of players who currently (next round) play defection.

$$A = \begin{pmatrix} & \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} & \mathbf{5} & \mathbf{6} & \mathbf{7} & \mathbf{8} \\ \mathbf{1} & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{2} & 0 & \frac{1}{7} & 0 & \frac{6}{7} & 0 & 0 & 0 & 0 \\ \mathbf{3} & 0 & 0 & 0 & \frac{3}{7} & 0 & \frac{4}{7} & 0 & 0 \\ \mathbf{4} & 0 & 0 & 0 & \frac{3}{35} & 0 & \frac{24}{35} & 0 & \frac{8}{35} \\ \mathbf{5} & 0 & 0 & 0 & 0 & 0 & \frac{3}{7} & 0 & \frac{4}{7} \\ \mathbf{6} & 0 & 0 & 0 & 0 & 0 & \frac{1}{7} & 0 & \frac{6}{7} \\ \mathbf{7} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \mathbf{8} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Then, we calculate  $V_d = z + \rho_d(h - z) + \delta * A_d V$ , the expected lifetime utility at the start of a round, before matching takes place, to a player who plays defection currently, given that  $d=1, 2, \dots, 8$  current number of defector in the group. Using the vector of probabilities  $\rho$  and the transition matrix  $A$ , we denote  $A_d$  the  $d^{th}$  row of matrix  $A$ . With  $z$  the payoff from a defection with a defective playmate, and  $h$  the payoff from a defection with a cooperative playmate.

$$V: \begin{bmatrix} V_1 = z + 1 * (h - z) + \delta(V_2) \\ V_2 = z + \frac{6}{7} * (h - z) + \delta(\frac{1}{7}V_2 + \frac{6}{7}V_4) \\ V_3 = z + \frac{5}{7} * (h - z) + \delta(\frac{3}{7}V_4 + \frac{4}{7}V_6) \\ V_4 = z + \frac{4}{7} * (h - z) + \delta(\frac{3}{35}V_4 + \frac{24}{35}V_6 + \frac{8}{35}V_8) \\ V_5 = z + \frac{3}{7} * (h - z) + \delta(\frac{3}{7}V_6 + \frac{4}{7}V_8) \\ V_6 = z + \frac{2}{7} * (h - z) + \delta(\frac{1}{7}V_6 + \frac{6}{7}V_8) \\ V_7 = z + \frac{1}{7} * (h - z) + \delta(V_8) \\ V_8 = z + 0 * (h - z) + \delta(V_8) \end{bmatrix}$$

$$V: \begin{bmatrix} V_1 = \frac{\delta \left( \frac{6\delta \left( (31\delta^2 + 56\delta + 105)z - 28(\delta^2 + 4\delta - 5)h \right)}{(35 - 3\delta)(1 - \delta)(7 - \delta)} + 6h + z \right)}{7 - \delta} + h \\ V_2 = \frac{6\delta \left( (31\delta^2 + 56\delta + 105)z - 28(\delta^2 + 4\delta - 5)h \right)}{(35 - 3\delta)(1 - \delta)(7 - \delta)} + 6h + z \\ V_3 = \frac{(75\delta^3 + 366\delta^2 + 413\delta + 490)z - 5(15\delta^3 + 69\delta^2 + 161\delta - 245)h}{7(35 - 3\delta)(1 - \delta)(7 - \delta)} \\ V_4 = \frac{(3\delta^2 + 11\delta + 28)z - 3(\delta^2 + 6\delta - 7)h}{7(\delta - 7)(\delta - 1)} \\ V_5 = \frac{3(\delta^2 + 6\delta - 7)h - (3\delta^2 + 11\delta + 28)z}{7(1 - \delta)(7 - \delta)} \\ V_6 = \frac{\frac{(\delta + 5)z}{1 - \delta} + 2h}{7 - \delta} \\ V_7 = \frac{1}{7} * \left( \frac{(\delta + 6)z}{1 - \delta} + h \right) \\ V_8 = \frac{z}{1 - \delta} \end{bmatrix}$$

Finally, we solve  $V_1 = \frac{y}{1 - \delta}$  to find the  $\delta^{SPE}$  equalizing the expected lifetime utility of defection and cooperation for a risk-neutral agent.



### B.1.2 Calculation of $\delta^{RD}$

Blonski and Spagnolo (2015) defines a minimum continuation rate that allows cooperation to be part of a risk-dominant equilibrium,  $\delta^{RD}$ . They defined  $\delta^{RD}$  as the minimum continuation rate allowing cooperation to be Risk Dominant, with the idea that the loss of payoff from a non-cooperative encounter should not be disproportionate to the additional payoff from a cooperative encounter. Their method of calculation does not require the number of participants playing the indefinitely repeated prisoner's dilemma game. Furthermore, they define how the encounter of two heterogeneous agents, resulting in an asymmetric gain matrix, affects the calculations.

For a symmetric payoff matrix,  $\delta^{RD} = \frac{b-a-(c-d)}{b-a}$ . With the following payoff matrix:

	Cooperate	Defect
Cooperate	c ; c	a ; b
Defect	b ; a	d ; d

For an asymmetric payoff matrix,  $\delta^{RD} = \frac{Y+Z}{2W} + \sqrt{(\frac{Y+Z}{2W})^2 - X/W}$ . where:

$$X = (d_1 - a_1)(d_2 - a_2) - (b_1 - c_1)(b_2 - c_2)$$

$$Y = (d_1 - a_1)(d_2 - a_2) - (b_1 - d_1)(b_2 - c_2)$$

$$Z = (d_1 - a_1)(d_2 - a_2) - (b_1 - c_1)(b_2 - d_2)$$

$$W = (d_1 - a_1)(d_2 - a_2) - (b_1 - d_1)(b_2 - d_2)$$

With the following payoff matrix:

	Cooperate	Defect
Cooperate	$c_1 ; c_2$	$a_1 ; b_2$
Defect	$b_1 ; a_2$	$d_1 ; d_2$

## B.2 Instructions

# Welcome to the Grenoble Applied Economics Laboratory

The experiment can last from 5 minutes to 75 minutes. Its duration is not defined in advance.

During this session, you can earn money. You'll receive a €5 participation fee. In addition to this 5€, you can earn extra money. Your earnings will depend on your decisions and the decisions of the other participants you interact with. They will also depend on the length of the experience. You will be paid in cash at the end of the experience in a separate room to respect confidentiality.

Your earnings in the experience will be expressed in points. At the end of the experience, your earnings will be converted into Euros to determine your payment, with the conversion rate: 55 points = 1€.

This experience has three parts. In the first and second parts of the experiment, you'll be asked to make decisions, which we'll explain in detail in the following screens. In the third part, you'll be asked to fill in a short questionnaire. Once all three parts have been completed for all participants in the room, we will call you one by one to indemnify you.

All your decisions during the experiment are anonymous.

# Instructions

At the beginning of the experiment, groups of **8 participants** are randomly formed. So you'll be in a group with 7 other participants. You won't know who they are, and they won't know who you are either.

\*\*\*\*\*

## Part 1

Part 1 is made up of several periods.

Each period, you will be randomly paired with another participant in your group. This means that you won't be paired with the same participant for every period in Part 1.

The participant with whom you are paired is called the "other participant".

## Description of the decision phase for each period

Each period, you will have to answer four questions and make a decision. **Three questions allow you to earn 3 points each if your answer is correct.** Only one question (with an asterisk \*) will not result in additional earnings.

Once you've answered these questions, you'll have to make your decision: you'll have to choose between K and P. **Your decision and the other participant's decision will determine your earnings and the other participant's earnings for the period..**

Your winnings will be presented in a table similar to the one below. In the table, the first number represents your winnings and the second number represents the other participant's winnings. Your winnings are shown in bold.

		Other participant's decision	
		K	P
Your decision	K	<b>32</b> ; 32	<b>16</b> ; 36
	P	<b>36</b> ; 16	<b>20</b> ; 20

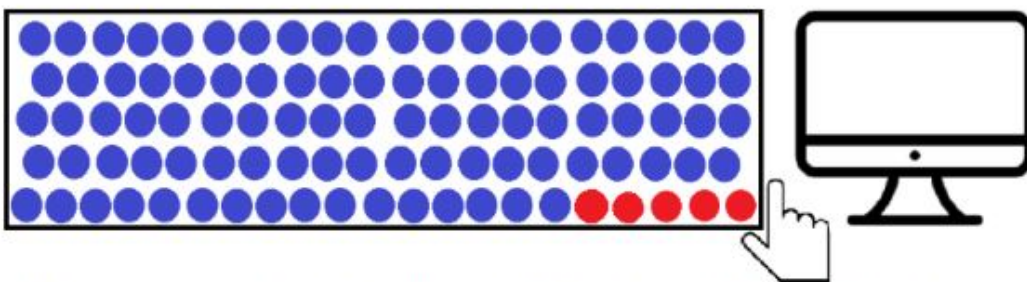
For example:

- If you and the other participant choose K, you will earn **32** points and the other participant will earn 32 points.
- If you choose K and the other participant chooses P, you will earn **16** points and the other participant will earn 36 points.
- If you choose P and the other participant chooses K, you will earn **36** points and the other participant will earn 16 points.
- If you and the other participant choose P, you will earn **20** points and the other participant will earn 20 points.

## Number of periods

**The number of periods is not fixed in advance, but is determined at random** : after each period, there is **95%** chances of another period starting. This probability is the same for all periods: for example, if you are in period 2, there are **95%** chances of another period starting, and if you're in period 9, there's also a **95%** chances of another period starting.

In other words, after each period, the computer randomly selects a ball from among 95 blue balls and 5 red balls, as shown in the picture below.



- If the computer draws a blue ball, all participants will start another period and you will be randomly paired with a participant from your group.
- If the computer draws a red ball, game 1 stops for all participants and game 2 may begin.

## Message

Every period, **before the decision phase**, you will be able to send a message to the other participant, and the other participant will also be able to send you a message.

The message is free but you must not send any text enabling the other participant to recognize you. Do not use smileys or abbreviations.

## Information at the end of the period

After the decision phase, you'll be reminded of your decision and you'll be informed of your earnings for the period, as well as the other participant's decision and earnings.

The amount of your earnings from your answers to the questions will be announced at the end of the experiment, once all the periods have been completed.

\*\*\*\*\*

## Part 2

Part 2 begins automatically once Part 1 ends. You remain in the same group as for Part 1, i.e. your group for Part 2 is made up of yourself and the 7 other participants who were in your group for Part 1.

Part 2 proceeds in the same way as Part 1.

\*\*\*\*\*

## Part 3

You will automatically be directed to a questionnaire when part 2 ends.

Once you have completed the questionnaire, your earnings from Part 1 and Part 2, as well as your earnings from answering the questions, will be announced.

\*\*\*\*\*

Before starting, you'll be asked to answer a questionnaire to help you understand the instructions. If you make a mistake, a message will appear telling you the correct answer. You can then modify your choice and validate the questionnaire. If you have any questions, raise your hand and an experimenter will come and answer you individually.



# Instructions

At the start of the experiment, groups of **8 participants** are randomly formed. So you'll be in a group with 7 other participants. In your group 4 participants will have the Lambda type and 4 other participants will have the Delta type. You won't know who they are, and they won't know who you are either.

At the start of the experiment, each person's type is drawn at random by the computer. You'll be randomly assigned either type Lambda or Delta. Your type will be announced at the start of Part 1, and you'll keep the same type for the duration of the experiment, i.e. for both Part 1 and Part 2.

\*\*\*\*\*

## Part 1

Part 1 is made up of several periods.

Each period, you'll be randomly paired with another participant in your group. This means that you won't be paired with the same participant for every period in Part 1.

The participant with whom you are paired is called the "other participant".

## Description of the decision phase for each period

Each period, you'll have to answer seven questions and make a decision. **Five questions will earn you 3 points each if your answer is correct.** Only one question (with an asterisk \*) will not result in additional earnings.

Once you've answered these questions, you'll have to make your decision: you'll have to choose between K and P. **Your decision and the other participant's decision will determine your earnings and the other participant's earnings for the period.**

Your earnings will be presented in a table similar to the one below. The amount you can earn depends on your type and the type of the other participant. In the table, the first number represents your winnings and the second number represents the other participant's winnings. Your winnings are shown in bold.

# Comprehension questionnaire

Please answer the following questions.

Don't hesitate to raise your hand or to call us if you need any further information.

How many participants (including you) are in your group?

- 2 participants
- 8 participants
- 24 participants

What is the composition of your group?

- 8 participants with the same type
- You will choose your type later
- 4 participants will be randomly assigned the Lambda type and 4 participants the Delta type

How many participants will you be interacting with each period?

- You will be interacting with 24 participants simultaneously
- You will be interacting with 8 participants simultaneously
- You will be randomly paired with another participant in your group each period.
- You will be randomly paired with another participant in your group for the duration of the experiment.

Which questions give you the opportunity to earn extra money?

- All questions
- No questions asked
- Questions without an asterisk (\*)

What do your earnings depend on?

- Your decision and the other participant's decision
- Your decision alone
- The decision of the other participant alone

What are the chances of starting a new period?

- Chances of starting a new period decrease over the course of the experiment
- Chances of starting a new period increase over the course of the experiment
- The chances of starting a new period are always 95%

Confirm

# Your type

Your type for the whole experience is **Lambda**.

Next

# Send message

You have the option of sending a message to the other participant (and the other participant has the same option).

Please do not send any text that would allow the other participant to recognize you. You must not use smileys or abbreviations.

If you wish to send a message to the other participant, please write it in the space below:

If you don't want to send a message, just click on "Send".

Send

Your payoff table is as follows:

		Other participant's decision	
		K	P
Your decision	K	32 ; 32	16 ; 36
	P	36 ; 16	20 ; 20



# Send message

You have the option of sending a message to the other participant (and the other participant has the same option).

Please do not send any text that would allow the other participant to recognize you. You must not use smileys or abbreviations.

If you wish to send a message to the other participant, please write it in the space below:

If you don't want to send a message, just click on "Send".

Send

Your payoff table is as follows:

		Decision de l'autre participant (Rôle Delta)	
		K	P
Votre decision (Rôle Lambda)	K	24 ; 24	16 ; 26
	P	28 ; 8	20 ; 10

- If you and the other participant choose K, you will earn **24** points and the other participant will earn 24 points.
- If you choose K and the other participant chooses P, you will earn **16** points and the other participant will earn 26 points.
- If you choose P and the other participant chooses K, you will earn **28** points and the other participant will earn 8 points.
- If you and the other participant choose P, you will earn **20** points and the other participant will earn 10 points.

# Message reception

The other participant has not sent you a message.

Next

Your payoff table is as follows:

		Other participant's decision	
		K	P
Your decision	K	32 ; 32	16 ; 36
	P	36 ; 16	20 ; 20

# Message reception

Your type is **Lambda** and the other participant's type is **Delta**.

**The other participant has not sent you a message.**

Next

Your payoff table is as follows:

		Decision de l'autre participant (Rôle Delta)	
		K	P
Votre decision (Rôle Lambda)	K	24 ; 24	16 ; 26
	P	28 ; 8	20 ; 10

- If you and the other participant choose K, you will earn **24** points and the other participant will earn 24 points.
- If you choose K and the other participant chooses P, you will earn **16** points and the other participant will earn 26 points.
- If you choose P and the other participant chooses K, you will earn **28** points and the other participant will earn 8 points.
- If you and the other participant choose P, you will earn **20** points and the other participant will earn 10 points.

# Decision phase

The following questions earn you 3 extra points each if you answer correctly (except those marked with an asterisk). They concern the current period.

## Questions

		Other participant's decision	
		K	P
Your decision	K	32 ; 32	16 ; 36
	P	36 ; 16	20 ; 20

1 - What do you think is the most appropriate decision a participant should make? \*:

- K
- P

2 - Regarding question 1, how many participants in your group (not including yourself) do you think will indicate K or P?

-- will indicate P  -- will indicate K

3 - How many participants in your group (not including you) do you think will actually make the K or P decision?

-- will choose P  -- will choose K

## Decisions

4 - Which decision do you think the other participant will choose? :

- K
- P

Please make your decision between K and P for this period (please click on the button corresponding to your choice; your choice will be automatically validated).

		Other participant's decision	
		K	P
Your decision	K	32 ; 32	16 ; 36
	P	36 ; 16	20 ; 20

- If you and the other participant choose K, you will earn **32** points and the other participant will earn 32 points.
- If you choose K and the other participant chooses P, you will earn **16** points and the other participant will earn 36 points.
- If you choose P and the other participant chooses K, you will earn **36** points and the other participant will earn 16 points.
- If you and the other participant choose P, you will earn **20** points and the other participant will earn 20 points.

# Decision phase

Your type is **Lambda** and the type of the other participant is **Delta**.

The following questions earn you 3 extra points each if you answer correctly (except those marked with an asterisk). They concern the current period.

## Questions about Lambda participants

Hide tables from a Lambda participant

		Participant Lambda	
		K	P
Participant Lambda	K	32 : 32	16 : 36
	P	36 : 16	20 : 20

		Participant Delta	
		K	P
Participant Lambda	K	24 : 24	16 : 26
	P	28 : 8	20 : 10

1 - What do you think is the most appropriate decision for a Lambda participant to make? \*:

- K
- P

2 - Regarding question 1, how many participants in your group (not including yourself) do you think will indicate K or P?

-- will indicate P  -- will indicate K

3 - How many Lambda participants in your group (not including you) do you think will actually make the K or P decision?

-- Lambda will choose P  -- Lambda will choose K

## Questions portant sur les participants Delta

Hide tables from a Delta participant

		Participant Delta	
		K	P
Participant Delta	K	16 ; 16	8 ; 18
	P	18 ; 8	10 ; 10

		Participant Lambda	
		K	P
Participant Delta	K	24 ; 24	8 ; 28
	P	26 ; 16	10 ; 20

4 - What do you think is the most appropriate decision for a Delta participant to make? \*:

- K  
 - P

5 - Regarding question 4, how many participants in your group (not including yourself) do you think will indicate K or P?

-- will indicate P  -- will indicate K

6 - How many Delta participants in your group do you think will actually make the K or P decision?

-- Delta will choose P  -- Delta will choose K

## Decisions

7 - Which decision do you think the other participant will choose? :

- K
- P

Please make your decision between K and P (please click on the button corresponding to your choice; your choice will be automatically validated).

		Decision de l'autre participant (Rôle Delta)	
		K	P
Votre decision (Rôle Lambda)	K	24 ; 24	16 ; 26
	P	28 ; 8	20 ; 10

- If you and the other participant choose K, you will earn **24** points et l'autre participant gagnera 24 points.
- If you choose K and the other participant chooses P, you will earn **16** points et l'autre participant gagnera 26 points.
- If you choose P and the other participant chooses K, you will earn **28** points et l'autre participant gagnera 8 points.
- If you and the other participant choose P, you will earn **20** points et l'autre participant gagnera 10 points.

## Information for the period

In this period, you have chosen **K** and the other participant has chosen **K**.  
Your earnings for this period are **32** points and the other participant's earnings are **32** points.

Next

## Information for the period

Your type is **Lambda** and the other participant's type is **Lambda**.

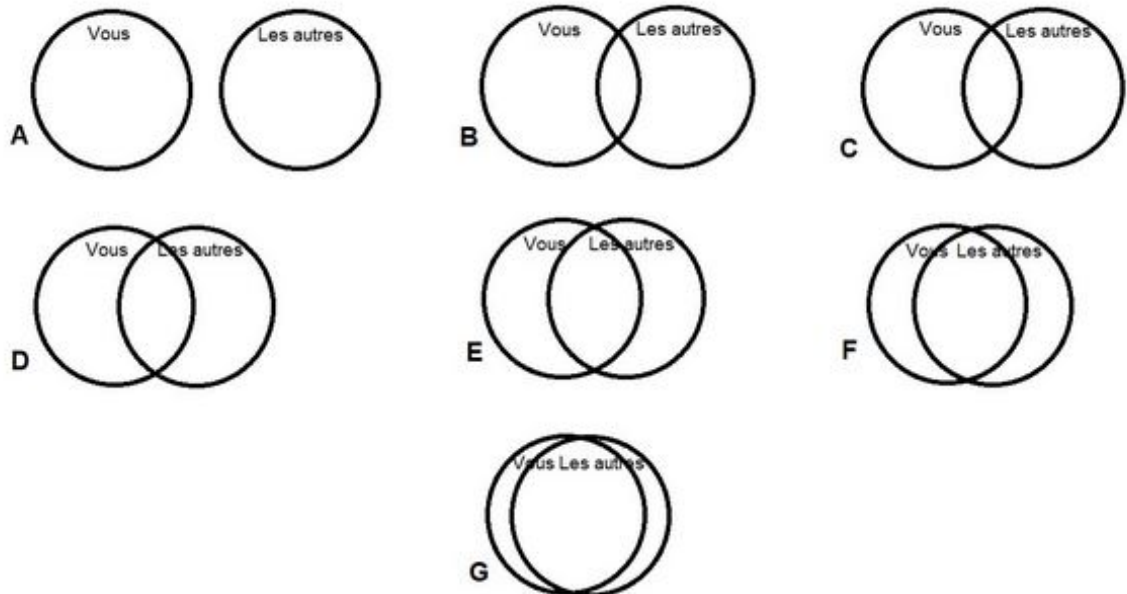
In this period, you have chosen **K** and the other participant has chosen **K**.  
Your earnings for this period are **32** points and the other participant's earnings are **32** points.

Next



# Survey

Please indicate how you feel about the other participants in your group in the previous parts of the experiment.

 ▾

Next

# Survey

Please answer the following questions.

How confident are you in your answers to the questions that could give you 3 extra points in the previous parts?

 ▾

When asked about the most appropriate decision a participant should make in the previous sections, what was the most important criterion in your choice?

- Maximizing individual profits
- Maximizing group profits
- Equity
- Ethics
- Equality of individual profits
- Justice

Next



## Survey

Please answer the following questions.

Considering a bat and a ball costing a total of €1.10. Knowing that the bat costs 1€ more than the ball. How much does the ball cost? (please give your answer in cents).

If it takes 5 minutes for 5 machines to produce 5 tools, how long does it take for 100 machines to produce 100 tools? (please give your answer in minutes) :

There is a patch of water lilies in a lake. Every day, the patch doubles in size. If it takes 48 days for the patch to fill the whole lake, how many days does it take for the patch to cover half the lake? (please give your answer in days).

Next

# Survey

Please answer the following questions.

How old are you?

What's your sex?

- Male
- Female

Do you have any brothers or sisters?

- Yes
- No

Are you a student?

- Yes
- No

What discipline are you majoring in (currently or previously)?

- Economics and management
- Law
- Political Science
- Psychology
- Other social science
- Mathematics
- Science
- Other

What is your level of education?

- No diploma
- Professional diploma (CAP, BEP)
- Baccalaureate
- BTS, DUT, Licence
- Master
- PhD
- Other

Generally speaking, would you say that the majority of people are trustworthy or, on the contrary, that you can never be too careful with people?

- I don't know
- You can never be too careful with people
- Most people are trustworthy

Next

### B.3 Regressions explaining cooperation for the first and last three rounds

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
INEQ	0.058 (.59)	0.141 (0.84)	0.175** (2.05)	0.466*** (2.77)	0.287* (1.93)
INEQ x Hetero.	-0.137** (-2.33)	-0.139** (-2.40)	-0.119* (-1.87)	-0.077 (-0.65)	-0.056 (-0.77)
High-type		0.090 (0.61)	0.065 (0.82)	0.079 (1.07)	0.077 (1.09)
High-type x INEQ		-0.103 (-0.59)	-0.109 (-1.04)	-0.109 (-1.12)	-0.100 (-1.02)
NB1			0.356*** (7.00)	0.672*** (6.25)	0.554*** (5.58)
NB1 x INEQ				-0.341*** (-2.74)	-0.374*** (-3.02)
NB1 x INEQ x Hetero.				-0.116 (-0.91)	0.014 (0.11)
NB2			0.130 (1.05)	-0.045 (-0.27)	-0.023 (-0.15)
NB2 x INEQ				0.008 (0.03)	-0.187 (-0.60)
NB2 x INEQ x Hetero.				0.205 (0.66)	0.284 (0.79)
DB			0.123 (1.16)	0.273 (1.20)	0.104 (0.41)
DB x INEQ				-0.040 (-0.13)	0.168 (0.49)
DB x INEQ x Hetero.				-0.135 (-0.70)	-0.252 (-1.12)
Playmate's Cooperation Expected					0.182*** (3.50)
Playmate's Cooperation Expected x INEQ					0.162 (1.29)
Playmate's Cooperation Expected x INEQ x Hetero.					-0.029 (-0.24)
Playmate's cooperation in the previous round	0.127** (2.51)	0.125** (2.36)	0.044 (0.99)	0.032 (0.74)	-0.036 (-0.75)
Number of rounds	0.001 (2.51)	0.003 (2.36)	0.00156 (0.99)	0.002 (0.74)	0.001 (-0.75)
Observations	400	400	400	400	400
Clusters	25	25	25	25	25
Log-pseudolikelihood	-268.99	-268.48	-210.24	-203.87	-182.54
Pseudo-R <sup>2</sup>	0.021	0.023	0.235	0.258	0.336

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Standard errors in parentheses.

Table B.1: Decision to cooperate in the first three rounds (Logit models - average marginal effects)

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
INEQ	-0.121 (-1.55)	-0.095 (-1.17)	-0.084 (-1.51)	0.069 (0.06)	-0.035 (-0.32)
INEQ x Hetero.	-0.086 (-1.23)	-0.086 (-1.23)	-0.064 (-1.12)	-0.089 (-0.90)	-0.129 (-1.26)
High-type		0.005 (0.06)	-0.042 (-0.91)	-0.068 (-1.34)	-0.063 (-1.25)
High-type x INEQ		-0.046 (-0.45)	-0.037 (-0.62)	-0.008 (-0.15)	0.002 (0.03)
NB1			0.325*** (8.01)	0.375*** (7.27)	0.351*** (7.80)
NB1 x INEQ				-0.143 (-1.23)	-0.159 (-1.47)
NB1 x INEQ x Hetero.				0.116 (0.92)	0.126 (1.13)
NB2			-0.012 (-0.12)	0.098 (0.91)	0.086 (0.81)
NB2 x INEQ				-0.138 (-0.76)	-0.138 (-0.75)
NB2 x INEQ x Hetero.				-0.217 (-0.94)	-0.260 (-1.45)
DB			0.364*** (3.46)	0.297** (2.37)	0.163 (1.43)
DB x INEQ				0.115 (0.58)	0.156 (0.80)
DB x INEQ x Hetero.				0.119 (0.60)	0.164 (1.11)
Playmate's Cooperation Expected					0.111*** (2.51)
Playmate's Cooperation Expected x INEQ					0.013 (0.17)
Playmate's Cooperation Expected x INEQ x Hetero.					0.122* (1.79)
Playmate's cooperation in the previous round	-0.092** (-2.17)	-0.092** (-2.19)	-0.075** (-2.25)	-0.076** (-2.24)	-0.083*** (-2.95)
Number of times playmate's cooperation in the 5 previous rounds	0.096*** (3.35)	0.097*** (3.35)	0.028 (1.33)	0.031 (1.61)	0.029 (1.53)
Number of times playmate's cooperation in a row	-0.027 (-1.61)	-0.028 (-1.60)	-0.019* (-1.87)	-0.020* (-1.83)	-0.023** (-2.34)
Number of rounds	-0.004** (-2.00)	-0.004* (-1.78)	-0.004*** (-2.87)	-0.006*** (-2.89)	-0.006*** (-3.20)
Observations	600	600	600	600	600
Clusters	25	25	25	25	25
Log-pseudolikelihood	-324.45	-324.10	-222.10	-217.66	-201.16
Pseudo- $R^2$	0.086	0.087	0.374	0.387	0.433

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Standard errors in parentheses.

Table B.2: Decision to cooperate in the last three rounds (Logit models - average marginal effects)

## B.4 Cooperation per matching in INEQ-EA and INEQ-EP

Rounds	First Three	Intermediate	Last Three	All
<b>INEQ-EA treatment</b>				
Homogeneous matching				
H-type	0.423 (0.483)	0.511 (0.451)	0.35 (0.405)	0.472 (0.460)
L-type	0.647 (0.482)	0.567 (0.451)	0.642 (0.415)	0.618 (0.460)
<i>Average</i>	<i>0.533 (0.482)</i>	<i>0.539 (0.451)</i>	<i>0.439 (0.410)</i>	<i>0.545 (0.460)</i>
Heterogeneous matching				
H-type	0.537 (0.481)	0.430 (0.447)	0.350 (0.392)	0.444 (0.459)
L-type	0.597 (0.481)	0.443 (0.447)	0.446 (0.386)	0.484 (0.459)
<i>Average</i>	<i>0.567 (0.481)</i>	<i>0.437 (0.447)</i>	<i>0.397 (0.389)</i>	<i>0.464 (0.459)</i>
<b>INEQ-EP treatment</b>				
Homogeneous matching				
H-type	0.680 (0.464)	0.485 (0.428)	0.449 (0.409)	0.554 (0.443)
L-type	0.486 (0.467)	0.509 (0.428)	0.472 (0.397)	0.483 (0.443)
<i>Average</i>	<i>0.580 (0.465)</i>	<i>0.497 (0.428)</i>	<i>0.461 (0.403)</i>	<i>0.518 (0.443)</i>
Heterogeneous matching				
H-type	0.526 (0.479)	0.399 (0.442)	0.367 (0.358)	0.413 (0.457)
L-type	0.505 (0.477)	0.400 (0.442)	0.431 (0.353)	0.457 (0.457)
<i>Average</i>	<i>0.516 (0.478)</i>	<i>0.400 (0.442)</i>	<i>0.399 (0.355)</i>	<i>0.435 (0.457)</i>

Table B.3: Cooperation per matching in INEQ-EA and INEQ-EP treatments

## B.5 Classification of messages in the communication treatments

The table below counts the number of messages of each category sent by each type of subject in the treatments with communication. The frequency of each category of messages is also given.

		EQ	INEQ	Total
<b><i>Ex ante communication</i></b>				
<i>Coordination</i>		0.603	0.609	0.606
	High-type	344	434	778
	Low-type	446	398	844
<i>Deliberation</i>		0.206	0.178	0.192
	High-type	119	118	237
	Low-type	152	125	277
<i>Relationship</i>		0.087	0.134	0.111
	High-type	53	80	133
	Low-type	61	104	165
<i>Trivia</i>		0.093	0.063	0.078
	High-type	61	42	103
	Low-type	62	45	107
<i>Information</i>		0.009	0.013	0.011
	High-type	4	11	15
	Low-type	8	7	15
<b><i>Ex post communication</i></b>				
<i>Coordination</i>		0.179	0.248	0.227
	High-type	39	113	152
	Low-type	110	106	216
<i>Deliberation</i>		0.116	0.114	0.122
	High-type	60	49	109
	Low-type	37	52	89
<i>Relationship</i>		0.443	0.431	0.463
	High-type	193	195	388
	Low-type	175	185	360
<i>Trivia</i>		0.071	0.139	0.112
	High-type	39	49	88
	Low-type	20	74	94
<i>Information</i>		0.073	0.065	0.073
	High-type	37	28	65
	Low-type	24	30	54

Table B.4: Proportion of messages per category of message and treatment





## Appendix C

# Dynamic Norms for Household Water Consumption - Supplementary

## Supplementary

### C.1 Household classification

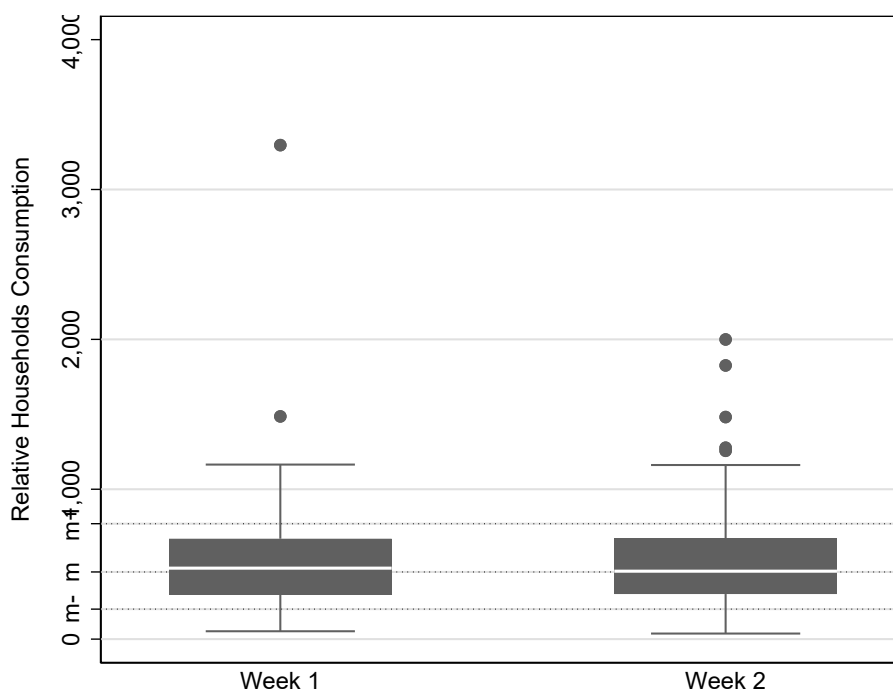


Figure C.1: Distribution of relative consumption for the first two weeks. Notes: Relative consumption refers to consumption per number of inhabitants. m-, m, and m+ represent the boundaries implemented to create categories.

	Efficient households	Medium-Low households	Medium-High households	High households
Nb . households	14	66	65	26
Households size	3.28	2.95	2.77	2.92
Living area size ( $m^2$ )	77.43	75.18	73.34	72.23
Dishwasher (%)	78.57	78.79	75.38	57.69
Paying for Water (%)	78.57	87.88	90.77	76.92
Nb. Men	0.86	1.12	1.14	0.96
Nb. Women	1.43	1.18	1.21	1.23
Nb. Son	0.64	0.41	0.2	0.38
Nb. Daughter	0.36	0.24	0.21	0.35
Income	2.21	2.14	2	1.61
Owner (%)	64.28	66.66	47.69	26.92
Phase 1 consumption	404.61	982.77	1567.42	2632.35
Phase 2 consumption	395.5	1044.629	1521.21	2550.46
Phase 3 consumption	436.61	1001.788	1433.41	2415.17
Phase 4 consumption	382	1037	1429	2182

Table C.1: Average household characteristics by categories. Notes: Income score thresholds 1=less than 30 000€, 2=between 30 000 and 40 000€, 3=between 40 000 and 50 000€, 4=between 50 000 and 60 000€, 5=more than 60 000€. Water consumption in liters per week.

## C.2 Long-term treatment effects

	Medium-Low (1)	Medium-High (2)
Treatment x Post	-58.94 (-0.94)	22.62 (0.29)
Treatment	-67.69 (-1.03)	18.21 (0.18)
Post	6.70 (0.16)	-123.5* (-1.79)
Living area size	0.34 (0.15)	4.30 (1.29)
Dishwasher	74.99 (0.71)	50.33 (0.48)
Paying for water	-40.31 (-0.39)	-350.6** (-2.17)
Nb. Men	136.1** (2.55)	292.4*** (3.53)
Nb. Women	230*** (5.63)	142.9 (1.47)
Nb. Son	299*** (6.31)	179.8 (1.64)
Nb. Daughter	218.1*** (3.52)	240*** (3.01)
Income	48.11* (1.72)	-3.93 (-0.10)
Owner	-128.7 (-1.25)	-53.84 (-0.36)
Observations	396	390
Cluster	66	65
R <sup>2</sup>	0.38	0.37

*t* statistics in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table C.2: Water consumption (OLS models - average marginal effects). Cluster at the household level. Notes: Regression was conducted on weeks 1-6. Men and Women are over 13 years old, while Son and Daughter are under 13 years old.

	Medium-Low (1)	Medium-High (2)
Treatment x Post	18.99 (0.22)	14.25 (0.24)
Treatment	-58.90 (-0.85)	33.38 (0.34)
Post	-9.67 (-0.25)	-106.7** (-2.37)
Living area size	0.06 (0.03)	3.24 (1.07)
Dishwasher	65.20 (0.60)	56.48 (0.53)
Paying for water	-43.37 (-0.39)	-344.9** (-2.17)
Nb. Men	152.9*** (2.70)	310.1*** (3.76)
Nb. Women	229.4*** (5.18)	181.4* (1.97)
Nb. Son	336.2*** (6.66)	203.9* (1.80)
Nb. Daughter	276*** (3.82)	260.4*** (3.32)
Income	30.10 (0.96)	-2.20 (-0.06)
Owner	-101.1 (-0.95)	-74.68 (-0.52)
Observations	462	455
Cluster	66	65
R <sup>2</sup>	0.33	0.39

*t* statistics in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table C.3: Water consumption (OLS models - average marginal effects). Cluster at the household level. Notes: Regression was conducted on week 1-7. Men and Women are over 13 years old, while Son and Daughter are under 13 years old.

### C.3 Cognitive and Psychological processes

	Medium-Low (1)	Medium-High (2)
Appropriateness of water conservation behavior	93.46 (71.90)	194.0** (83.06)
Appropriateness of water wasting behavior	-42.04 (78.18)	62.55 (148.1)
Positive valence from phase 3 information provision	-111.2 (84.27)	129.9 (109.7)
Negative valence from phase 3 information provision	40.01 (72.16)	-18.33 (96.72)
Moral/Personal norms	-331.6*** (104.1)	71.78 (170.7)
Expected results	141.7 (98.30)	-225.6** (106.0)
Attitude toward water conservation	-38.97 (200.6)	352.7 (306.9)
Perceived behavioral control	50.14 (71.09)	56.21 (119.1)
Change of conservation behavior	-318.4 (234.0)	-451.9 (478.7)
Change of conservation equipment	201.5 (321.2)	-471.0 (308.1)
Observations	66	65
$R^2$	0.23	0.24

Standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table C.4: Reduction in per capita water consumption in phase 3 (OLS models - average marginal effects). Notes: All scores range between -1 and 1.

### C.3.1 Theory of planned behavior

	Efficient participants	Medium-Low participants	<i>p-value</i> of Mediums	Medium-High participants	High participants
Theory of planned behavior					
Perceived control with DI	0.325	0.193	0.247	0.314	0.028
<i>p-value</i> of DI vs SI		0.102		0.002***	
Perceived control with SI	0.257	0.028	0.997	0.027	0.143
Attitude toward water conservation with DI	0.633	0.612	0.982	0.615	0.573
<i>p-value</i> of DI vs SI		0.301		0.540	
Attitude toward water conservation with SI	0.62	0.576	0.059*	0.641	0.590
Moral/Personal norms with DI	0.789	0.685	0.246	0.642	0.567
<i>p-value</i> of DI vs SI		0.041**		0.379	
Moral/Personal norms with SI	0.84	0.570	0.058*	0.653	0.554
Related components					
Expected Result with DI	0.630	0.576	0.196	0.495	0.544
<i>p-value</i> of DI vs SI		0.148		0.071*	
Expected Result with SI	0.567	0.470	0.040**	0.614	0.409
Anterior Conservation Behaviors with DI	0.839	0.761	0.326	0.731	0.548
<i>p-value</i> of DI vs SI		0.857		0.800	
Anterior Conservation Behaviors with SI	0.866	0.751	0.384	0.719	0.616
Additional information					
Current Conservation Behaviors with DI	0.852	0.882	0.270	0.848	0.763
<i>p-value</i> of DI vs SI		0.042**		0.755	
Current Conservation Behaviors with SI	0.911	0.804	0.623	0.833	0.747
Anterior Conservation Equipment with DI	0.204	0.253	0.643	0.535	0.389
<i>p-value</i> of DI vs SI		0.932		0.804	
Anterior Conservation Equipment with SI	0.253	0.495	0.809	0.521	0.379
Current Conservation Equipment with DI	0.537	0.561	0.956	0.566	0.4
<i>p-value</i> of DI vs SI		0.446		0.723	
Current Conservation Equipment with SI	0.633	0.505	0.660	0.542	0.394

Table C.5: Direct and related components to Theory of planned behavior. Notes: Score between -1 and 1, with -1 referring to low and 1 to high levels. DI refers to Dynamic Information, and SI refers to Static information.

### C.3.2 Normative evaluation

Figure C.2 regroups answers of both Medium categories. In the Figure, the more a participant is on the top of the ordinate, the more he is perceived as socially appropriate water conservation behavior. Similarly, the more a participant is on the left of the abscissa, the more he is perceived as socially inappropriate water-wasting behavior. As a consequence, the top-left part represents a perception of a strong norm in favor of water conservation behavior; the bottom-right part represents a perception of a strong norm in favor of water wasting behavior; the middle-center and the top-right represent a perception of an absence of norm; the bottom-left represent nothing logical; the other intermediary parts represent ambiguous norms.

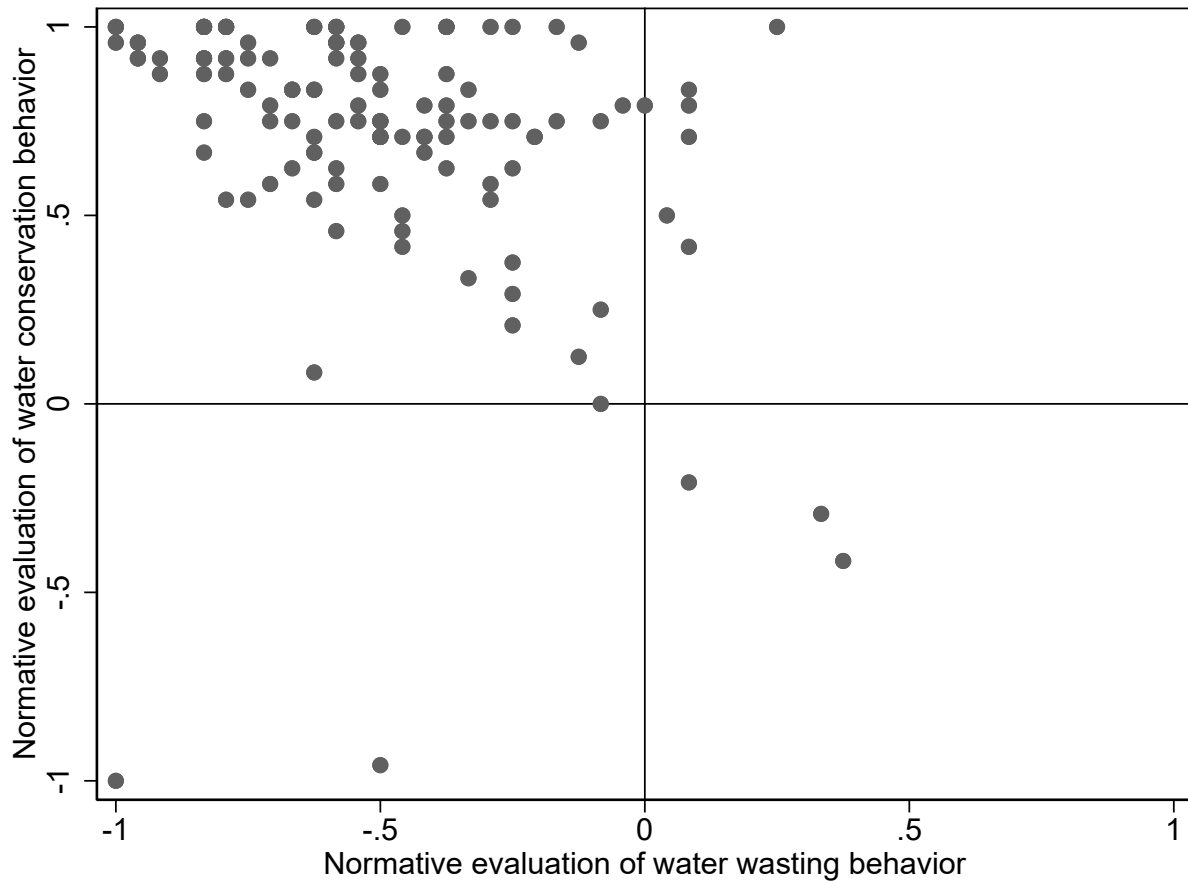


Figure C.2: Normative evaluation of water behavior

## C.4 Instruction

### C.4.1 Weekly-email

In this section, text tags represent variations according to weeks and participants. In addition, the texts in red only concern the second weeks of the phases with information provision (phases 2 and 3).

#### Beginning of the study

Subject: Water consumption study

Good morning,

Week 1

The water consumption survey you're taking part in starts today. We want to remind you that you must take a photo of your water meter(s) every Tuesday (the photo must not have been taken in advance) and submit it to us via the link provided (available below). If you forget to send a photo or make a mistake in the file you submit, you'll receive a reminder at this email address. You will be compensated 20€ for submitting your photos for the 7-week study

period. In addition, the final questionnaire will allow you to earn an extra 5€ depending on your answers to clearly marked questions.

We invite you to take a photo of your water meter now and send it to us at <https://cloud.univ-grenoble-alpes.fr/s/{{ Code }}>. Please do not share this link. Your photo must have been taken today to be eligible for payment.

We thank you in advance for the photo,

If you have any problems, please get in touch with us at the following email address: [gael-ecoexp@univ-grenoble-alpes.fr](mailto:gael-ecoexp@univ-grenoble-alpes.fr).

Best regards

The Experimental Economics Team

### **First weekly e-mail**

Subject: Water consumption study

Good morning,

Week {{ Week }}

We invite you to take a photo of your water meter and send it to us at <https://cloud.univ-grenoble-alpes.fr/s/{{ Code }}>. This link is unique to you; please do not share it. Your photo must have been taken today to be eligible for payment.

**We also remind you to consult the first results of the study at the following link {{ Link }}. Please do not share this link with others.**

We thank you in advance for the photo,

If you have any problems, please get in touch with us at the following email address: [gael-ecoexp@univ-grenoble-alpes.fr](mailto:gael-ecoexp@univ-grenoble-alpes.fr).

Best regards

The Experimental Economics Team

### **Second weekly e-mail**

Subject: Reminder of water consumption study

Good morning,

Week {{ Week }}

We have not received the photo of your water meter.

Please take a picture of your water meter and send it to us at <https://cloud.univ-grenoble-alpes.fr/s/{{ Code }}>. This is your own link, so please don't share it. Your photo must have been taken today to be eligible for payment.

**We also remind you to consult the first results of the study at the following link {{ Link }}. Please do not share this link with others.**



We thank you in advance for the photo,

If you have any problems, please get in touch with us at the following email address: [gael-ecoexp@univ-grenoble-alpes.fr](mailto:gael-ecoexp@univ-grenoble-alpes.fr) .

Best regards

The Experimental Economics Team

### **Third weekly e-mail**

Subject: Latest water consumption study reminder

Good morning,

Week {{ Week }}

This is the last reminder for the water consumption study.

We have not received your water meter photo. You must take a picture of your meter today to be eligible for payment.

If you don't, we'll remove you from the study, and you won't receive any more emails about it.

We invite you to take a photo of your water meter and send it to us using the link <https://cloud.univ-grenoble-alpes.fr/s/{{ Code }}>. This link is unique to you, so please don't share it. Your photo must have been taken today to be eligible for payment.

**We also remind you to consult the first results of the study at the following link {{ Link }}. Please do not share this link with others.**

We thank you in advance for the photo,

If you have any problems, please get in touch with us at the following email address: [gael-ecoexp@univ-grenoble-alpes.fr](mailto:gael-ecoexp@univ-grenoble-alpes.fr) .

Best regards

The Experimental Economics Team

### **Information provision email**

Subject: Water consumption study

Good morning,

Week {{ Week }}

After these {{ Week }} weeks of study, we have some results to share with you.

{{ Information\_Provision }}

If you don't see an image displayed, you will find it at the link provided below.

We invite you to visit the following page, which includes information on household water consumption {{ Code }}.

Please do not share this link with others.

You can visit this page as often as you like over the next two weeks.

Apart from consulting the page provided below, you have no further action to take until next Tuesday.

If you have any problems, please get in touch with us at the following email address: [gael-ecoexp@univ-grenoble-alpes.fr](mailto:gael-ecoexp@univ-grenoble-alpes.fr) .

Best regards

The Experimental Economics Team

## C.4.2 Web page information

### Water is essential, let's save it!

Information about your water consumption



Information about French water consumption

#### Water consumption in France

- French households consume an average of 220 m<sup>3</sup> of water per year. This represents a major source of expenditure, as well as a significant environmental impact.
- Adopting the right habits, combined with the installation of certain inexpensive devices, can save up to 40% of water consumption costs.
- Household water consumption breaks down as follows: 39% for hygiene, 20% for sanitary, 12% for laundry, 10% for washing up, 6% for cooking, 6% for gardening, 6% for miscellaneous, and 1% for drinks.

#### Environmental issues

- Rising temperatures are reducing river and groundwater flows, drying out soils and making water less available.
- At the same time, water requirements are increasing, whether for agriculture, for the growing number of households, or for increased industrial production.

## Best practice

- Turning off water faucets as soon as they are no longer needed prevents unnecessary water loss.
- Watch out for water leaks and dripping taps, to avoid unnecessary water loss.

## Reducing water consumption for laundry and dishwashing

- Using machines only when they are full reduces overall water consumption.
- Use "eco" modes or short cycles to reduce water consumption for each machine.
- Using rinse trays for hand-wash dishes rather than leaving the water running for each item, as well as using the rinse water from one group of items as wash water for the next group, saves a significant amount of water.

## Water reduction in the bathroom

- Turning off the water when soaping your hands, head or body, as well as when shaving or removing make-up, saves a significant amount of water.
- Using a timer or music to control your showering time from 3 to 5 minutes (i.e. the time of less than two songs), saves a significant amount of water.

## Equip yourself with energy-saving appliances

- For laundry and dishwashing. Installing appliances that consume less water (A, A+, A++ energy label) saves water in the long term.
- For sinks and showers. Fitting a mixer tap, thermostatic mixer tap, aerator, economizer, mousseur or eco-showerhead can reduce the flow of water out of the faucet, saving up to 70% water each time you use it.
- For toilets. Fitting a dual-flush system, Eco-WC pads or Eco-WC bags can reduce the amount of water per flush, saving up to 50% water per use.

Envoyer

### C.4.3 Survey

## Assessment of household water consumption

This questionnaire concludes the household consumption survey. You must complete it to the end in order to validate your participation and allow us to pay your compensation for this survey.

During the survey, you will have the opportunity to win an additional 5 euros for questions that are distinctly marked. Please answer all questions honestly.

Suivant

### Bonus Questions 1/2

This page is one of two that will allow you to win an extra 5€.

If this page is selected in the random draw, one of the questions below will be randomly selected as the one to win the extra 5€.

To win the extra 5€, you must find the answer most frequently given by participants in this survey.

\*On this page, you will read about various decisions facing a person, "individual A".

After reading the description of the situation, you will be asked to evaluate the different choices available to individual A and to decide, for each of the possible actions, whether that action would be "socially appropriate" and "consistent with moral or proper social behavior" or "socially inappropriate" and "inconsistent with moral or proper social behavior". By "socially appropriate", we mean an action that most people would consider "correct" or "ethical". Another way of understanding what we mean is that if individual A were to make a socially inappropriate choice, someone else might be angry with individual A for doing so.

In each of your responses, we would like you to answer as truthfully as possible, based on your opinions of what constitutes socially appropriate or socially inappropriate behavior.

If this page is selected in the random draw, we will also randomly select one of the possible choices that individual A might make. For the selected choice, we will determine which answer was chosen by the greatest number of people among the participants in this study. If you give the same answer as the one chosen most often by other people, you'll receive an extra €5.

*Let's assume that Individual A is a well-off person living in an apartment in the Auvergne-Rhône-Alpes region.*

*The table below lists the possible choices available to Individual A. For each choice, please indicate whether you think this option is very socially inappropriate, somewhat socially inappropriate, somewhat socially appropriate or very socially appropriate. To indicate your answer, select the corresponding box. Don't forget that you'll earn money (€5) if your answer to a randomly chosen question is the same as the most common answer given by participants in this study.*

	Very socially inappropriate	Somewhat socially inappropriate	Somewhat socially appropriate	Very socially appropriate
Turn off faucet while brushing teeth or soaping face and hands	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Hand-wash several kitchen utensils at the same time and use the rinse water from one group of utensils as wash water for the next group.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Rinse all your dishes in the same water	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Turn off tap immediately after use	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Take showers lasting less than 5 minutes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Avoid letting the water run in the shower while he shampoos his hair and soaps his body.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Avoid launching under-filled laundry machines	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Avoid launching under-filled dishwashers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Use "eco" modes and short cycles	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Use energy-saving appliances for laundry and dishes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Use energy-saving appliances for sinks and showers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Use energy-saving toilet fixtures	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Suivant

## Bonus Questions 2/2

This page is one of two that will allow you to win an extra 5€.

If this page is selected in the random draw, one of the questions below will be randomly selected as the one to win the extra 5€.

To win the extra 5€, you must find the answer most frequently given by participants in this survey.

\*On this page, you will read about various decisions facing a person, "individual A".

After reading the description of the situation, you will be asked to evaluate the different choices available to individual A and to decide, for each of the possible actions, whether that action would be "socially appropriate" and "consistent with moral or proper social behavior" or "socially inappropriate" and "inconsistent with moral or proper social behavior". By "socially appropriate", we mean an action that most people would consider "correct" or "ethical". Another way of understanding what we mean is that if individual A were to make a socially inappropriate choice, someone else might be angry with individual A for doing so.

In each of your responses, we would like you to answer as truthfully as possible, based on your opinions of what constitutes socially appropriate or socially inappropriate behavior.

If this page is selected in the random draw, we will also randomly select one of the possible choices that individual A might make. For the selected choice, we will determine which answer was chosen by the greatest number of people among the participants in this study. If you give the same answer as the one chosen most often by other people, you'll receive an extra €5.



Let's assume that Individual A is a well-off person living in an apartment in the Auvergne-Rhône-Alpes region.

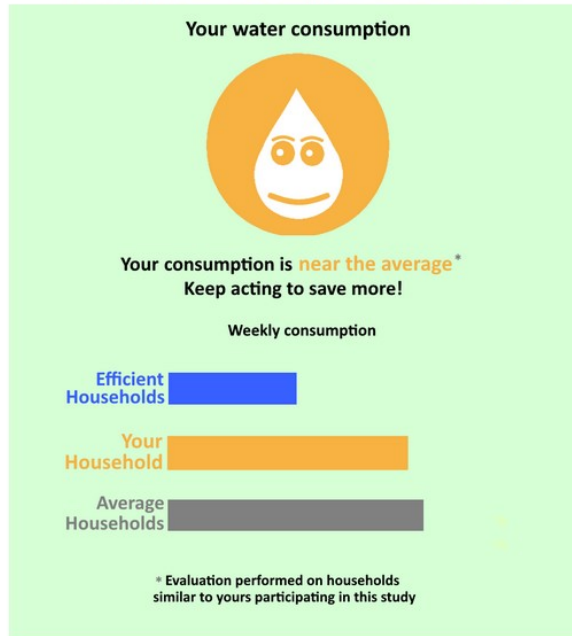
The table below lists the possible choices available to Individual A. For each choice, please indicate whether you think this option is very socially inappropriate, somewhat socially inappropriate, somewhat socially appropriate or very socially appropriate. To indicate your answer, select the corresponding box. Don't forget that you'll earn money (€5) if your answer to a randomly chosen question is the same as the most common answer given by participants in this study.

	Very socially inappropriate	Somewhat socially inappropriate	Somewhat socially appropriate	Very socially appropriate
Do not turn off the tap while brushing teeth or soaping face and hands	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Do not hand-wash several kitchen utensils at the same time and use the rinse water from one group of utensils as wash water for the next group.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Do not rinse all your dishes in the same water	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Do not close the tap immediately after use	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Do not take showers lasting less than 5 minutes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Let the water run in the shower while he shampoos his hair and soaps his body.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Launching underfilled laundry machines	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Launching shallow dishwashers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Do not use "eco" or short-cycle modes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Do not use energy-saving appliances for laundry and dishes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Do not use energy-saving appliances for sinks and showers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Do not use energy-saving toilets	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Suivant

## Visual of the third week of study

You received this visual during the third week of study:



\*Which of the following statements best describes your experience?

📌 Veuillez sélectionner une réponse ci-dessous

- I remember this visual
- I do not remember this visual
- I am sure I have not seen this visual



\*How did you feel about the visual above?

Use the following scale to answer:

- Strongly disagree
- Somewhat disagree
- Undecided
- Somewhat agree
- Strongly agree

	Strongly disagree	Somewhat disagree	Undecided	Somewhat agree	Strongly agree
Interested	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Jittery	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Excited	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Distressed	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Strong	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ashamed	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Scared	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Hostile	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Enthusiastic	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Proud	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Upset	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Alert	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Guilty	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Inspired	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Nervous	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Determined	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Attentive	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Irritable	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Active	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Afraid	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Suivant

## Visual of the fifth week of study

You received this visual during the fifth week of study:



\*Which of the following sentences best describes your experience?

📌 Veuillez sélectionner une réponse ci-dessous

- I remember this visual
- I do not remember this visual
- I am sure I have not seen this visual

\*How did you feel about the visual above?

Use the following scale to answer:

• Strongly disagree • Somewhat disagree • Undecided • Somewhat agree • Strongly agree

	Strongly disagree	Somewhat disagree	Undecided	Somewhat agree	Strongly agree
Interested	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Jittery	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Excited	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Distressed	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Strong	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ashamed	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Scared	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Hostile	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Enthusiastic	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Proud	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Upset	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Alert	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Guilty	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Inspired	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Nervous	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Determined	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Attentive	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Irritable	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Active	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Afraid	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Suivant

## What do you think of water conservation practices?

\*Implementing water conservation practices is :

📌 Veuillez sélectionner une réponse ci-dessous

- Not possible for me     Somewhat not possible for me     Undecided     Somewhat possible for me     Possible for me

\*Implementing water conservation practices is :

📌 Veuillez sélectionner une réponse ci-dessous

- Not easy for me     Somewhat not easy for me     Undecided     Somewhat easy for me     Easy for me

\*Implementing water conservation practices is :

📌 Veuillez sélectionner une réponse ci-dessous

- Not under my control     Somewhat not under my control     Undecided     Somewhat under my control     Under my control

\*Implementing water conservation practices is :

📌 Veuillez sélectionner une réponse ci-dessous

- Does not depend on me     Somewhat does not depend on me     Undecided     Somewhat depend on me     Depends on me

\*Implementing water conservation practices is :

📌 Veuillez sélectionner une réponse ci-dessous

- Not practical for me     Somewhat not practical for me     Undecided     Somewhat practical for me     Practical for me

Suivant

## What do you think of water conservation practices?

\*Water conservation practices won't change my daily life:

📌 Veuillez sélectionner une réponse ci-dessous

- Strongly disagree     Somewhat disagree     Undecided     Agree     Strongly agree

\*Water conservation practices do not require extra time or physical effort:

📌 Veuillez sélectionner une réponse ci-dessous

- Strongly disagree     Somewhat disagree     Undecided     Somewhat agree     Strongly agree

Suivant

## What do you think of the environment?

\*The natural environment is very important and fragile:

📌 Veuillez sélectionner une réponse ci-dessous

- Strongly disagree     Somewhat disagree     Undecided     Somewhat agree     Strongly agree

\*In recent years, natural disasters have become more frequent and more severe:

📌 Veuillez sélectionner une réponse ci-dessous

- Strongly disagree     Somewhat disagree     Undecided     Somewhat agree     Strongly agree

\*Human beings have the right to modify nature to meet their daily needs:

📌 Veuillez sélectionner une réponse ci-dessous

- Strongly disagree     Somewhat disagree     Undecided     Somewhat agree     Strongly agree

\*Water is a very important natural resource:

📌 Veuillez sélectionner une réponse ci-dessous

- Strongly disagree     
  Somewhat disagree     
  Undecided     
  Somewhat agree     
  Strongly agree

\*Reducing water consumption is a way of protecting the environment:

📌 Veuillez sélectionner une réponse ci-dessous

- Strongly disagree     
  Somewhat disagree     
  Undecided     
  Somewhat agree     
  Strongly agree

Suivant

## Water-saving behaviors and devices

\*Please indicate the water conservation behaviors that you had before starting this study and that you realized during this study:

	Comportements antérieurs à l'étude			Comportements durant l'étude	
	Yes	No		Yes	No
I cook with as little water as possible	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>	<input type="radio"/>
I used to soak pots and pans instead of letting the water run when I washed the dishes.	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>	<input type="radio"/>
When I hand-washed several kitchen utensils at the same time, I used the rinse water from one group of utensils as wash water for the next.	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>	<input type="radio"/>
I turned off the tap while brushing my teeth or soaping my face and hands.	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>	<input type="radio"/>
I avoided letting the water run in the shower while I shampooed my hair and soaped my body.	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>	<input type="radio"/>
I filled my washing machine as much as possible	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>	<input type="radio"/>
I filled my dishwasher as much as possible	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>	<input type="radio"/>
I was thinking of "eco" modes and short cycles	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>	<input type="radio"/>
I took showers lasting less than 5 minutes	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>	<input type="radio"/>

\*Please indicate which water conservation devices you had before starting this study and which you purchased during this study:

	Appareils antérieurs à l'étude		Appareils achetés durant l'étude	
	Yes	No	Yes	No
Water-saving washing machines	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Low-water consumption dishwasher	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sink flow-reducing mixer valve	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Shower head to reduce shower flow	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Dual-flush system	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Installation to reduce the amount of water used per flush	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Suivant

## What do you think of water conservation?

\*Water conservation practices can solve water scarcity problems:

📍 Veuillez sélectionner une réponse ci-dessous

- Strongly disagree   
  Somewhat disagree   
  Undecided   
  Somewhat agree   
  Strongly agree

\*Water conservation significantly reduces water bills:

📍 Veuillez sélectionner une réponse ci-dessous

- Strongly disagree   
  Somewhat disagree   
  Undecided   
  Somewhat agree   
  Strongly agree

\*Water conservation will considerably transform the environment:

📍 Veuillez sélectionner une réponse ci-dessous

- Strongly disagree   
  Somewhat disagree   
  Undecided   
  Somewhat agree   
  Strongly agree

## What is your position on water resources?

**\*I personally feel obliged to explore ways of reducing my impact on water resources:**

📌 Veuillez sélectionner une réponse ci-dessous

- Strongly disagree     Somewhat disagree     Undecided     Somewhat agree     Strongly agree

**\*I personally feel obliged to minimize my impact on local water resources:**

📌 Veuillez sélectionner une réponse ci-dessous

- Strongly disagree     Somewhat disagree     Undecided     Somewhat agree     Strongly agree

**\*It is important to use as little water as possible:**

📌 Veuillez sélectionner une réponse ci-dessous

- Strongly disagree     Somewhat disagree     Undecided     Somewhat agree     Strongly agree

**\*It is important to encourage my friends and family to protect our water resources:**

📌 Veuillez sélectionner une réponse ci-dessous

- Strongly disagree     Somewhat disagree     Undecided     Somewhat agree     Strongly agree

**\*I should be responsible for doing everything in my power to protect water resources:**

📌 Veuillez sélectionner une réponse ci-dessous

- Strongly disagree     Somewhat disagree     Undecided     Somewhat agree     Strongly agree

Suivant



## Information about you and your household

\*Please indicate the status that best corresponds to you?

📌 Veuillez sélectionner une réponse ci-dessous

Veillez choisir ...

\*Please indicate your highest diploma :

📌 Veuillez sélectionner une réponse ci-dessous

📌 Au besoin, veuillez préciser le champ 'Autre' :

Veillez choisir ...

\*Please indicate the field of study of your highest degree:

📌 Veuillez sélectionner une réponse ci-dessous

Veillez choisir ...

\*Please indicate your approximate annual household income :

📌 Veuillez sélectionner une réponse ci-dessous

- Less than 30 000€       Between 30 000€ and 40 000€       Between 40 000€ and 50 000€       Between 50 000€ and 60 000€       More than 60 000€

\*How do you feel financially :

📌 Veuillez sélectionner une réponse ci-dessous

- Very limited       Limited       comfortable       Very comfortable

\*Do you own your home?

📌 Veuillez sélectionner une réponse ci-dessous

- Yes  
 No

**\*On average over the entire study, how many days a week did you stay at home most of the day?**

📌 Votre réponse doit être comprise entre 0 et 7

📌 Seul un nombre entier peut être inscrit dans ce champ.

Jour(s)

**\*On average over the entire study, how many days a week were you completely absent from your home?**

📌 Votre réponse doit être comprise entre 0 et 7

📌 Seul un nombre entier peut être inscrit dans ce champ.

Jour(s)

**\*Please indicate the living area of your home (in square meters)?**

📌 Seul un nombre entier peut être inscrit dans ce champ.

m<sup>2</sup>

**\*Please enter your postal code :**

📌 Seul un nombre entier peut être inscrit dans ce champ.

**\*Please indicate the type of your accommodation :**

📌 Veuillez sélectionner une réponse ci-dessous

Apartment

House

**\*Do you have a swimming pool or Jacuzzi?**

📌 Veuillez sélectionner une réponse ci-dessous

Yes - I have a swimming pool and jacuzzi

Yes - I have a pool or jacuzzi

No - I have neither

**\*Do you have a washing machine?**

📌 Veuillez sélectionner une réponse ci-dessous

Yes

No

**\*Do you have a dishwasher?**

📌 Veuillez sélectionner une réponse ci-dessous

- Yes
- No

**\*Has the number of people living in your household changed over the weeks?**

📌 Veuillez sélectionner une réponse ci-dessous

- Yes - Some weeks, there were more or fewer residents in my home.
- No - No one stayed at my house for more than 4 days a week during the 7-week study.

Suivant

Composition of your household - including yourself

**\*How many men over 13 are there in your household?**

📌 Seul un nombre entier peut être inscrit dans ce champ.

**\*How many women over the age of 13 are there in your household?**

📌 Seul un nombre entier peut être inscrit dans ce champ.

**\*How many males under 13 are there in your household?**

📌 Seul un nombre entier peut être inscrit dans ce champ.

**\*How many women under 13 are there in your household?**

📌 Seul un nombre entier peut être inscrit dans ce champ.

Suivant

Change in household composition during the first week (May 30 to June 6)

\*During the first week, did your household count one or more people more or less than usual? (For at least 4 days of the week)

ⓘ Veuillez sélectionner une réponse ci-dessous

- No
- Yes - One or more additional persons
- Yes - One or more less persons

\*How many more or fewer people are affected by the previous question?

ⓘ Veuillez sélectionner une réponse ci-dessous

2 person ▾

\*The first person concerned is :

ⓘ Veuillez sélectionner une réponse ci-dessous

A man over 13 years old ▾

\*The second person concerned is :

ⓘ Veuillez sélectionner une réponse ci-dessous

A woman under 13 ▾

Suivant

Change in household composition during second week (June 6 to June 13)

\*During the second week, did your household count one or more people more or less than usual? (For at least 4 days of the week)

ⓘ Veuillez sélectionner une réponse ci-dessous

- No
- Yes - One or more additional persons
- Yes - One or more less persons

Suivant

Thank you for completing this questionnaire. You have completed the household water consumption survey. Your answers to the different phases of the study will soon be studied in order to proceed with your compensation.

You will receive an e-mail and a text message within two weeks, allowing you to receive your compensation.

The Grenoble Applied Economics Laboratory thanks you for your participation.

#### **C.4.4 Preregistration**

**CONFIDENTIAL - FOR PEER-REVIEW ONLY****DynEau - Dynamic norms on the field of water consumption, Grenoble, June 2023 (#131119)**

Created: 05/04/2023 10:07 AM (PT)

---

This is an anonymized copy (without author names) of the pre-registration. It was created by the author(s) to use during peer-review. A non-anonymized version (containing author names) should be made available by the authors when the work it supports is made public.

---

**1) Have any data been collected for this study already?**

No, no data have been collected for this study yet.

**2) What's the main question being asked or hypothesis being tested in this study?**

Dynamic information-based policies induce a greater reduction in water consumption than Statics information-based policies.

**3) Describe the key dependent variable(s) specifying how they will be measured.**

Water consumption calculated from participants' water meters. Each participant will send a picture of their water meter on a weekly basis.

**4) How many and which conditions will participants be assigned to?**

On the second week, participants will be assigned to treatment and baseline group with stratified random sampling technique taking into account Household characteristics and weekly water consumption. On the fourth week, the treatment group will receive a dynamic information-based policy and baseline group will receive a static information-based policy.

**5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.**

- Linear regression predicting the water consumption after the intervention with a dummy variable indicating whether the participant received a Dynamic information-based policy or not. We will also conduct the same regression controlling for individual household characteristics and water consumption categories.
- Difference-in-difference models predicting the water consumption after the intervention between treatment and baseline. We will also conduct the same models controlling for individual household characteristics and water consumption categories.
- Independent sample t-tests to test the equality of mean in weekly water consumption between treatment and baseline, as well as between water consumption categories.
- One-sample Wilcoxon rank sum test to test equality of distributions in weekly water consumption between treatment and baseline, as well as between water consumption categories.

**6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.**

We will exclude participants who do not provide their individual water meter image each week as well as those who do not complete the final survey. Participants who fully comply with the procedures explained before will be considered regardless of their water consumption level.

**7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.**

Our goal is to recruit 300 to 400 participants for this experiment. This range is given due to the uncertainty of number of participating households with similar characteristic.

**8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)**

We will include an emotional evaluation PANAS (Watson et al., 1988 ; French version by Caci & Baylé, 2007) on each information-based policy to measure the emotional response of participants.

We will include a normative evaluation of water conservation behavior (Krupka and Weber, 2013) to measure the impact of information-based policies on normative assessment of the norms.

Final survey contains several questions for exploratory purposes, assessing Perceived Behavioral control, Attitude toward conservation, Expected results, Past water conservation behavior, Actual water conservation behavior, Personal/Moral norms, Social-demographics information.

Common to two treatments, each participant will received a Statics information-based policy at the week two (built similarly to Bhanot, 2018) in order to induce the first behavioral change needed for the creation of the Dynamic information-based policy. Four Statics information-based policies will be implemented according to the relative position of the participant's consumption to the panel's average and efficient consumption. These four Statics information-based policies constitute the four water consumption categories.

## **Appendix D**

# **Disentangling Individual Attitudes, Expectations, and Experience in the Adoption of Proenvironmental Behavior - Supplementary**

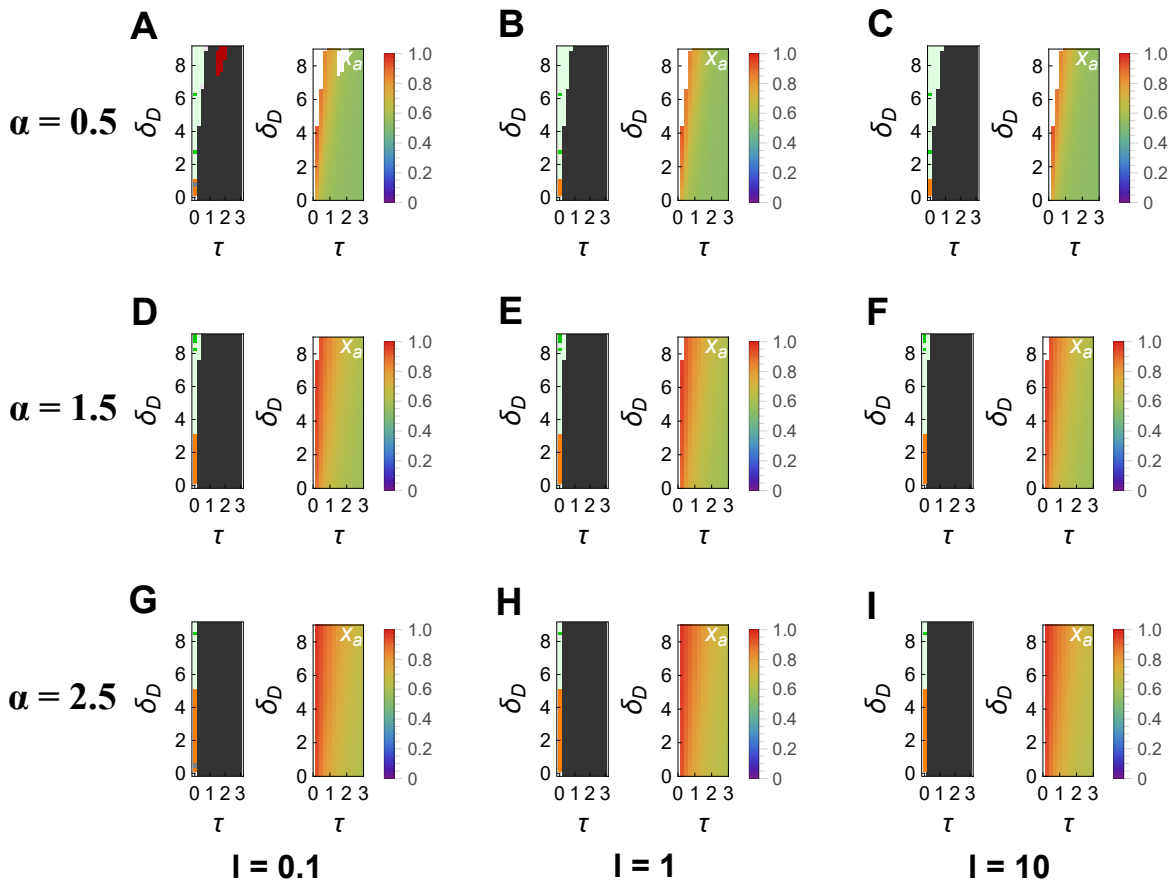


Figure D.1: **Effect of attitudes and attitude strength on the adoption of the active behavior.** For each value of the environmental reactivity,  $l$ , and attitude strength,  $\alpha_a = \alpha_b = \alpha$ , the equilibrium frequency of the active behavior is shown as a function of the environmental sensitivity,  $\tau$ , and strength of the descriptive component of the social norm,  $\delta_D$ . **A**,  $l = 0.1$ ,  $\alpha = 0.5$ . **B**,  $l = 1$ ,  $\alpha = 0.5$ . **C**,  $l = 10$ ,  $\alpha = 0.5$ . **D**,  $l = 0.1$ ,  $\alpha = 1.5$ . **E**,  $l = 1$ ,  $\alpha = 1.5$ . **F**,  $l = 10$ ,  $\alpha = 1.5$ . **G**,  $l = 0.1$ ,  $\alpha = 2.5$ . **H**,  $l = 1$ ,  $\alpha = 2.5$ . **I**,  $l = 10$ ,  $\alpha = 2.5$ . See Fig. 5.1 for details. Parameters are set to their default values (Table 5.1). The strength of the normative component of the social norm is  $\delta_N = 2$ . Proenvironmental attitude frequency,  $P_a = 1$ .



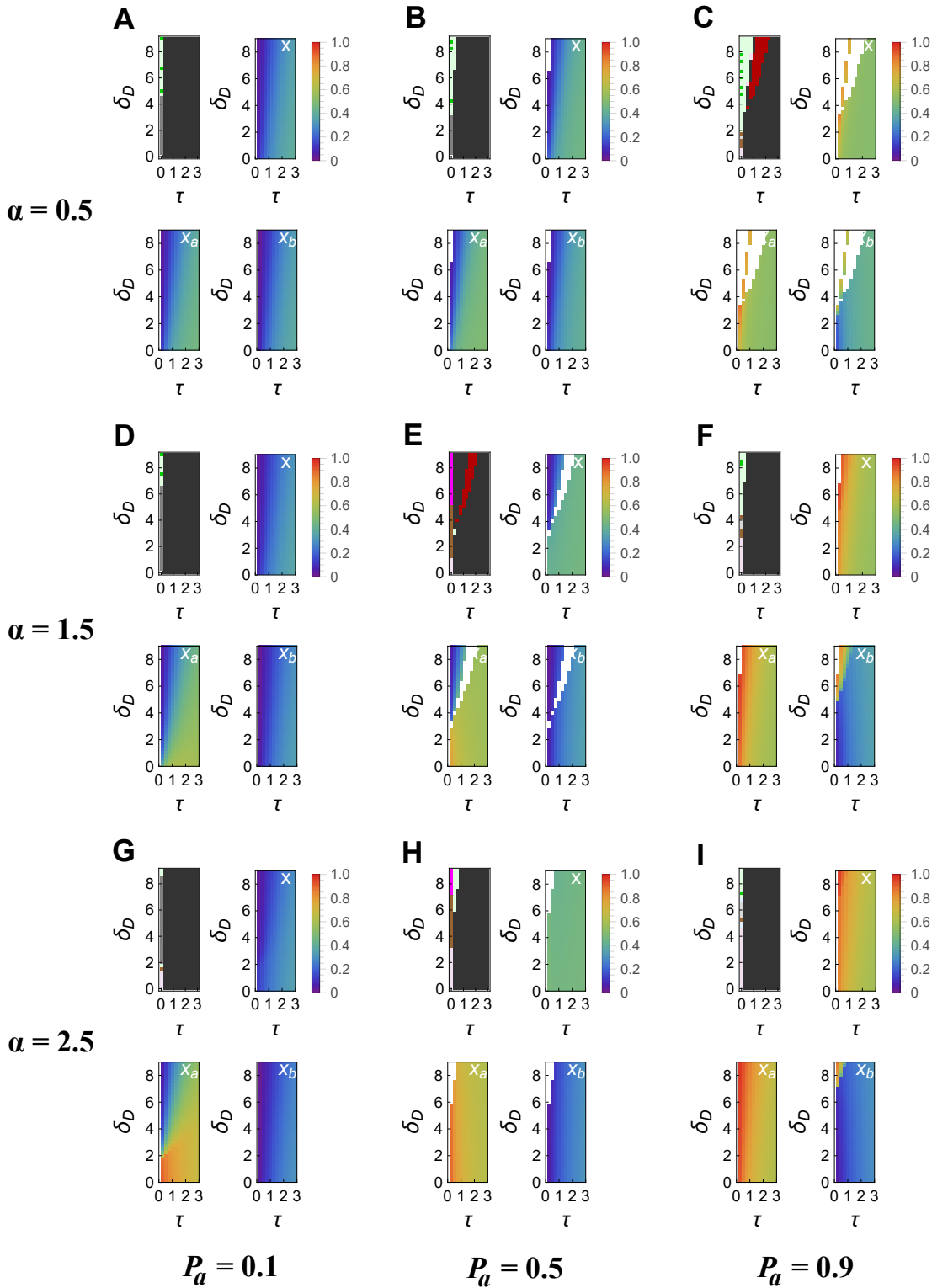


Figure D.2: **Effect of attitudes and attitude strength on the adoption of the active behavior.** For each value of proenvironmental attitude frequency,  $P_a$ , and attitude strength,  $\alpha_a = \alpha_b = \alpha$ , the equilibrium frequency of the active behavior is shown as a function of the environmental sensitivity,  $\tau$ , and strength of the descriptive component of the social norm,  $\delta_D$ . **A**,  $P_a = 0.1$ ,  $\alpha = 0.5$ . **B**,  $P_a = 0.5$ ,  $\alpha = 0.5$ . **C**,  $P_a = 0.9$ ,  $\alpha = 0.5$ . **D**,  $P_a = 0.1$ ,  $\alpha = 1.5$ . **E**,  $P_a = 0.5$ ,  $\alpha = 1.5$ . **F**,  $P_a = 0.9$ ,  $\alpha = 1.5$ . **G**,  $P_a = 0.1$ ,  $\alpha = 2.5$ . **H**,  $P_a = 0.5$ ,  $\alpha = 2.5$ . **I**,  $P_a = 0.9$ ,  $\alpha = 2.5$ . See Fig. 5.1 for details. Parameters are set to their default values (Table 5.1). The strength of the normative component of the social norm is  $\delta_N = 2$ . Environmental reactivity is  $l = 0.1$ .

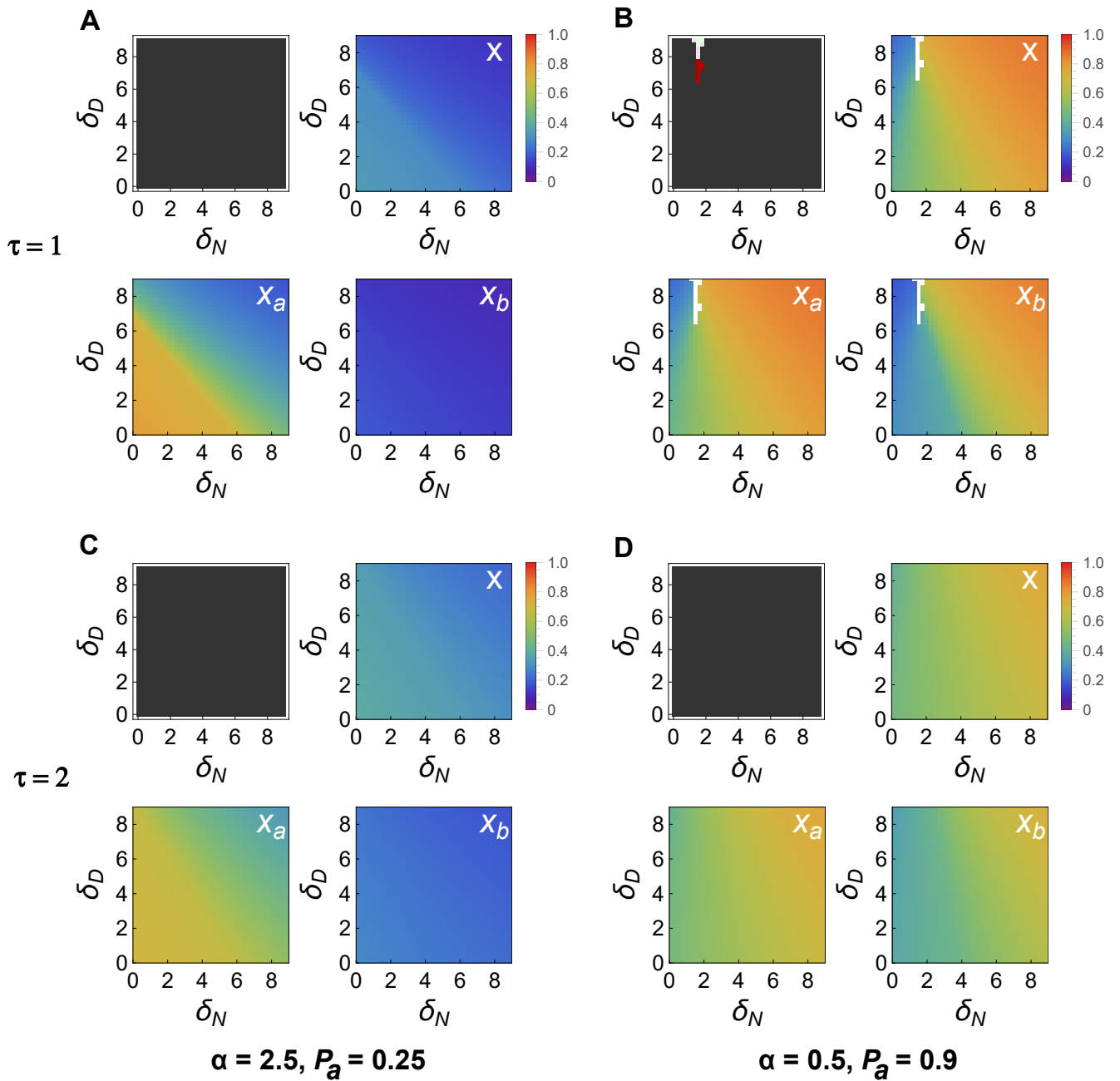


Figure D.3: **Effect of the strength of the descriptive and normative components of the social norm,  $\delta_D$  and  $\delta_N$ , on the adoption of the active behavior.** **A**, environmental sensitivity is  $\tau = 1$ , attitude strength is  $\alpha_a = \alpha_b = \alpha = 2.5$  and the proenvironmental attitude frequency is  $P_a = 0.1$ . **B**, environmental sensitivity is  $\tau = 1$ , attitude strength is  $\alpha_a = \alpha_b = \alpha = 0.5$  and the proenvironmental attitude frequency is  $P_a = 0.9$ . **C**, environmental sensitivity is  $\tau = 2$ , attitude strength is  $\alpha_a = \alpha_b = \alpha = 2.5$  and the proenvironmental attitude frequency is  $P_a = 0.1$ . **D**, environmental sensitivity is  $\tau = 2$ , attitude strength is  $\alpha_a = \alpha_b = \alpha = 0.5$  and the proenvironmental attitude frequency is  $P_a = 0.9$ . See Fig. 5.1 for details. Parameters are set to their default values (Table 5.1). Other parameters: environmental reactivity is  $l = 1$ .

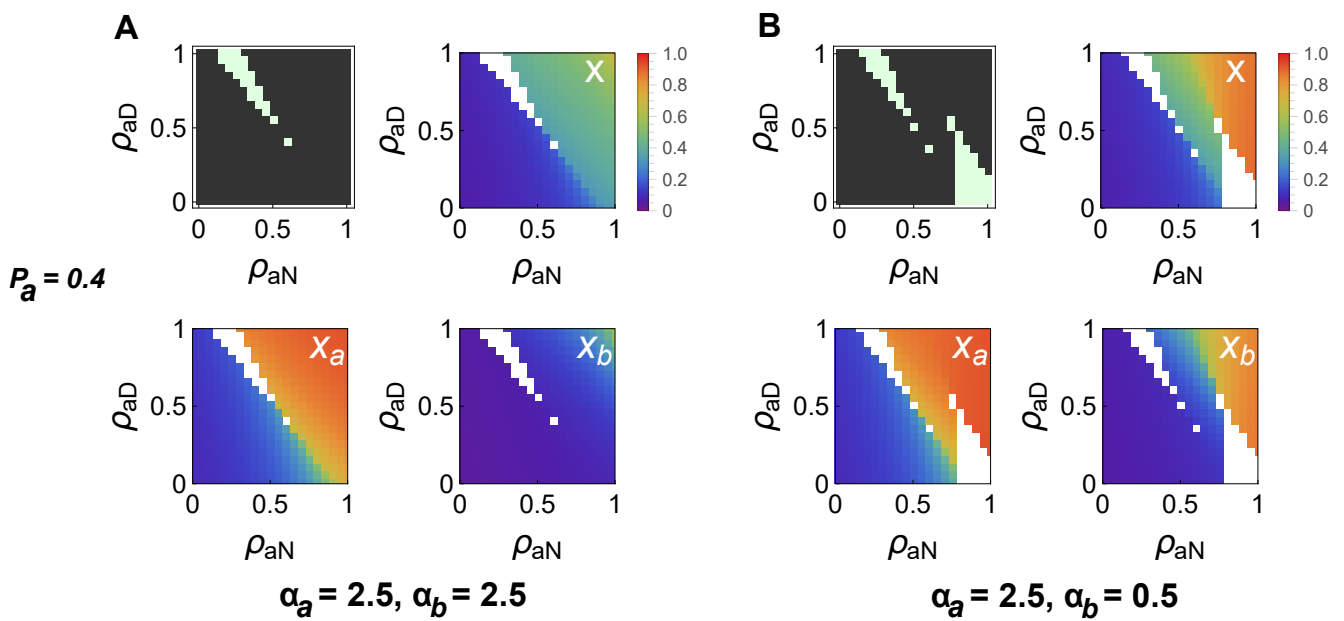


Figure D.4: **Effect of descriptive focus,  $\rho_{aD}$ , and normative focus,  $\rho_{aN}$ , on the adoption of the active behavior.** **A**,  $P_a = 0.1$ ,  $\alpha_a = 2.5$ ,  $\alpha_b = 2.5$ . **B**,  $P_a = 0.1$ ,  $\alpha_a = 2.5$ ,  $\alpha_b = 0.5$ . **C**,  $P_a = 0.5$ ,  $\alpha_a = 2.5$ ,  $\alpha_b = 2.5$ . **D**,  $P_a = 0.5$ ,  $\alpha_a = 2.5$ ,  $\alpha_b = 0.5$ . See Fig. 5.1 for details. Parameters are set to their default values (Table 5.1). Other parameters: strength of descriptive expectations is  $\delta_D = 6$ , strength of normative expectations is  $\delta_N = 4.75$ , environmental reactivity is  $l = 1$ , environmental sensitivity is  $\tau = 0.5$ . Proenvironmental attitude frequency,  $P_a = 0.4$ .

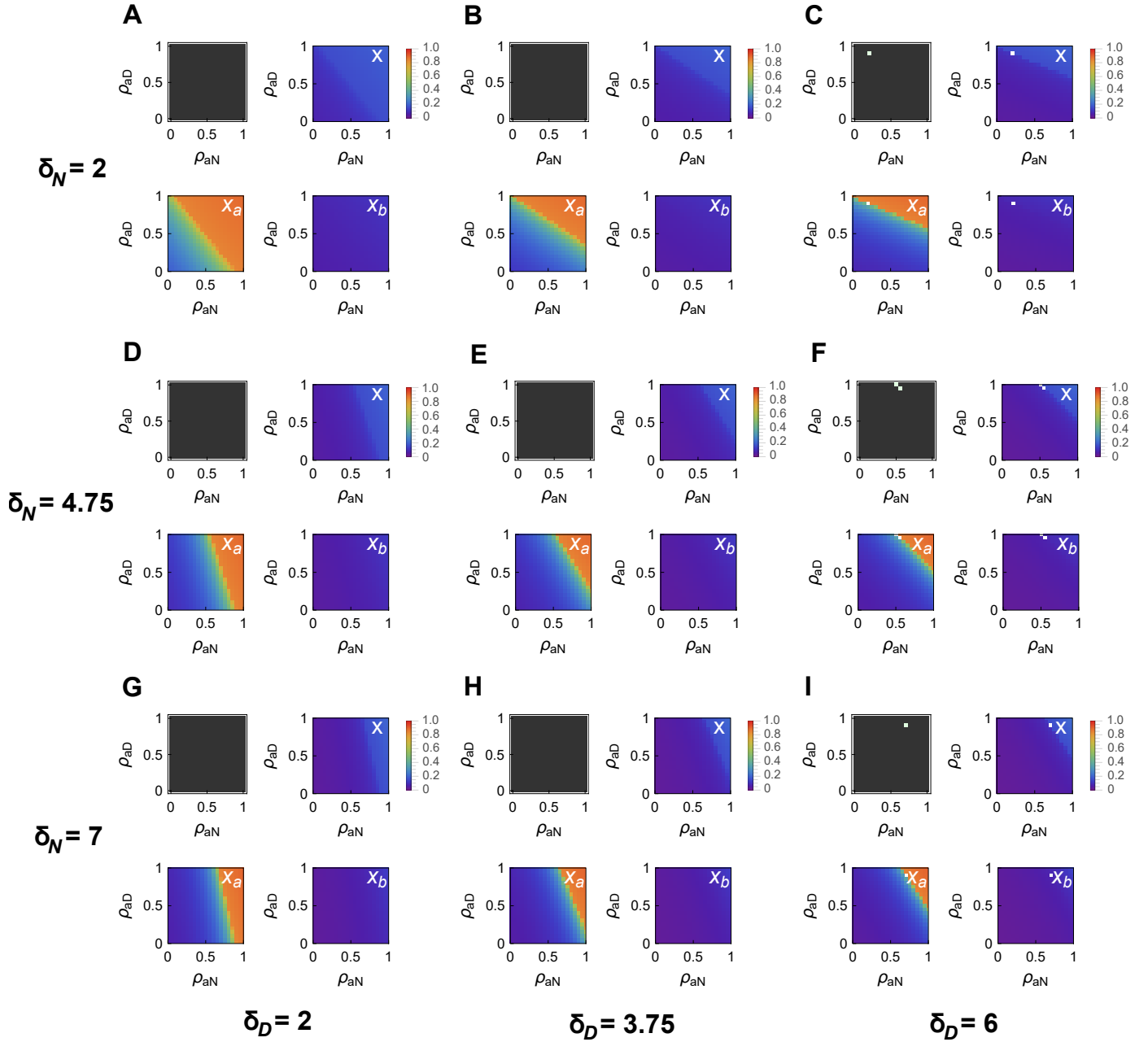


Figure D.5: **Effect of descriptive focus,  $\rho_{aD}$ , and normative focus,  $\rho_{aN}$ , on the adoption of the active behavior.** **A**,  $\delta_N = 2$ ,  $\delta_D = 2$ . **B**,  $\delta_N = 2$ ,  $\delta_D = 3.75$ . **C**,  $\delta_N = 2$ ,  $\delta_D = 6$ . **D**,  $\delta_N = 4.75$ ,  $\delta_D = 2$ . **E**,  $\delta_N = 4.75$ ,  $\delta_D = 3.75$ . **F**,  $\delta_N = 4.75$ ,  $\delta_D = 6$ . **G**,  $\delta_N = 7$ ,  $\delta_D = 2$ . **H**,  $\delta_N = 7$ ,  $\delta_D = 3.75$ . **I**,  $\delta_N = 7$ ,  $\delta_D = 6$ . See Fig. 5.1 for details. Parameters are set to their default values (Table 5.1). Other parameters: proenvironmental attitude frequency is  $P_a = 0.1$ , attitude strengths are  $\alpha_a = 2.5$  and  $\alpha_b = 2.5$ .

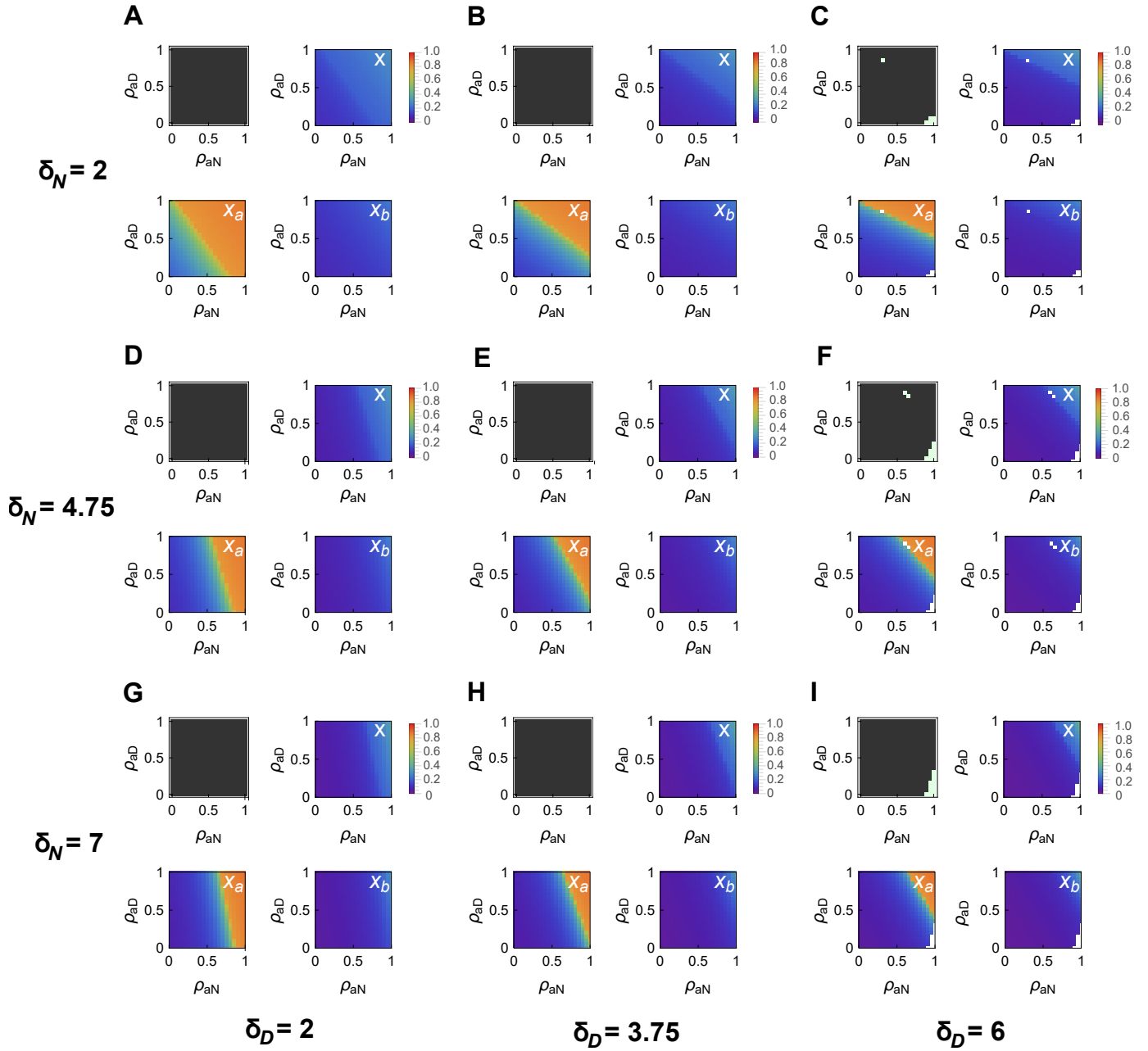


Figure D.6: **Effect of descriptive focus,  $\rho_{aD}$ , and normative focus,  $\rho_{aN}$ , on the adoption of the active behavior.** **A**,  $\delta_N = 2$ ,  $\delta_D = 2$ . **B**,  $\delta_N = 2$ ,  $\delta_D = 3.75$ . **C**,  $\delta_N = 2$ ,  $\delta_D = 6$ . **D**,  $\delta_N = 4.75$ ,  $\delta_D = 2$ . **E**,  $\delta_N = 4.75$ ,  $\delta_D = 3.75$ . **F**,  $\delta_N = 4.75$ ,  $\delta_D = 6$ . **G**,  $\delta_N = 7$ ,  $\delta_D = 2$ . **H**,  $\delta_N = 7$ ,  $\delta_D = 3.75$ . **I**,  $\delta_N = 7$ ,  $\delta_D = 6$ . See Fig. 5.1 for details. Parameters are set to their default values (Table 5.1). Other parameters: proenvironmental attitude frequency is  $P_a = 0.1$ , attitude strengths are  $\alpha_a = 2.5$  and  $\alpha_b = 0.5$ .

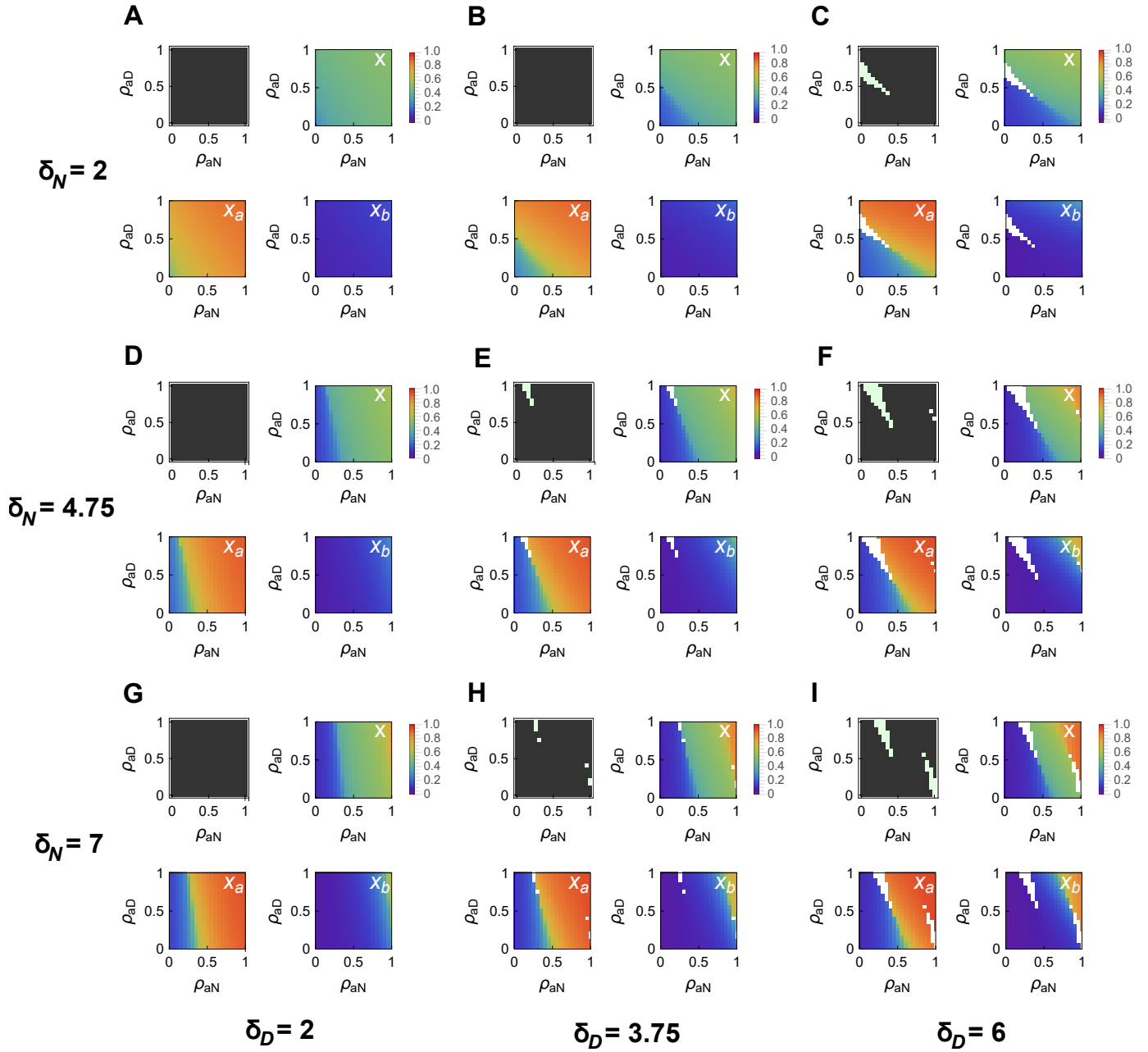


Figure D.7: **Effect of descriptive focus,  $\rho_{aD}$ , and normative focus,  $\rho_{aN}$ , on the adoption of the active behavior.** **A**,  $\delta_N = 2$ ,  $\delta_D = 2$ . **B**,  $\delta_N = 2$ ,  $\delta_D = 3.75$ . **C**,  $\delta_N = 2$ ,  $\delta_D = 6$ . **D**,  $\delta_N = 4.75$ ,  $\delta_D = 2$ . **E**,  $\delta_N = 4.75$ ,  $\delta_D = 3.75$ . **F**,  $\delta_N = 4.75$ ,  $\delta_D = 6$ . **G**,  $\delta_N = 7$ ,  $\delta_D = 2$ . **H**,  $\delta_N = 7$ ,  $\delta_D = 3.75$ . **I**,  $\delta_N = 7$ ,  $\delta_D = 6$ . See Fig. 5.1 for details. Parameters are set to their default values (Table 5.1). Other parameters: proenvironmental attitude frequency is  $P_a = 0.5$ , attitude strengths are  $\alpha_a = 2.5$  and  $\alpha_b = 2.5$ .

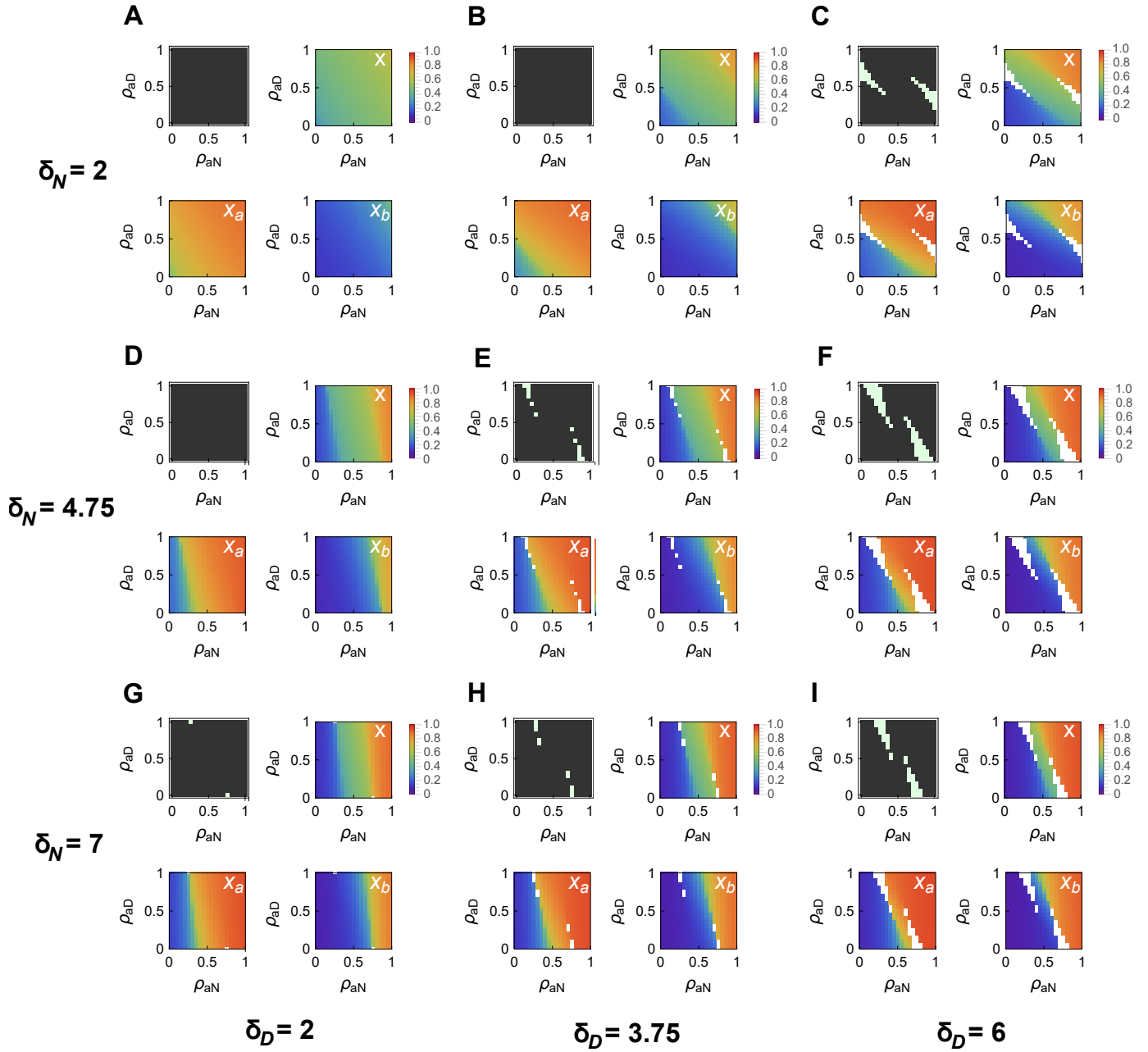


Figure D.8: **Effect of descriptive focus,  $\rho_{aD}$ , and normative focus,  $\rho_{aN}$ , on the adoption of the active behavior.** **A**,  $\delta_N = 2$ ,  $\delta_D = 2$ . **B**,  $\delta_N = 2$ ,  $\delta_D = 3.75$ . **C**,  $\delta_N = 2$ ,  $\delta_D = 6$ . **D**,  $\delta_N = 4.75$ ,  $\delta_D = 2$ . **E**,  $\delta_N = 4.75$ ,  $\delta_D = 3.75$ . **F**,  $\delta_N = 4.75$ ,  $\delta_D = 6$ . **G**,  $\delta_N = 7$ ,  $\delta_D = 2$ . **H**,  $\delta_N = 7$ ,  $\delta_D = 3.75$ . **I**,  $\delta_N = 7$ ,  $\delta_D = 6$ . See Fig. 5.1 for details. Parameters are set to their default values (Table 5.1). Other parameters: proenvironmental attitude frequency is  $P_a = 0.5$ , attitude strengths are  $\alpha_a = 2.5$  and  $\alpha_b = 0.5$ .





# Bibliography

- Aghion, Philippe, and Jeffrey G. Williamson.** 1998. *Growth, Inequality, and Globalization: Theory, History, and Policy*. Cambridge University Press.
- Agthe, Donald E., and R. Bruce Billings.** 1987. "Equity, Price Elasticity, and Household Income Under Increasing Block Rates for Water." *The American Journal of Economics and Sociology*, 46(3): 273–286.
- Ahn, Insook, Soo Hyun Kim, and Munyoung Kim.** 2020. "The Relative Importance of Values, Social Norms, and Enjoyment-Based Motivation in Explaining Pro-Environmental Product Purchasing Behavior in Apparel Domain." *Sustainability*, 12(17): 6797.
- Ahn, T. K., Myungsuk Lee, Lore Ruttan, and James Walker.** 2007. "Asymmetric payoffs in simultaneous and sequential prisoner's dilemma games." *Public Choice*, 132(3): 353–366.
- Ajzen, Icek.** 1991. "The theory of planned behavior." *Organizational Behavior and Human Decision Processes*, 50(2): 179–211.
- Akerlof, George A.** 1980. "A Theory of Social Custom, of which Unemployment may be One Consequence\*." *The Quarterly Journal of Economics*, 94(4): 749–775.
- Allcott, Hunt.** 2011. "Social norms and energy conservation." *Journal of Public Economics*, 95(9): 1082–1095.
- Amarasinghe, Pradeep, An Liu, Prasanna Egodawatta, Paul Barnes, James McGree, and Ashantha Goonetilleke.** 2016. "Quantitative assessment of resilience of a water supply system under rainfall reduction due to climate change." *Journal of Hydrology*, 540: 1043–1052.
- Anderson, Lisa R., Jennifer M. Mellor, and Jeffrey Milyo.** 2008. "Inequality and public good provision: An experimental analysis." *The Journal of Socio-Economics*, 37(3): 1010–1028.
- Andreoni, James, Nikos Nikiforakis, and Simon Siegenthaler.** 2021. "Predicting social tipping and norm change in controlled experiments." *Proceedings of the National Academy of Sciences*, 118(16): e2014893118.
- Andreoni, James, Paul M. Brown, and Lise Vesterlund.** 2002. "What Makes an Allocation Fair? Some Experimental Evidence." *Games and Economic Behavior*, 40(1): 1–24.

- Ayres, Ian, Sophie Raseman, and Alice Shih.** 2013. "Evidence from Two Large Field Experiments that Peer Comparison Feedback Can Reduce Residential Energy Usage." *The Journal of Law, Economics, and Organization*, 29(5): 992–1022.
- Baland, Jean-Marie, Pranab Bardhan, and Samuel Bowles.** 2007. *Inequality, Cooperation, and Environmental Sustainability*. Princeton University Press.
- Bandiera, Oriana, Iwan Barankay, and Imran Rasul.** 2006. "The Evolution of Cooperative Norms: Evidence from a Natural Field Experiment." *The B.E. Journal of Economic Analysis & Policy*, 6(2).
- Banerjee, Abhijit, Dilip Mookherjee, Kaivan Munshi, and Debraj Ray.** 2001. "Inequality, Control Rights, and Rent Seeking: Sugar Cooperatives in Maharashtra." *Journal of Political Economy*, 109(1): 138–190.
- Banerjee, Abhijit V., and Esther Duflo.** 2005. "Chapter 7 Growth Theory through the Lens of Development Economics." In *Handbook of Economic Growth*. Vol. 1, , ed. Philippe Aghion and Steven N. Durlauf, 473–552. Elsevier.
- Banerjee, Ritwik.** 2016a. "Corruption, norm violation and decay in social capital." *Journal of Public Economics*, 137: 14–27.
- Banerjee, Ritwik.** 2016b. "On the interpretation of bribery in a laboratory corruption game: moral frames and social norms." *Experimental Economics*, 19(1): 240–267.
- Barr, Abigail, Tom Lane, and Daniele Nosenzo.** 2018. "On the social inappropriateness of discrimination." *Journal of Public Economics*, 164: 153–164.
- Battigalli, Pierpaolo, Emiliano Catonini, and Nicodemo De Vito.** n.d.. "GAME THEORY : Analysis of Strategic Thinking."
- Bašić, Zvonimir, and Eugenio Verrina.** 2023. "Personal Norms — and Not Only Social Norms — Shape Economic Behavior." *MPI Collective Goods Discussion Paper, (2020/25)*.
- Beckenkamp, Martin, Heike Hennig-Schmidt, and Frank P. Maier-Rigaud.** 2007. "Cooperation in Symmetric and Asymmetric Prisoner's Dilemma Games." *MPI Collective Goods Preprint, (2006/25)*.
- Bentham, Jeremy,** ed. 1970. *The Collected Works of Jeremy Bentham: An Introduction to the Principles of Morals and Legislation*. New York:Clarendon Press.
- Bernheim, B. Douglas.** 1994. "A Theory of Conformity." *Journal of Political Economy*, 102(5): 841–877.
- Bhanot, Syon P.** 2021. "Isolating the effect of injunctive norms on conservation behavior: New evidence from a field experiment in California." *Organizational Behavior and Human Decision Processes*, 163: 30–42.

- Bicchieri, Cristina.** 2005. *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge University Press.
- Bicchieri, Cristina.** 2016. *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms*. Oxford University Press.
- Bicchieri, Cristina.** 2023. "Norm nudging and twisting preferences." *Behavioural Public Policy*, 7(4): 914–923.
- Bicchieri, Cristina, and Alex Chavez.** 2010. "Behaving as expected: Public information and fairness norms." *Journal of Behavioral Decision Making*, 23(2): 161–178.
- Bicchieri, Cristina, and Erte Xiao.** 2009. "Do the right thing: but only if others do so." *Journal of Behavioral Decision Making*, 22(2): 191–208.
- Bicchieri, Cristina, and Eugen Dimant.** 2022. "Nudging with care: the risks and benefits of social information." *Public Choice*, 191(3): 443–464.
- Bicchieri, Cristina, and Eugen Dimant.** 2023. "Norm-Nudging: Harnessing Social Expectations for Behavior Change."
- Bicchieri, Cristina, and Hugo Mercier.** 2013. "Self-serving biases and public justifications in trust games." *Synthese*, 190(5): 909–922.
- Bicchieri, Cristina, Eugen Dimant, and Silvia Sonderegger.** 2023. "It's not a lie if you believe the norm does not apply: Conditional norm-following and belief distortion." *Games and Economic Behavior*, 138: 321–354.
- Bifulco, Maurizio, and Roberta Ranieri.** 2017. "Impact of drought on human health." *European Journal of Internal Medicine*, 46: e9–e10.
- Binmore, Ken.** 2005. *Natural Justice*. Oxford University Press.
- Blonski, Matthias, and Giancarlo Spagnolo.** 2015. "Prisoners' other Dilemma." *International Journal of Game Theory*, 44(1): 61–81.
- Blonski, Matthias, Peter Ockenfels, and Giancarlo Spagnolo.** 2011. "Equilibrium Selection in the Repeated Prisoner's Dilemma: Axiomatic Approach and Experimental Evidence." *American Economic Journal: Microeconomics*, 3(3): 164–192.
- Blume, Andreas, and Andreas Ortmann.** 2007. "The effects of costless pre-play communication: Experimental evidence from games with Pareto-ranked equilibria." *Journal of Economic Theory*, 132(1): 274–290.
- Bolton, Gary, Eugen Dimant, and Ulrich Schmidt.** 2021. "Observability and social image: On the robustness and fragility of reciprocity." *Journal of Economic Behavior & Organization*, 191: 946–964.

- Borg, Kim, Jim Curtis, and Jo Lindsay.** 2020. "Social norms and plastic avoidance: Testing the theory of normative social behaviour on an environmental behaviour." *Journal of Consumer Behaviour*, 19(6): 594–607.
- Buckley, Edward, and Rachel Croson.** 2006. "Income and wealth heterogeneity in the voluntary provision of linear public goods." *Journal of Public Economics*, 90(4): 935–955.
- Burke, Mary A., and H. Peyton Young.** 2011. "Chapter 8 - Social Norms." In *Handbook of Social Economics*. Vol. 1, , ed. Jess Benhabib, Alberto Bisin and Matthew O. Jackson, 311–338. North-Holland.
- Burks, Stephen V., and Erin L. Krupka.** 2012. "A Multimethod Approach to Identifying Norms and Normative Expectations Within a Corporate Hierarchy: Evidence from the Financial Services Industry." *Management Science*, 58(1): 203–217.
- Bursztyn, Leonardo, Alessandra L. González, and David Yanagizawa-Drott.** 2020. "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia." *American Economic Review*, 110(10): 2997–3029.
- Bénabou, Roland, and Jean Tirole.** 2006. "Incentives and Prosocial Behavior." *American Economic Review*, 96(5): 1652–1678.
- Bénabou, Roland, Armin Falk, and Jean Tirole.** 2018. "Narratives, Imperatives, and Moral Reasoning."
- Caci, H, and F Baylé.** 2007. "L'échelle d'affectivité positive et d'affectivité négative." *Première traduction en français. Congrès de l'Encéphale, Paris, 25–27.*
- Camera, Gabriele, and Marco Casari.** 2009. "Cooperation among Strangers under the Shadow of the Future." *American Economic Review*, 99(3): 979–1005.
- Camera, Gabriele, Cary Deck, and David Porter.** 2020. "Do economic inequalities affect long-run cooperation and prosperity?" *Experimental Economics*, 23(1): 53–83.
- Camera, Gabriele, Marco Casari, and Maria Bigoni.** 2012. "Cooperative strategies in anonymous economies: An experiment." *Games and Economic Behavior*, 75(2): 570–586.
- Camerer, Colin F.** 2011. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press.
- Campisi, Tiziana, Giovanni Tesoriere, Antonino Canale, Socrates Basbas, Panagiotis Vaitsis, Andreas Niki-foriadis, and Miltiadis Nikolaidis.** 2020. "Comparison of Red-Light Running (RLR) and Yellow-Light Running (YLR) traffic violations in the cities of Enna and Thessaloniki." *Transportation Research Procedia*, 45: 947–954.
- Carlsson, Fredrik, Christina Annette Gravert, Verena Kurz, and Olof Johansson-Stenman.** 2019. "Nudging as an Environmental Policy Instrument."

- Centola, Damon.** 2010. "The Spread of Behavior in an Online Social Network Experiment." *Science*, 329(5996): 1194–1197.
- Centola, Damon.** 2021. *Change: How to Make Big Things Happen*. Hachette UK.
- Chancel, Lucas, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman.** 2022. *World Inequality Report 2022*. Harvard University Press.
- Chan, Kenneth S., Stuart Mestelman, Rob Moir, and R. Andrew Muller.** 1996. "The Voluntary Provision of Public Goods under Varying Income Distributions." *The Canadian Journal of Economics / Revue canadienne d'Economique*, 29(1): 54–69.
- Charness, Gary.** 2000. "Self-Serving Cheap Talk: A Test Of Aumann's Conjecture." *Games and Economic Behavior*, 33(2): 177–194.
- Chaudhary, Anil Kumar, Laura A. Warner, Alexa J. Lamm, Glenn D. Israel, Joy N. Rumble, and Randall A. Cantrell.** 2017. "Using the Theory of Planned Behavior to Encourage Water Conservation among Extensin Clients." *Journal of Agricultural Education*, 58(3): 185–202.
- Chen, Daniel L., Martin Schonger, and Chris Wickens.** 2016. "oTree—An open-source platform for laboratory, online, and field experiments." *Journal of Behavioral and Experimental Finance*, 9: 88–97.
- Cherry, Todd L., Stephan Kroll, and Jason F. Shogren.** 2005. "The impact of endowment heterogeneity and origin on public good contributions: evidence from the lab." *Journal of Economic Behavior & Organization*, 57(3): 357–365.
- Cialdini, Robert B.** 2003. "Crafting Normative Messages to Protect the Environment." *Current Directions in Psychological Science*, 12(4): 105–109.
- Cialdini, Robert B, and Ryan P Jacobson.** 2021. "Influences of social norms on climate change-related behaviors." *Current Opinion in Behavioral Sciences*, 42: 1–8.
- Cialdini, Robert B., Linda J. Demaine, Brad J. Sagarin, Daniel W. Barrett, Kelton Rhoads, and Patricia L. Winter.** 2006. "Managing social norms for persuasive impact." *Social Influence*, 1(1): 3–15.
- Cialdini, Robert B., Raymond R. Reno, and Carl A. Kallgren.** 1990. "A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places." *Journal of Personality and Social Psychology*, 58(6): 1015–1026.
- Clark, William A., and James C. Finley.** 2007. "Determinants of Water Conservation Intention in Blagoevgrad, Bulgaria." *Society & Natural Resources*, 20(7): 613–627.

- Collotta, Mario, Giovanni Pau, Gianfranco Scatà, and Tiziana Campisi.** 2014. "A dynamic traffic light management system based on wireless sensor networks for the reduction of the red-light running phenomenon." *Transport and Telecommunication Journal*, 15(1): 1–11.
- Constantino, Sara M., Gregg Sparkman, Gordon T. Kraft-Todd, Cristina Bicchieri, Damon Centola, Bettina Shell-Duncan, Sonja Vogt, and Elke U. Weber.** 2022. "Scaling Up Change: A Critical Review and Practical Guide to Harnessing Social Norms for Climate Action." *Psychological Science in the Public Interest*, 23(2): 50–97.
- Cooper, Russell, Douglas V. DeJong, Robert Forsythe, and Thomas W. Ross.** 1992. "Communication in Coordination Games\*." *The Quarterly Journal of Economics*, 107(2): 739–771.
- Corno, Lucia, Eliana La Ferrara, and Alessandra Voena.** 2020. "Female Genital Cutting and the Slave Trade."
- d'Adda, Giovanna, Martin Dufwenberg, Francesco Passarelli, and Guido Tabellini.** 2020. "Social norms with private values: Theory and experiments." *Games and Economic Behavior*, 124: 288–304.
- Dal Bó, Pedro, and Guillaume R. Fréchette.** 2011. "The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence." *American Economic Review*, 101(1): 411–429.
- Dal Bó, Pedro, and Guillaume R. Fréchette.** 2018. "On the Determinants of Cooperation in Infinitely Repeated Games: A Survey." *Journal of Economic Literature*, 56(1): 60–114.
- Dal Bó, Pedro, and Guillaume R. Fréchette.** 2019. "Strategy Choice in the Infinitely Repeated Prisoner's Dilemma." *American Economic Review*, 109(11): 3929–3952.
- Deb, Joyee, Julio González-Díaz, and Jérôme Renault.** 2016. "Uniform folk theorems in repeated anonymous random matching games." *Games and Economic Behavior*, 100: 1–23.
- Deb, Joyee, Takuo Sugaya, and Alexander Wolitzky.** 2020. "The Folk Theorem in Repeated Games With Anonymous Random Matching." *Econometrica*, 88(3): 917–964.
- Dimant, Eugen.** 2024. "Hate Trumps Love: The Impact of Political Polarization on Social Preferences." *Management Science*, 70(1): 1–31.
- Dimant, Eugen, and Shaul Shalvi.** 2022. "Meta-nudging honesty: Past, present, and future of the research frontier." *Current Opinion in Psychology*, 47: 101426.
- Dimant, Eugen, and Tobias Gesche.** 2023. "Nudging enforcers: how norm perceptions and motives for lying shape sanctions." *PNAS Nexus*, 2(7): pgad224.
- Dimant, Eugen, Gerben A. van Kleef, and Shaul Shalvi.** 2020. "Requiem for a Nudge: Framing effects in nudging honesty." *Journal of Economic Behavior & Organization*, 172: 247–266.

- Duffy, John, and Jack Ochs.** 2009. "Cooperative behavior and the frequency of social interaction." *Games and Economic Behavior*, 66(2): 785–812.
- Dvorak, Fabian, and Sebastian Fehrler.** 2018. "Negotiating Cooperation Under Uncertainty: Communication in Noisy, Indefinitely Repeated Interactions."
- d'Adda, Giovanna, Donja Darai, Nicola Pavanini, and Roberto A Weber.** 2017. "Do Leaders Affect Ethical Conduct?" *Journal of the European Economic Association*, 15(6): 1177–1213.
- Eckel, Catherine C., and Philip J. Grossman.** 1996. "Altruism in Anonymous Dictator Games." *Games and Economic Behavior*, 16(2): 181–191.
- Ecotièrre, Claire, Sylvain Billiard, Jean-Baptiste André, Pierre Collet, Régis Ferrière, and Sylvie Méléard.** 2023. "Human-environment feedback and the consistency of proenvironmental behavior." *PLOS Computational Biology*, 19(9): e1011429.
- Ellison, Glenn.** 1994. "Cooperation in the Prisoner's Dilemma with Anonymous Random Matching." *The Review of Economic Studies*, 61(3): 567–588.
- Engle-Warnick, Jim, and Robert L. Slonim.** 2006. "Inferring repeated-game strategies from actions: evidence from trust game experiments." *Economic Theory*, 28(3): 603–632.
- Engle-Warnick, Jim, William J. McCausland, and John H. Miller.** 2004. "The Ghost in the Machine: Inferring Machine-Based Strategies from Observed Behavior." Université de Montréal. Département de sciences économiques.
- Eriksson, Kimmo, Pontus Strimling, Michele Gelfand, Junhui Wu, Jered Abernathy, Charity S. Akotia, Alisher Aldashev, Per A. Andersson, Giulia Andrighetto, Adote Anum, Gizem Arıkan, Zeynep Aycan, Fatemeh Bagherian, Davide Barrera, Dana Basnight-Brown, Birzhan Batkeyev, Anabel Belaus, Elizaveta Berezina, Marie Björnstjerna, Sheyla Blumen, Paweł Boski, Fouad Bou Zeineddine, Inna Bovina, Bui Thi Thu Huyen, Juan-Camilo Cardenas, Đorđe Čekrlija, Hoon-Seok Choi, Carlos C. Contreras-Ibáñez, Rui Costa-Lopes, Mícheál de Barra, Piyanjali de Zoysa, Angela Dorrough, Nikolay Dvoryanchikov, Anja Eller, Jan B. Engelmann, Hyun Euh, Xia Fang, Susann Fiedler, Olivia A. Foster-Gimbel, Márta Fülöp, Ragna B. Gardarsdóttir, C. M. Hew D. Gill, Andreas Glöckner, Sylvie Graf, Ani Grigoryan, Vladimir Gritskov, Katarzyna Growiec, Peter Halama, Andree Hartanto, Tim Hophthrow, Martina Hřebíčková, Dzintra Iliško, Hirotaka Imada, Hansika Kapoor, Kerry Kawakami, Narine Khachatryan, Natalia Kharchenko, Ninetta Khoury, Toko Kiyonari, Michal Kohút, Lê Thuỳ Linh, Lisa M. Leslie, Yang Li, Norman P. Li, Zhuo Li, Kadi Liik, Angela T. Maitner, Bernardo Manhique, Harry Manley, Imed Medhioub, Sari Mentser, Linda Mohammed, Pegah Nejat, Orlando Nipassa, Ravit Nussinson, Nneoma G. Onyedire, Ike E. Onyishi, Seniha Özden, Penny Panagiotopoulou, Lorena R.**

- Perez-Floriano, Minna S. Persson, Mpho Pheko, Anna-Maija Pirttilä-Backman, Marianna Pogosyan, Jana Raver, Cecilia Reyna, Ricardo Borges Rodrigues, Sara Romanò, Pedro P. Romero, Inari Sakki, Alvaro San Martin, Sara Sherbaji, Hiroshi Shimizu, Brent Simpson, Erna Szabo, Kosuke Takemura, Hassan Tieffi, Maria Luisa Mendes Teixeira, Napoj Thanomkul, Habib Tiliouine, Giovanni A. Travaglino, Yannis Tsirobas, Richard Wan, Sita Widodo, Rizqy Zein, Qing-peng Zhang, Lina Zirganou-Kazolea, and Paul A. M. Van Lange.** 2021. "Perceptions of the appropriate response to norm violation in 57 societies." *Nature Communications*, 12(1): 1481.
- Eslamian, Saeid, and Faezeh A. Eslamian.** 2017. *Handbook of Drought and Water Scarcity: Environmental Impacts and Analysis of Drought and Water Scarcity*. CRC Press.
- Fallucchi, Francesco, and Daniele Nosenzo.** 2022. "The coordinating power of social norms." *Experimental Economics*, 25(1): 1–25.
- Fehr, Ernst, and Ivo Schurtenberger.** 2018. "Normative foundations of human cooperation." *Nature Human Behaviour*, 2(7): 458–468.
- Fehr, Ernst, and Klaus M. Schmidt.** 1999. "A Theory of Fairness, Competition, and Cooperation\*." *The Quarterly Journal of Economics*, 114(3): 817–868.
- Fehr, Ernst, and Simon Gächter.** 1998. "Reciprocity and economics: The economic implications of Homo Reciprocans<sup>1</sup>This paper is part of a research project on the impact of social norms on wage formation. It is financed by the Swiss National Science Foundation under the project no. 12-43590.95.1." *European Economic Review*, 42(3): 845–859.
- Fehr, Ernst, and Simon Gächter.** 2000a. "Cooperation and Punishment in Public Goods Experiments." *American Economic Review*, 90(4): 980–994.
- Fehr, Ernst, and Simon Gächter.** 2000b. "Fairness and Retaliation: The Economics of Reciprocity." *Journal of Economic Perspectives*, 14(3): 159–181.
- Feldman, David L.** 2013. *Water*. John Wiley & Sons.
- Ferraro, Paul J., and Michael K. Price.** 2013. "Using Nonpecuniary Strategies to Influence Behavior: Evidence from a Large-Scale Field Experiment." *The Review of Economics and Statistics*, 95(1): 64–73.
- Festinger, Leon.** 1962. "Cognitive Dissonance." *Scientific American*, 207(4): 93–106.
- Fischbacher, Urs, Simeon Schudy, and Sabrina Teyssier.** 2014. "Heterogeneous reactions to heterogeneity in returns from public goods." *Social Choice and Welfare*, 43(1): 195–217.



- Foxcroft, David R., Maria Teresa Moreira, Nerissa ML Almeida Santimano, and Lesley A. Smith.** 2015. "Social norms information for alcohol misuse in university and college students." *Cochrane Database of Systematic Reviews*, (12). Publisher: John Wiley & Sons, Ltd.
- Gadgil, Ashok.** 1998. "Drinking Water in Developing Countries." *Annual Review of Energy and the Environment*, 23(1): 253–286.
- Galbiati, Roberto, Emeric Henry, and Nicolas Jacquemet.** 2019. "Learning to cooperate in the shadow of the law."
- Gangadharan, Lata, Nikos Nikiforakis, and Marie Claire Villeval.** 2017. "Normative conflict and the limits of self-governance in heterogeneous populations." *European Economic Review*, 100: 143–156.
- Gangadharan, Lata, Tarun Jain, Pushkar Maitra, and Joseph Vecci.** 2016. "Social identity and governance: The behavioral response to female leaders." *European Economic Review*, 90: 302–325.
- Gardner, Roy, Elinor Ostrom, and James M. Walker.** 1990. "The Nature of Common-Pool Resource Problems." *Rationality and Society*, 2(3): 335–358.
- Gavrilets, Sergey.** 2021. "Coevolution of actions, personal norms and beliefs about others in social dilemmas." *Evolutionary Human Sciences*, 3: e44.
- Geiger, Nathaniel, and Janet K. Swim.** 2016. "Climate of silence: Pluralistic ignorance as a barrier to climate change discussion." *Journal of Environmental Psychology*, 47: 79–90.
- Goldstein, Noah J., Robert B. Cialdini, and Vladas Griskevicius.** 2008. "A Room with a Viewpoint: Using Social Norms to Motivate Environmental Conservation in Hotels." *Journal of Consumer Research*, 35(3): 472–482.
- Goodchild, C. W.** 2003. "Modelling the Impact of Climate Change on Domestic Water Demand." *Water and Environment Journal*, 17(1): 8–12.
- Granovetter, Mark.** 2011. "Economic Action and Social Structure: The Problem of Embeddedness." In *The Sociology of Economic Life*. . 3 ed. Routledge.
- Guilbeault, Douglas, and Damon Centola.** 2021. "Topological measures for identifying and predicting the spread of complex contagions." *Nature Communications*, 12(1): 4430.
- Gulesci, Selim, Sam Jindani, Eliana La Ferrara, David Smerdon, Munshi Sulaiman, and H. Young.** 2021. "A Stepping Stone Approach to Understanding Harmful Norms."
- Gächter, Simon, Daniele Nosenzo, and Martin Sefton.** 2013. "Peer Effects in Pro-Social Behavior: Social Norms or Social Preferences?" *Journal of the European Economic Association*, 11(3): 548–573.

- Gächter, Simon, Leonie Gerhards, and Daniele Nosenzo.** 2017. "The importance of peers for compliance with norms of fair sharing." *European Economic Review*, 97: 72–86.
- Haaland, Ingar, Christopher Roth, and Johannes Wohlfart.** 2023. "Designing Information Provision Experiments." *Journal of Economic Literature*, 61(1): 3–40.
- Halbesleben, Jonathon R. B., Anthony R. Wheeler, and M. Ronald Buckley.** 2005. "Everybody Else is Doing it, So Why Can't We? Pluralistic Ignorance and Business Ethics Education." *Journal of Business Ethics*, 56(4): 385–398.
- Han, Heesup, Li-Tzang (Jane) Hsu, and Chwen Sheu.** 2010. "Application of the Theory of Planned Behavior to green hotel choice: Testing the effect of environmental friendly activities." *Tourism Management*, 31(3): 325–334.
- Harrison, Glenn W., and John A. List.** 2004. "Field Experiments." *Journal of Economic Literature*, 42(4): 1009–1055.
- Harsanyi, John C., and Reinhard Selten.** 1988. "A General Theory of Equilibrium Selection in Games." *MIT Press Books*, 1.
- Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, and Herbert Gintis.** 2004. *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. OUP Oxford.
- Horsevad, Nikolaj, David Mateo, Robert E. Kooij, Alain Barrat, and Roland Bouffanais.** 2022. "Transition from simple to complex contagion in collective decision-making." *Nature Communications*, 13(1): 1442.
- Howe, Charles W., and F. P. Linaweaver Jr.** 1967. "The impact of price on residential water demand and its relation to system design and price structure." *Water Resources Research*, 3(1): 13–32.
- Howley, Peter, and Neel Ocean.** 2022. "Can nudging only get you so far? Testing for nudge combination effects." *European Review of Agricultural Economics*, 49(5): 1086–1112.
- Hubbard, Timothy L.** 2015. "The varieties of momentum-like experience." *Psychological Bulletin*, 141(6): 1081–1119.
- Inclusion Matters: The Foundation for Shared Prosperity.** 2013. *Inclusion Matters: The Foundation for Shared Prosperity. New Frontiers of Social Policy*, The World Bank.
- Intergovernmental Panel On Climate Change.** 2023. *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. . 1 ed., Cambridge University Press.

- Jorgensen, Bradley, Michelle Graymore, and Kevin O'Toole.** 2009. "Household water use behavior: An integrated model." *Journal of Environmental Management*, 91(1): 227–236.
- Kandori, Michihiro.** 1992. "Social Norms and Community Enforcement." *The Review of Economic Studies*, 59(1): 63–80.
- Kimbrough, Erik O., Erin L. Krupka, Rajnish Kumar, Jennifer Murray, Abhijit Ramalingam, Sharon Sánchez-Franco, Olga Lucía Sarmiento, Frank Kee, and Ruth Hunter.** 2022. "On the Stability of Norms and Norm-Following Propensity: A Cross-Cultural Panel Study with Adolescents."
- Kitamori, Kumi, Ton Manders, Rob Dellink, and A. A. Tabeau.** 2012. "OECD environmental outlook to 2050 : the consequences of inaction." OECD.
- Klemeš, Jiří Jaromír.** 2012. "Industrial water recycle/reuse." *Current Opinion in Chemical Engineering*, 1(3): 238–245.
- Koch, Christian, Nikos Nikiforakis, and Charles N. Noussair.** 2021. "Covenants before the swords: The limits to efficient cooperation in heterogeneous groups." *Journal of Economic Behavior & Organization*, 188: 307–321.
- Kollmuss, Anja, and Julian Agyeman.** 2002. "Mind the Gap: Why do people act environmentally and what are the barriers to pro-environmental behavior?" *Environmental Education Research*, 8(3): 239–260.
- Konapala, Goutam, Ashok K. Mishra, Yoshihide Wada, and Michael E. Mann.** 2020. "Climate change will affect global water availability through compounding changes in seasonal precipitation and evaporation." *Nature Communications*, 11(1): 3044.
- Krupka, Erin L., and Roberto A. Weber.** 2013. "Identifying Social Norms Using Coordination Games: Why Does Dictator Game Sharing Vary?" *Journal of the European Economic Association*, 11(3): 495–524.
- Krupka, Erin L., Stephen Leider, and Ming Jiang.** 2017. "A Meeting of the Minds: Informal Agreements and Social Norms." *Management Science*, 63(6): 1708–1729.
- Lam, San-Pui.** 2006. "Predicting Intention to Save Water: Theory of Planned Behavior, Response Efficacy, Vulnerability, and Perceived Efficiency of Alternative Solutions1." *Journal of Applied Social Psychology*, 36(11): 2803–2824.
- Landon, Adam C., Richard T. Woodward, Gerard T. Kyle, and Ronald A. Kaiser.** 2018. "Evaluating the efficacy of an information-based residential outdoor water conservation program." *Journal of Cleaner Production*, 195: 56–65.
- Lane, Tom, Daniele Nosenzo, and Silvia Sonderegger.** 2023. "Law and Norms: Empirical Evidence." *American Economic Review*, 113(5): 1255–1293.

- Loewenstein, George.** 2000. "Emotions in Economic Theory and Economic Behavior." *American Economic Review*, 90(2): 426–432.
- Markman, Keith D., and Corey L. Guenther.** 2007. "Psychological Momentum: Intuitive Physics and Naive Beliefs." *Personality and Social Psychology Bulletin*, 33(6): 800–812.
- Masclet, David, Charles Noussair, Steven Tucker, and Marie-Claire Villeval.** 2003. "Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism." *American Economic Review*, 93(1): 366–380.
- Mortensen, Chad R., Rebecca Neel, Robert B. Cialdini, Christine M. Jaeger, Ryan P. Jacobson, and Megan M. Ringel.** 2019. "Trending Norms: A Lever for Encouraging Behaviors Performed by the Minority." *Social Psychological and Personality Science*, 10(2): 201–210.
- Nauges, Céline, and Alban Thomas.** 2003. "Long-run Study of Residential Water Consumption." *Environmental and Resource Economics*, 26(1): 25–43.
- Nikiforakis, Nikos.** 2008. "Punishment and counter-punishment in public good games: Can we really govern ourselves?" *Journal of Public Economics*, 92(1): 91–112.
- Nikiforakis, Nikos, Charles N. Noussair, and Tom Wilkening.** 2012. "Normative conflict and feuds: The limits of self-enforcement." *Journal of Public Economics*, 96(9): 797–807.
- Nikiforakis, Nikos, Jörg Oechssler, and Anwar Shah.** 2014. "Hierarchy, coercion, and exploitation: An experimental analysis." *Journal of Economic Behavior & Organization*, 97: 155–168.
- Nyborg, Karine, John M. Anderies, Astrid Dannenberg, Therese Lindahl, Caroline Schill, Maja Schlüter, W. Neil Adger, Kenneth J. Arrow, Scott Barrett, Stephen Carpenter, F. Stuart Chapin, Anne-Sophie Crépin, Gretchen Daily, Paul Ehrlich, Carl Folke, Wander Jager, Nils Kautsky, Simon A. Levin, Ole Jacob Madsen, Stephen Polasky, Marten Scheffer, Brian Walker, Elke U. Weber, James Wilen, Anastasios Xepapadeas, and Aart de Zeeuw.** 2016. "Social norms as solutions." *Science*, 354(6308): 42–43.
- Oechssler, Jörg, Andreas Roider, and Patrick W. Schmitz.** 2009. "Cognitive abilities and behavioral biases." *Journal of Economic Behavior & Organization*, 72(1): 147–152.
- Ostrom, Elinor.** 2009. "A General Framework for Analyzing Sustainability of Social-Ecological Systems." *Science*, 325(5939): 419–422.
- Ostry, Mr Jonathan David, Mr Andrew Berg, and Mr Charalambos G. Tsangarides.** 2014. *Redistribution, Inequality, and Growth*. International Monetary Fund.
- Otaki, Yurina, Kazuhiro Ueda, and Osamu Sakura.** 2017. "Effects of feedback about community water consumption on residential water conservation." *Journal of Cleaner Production*, 143: 719–730.

- Otto, Betsy, and Leah Schleifer.** 2020. "Domestic Water Use Grew 600% Over the Past 50 Years."
- Owens, Susan, and Louise Driffill.** 2008. "How to change attitudes and behaviours in the context of energy." *Energy Policy*, 36(12): 4412–4418.
- Perkins, H Wesley.** 2002. "Social norms and the prevention of alcohol misuse in collegiate contexts." *Journal of Studies on Alcohol, Supplement*, (s14): 164–172.
- Persson, Linn, Bethanie M. Carney Almroth, Christopher D. Collins, Sarah Cornell, Cynthia A. de Wit, Miriam L. Diamond, Peter Fantke, Martin Hassellöv, Matthew MacLeod, Morten W. Ryberg, Peter Søgaard Jørgensen, Patricia Villarrubia-Gómez, Zhanyun Wang, and Michael Zwicky Hauschild.** 2022. "Outside the Safe Operating Space of the Planetary Boundary for Novel Entities." *Environmental Science & Technology*, 56(3): 1510–1521.
- Piketty, Thomas.** 2013. *Le Capital au XXI<sup>e</sup> siècle*. Média Diffusion.
- Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman.** 2018. "Distributional National Accounts: Methods and Estimates for the United States\*." *The Quarterly Journal of Economics*, 133(2): 553–609.
- Pimentel, David, James Houser, Erika Preiss, Omar White, Hope Fang, Leslie Mesnick, Troy Barsky, Stephanie Tariche, Jerrod Schreck, and Sharon Alpert.** 1997. "Water Resources: Agriculture, the Environment, and Society." *BioScience*, 47(2): 97–106.
- Putnam, Robert D.** 2000. *Bowling Alone: The Collapse and Revival of American Community*. Simon and Schuster.
- Pörtner, Hans-Otto, Debra Roberts, Melinda Tignor, Elvira Poloczanska, Katja Mintenbeck, Andrés Alegría, Marlies Craig, Stefanie Langsdorf, Sina Löschke, Vincent Möller, Andrew Okem, Bard Rama Phd, Daniel Belling, Wolfgang Dieck, Sandra Götze, Tijama Kersher, Philisiwe Mangele, Bastian Maus, Anka Mühle, and Nora Weyer.** 2022. *Climate Change 2022: Impacts, Adaptation and Vulnerability Working Group II Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*.
- Reuben, Ernesto, and Arno Riedl.** 2013. "Enforcement of contribution norms in public good games with heterogeneous populations." *Games and Economic Behavior*, 77(1): 122–137.
- Rockström, Johan, Will Steffen, Kevin Noone, Åsa Persson, F. Stuart Chapin, Eric F. Lambin, Timothy M. Lenton, Marten Scheffer, Carl Folke, Hans Joachim Schellnhuber, Björn Nykvist, Cynthia A. de Wit, Terry Hughes, Sander van der Leeuw, Henning Rodhe, Sverker Sörlin, Peter K. Snyder, Robert Costanza, Uno Svedin, Malin Falkenmark, Louise Karlberg, Robert W. Corell, Victoria J. Fabry, James Hansen, Brian Walker, Diana Liverman, Katherine Richardson, Paul Crutzen, and Jonathan A. Foley.** 2009. "A safe operating space for humanity." *Nature*, 461(7263): 472–475.

- Romero, Julian, and Yaroslav Rosokha.** 2018. "Constructing strategies in the indefinitely repeated prisoner's dilemma game." *European Economic Review*, 104: 185–219.
- Roth, Alvin E, and J. Keith Murnighan.** 1978. "Equilibrium behavior and repeated play of the prisoner's dilemma." *Journal of Mathematical Psychology*, 17(2): 189–198.
- Scheierling, Susanne M., John B. Loomis, and Robert A. Young.** 2006. "Irrigation water demand: A meta-analysis of price elasticities." *Water Resources Research*, 42(1).
- Schill, Caroline, John M. Anderies, Therese Lindahl, Carl Folke, Stephen Polasky, Juan Camilo Cárdenas, Anne-Sophie Crépin, Marco A. Janssen, Jon Norberg, and Maja Schlüter.** 2019. "A more dynamic understanding of human behaviour for the Anthropocene." *Nature Sustainability*, 2(12): 1075–1082.
- Schram, Arthur, and Aljaž Ule.** 2019. *Handbook of Research Methods and Applications in Experimental Economics*. Edward Elgar Publishing.
- Schultz, P. Wesley, Jessica M. Nolan, Robert B. Cialdini, Noah J. Goldstein, and Vladas Griskevicius.** 2007. "The Constructive, Destructive, and Reconstructive Power of Social Norms." *Psychological Science*, 18(5): 429–434.
- Schwartz, Steven T., Richard A. Young, and Kristina Zvinakis.** 2000. "Reputation Without Repeated Interaction: A Role for Public Disclosures." *Review of Accounting Studies*, 5(4): 351–375.
- Sheposh, J.P., and P.S. Gallo.** 1973. "Asymmetry of Payoff Structure and Cooperative Behavior in the Prisoner's Dilemma Game." *Journal of Conflict Resolution*, 17(2): 321–333.
- Shukla, Priyadarshi R, Jim Skea, and Raphael Slade.** 2022. "Working Group III Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change."
- Singh, Anita, and Madhoolika Agrawal.** 2008. "Acid rain and its ecological consequences." *Journal of Environmental Biology*, 29(1): 15–24.
- Smith, Adam B.** 2022. "U.S. Billion-dollar Weather and Climate Disasters, 1980 - present."
- Sparkman, Gregg, and Gregory M. Walton.** 2017. "Dynamic Norms Promote Sustainable Behavior, Even if It Is Counternormative." *Psychological Science*, 28(11): 1663–1674.
- Spencer, Theo, and Peter Altman.** 2010. "Climate Change, Water, and Risk: Current Water Demands Are Not Sustainable." *Climate Change*.
- Steffen, Will, Katherine Richardson, Johan Rockström, Sarah E. Cornell, Ingo Fetzer, Elena M. Bennett, Reinette Biggs, Stephen R. Carpenter, Wim de Vries, Cynthia A. de Wit, Carl Folke, Dieter Gerten,**

- Jens Heinke, Georgina M. Mace, Linn M. Persson, Veerabhadran Ramanathan, Belinda Reyers, and Sverker Sörlin.** 2015. "Planetary boundaries: Guiding human development on a changing planet." *Science*, 347(6223): 1259855.
- Stiglitz, Joseph E.** 2015. "The Price of Inequality: How Today's Divided Society Endangers Our Future." 379–399.
- Suchon, Rémi, and Vincent Theroude.** 2022. "Inequality and cooperation: meta-analytical evidence from Public Good Experiments."
- S. Mauter, Meagan, and Peter S. Fiske.** 2020. "Desalination for a circular water economy." *Energy & Environmental Science*, 13(10): 3180–3184.
- Tajfel, Henri.** 1970. "Experiments in Intergroup Discrimination." *Scientific American*, 223(5): 96–103.
- Tajfel, Henri, M. G. Billig, R. P. Bundy, and Claude Flament.** 1971. "Social categorization and intergroup behaviour." *European Journal of Social Psychology*, 1(2): 149–178.
- Thaler, Richard H., and Cass R. Sunstein.** 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. Yale University Press.
- Thomson, W I, and Alfred A Knopf.** 1928. "THE CHILD IN AMERICA: BEHAVIOR PROBLEMS AND PROGRAM."
- Thöni, Christian, Simon Gächter.** 2012. "Micromotives, Microstructure, and Macrobehavior: The Case of Voluntary Cooperation." In *Micro-Macro Links and Microfoundations in Sociology*. Routledge.
- Tol, Richard S. J.** 2018. "The Economic Impacts of Climate Change." *Review of Environmental Economics and Policy*, 12(1): 4–25.
- UNEP, United Nations Environment Programme.** 2023. "Measuring Progress: Water-related ecosystems and the SDGs." *UNEP Publishing Services Section, Nairobi*.
- United Nations Development Programme.** 2013. "Umanity divided: Confronting inequality in developing countries."
- Villaveces, Andrés, Luis Alfonso Nieto, Delia Ortega, José Fernando Ríos, John Jairo Medina, María Isabel Gutiérrez, and Daniel Rodríguez.** 2012. "Pedestrians' perceptions of walkability and safety in relation to the built environment in Cali, Colombia, 2009–10." *Injury Prevention*, 18(5): 291–297.
- Visser, M., and J. Burns.** 2015. "Inequality, social sanctions and cooperation within South African fishing communities." *Journal of Economic Behavior & Organization*, 118: 95–109.

- Wada, Y., L. P. H. van Beek, and M. F. P. Bierkens.** 2011. "Modelling global water stress of the recent past: on the relative importance of trends in water demand and climate variability." *Hydrology and Earth System Sciences*, 15(12): 3785–3808.
- Wada, Yoshihide, Dominik Wisser, Stephanie Eisner, Martina Flörke, Dieter Gerten, Ingjerd Haddeland, Naota Hanasaki, Yoshimitsu Masaki, Felix T. Portmann, Tobias Stacke, Zachary Tessler, and Jacob Schewe.** 2013. "Multimodel projections and uncertainties of irrigation water demand under climate change." *Geophysical Research Letters*, 40(17): 4626–4632.
- Watson, David, Lee Anna Clark, and Greg Carey.** 1988. "Positive and negative affectivity and their relation to anxiety and depressive disorders." *Journal of Abnormal Psychology*, 97(3): 346–353.
- Weitz, Joshua S., Ceyhun Eksin, Keith Paarporn, Sam P. Brown, and William C. Ratcliff.** 2016. "An oscillating tragedy of the commons in replicator dynamics with game-environment feedback." *Proceedings of the National Academy of Sciences*, 113(47): E7518–E7525.
- World Bank.** 2005. *World Development Report 2006: Equity and Development*. The World Bank.
- World Health Organization.** 2015. *Progress on Sanitation and Drinking Water : 2015 Update and MDG Assessment*. World Health Organization.
- Xiao, Erte, and Cristina Bicchieri.** 2010. "When equality trumps reciprocity." *Journal of Economic Psychology*, 31(3): 456–470.
- Young, H. Peyton.** 2015. "The Evolution of Social Norms." *Annual Review of Economics*, 7(1): 359–387.
- Zelmer, Jennifer.** 2003. "Linear Public Goods Experiments: A Meta-Analysis." *Experimental Economics*, 6(3): 299–310.
- Zhang, David D., Peter Brecke, Harry F. Lee, Yuan-Qing He, and Jane Zhang.** 2007. "Global climate change, war, and population decline in recent human history." *Proceedings of the National Academy of Sciences*, 104(49): 19214–19219.





**Titre:** Application par la Communauté des Comportements Pro-Sociaux et Pro-Environnementaux face aux Inégalités : Une Voie vers l'Atténuation du Changement Climatique

**Mots clés:** Economie comportementale - Economie expérimentale - Changement climatique - Dilemme social - Normes - Inégalités

**Résumé:** Le changement climatique est le plus grand défi de nos jours. Il nuit à l'environnement, damages la santé et compromet la croissance et la stabilité mondiales. Pour l'atténuer, nous étudions comment les normes sociales peuvent renforcer les comportements coopératifs pour le bien commun. Pour ce faire, nous étudions comment les croyances sur les normes sociales varient en fonction du contexte. Comment les normes influencent le comportement. Comment les normes peuvent être utilisées pour promouvoir de bons comportements.

Nous montrons que les croyances sur les normes sociales de coopération influencent les comportements et les stratégies. La présence d'inégalités réduit ces croyances. De plus, en absence de communication, ces croyances diminuent au fil du temps. Lorsque qu'elles sont faibles, les individus privilégient les comportements égoïstes. Néanmoins, les politiques d'information permettent de renforcer ces croyances et ainsi les comportements pro-sociaux et pro-environnementaux.

**Title:** Community Enforcement of Pro-Social and Pro-Environmental Behavior in the Face of Inequalities: A Path to Climate Change Mitigation

**Keywords:** Behavioral economics - Experimental economics - Climate change - Social dilemma - Norms - Inequalities

**Abstract:** Climate change is the most challenging issues we are facing today. It damages the environment, harms humankind, and undermines global growth and stability. To contribute to climate change mitigation, we investigate how social norms can reinforce cooperative behavior to benefit the common good. For this purpose, we investigate how beliefs about social norms vary according to context, how norms influence behavior, and how social norms can promote pro-environmental behaviors.

We show that social norms' beliefs influence both direct and strategic decision-making. However, without communication, these beliefs decline over time. In addition, the mere presence of inequality decreases social norms' beliefs. When beliefs about the social norms of cooperation are low, individuals tend to privilege selfish behavior. Nevertheless, information-based policies can be implemented to change beliefs and hence reinforce pro-social and pro-environmental behaviors.