IMT School for Advanced Studies Lucca
Lucca, Italy

Essays on Public Good Game Experiments

PhD Program in Systems Science
Track in Economics, Networks and Business Analytics
XXXV Cycle

By

Veronica Pizziol

2023

The dissertation of Veronica Pizziol is approved.

PhD Program Coordinator: Prof. Dr. Ennio Bilancini, IMT School for Advanced Studies Lucca

Advisor: Prof. Dr. Ennio Bilancini, IMT School for Advanced Studies Lucca

Co-Advisor: Dr. Chiara Nardi, IMT School for Advanced Studies Lucca

The dissertation of Veronica Pizziol has been reviewed by:

Prof. Dr. Luca Corazzini, Department of Economics, University of Milan-Bicocca

Prof. Dr. Alessandro Tavoni, Department of Economics, University of Bologna

IMT School for Advanced Studies Lucca
2023

To Maria and Valentino.

Contents

List of Figures **x**

List of Tables **xiii**

Vita and Publications **xvii**

Acknowledgements **xxii**

Abstract **xxiv**

1 Introduction **1**

2 The role of personal and social norms in the public good game **9**
2.1 Introduction 9
2.2 Experimental design and procedures 11
2.3 Results 12
2.4 Checking for self-confirmation bias 15
2.5 Discussion and conclusions 17

3 Efficiency-driven effects in the multilevel public goods game **19**
3.1 Introduction 19
3.2 Methods 21
3.2.1 Review of related studies 23
3.2.2 Experimental design 26
3.2.3 Hypotheses 28
3.2.4 Implementation 29
3.3 Results 31

3.3.1 Contribution to the local and global public goods . 33
3.3.2 Total contribution 36
3.4 Discussion and conclusions 39

4 A cross-country experiment on the multilevel public goods game 43
4.1 Introduction 43
4.2 Methods 47
4.2.1 Experimental design 47
4.2.2 Groups formation 50
4.2.3 The post-experimental questionnaire 52
4.2.4 Procedures 54
4.3 Sample characteristics 55
4.3.1 Demographics 55
4.3.2 The post-experimental questionnaire 58
4.4 Results 63
4.4.1 Contributing behavior across countries 63
4.4.2 Efficiency-related effects 64
4.4.3 Additional analysis on migrants 69
4.5 Discussion and conclusions 72

5 Cooperation under the threat of severe adverse events 75
5.1 Introduction 75
5.2 Related literature 78
5.3 Methods 82
5.3.1 Experimental design 82
5.3.2 Procedures 83
5.3.3 Participants 85
5.4 The role of uncertainty 87
5.5 Results 87
5.5.1 Descriptive and non-parametric analyses 88
5.5.2 Treatment effects on contributions 91
5.5.3 Beliefs 92
5.6 Discussion and conclusions 95

6 Conclusion 98

A Supplementary materials for Chapter 2 **102**
A.1 Additional tables 102
A.2 Experimental instructions (Experiment 1) 109
A.3 Experimental instructions (Experiment 2) 114

B Supplementary materials for Chapter 3 **119**
B.1 Additional tables 119
B.2 Experimental instructions 126

C Supplementary materials for Chapter 4 **132**
C.1 Additional tables 133
C.2 List of variables from the post-experimental questionnaire 139
C.3 Experimental instructions 141

D Supplementary materials for Chapter 5 **148**
D.1 Additional tables 148
D.2 Simulations on best-response analysis 150
D.3 Experimental instructions 157

List of Figures

1 Mean (and standard deviation) of contribution, personal norm, empirical expectation, and normative expectation. Confidence intervals at the 95% level. 13

2 Mean (and standard deviation) of personal norm, empirical expectation, and normative expectation in Experiment 1 vis-à-vis Experiment 2. Confidence intervals at the 95% level. 16

3 Configuration of the MLPGG in this study. 22

4 Average contribution (and standard deviation) to the local public good by treatment. Standard deviations in parentheses. Confidence intervals at the 95% level. 32

5 Average contribution (and standard deviation) to the global public good by treatment. Standard deviations in parentheses. Confidence intervals at the 95% level. 32

6 Average total contribution (and standard deviation) by treatment. Standard deviations in parentheses. Confidence intervals at the 95% level. 33

7 Difference between average contributions to the local and global goods per treatment. Confidence intervals at the 95% level. 36

8 Configuration of the MLPGG in this study. 48

9 Mean answers to Feel questions by country. Confidence intervals at the 95% level. 59
10 Mean answers to COVID-19 questions by country. Confidence intervals at the 95% level. 61
11 Mean answers to war questions by country. The Military Spending question is standardized to vary between 0 and 5 for graphical comparability. Confidence intervals at the 95% level. 62
12 Mean contributions to the Country Budget by country and treatment. Confidence intervals at the 95% level. 65
13 Mean contributions to the EU Budget by country and treatment. Confidence intervals at the 95% level. 65
14 Mean Total contribution by country and treatment. Confidence intervals at the 95% level. 66

15 Mean contributions by treatment. Standard deviations in parentheses. Confidence intervals at the 95% level. 88
16 Distribution of contribution choices in the Control treatment (C). 89
17 Distribution of contribution choices in the Independent Risk treatment (IR). 90
18 Distribution of contribution choices in the Positively Correlated Risk treatment (PCR). 90
19 Distribution of contribution choices in the Negatively Correlated Risk treatment (NCR). 91
20 Mean beliefs by treatment. Standard deviations in parentheses. Confidence intervals at the 95% level. 94

D.1 Best-response analysis, simulation #1 results. 152
D.2 Best-response analysis, simulation #2 results. 152
D.3 Best-response analysis, simulation #3 results. 153
D.4 Best-response analysis, simulation #4 results. 153
D.5 Best-response analysis, simulation #5 results. 154
D.6 Best-response analysis, simulation #6 results. 154
D.7 Best-response analysis, simulation #7 results. 155

D.8 Best-response analysis, simulation #8 results. 155
D.9 Best-response analysis, simulation #9 results. 156
D.10 Best-response analysis, simulation #10 results. 156

List of Tables

1 Mean (and standard deviation) of participants' characteristics. 11
2 Tobit regressions examining the contribution choices in the PGG. 16

3 Summary of experimental designs in the MLPGG literature. 25
4 Summary of the treatments parameters. 27
5 Sample sizes and participants' average characteristics by treatment. 30
6 OLS regressions examining the local and global contribution choices in the MLPGG. 34
7 OLS regressions examining total contributions in the MLPGG. 37
8 Relative Efficiency Index (REI) per treatment. 39
9 Summary of the main results in the MLPGG literature (including our study). 40

10 Summary of the treatments parameters. 49
11 Means (and standard deviations) of participants' characteristics per treatment. 56
12 Means (and standard deviations) of participants' characteristics per country. 57
13 OLS regressions examining contribution choices to the Country Budget (Columns 1, 4), the EU Budget (Columns 2, 5), and the sum of the two (Columns 3, 6) in the MLPGG. . . . 67

14 OLS regressions examining the contribution decisions of the subsample of migrants to the Country Budget, to the EU Budget, and the sum of contributions to both budgets. 71

15 Summary of experimental designs and results employed to manipulate the role of risk correlations. 81
16 Means (and standard deviations) of participants' characteristics and preferences. 86
17 Tobit regressions examining the contribution choices in the PGG. 93

A.1 Participants' average characteristics by experiment. 102
A.2 Tobit regressions examining the contribution choices in the PGG with controls. 103
A.3 Tobit regressions examining different types of norms with the pooled data from the two experiments. 104
A.4 Tobit regressions examining the contribution choices in the PGG with controls and interaction terms with Male. 105
A.5 Tobit regressions examining the contribution choices in the PGG with controls and interaction terms with Education. . 106
A.6 Tobit regressions examining the contribution choices in the PGG with controls and interaction terms with Student Status. 107
A.7 Tobit regressions examining the contribution choices in the PGG with controls and interaction terms with Socioeconomic Status. 108

B.1 Number of subjects, mean (and standard deviation) of the contributions to local and global public goods, and total contributions by treatment. 119
B.2 P-value results of the Wilcoxon rank-sum test over the full sample for each pairwise comparison. 120
B.3 OLS regressions examining the contributions in the MLPGG with the inclusion of controls. 121

B.4 Tobit models examining the contribution choices in the MLPGG. 122
B.5 Seemingly Unrelated Regressions (SUR) examining contribution choices in the MLPGG. 123
B.6 OLS regressions examining contribution choices in the MLPGG with treatment dummy variables. 124
B.7 OLS regressions examining contribution choices in the MLPGG with a restricted sample of people who gave at least two correct answers out of the three control questions. 125
B.8 OLS regressions examining contribution choices in the MLPGG with a restricted sample of people who gave three correct answers out of three control questions. 125

C.1 Mean (and standard deviation) of contribution choices by treatment. 133
C.2 Mean (and standard deviation) of contribution choices by treatment and country. 133
C.3 Mean (and standard deviation) of participants' characteristics by country and treatment. 134
C.4 Mean (and standard deviation) of the answers to the post-experimental questionnaire by country. 135
C.5 Kruskal–Wallis tests on the answers to the post-experimental questionnaire by country. 135
C.6 Wilcoxon rank-sum tests results of pairwise comparisons of answers to the post-experimental questionnaire between countries. 136
C.7 Post-estimation tests on the equality of coefficients. 137
C.8 OLS regressions examining the contribution decisions of the subsample of migrants to the Country Budget, to the EU Budget, and the sum of contributions to both budgets. 138

D.1 Wilcoxon rank-sum tests results over the contribution variable. 148
D.2 Hurdle model examining contribution choices in the PGG. 149

D.3 Summary of simulations' results from best-response analyses. 151

Vita

January 23, 1993 Born, Rome (RM) Italy

Education

2019- Ph.D. Student in Economics, Networks and Business Analytics
IMT School for Advanced Studies Lucca
(Lucca) Italy

2022 Visiting Ph.D. Student
University of Amsterdam, CREED
(Amsterdam) The Netherlands

2018 M.Sc. Degree in Economics
Final mark: 110/110 magna cum laude
Università degli Studi di Roma Tor Vergata
(Rome) Italy

2014 B.Sc. Degree in Economics
Final mark: 110/110 magna cum laude
Università degli Studi di Roma Tor Vergata
(Rome) Italy

2013 Erasmus Bachelor Student
University of Malta
(Msida) Malta

2011	High School Diploma Final mark: 100/100 magna cum laude Liceo Scientifico Statale Vito Volterra (Rome) Italy

Work experience

2023-	Post-Doctoral Researcher Department of Economics, University of Bologna (Bologna) Italy
2018-2019	Solutions Assistant Everis Italia S.p.A. (Rome) Italy
2015-2018	Assistant Librarian Università degli Studi di Roma Tor Vergata (Rome) Italy
2017	Trainee European Commission, DG for Competition Chief Economist Team (Brussels) Belgium
2015-2017	Office Assistant Università degli Studi di Roma Tor Vergata (Rome) Italy

Publications

Publications

1. M. Catola, S. D'Alessandro, P. Guarnieri, V. Pizziol "Personal norms in the online public good game," in *Economics Letters*, vol. 207, pp. 110024, 2021.
2. E. Bilancini, L. Boncinelli, R. Di Paolo, D. Menicagli, V. Pizziol, E. Ricciardi, F. Serti, "Prosocial behavior in emergencies: Evidence from blood donors recruitment and retention during the COVID-19 pandemic," in *Social Science & Medicine*, vol. 314, pp. 115438, 2022.
3. M. Bisanti, R. Di Paolo, V. Pizziol, S. Accardi, F. Maggi, G. Paladino, E. Ricciardi, E. Bilancini, "Digital Escape Games in Educational Programs for Financial Literacy," in *European Conference on Games Based Learning*, vol. 16, pp. 666-674, 2022.
4. V. Pizziol, X. Demaj, R. Di Paolo, V. Capraro, "Political ideology and generosity around the globe," in *Proceedings of the National Academy of Sciences*, vol. 120, pp. e2219676120, 2023.
5. M. Catola, S. D'Alessandro, P. Guarnieri, V. Pizziol, "Multilevel public goods game: Levelling up, substitution and crowding-in effects," in *Journal of Economic Psychology*, vol. 97, pp. 102626, 2023.
6. R. Di Paolo & V. Pizziol, "Gamification and Sustainable Water Use: The Case of the BLUTUBE Educational Program," in *Simulation & Gaming*, vol. 0, pp. 10468781231181652, 2023.

Working papers

1. E. Bilancini, L. Boncinelli, C. Nardi, V. Pizziol "Cooperation is unaffected by the threat of severe adverse events in Public Goods Games," OSF Preprint, 2023. Revise & Resubmit at *Journal of Behavioral and Experimental Economics*.
2. M. Catola, S. D'Alessandro, P. Guarnieri, V. Pizziol "Norms and efficiency in a multi-group society: an online experiment," Dipartimento di Economia e Management (DEM), University of Pisa, Pisa, Italy, 2021. Revise & Resubmit at *Journal of Public Economic Theory*.
3. M. Catola, P. Guarnieri, V. Pizziol, C. Rapallini "Measuring the attitude towards a European public budget: A cross-country experiment." Dipartimento di Economia e Management (DEM), University of Pisa, Pisa, Italy, 2023.

4. V. Capraro, R. Di Paolo, V. Pizziol "Predict-AI-bility of how humans balance self-interest with the interest of others." ArXiv Preprint, 2023.
5. N. Toccafondi, V. Pizziol, E. Santin, T. Antognozzi, R. Di Paolo, E. Bilancini "Delocalized events management: The experience of hobby stores hosting the Campfire event during LCG20."

Presentations

1. V. Pizziol, "Cooperation under Threats of Catastrophic Events," at Sixth Meeting of the Behavioral and Experimental Economics Network, *University of Bologna*, Bologna, Italy, 2021.
2. V. Pizziol, "Prosocial behavior in emergencies: Evidence from blood donors recruitment and retention during the COVID-19 pandemic," at AXES Internal Seminar, *IMT School for Advanced Studies Lucca*, Lucca, Italy, 2022.
3. V. Pizziol, "Cooperation is unaffected by the threat of severe adverse events in Public Goods Games," at European Economic Science Association Conference 2022, *University of Bologna*, Bologna, Italy, 2022.
4. V. Pizziol, "Contributing to the European public budget? An experimental comparison across countries," at Eight Meeting of the Behavioral and Experimental Economics Network, *Luiss Guido Carli & Sapienza Università di Roma*, Rome, Italy, 2022.
5. V. Pizziol, "Digital Escape Games in Educational Programs for Financial Literacy," at 16th European Conference on Games Based Learning, *Lusófona University*, Lisbon, Portugal, 2022.
6. V. Pizziol, "Measuring the attitude towards a European public budget: A cross-country experiment," at *Masaryk University*, Brno, Czech Republic, 2022 (Invited Seminar).
7. V. Pizziol, "Measuring the attitude towards a European public budget: A cross-country experiment," at *WU Vienna Institute for Markets and Strategy*, Vienna, Austria, 2022 (Invited Seminar).
8. V. Pizziol, "Measuring the attitude towards a European public budget: A cross-country experiment," at 2023 Workshop Southern Europe Experimental Team's, *Fundación Universidad-Empresa de Valencia*, Valencia, Spain, 2023.
9. V. Pizziol, "Cooperation is unaffected by the threat of severe adverse events in Public Goods Games," at Young Economists' Meeting (YEM) 2023, *Masaryk University*, Brno, Czech Republic, 2023.

10. V. Pizziol, "Measuring the attitude towards a European public budget: A cross-country experiment," at Economic Science Association World meeting 2023, *University of Lyon*, Lyon, France, 2023.
11. V. Pizziol, "Cooperation is unaffected by the threat of severe adverse events in Public Goods Games," at 64th Annual Conference of the Italian Economic Association, *Gran Sasso Science Institute*, L'Aquila, Italy, 2023.

Acknowledgements

I am deeply grateful to my supervisor, *Ennio Bilancini*, and co-advisor, *Chiara Nardi*, for their valuable guidance and support throughout my Ph.D. journey: I consider myself lucky and proud to have been your Ph.D. student during these years.

I extend my gratitude to the external referees, *Luca Corazzini* and *Alessandro Tavoni*, for their insightful comments: they undoubtedly helped improve this thesis.

Also, I would like to express my sincere appreciation to other researchers who have contributed to my professional and personal development with their comments, opportunities, and conversations: *Andrea Albertazzi, Leonardo Boncinelli, Nicola Campigotto, Valerio Capraro, Marco Catola, Tatiana Celadin, Gustavo Cevolani, Simone D'Alessandro, Sibilla Di Guida, Roberto Di Paolo, Pietro Guarnieri, Matteo M. Marini, Jesus Erubiel Ordaz Cuevas, Chiara Rapallini, Francesco Serti, Federico Vaccari*.

I am also thankful to *Theo Offerman* and *Ivan Soraperra* for the opportunity they gave me to pursue a visiting period at CREED, University of Amsterdam, and for their supervision.

I am grateful to all the members of the "BEE group" (special mention to *Lina Rinaldi, Bianca Sanesi, Niccolò Toccafondi*), the "Reading Group on Cognition, Decisions, and Economic Behavior" and the "GAME Science Research Center", and to the amazing people I met at the conferences I attended, as well as all the IMT people that I met during these years (especially my roommate *Damiana Bergamo*) and with whom I shared so much. I also want to thank my friends for being there for me: *Deianira Bellitto, Fabrizio Iuliano, Francesco Lucantoni, Maria Chiara Nicolosi, Francesco Olivanti, Martina*

Padovano, Federica Pennarola, Annalivia Polselli, and all the friends from the *Hunters of the odd socks*.

Special thanks go to *Marzio* for your love and patience during this journey; to my elder brothers *Marco* and *Sergio* and my twin *Roberto* for the inspiration you constantly give me: you are truly the best brothers I could ask for; to my aunt *Melina* and grandma *Pierina*, thank you for your unconditional affection, sustaining me in every step I took on the way; and to *Gabriel*, for the light you brought into our lives. Each of you has been an indispensable part of my achievements.

Last but not least, I am immensely grateful to my parents, *Maria* and *Valentino*, to whom I dedicate this thesis.

Abstract

Cooperation, i.e., paying a cost to benefit others, is a recurring phenomenon in human interactions and a fundamental principle of our societies. Hence, it is of great interest to understand under what conditions this behavior can be promoted. In the context of public good games and multilevel public goods games, I behaviorally and experimentally investigate if and how cooperation varies along with or as a response to other factors, namely norms, social efficiency, group identity, and risk. First, I find that personal norms, i.e., what one unconditionally believes to be the right thing to do, have major explanatory power over cooperation than social norms, i.e., what one believes others will do and think is the right thing to do. Moreover, I find that individuals positively react to social efficiency increases related to an upper-level (global) public good. The documented increase in contributions toward the global good comes at the expense of the contributions to a lower-level (local) public good, with the total contribution remaining unvaried. Furthermore, I obtain evidence that this result is robustly replicated in the context of groups primed with a strong sense of national identity and facing a task framed to recall real-world institutions (national and European Union public budgets). Lastly, I document that the presence of a probability of facing significant losses - whether independent or correlated among group members - does not impact contributing behavior in the public good compared to deterministic scenarios. These results, while building on recent cutting-edge experimental literature, suggest interesting avenues for new research.

Chapter 1

Introduction

Humans have the remarkable ability to live in large societies of many unrelated individuals (Capraro, 2019). This fact often creates *social dilemmas*: situations in which individuals must choose between acting in their own best interest and acting in the best interest of the collective (Dawes, 1980; Kerr, 1983; Samuelson and Messick, 1986). Social dilemmas are characterized by the fact that (i) each individual in a group, regardless of what others choose, receives a higher monetary payoff for making a socially defective choice than for making a socially cooperative one, (ii) while all individuals would receive a higher monetary payoff if everyone makes the socially cooperative choice rather than if everyone defects.

A specific type of social dilemma known as the *public good game* (henceforth, PGG) serves as an ideal framework for examining individuals' willingness to cooperate, which is a crucial factor for the stability of human societies organized in large groups, from ancient hunter-gatherer societies to modern states and nations (Nowak, 2006). Cooperation involves individuals making decisions to incur costs for the benefit of others or society as a whole. The feature of providing a benefit to others makes cooperative behaviors belonging to the broader category of *prosocial behaviors* alongside, for instance, volunteering, donating money to charities, or donating blood, which have been extensively studied in economics (e.g., Bilancini, Boncinelli, Di Paolo, et al., 2022; Di Paolo and

Pizziol, 2023; Goette and Tripodi, 2020; Iajya et al., 2013; Lorko et al., 2023; Pizziol et al., 2023). However, the reasons why some individuals choose to cooperate while others do not and why individuals cooperate in certain situations but not in others are still topics of considerable debate in the academic community.

The PGG involves individuals being matched in groups and individually receiving an initial amount of resources, referred to as endowment, which they individually must decide how much to spend on the group's public good. All contributions from all group members are then multiplied by an efficiency factor and divided evenly among the group members. This game has been widely used to study cooperation as it well captures the tension between self-interest, which ultimately leads to free-riding, and the common good, which drives toward maximizing the group payoff.

The individual payoff from the game can be formalized as follows:

$$\pi_i = e_i - c_i + \frac{\delta}{N} \sum_{j=1}^{N} c_j \tag{1.1}$$

where e_i is the individual endowment; c_i is the individual contribution to the public good; δ is the efficiency factor; N is the group size, i.e., the number of members composing the group representing the societal context; and $\sum_{j=1}^{N} c_j$ denotes total contribution of the group. The ratio $\alpha = \frac{\delta}{N}$ is known as the *marginal per capita return* (MPCR), and represents the benefit that each individual receives from contributing 1 unit to the public good. Thus, $(1 - \alpha)$ represents the actual cost that player i incurs by contributing 1 unit to the public good. Why this game nicely captures the tension between self-interest and collective interest relies on how it is parameterized. That is: $1 < \delta < N$ (or alternatively, given $\alpha = \frac{\delta}{N}$, $\frac{1}{N} < \alpha < 1$). Since $\delta < N$, a group member maximizes his/her own monetary payoff by contributing zero to the public good. However, given that $\delta > 1$, then it is socially efficient for a group member to contribute his/her entire endowment. Hence, the Pareto optimal outcome is one in which everyone contributes their entire endowment, but each

individual has the incentive to contribute less, given that if everyone else contributed their entire endowment, one could receive the highest payoff by contributing nothing. Since this holds for everyone in the game, the game-theoretical prediction based on Nash Equilibrium (NE), assuming rationality in combination with self-regarding preferences, is that everyone contributes zero to the public good.

However, over the past few decades, the economic literature has moved away from the rationality and self-interest paradigm that characterized early studies on this topic (e.g., Bergstrom et al. (1986)), as many real-world situations and experiments have revealed a wide range of behaviors that are not accounted for by traditional models and assumptions in this game (as well as in many other decision-making scenarios). Experimental research on the public good game has produced a large body of evidence showing that only a small portion of players typically plays the NE strategy, while most people actually tend to contribute more than the game-theoretical prediction, despite usually still at suboptimal levels (e.g., Chaudhuri, 2011; Fiala and Suetens, 2017; Ledyard, 1995; Zelmer, 2003). In one-shot trials and in the initial stages of finitely repeated trials, subjects generally provide contributions halfway between the socially efficient and free-riding levels, a phenomenon that has been labeled as *overcontribution* (Guala, 2005). Recently, it has been demonstrated that confusion about the game's incentive structure and the dominant strategy cannot explain this evidence (Granulo et al., 2023). Also, theoretical approaches started incorporating findings from experimental economics (e.g., Ordaz-Cuevas and Sánchez-Pérez, 2023), also going beyond the *homo economics* and allowing for rationality with cognitive bounds.

Public goods provisions, such as national defense, clean air, public health, and the environment, are just some of the many real-world scenarios where the PGG applies. Understanding what people decide and how they make decisions about public goods provision is crucial not only for scholars willing to uncover the unknown about human cooperative behavior but also for policymakers planning to design effective policies and institutions aimed at achieving optimal outcomes for society.

The focus of this doctoral dissertation is on cooperation. The topic is approached from multiple perspectives, with a particular emphasis on norms, efficiency, group identity, and risk, utilizing Behavioral and Experimental approaches. More specifically, this dissertation comprises four essays that study cooperative behavior in the PGG and in a variant of the game known as the multilevel public goods game (MLPGG). Chapter 2, Chapter 3, Chapter 4, and Chapter 5 contain the four essays, while Chapter 6 concludes the dissertation.

From a methodological standpoint, all experiments presented in this dissertation are one-shot anonymous conducted online using Prolific (Palan and Schitter, 2018). Online experiments offer advantages over traditional laboratory experiments with undergraduate and graduate students, but they also present new challenges that require careful consideration (Birnbaum, 2004). On the one hand, online experiments enable rapid recruitment of larger and more diverse samples, reducing the likelihood of demand effects as experimenters have limited interaction with participants. Additionally, they facilitate standardized procedures, enhancing replicability and cost-effectiveness compared to traditional lab experiments, and they remove the need for participants to travel to a physical location. On the other hand, online experiments entail less control over participants, necessitating the use of simplified instructions with carefully chosen wording to ensure comprehension. Thus, attention checks and control questions are crucial to maintaining data quality. Dropout rates tend to be higher compared to lab studies, and self-selection of participants based on features such as study length can occur. Furthermore, participants may vary in the device or browser they use, requiring experimenters to test their protocols on different platforms and recommend or enforce specific device usage.[1] It is worth noting that potential distortions in sample composition can arise from the virality of social media content, as demonstrated by the case of Prolific in the summer of 2021 due to a viral TikTok video, which promoted "easy gains" through subscription to the platform. Nevertheless, platforms such as Prolific have

[1]The advantages and disadvantages of online experiments are further discussed by Arechar et al. (2018a), Buso et al. (2021), and Horton et al. (2011).

addressed the issue by implementing filters that enable researchers to select participants according to key characteristics such as gender and registration date, providing researchers with greater control over the recruited sample.

The first essay is presented in Chapter 2. It investigates the roles played by personal moral judgments and social norms on cooperation, focusing on which drives contribution choices more using the one-shot anonymous PGG as the relevant strategic setting where to study behaviors. In summary, this work aims to reply to the following research questions:

> *Do and to what extent personal norms drive contribution decisions? What is their comparative role with respect to social norms?*

We consider *personal norms* to be a measure of one's unconditional normative conviction of what is the right thing to do – i.e., the right amount to be contributed by a group member to the group's public good –, while the combination of individuals' expectations of others' behavior and their expectations of others' normative judgments (i.e., the second-order belief about others' personal norms) represents for us what people believe the *social norm is* (in this, we follow the theory and procedures developed by Bicchieri and Xiao, 2009). We find that personal norms are a dominant force driving contributions in the one-shot online anonymous PGG compared to social norms. In doing so, we contribute to a recent but growing literature that focuses on the relevance of personal norms in motivating behaviors (Bašić and Verrina, 2021; Capraro and Rand, 2018). Also, we mitigate concerns about the potential endogeneity of subjects' responses by eliciting norms also with an out-of-sample group of individuals and showing that their norms are not statistically different from those elicited in the in-sample group. Thus, our work also has a methodological connotation in the field of norms elicitation, offering a tool to overcome the endogeneity issue in the elicitation of norms

within the PGG framework (borrowing from Krupka and Weber (2013) the idea to use an independent sample while differing from them in the kind of norm elicitations).

Chapter 3 presents the second essay, which focuses on the MLPGG variant of the standard PGG. In this game, each individual is placed in a local group and a global group, the former being nested in the latter to form a hierarchical structure. This work explores the effects of changes in the relative efficiency of the two nested goods, with efficiency being the total amount of benefits produced by the respective public good when every group member makes a 1-unit contribution to it. In doing so, we aim to reply to the following questions:

> *Do changes in the relative social efficiency of two public goods in a nested structure affect people's contributing decisions to them? If so, in which way?*

With our design, we offer a comprehensive investigation of the main effects driven by relative efficiency changes by setting up a series of incremental efficiency treatments to clear up the mixed evidence in the literature (Blackwell and McKee, 2003; Chakravarty and Fonseca, 2017; Fellner and Lünser, 2014; Gallier et al., 2019). Our results show that individuals tend to increase their contributions to the global good as its relative efficiency increases (*leveling-up* effect) while decreasing their average contributions to the local public good (*substitution* effect). We do not find evidence of an overall increase in total contributions (i.e., the sum of the contributions to the local and global public goods). Additionally, by contrasting the data from the MLPGG with that of a standard linear PGG, which serves as the control treatment, we are able to verify our *categorical crowding-in* hypothesis, that is, average total contributions increase as a consequence of the addition of a global good *per se*.

In Chapter 4, the third essay is reported. In it, we rely again on a one-shot MLPGG. This work is based on an experimental task framed

in a non-neutral way and on groups formed to prime a sense of group identity, both locally and globally. We make experimental subjects face a trade-off between contributing to a global public good referred to as the *European Union (EU) public budget* or to a local public good labeled as *Country public budget*. Local-level groups are homogeneous in terms of the EU country where subjects actually come from, while global-level groups are made up of three local groups from three different EU countries. Also, in the spirit of the previous essay, we set up two efficiency treatments to measure how EU participants respond to an increase in the efficiency of the EU public good. This change represents the higher expected returns from a public expenditure dealt at the transnational level in sectors like the environment, energy, defense, or public health in the face of pandemics. In this work, hence, we want to reply to the following research questions:

> *To what extent are EU citizens willing to contribute to their own potential Country public good as opposed to a potential EU public good? To what extent do changes in the relative efficiency between the two nested goods affect contribution decisions? Are there cross-country differences in the contribution decisions?*

We document that participants in our experiment are willing to contribute to both the Country and the European budget up to around 70% of their initial endowment. We confirm the absence of a marginal crowding-in effect while verifying the presence of both leveling-up and substitution effects in our framed context. In doing so, this work has the flavor of a replication study, confirming the robustness of our previous results (Chapter 3) in a very different setting. As for cross-country differences, we do not find evidence of statistically significant differences in the contributing behaviors, showing rather homogenous responses in our pool of EU citizens.

The fourth essay, presented in Chapter 5, is based on an experiment of a one-shot linear PGG. This essay explores the role of risk in the PGG

to detect if and how cooperation is affected by the presence of risk or how risk is correlated among the group members. The type of risk we consider is an "environmental risk," i.e., the existence of an exogenous stochastic process that can potentially produce adverse events negatively affecting individuals' payoffs. In summary, this essay aims to reply to the following research questions:

> *Does it and to what extent does the presence of a slight chance of severe adverse events impact cooperative behavior? Do different risk correlations across individuals play a role?*

While documenting a considerable level of cooperation (about 60% of resources available for contribution) despite the low marginal return of contributing, we find that environmental risk does not change cooperative behaviors compared to deterministic scenarios. Additionally, we find that the nature of environmental risk – i.e., whether it is independent, positively, or negatively correlated across group members – does not significantly affect cooperation either. Thus, our results support standard choice models that rely on expected utility theory with other-regarding preferences, excluding specific effects of over-weighting low probabilities or risk correlations.

Supplementary materials in support of each Chapter are reported in the Appendixes attached to this dissertation. All include further details provided in table format on the experimental samples' descriptive statistics or on the analyses based on the experimental data. In addition, Chapter 5's Appendix provides a summary of a series of simulations based on best-response analyses. Lastly, each Appendix reproduces the full set of experimental instructions that the subjects faced in the experiments.

Chapter 2

The role of personal and social norms in the public good game

This essay is based on the paper: M. Catola, S. D'Alessandro, P. Guarnieri, V. Pizziol "Personal norms in the online public good game," Economics Letters, vol. 207, pp. 110024, 2021.

2.1 Introduction

The evidence that individuals tend to make significant contributions in the public good game (PGG), even in one-shot anonymous interactions, has been replicated across various experimental designs (Chaudhuri, 2011). The literature offers several explanations for such high levels of contributions, mainly in terms of behaviors conditioned on social expectations. Examples include the theory of conditional cooperators (Fischbacher, Gächter, et al., 2001; Thöni and Volk, 2018), or experiments where the possibility of punishment triggers injunctive norms (Herrmann et al., 2008). However, some other recent studies have focused on how individuals, in one-shot games or highly anonymous interactions, rely on their personal norms and comply with what they personally and un-

conditionally regard as the right thing to do (Bašić and Verrina, 2021; Biziou-van-Pol et al., 2015; Capraro and Perc, 2021; Capraro and Rand, 2018; Eriksson et al., 2017).

In this paper, we present an online experiment investigating the role of personal norms as compared to social norms in motivating contributions to a public good. To measure both personal and social norms, we apply the procedure developed by Bicchieri and Xiao (2009), eliciting Personal Normative Beliefs (PN), Empirical Expectations (EE), and Normative Expectations (NE). While PN measures one's unconditional normative conviction and thus represents the subjects' personal norms, the combination of EE (i.e., subjects' expectations of others' behavior) and NE (i.e., subjects' expectations of others' normative judgments) represents what subjects believe the social norm is. This methodology differs from that of Bašić and Verrina (2021), which also studies a PGG but only elicits PN. Our results bring evidence in favor of personal norms being a stronger predictor of the contribution choice, providing additional evidence that, at least in one-shot online interactions, people tend to follow their individual normative judgments more than social norms.

Additionally, we investigate a potential self-justification bias in the elicitation of personal and social norms. Since norms are elicited after the decision task, subjects may be responding to the norm-elicitation questions by justifying *ex-post* their decision. To assess whether this is the case, we conduct an additional online experiment with an independent sample of subjects where they only face the norm-elicitation task without performing the decision task (in the spirit of Krupka and Weber, 2013). With this approach, we can compare the norms elicited in the first experiment with those of the external sample and evaluate the reliability of our main result. We find no statistically significant difference between the norms elicited in the two experiments, reinforcing the reliability of our findings and mitigating any potential concern about endogeneity issues.

2.2 Experimental design and procedures

Our primary study was a one-shot linear public good game. We randomly assigned 164 UK nationals to groups of 4 members. Table 1 reports on our sample's demographics.[1] Overall, the average age is around 36, and about twenty-three percent of the participants are students. About two-thirds of the participants are female, while seventy percent have a full- or part-time job.

Table 1: Mean (and standard deviation) of participants' characteristics.

Age	Male	Student	Soc. Status	Education	Employed
36.38	0.33	0.23	5.40	3.69	0.71
(12.74)	(0.47)	(0.42)	(1.54)	(1.00)	(0.46)

Notes: Socioeconomic status refers to what participants self-reported as their place on a ladder representing society that goes from 1 to 10. Education is coded as: 1 "no formal qualifications", 2 "secondary education", 3 "high school diploma", 4 "undergraduate degree", 5 "graduate degree", 6 "doctorate degree".

Each participant was endowed with 10 Points. The contributions to the public good could be of any integer from 0 to 10 Points. The monetary payoff of each individual was determined by the amount privately kept plus the earnings from the group common pool, with the marginal per capita return (MPCR) being equal to 0.6. After making their contribution decisions, subjects were presented with questions aimed at eliciting their norms. Personal Normative Beliefs (PN) were elicited by asking participants their opinions on how much one group member should contribute. Empirical Expectations (EE) were identified by querying subjects on what they believed was the average contribution of the other members, while Normative Expectations (NE) were obtained by asking for their opinions on how much the other members believed one group member should contribute. As is standard, questions on EE and NE were

[1] This information was retrieved from Prolific.

incentive-based, but the ones eliciting PN were not. Participants received an additional £ 0.10 for each correct answer, i.e., when EE matched the other members' average contribution and when NE matched the other members' average PN. At the end of the experimental task, subjects replied to three control questions to test their understanding of the game. We also administered a three-item Cognitive Reflection Test (Frederick, 2005) and selected questions from the Global Preference Survey (GPS) (Falk, A. Becker, et al., 2018) (namely, items AF.1.2, AF.2.1, AF.3.2, AF.4.3, AF.5.1 and AF.6) to gather subjects' social and risk preferences.

In our second experiment, we recruited 104 UK nationals who had not participated in the first experiment to express their beliefs about personal and social norms. We explained the contribution task from the main experiment to these subjects and asked them what they believed a group member ought to contribute, what was the average contribution of the group members in the main experiment, and what they believed group members in the main experiment thought others ought to contribute. We set the conversion rate for both experiments at 1 Point = £ 0.025. In the first experiment, subjects earned an average of £ 1.13 (with £ 0.50 being show-up fees).[2] In the second experiment, the average payment was £ 0.16 (including a show-up fee of £ 0.10). Both experiments were programmed in oTree (Chen et al., 2016) and conducted online using the Prolific platform (Palan and Schitter, 2018).

2.3 Results

Figure 1 shows the average contribution choice to the common pool as well as the average values of each elicited norm.[3]

The average contribution to the public good and the average personal normative beliefs are clearly both remarkably higher than empirical and

[2]The overall absolute level of payments was in line with the standard for online experiments, which typically pay less compared to lab experiments. However, the average reward per hour was still valuable (about £10) and higher than the £5 per hour minimum requirement of the platform we used to recruit participants (Prolific). Moreover, the proportion between the fixed show-up fee and the payoff earned from the experiment appears consistent with the standard both in the lab and online.

[3]We discarded 5 observations due to implausible answers.

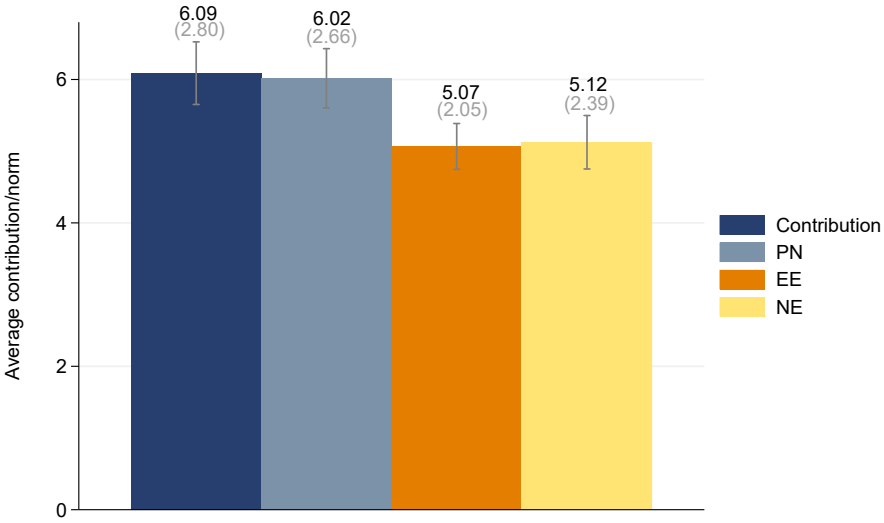

Figure 1: Mean (and standard deviation) of contribution, personal norm, empirical expectation, and normative expectation. Confidence intervals at the 95% level.

normative expectations and reach about the same level. This evidence suggests that in the context of our experiment, PN are highly aligned with individuals' contribution decisions. The analysis of a set of non-parametric Wilcoxon signed-ranks tests confirms this intuition. Indeed, on the one hand, we do not find any statistically significant difference between contribution choices and PN ($p = 0.3773$) as well as between EE and NE ($p = 0.8787$). On the other hand, the difference is statistically significant when it comes to comparing contribution with EE and NE ($p < 0.001$) and PN with EE and NE, respectively (both $p's < 0.001$). Furthermore, it is worth mentioning that the proportion of people whose PN is equal to their contribution choice is considerably higher (65.41%) than those whose EE or NE are, respectively, equal to their contribution choice (namely, 26.42% and 25.16%).

In addition, in Table 2 we also run Tobit regressions where the contribution is the dependent variable and norms are the main regressors. We consider eight specifications of the model: the first three in which we include only two norms, the fourth in which we include all three norms; the last four follow the same logic but with the inclusion of a battery of control variables taking into account socio-demographic and individual-specific characteristics about preferences. Specifically, the controls are: Altruism, Patience, Risk tolerance, Trust, Negative reciprocity, Positive reciprocity, the score in the comprehension questions, the CRT score, Age, Male, Student status, Socioeconomic status, Education. The full set of results for columns 5-8 is reported in Table A.2 in Appendix A.

Table 2 confirms our result about the relative importance of personal norms compared to social norms. In all specifications in which personal norms are included, the attached coefficients are always highly significant, and their magnitudes are the strongest. This result is also very robust to the inclusion of all the controls. Additionally, we conducted supplementary analyses that explored the heterogeneous effects of norms across various socio-demographic characteristics of the participants. The results of these analyses can be found in Table A.4, Table A.5, Table A.6, and Table A.7 in Appendix A. To prevent overfitting, we interacted different socio-demographic variables with the norms in each regression.

Notably, in all specifications, the coefficient associated with PN is consistently positive and statistically significant. Furthermore, certain effects emerge from the interaction between EE and Education, as well as between all norms and Socioeconomic status. Specifically, as shown in Table A.5, contributions appear to increase at a higher rate with an increase in EE for individuals with education levels higher than the second level, except for those with a Ph.D. level of education, for which the coefficient is not statistically significant, likely due to a limited number of relevant observations. From the results presented in Table A.7, it becomes evident that the Socioeconomic status variable exhibits a positive effect on contributions only when interactions with norms are included, with the exception of the highest level of self-perceived socioeconomic status. Another noteworthy observation is the amplified magnitude of the coefficients associated with norms compared to the main regression findings (Table 2). This amplification can be attributed to the coupling of norms with the socioeconomic status variable. Specifically, for PN, EE, and NE, the coefficients linked to the interaction with socioeconomic status predominantly display a negative trend (except for the case of PN with the lowest perceived economic status) in comparison to the baseline. In conclusion, we find that a portion of the variability in norms is intertwined with the socioeconomic status variable.

2.4 Checking for self-confirmation bias

We address the potential endogeneity problem in the answers to the beliefs' elicitation questions using the data collected in the second experiment.

First, we check for the homogeneity of the two samples of participants employed in the experiments, as we want to exclude that the two samples are drawn from populations with different distributions for some pivotal demographic and socio-economic characteristics. Altogether, we found that there is no statistically significant difference at the 5% level of significance for any characteristic. Results are reported in Table A.1 in Appendix A.

Table 2: Tobit regressions examining the contribution choices in the PGG.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PN	0.878***	1.180***		0.992***	0.828***	1.114***		0.944***
	(0.105)	(0.098)		(0.106)	(0.0986)	(0.104)		(0.109)
EE	0.486**		1.174***	0.616***	0.423**		1.098***	0.591***
	(0.160)		(0.163)	(0.163)	(0.145)		(0.159)	(0.139)
NE		-0.084	0.233	-0.292*		-0.0969	0.128	-0.336*
		(0.104)	(0.122)	(0.118)		(0.130)	(0.144)	(0.140)
Constant	-1.360**	-0.307	-0.691	-1.214*	-3.409**	-2.919*	-3.757**	-3.048**
	(0.512)	(0.481)	(0.747)	(0.479)	(1.120)	(1.201)	(1.386)	(1.097)
Controls	✗	✗	✗	✗	✓	✓	✓	✓
Observations	159	159	159	159	158	158	158	158
Pseudo R^2	0.294	0.267	0.165	0.306	0.309	0.307	0.183	0.334

Notes: The controls include Altruism, Patience, Risk tolerance, Trust, Negative Reciprocity, Positive Reciprocity, a score variable for Comprehension, a score variable for Cognitive Reflection Test, Age, Male, Student Status, Socioeconomic Status, and Education. Robust standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

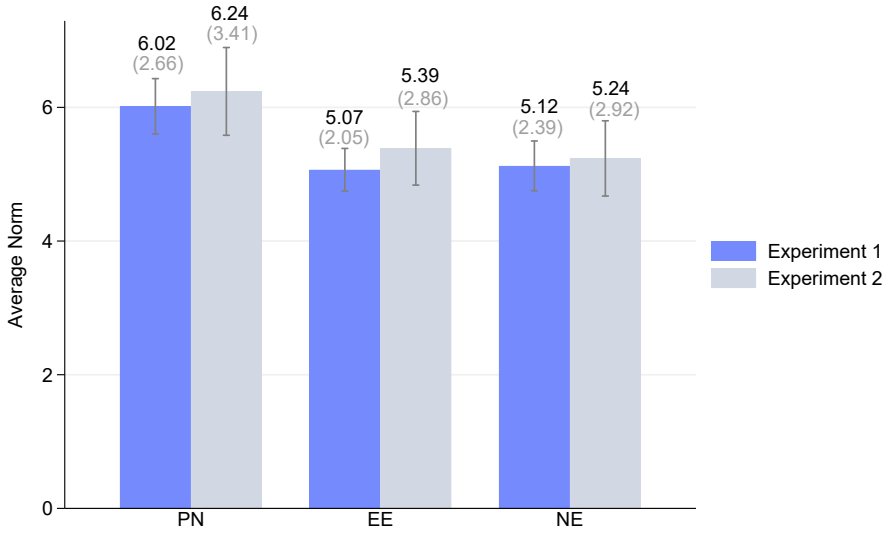

Figure 2: Mean (and standard deviation) of personal norm, empirical expectation, and normative expectation in Experiment 1 vis-à-vis Experiment 2. Confidence intervals at the 95% level.

The comparison between the average value of each norm between the two experiments shows that subjects have similar personal and social norms irrespective of whether they have or not performed the contribution task (see Figure 2). We confirm this intuition by pairwise comparisons of each norm in the two experiments through Wilcoxon rank-sum tests, and we find no statistically significant difference between any of the elicited norms (PN: $p = 0.3687$; EE: $p = 0.4201$; NE: $p = 0.9033$).[4]

2.5 Discussion and conclusions

This paper presents two main findings. First, we find that personal norms are the dominant force driving contributions in a one-shot online public good game. This result is in line with previous recent research, such as Capraro and Rand (2018), which showed that personal norms are stronger than descriptive social norms (EE in our framework) in the context of a Prisoner's Dilemma. Second, we find that the norms collected after the task are not statistically different from the norms of external subjects who did not perform the task. This result reinforces the reliability of our findings and suggests that the norms reported are not merely a *post hoc* justification of the decisions made.

Our contribution is particularly relevant in light of recent literature highlighting the potential role of manipulating social expectations in inducing pro-social behavior (Bicchieri and Dimant, 2019). The fact that personal norms are not contingent on social expectations makes them appear more stable, less susceptible to be influenced by contingent information, and deeply rooted in subjects' remote experience and education. Although this paper does not address how personal norms are formed, future research could investigate whether acting on the framework of the decision or the efficiency of the public good can affect them, potentially sustaining pro-sociality.

Given the possible long-term social and behavioral consequences of

[4]For robustness, we also pooled the data from the two samples and included a dummy for the experiment subjects participated in. Regressing every norm against such a dummy variable, we never find any statistical significance of this variable (all $p's > 0.1$). See Table A.3 in Appendix A.

the COVID-19 pandemic, the ability to form, sustain, and, if necessary, change personal norms held by isolated and digitalized individuals is of utmost importance in future policy-making.

Finally, it is important to acknowledge the limitations of this study. First, the study was conducted as an online experiment, which may restrict the generalizability of the findings to real-world, offline, and more intricate interactions. Additionally, the study primarily focuses on one-shot interactions, where participants make decisions without repeated interactions. This scope may not fully encompass the complexities of repeated interactions or the influence of learning and reputation on levels of cooperation. As a result, the findings may not apply to situations where individuals have the chance to establish cooperative strategies over time. Lastly, the study relies on self-reported measures to assess personal norms and social norms. While social norms were elicited with incentive-based questions, personal norms, as common in this literature, were not. Future works could explore incentive-compatible methods for eliciting personal norms.

Chapter 3

Efficiency-driven effects in the multilevel public goods game

This essay is based on the paper: M. Catola, S. D'Alessandro, P. Guarnieri, V. Pizziol, "Multilevel public goods game: Levelling up, substitution and crowding-in effects," Journal of Economic Psychology, vol. 97, pp. 102626, 2023.

3.1 Introduction

The Multilevel Public Goods Game (MLPGG) is an experimental design characterized by multiple public goods in a nested structure. Decision makers are assigned to one of several groups and asked to allocate their endowment among their private accounts, the public good provided only to their group (namely, the local public good), and the public good provided to all the subjects in the game (namely, the global public good).

This design has often been applied to investigate the tension between the tendency to favor their own groups (in-group favoritism) and the pro-sociality that leads individuals to contribute to the overall social benefit. This line of research typically acts on group composition to elicit

identity. Buchan, Brewer, et al. (2011) and Buchan, Grimalda, et al. (2009) apply the MLPGG to investigate the impact of globalization on the willingness of nationality-based groups to cooperate at the international level. Chakravarty and Fonseca (2017) study whether contributing to the local public good can be used to exclude members of other groups because of their lack of cooperation or to reward group members for their cooperation. Beekman et al. (2017) induce strong group identity by making groups conflict with each other in a pre-task. Gallier et al. (2019) measure in-group favoritism by eliciting group identity in subjects living in the same region of Germany. Finally, building on the established literature in the public good game (e.g., Martinangeli, 2021), Lange et al. (2022) differentiate between high- and low-endowment local groups to explore the effect of income heterogeneity on contributions.

Furthermore, the manipulation of the marginal per capita return (MPCR) – i.e., the return of a unitary contribution – has allowed scholars to study to what extent changes in the relative efficiency of the local and the global public goods affect contribution decisions in the MLPGG. In fact, while it is an established result that an increase in the MPCR has a positive effect on contribution in the standard public good game (Chaudhuri, 2011; Isaac and Walker, 1988; Isaac, Walker, and Williams, 1994; Ledyard, 1995; Zelmer, 2003), efficiency changes in the nested structure of the MLPGG entail additional trade-offs with several potential effects that make predictions on contributions less straightforward.

In this study, we exclusively focus on efficiency effects and sterilize group identity by running our experiment online, thus obtaining complete anonymity and excluding any feedback on group composition. The main objective is to add robustness to the evidence collected in the MLPGG literature and systematize the mixed and non-conclusive findings. To this end, we perform a set of treatments that investigate how subjects' allocation decisions are affected by the increase in the relative efficiency of the global public good. In particular, we investigate $i)$ to what extent this increase levels up the contribution to the global good itself (*leveling-up effect*), $ii)$ whether it decreases the contribution to the local public good – thus producing a substitution in the allocation between the local and the

global goods – (*substitution effect*), iii) or whether it crowds in the overall amount contributed to the two public goods (*marginal crowding-in effect*). Furthermore, we follow Bowles and Polania-Reyes (2012) and Bowles (2016) and investigate the presence of a *categorical crowding-in effect* by adding a treatment where only the local public good is provided in order to single out the impact on the total contribution of the mere addition of the global good.

Our results provide robust evidence of a leveling-up effect. While we find no evidence of marginal crowding in, we observe a decrease in the contribution to the local public good that enables us to confirm the substitution effect. Moreover, the mere introduction of a global public good significantly increases total contribution, thus verifying the categorical crowding-in effect. Finally, we observe that subjects contribute to one of the public goods even when it is dominated by the other public good both in terms of costs and returns. This evidence reinforces the argument that in the context of the MLPGG, subjects' decisions can be inconsistent with the narrow preference for maximizing either individual or group payoffs and may be driven, for instance, by preferences for allocations revealing inequity aversion or fairness criteria.

3.2 Methods

In the MLPGG framework, subjects are placed both in a local and a global group, the former being nested in the latter to form a hierarchical structure. In fact, the nested structure is what distinguishes the MLPGG from other multiple linear public goods designs (e.g., Bernasconi et al., 2009; Todd L. Cherry and David L. Dickinson, 2008; Falk, Fischbacher, et al., 2013; McCarter et al., 2014) or multiple threshold ones (e.g., Abraham et al., 2021; Corazzini, Cotton, and Reggiani, 2020; Corazzini, Cotton, and Valbonesi, 2015; Corazzini and Marini, 2022). An alternative approach to MLPGG design consists in keeping the standard single public good set up while allowing for different spillovers between the local and the global groups (Engel and Rockenbach, 2011; Güth and Sääksvuori, 2012).

We illustrate the specific settings of our design to introduce the main

features of the MLPGG structure. As depicted in Figure 3, we set two local groups of 4 members each, forming a global group of 8.

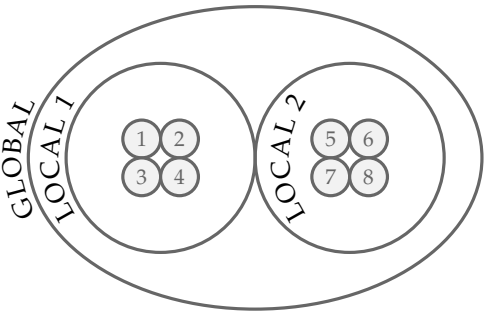

Figure 3: Configuration of the MLPGG in this study.

Each subject has to decide how to allocate an initial endowment of 10 tokens among three alternatives: their private account, a local public good, and a global public good. Every token contributed to the local good is multiplied by a local-specific factor and then redistributed equally among all 4 members of the subject's group, while every token allocated to the global good is multiplied by a global-specific factor and then redistributed equally among the 8 subjects. Finally, the tokens allocated to the private account are simply retained by the subjects.

Given the structure of the game, the payoff of player i is equal to:

$$\pi_i = 10 - c_i - C_i + \alpha \sum_{j=1}^{M} c_j + \beta \sum_{k=1}^{N} C_k. \tag{3.1}$$

where c is the individual contribution to the local public good, and C is the individual contribution to the global good; α and β are the MPCRs of the local and global public goods, respectively; M and N represent the sizes of the local and global groups, respectively. Also, we will refer to T as the total contribution defined as the sum of c and C.

3.2.1 Review of related studies

In recent years, several scholars have studied the efficiency effects in the MLPGG (Blackwell and McKee, 2003; Chakravarty and Fonseca, 2017; Fellner and Lünser, 2014; Gallier et al., 2019). However, while the leveling-up effect has been confirmed in all available studies, the debate concerning the substitution and marginal crowding-in effects is far from settled. On the one hand, Blackwell and McKee (2003) does not find any supporting evidence for the substitution effect and concludes that a rise in efficiency increases total contribution. On the other hand, more recent studies (Chakravarty and Fonseca, 2017; Fellner and Lünser, 2014; Gallier et al., 2019) obtain instead a strong substitution effect, which in the case of Chakravarty and Fonseca (2017) and Gallier et al. (2019) fully balances the leveling up, leaving the total contribution unchanged. The differences in the results are accompanied by a high degree of heterogeneity in the experimental designs, involving the manipulation of group identity and the relative efficiency.[1]

Group identity manipulation serves the purpose of inducing in-group bias in the context of the MLPGG structure. While Blackwell and McKee (2003) apply a minimal identity approach (H. Tajfel, 1974; Henri Tajfel, 1970, 1982) and Fellner and Lünser (2014) rely on random assignments of individuals to different groups, Chakravarty and Fonseca (2017) implement an endogenous reinforced procedure to form groups and make group identity more salient before subjects play the game. In contrast, Gallier et al. (2019) set up an artefactual field experiment exploiting the fact that participants belong to municipalities within the same region to bring out localism in a natural way. These differences are bound to impact on the efficiency effects as they affect the trade-off between the contribution to the subjects' own group and the global public good differently.

There are also significant differences in terms of efficiency manipula-

[1] While the MLPGG is usually implemented in a lab setting and with repeated interactions, Gallier et al. (2019) rely on a one-shot field experiment. However, this does not seem to account for the differences in the empirical results in terms of the impact of efficiency changes and in-group bias.

tion. Indeed, while Blackwell and McKee (2003) employs four different efficiency treatments, the subsequent studies only rely on two. In particular, both Chakravarty and Fonseca (2017) and Gallier et al. (2019) rely on a simplified design where only two critical treatments are compared. In the first treatment, the MPCRs of the public goods are normalized for group size (i.e., $\beta = \frac{M}{N}\alpha$), while in the second, the MPCRs are equal (i.e., $\alpha = \beta$). This experimental setting eliminates the trade-offs between returns, strategic risk, and costs, and it is likely to work in favor of a leveling up and against the marginal crowding-in effect. Indeed, in the normalized case, the goods' total returns are equal (as $\alpha M = \beta N$), but the local public good is safer in terms of strategic uncertainty and less costly, thus undermining the incentive to contribute to the global good. Conversely, when $\alpha = \beta$, the two goods are equally costly for the player but the potential returns for the global good are higher, providing a strong incentive to choose the global good. Consider, for example, the setup of Gallier et al. (ibid.) with 2 local groups of 4 members. In the first treatment, where $\alpha = 0.5$ and $\beta = 0.25$, the revenue generated by a token contributed to the local public good is twice the revenue generated by a token contributed to the global good, but only half of the players enjoy it. In the second treatment, where $\alpha = \beta = 0.5$, the revenue generated by the public goods is the same, but in the case of the global public good, it is enjoyed by all 8 players rather than just 4. Therefore, while the evidence of leveling up obtained by comparing only the two critical cases might be overestimated and hardly generalizable, Chakravarty and Fonseca (2017) and Gallier et al. (2019) are, nevertheless, the only two studies that do not find any evidence of marginal crowding-in.

Table 3 provides a summary of the differences in terms of efficiency treatments and group identity elicitation in the previously mentioned studies. Scholars have exploited the characteristics of the normalized efficiency treatment mentioned above to test in-group favoritism in the MLPGG setup. Indeed, the two public goods produce the same expected gain (in the case of equal contribution by each local-group member) and thus, the evidence that people tend to contribute more to the local public good than to the global public good has been interpreted as revealing a

Table 3: Summary of experimental designs in the MLPGG literature.

Paper	Type	Iterations	α	β	M, N	Group identity
Gallier et al. (2019)	Field	One-shot	0.5	0.25, 0.5	4, 8	Neighbourhood
Chakravarty, Fonseca (2017)	Lab	Repeated	0.4, 0.8	0.4	3, 6	Klee-Kandinsky task
Fellner, Lünser (2014)	Lab	Repeated	0.4	0.2, 0.3	4, 8	No manipulation
Blackwell and McKee (2003)	Lab	Repeated	0.3	0.1, 0.15, 0.2, 0.3	4, 12	Group colours

Notes: Type is whether the experiment was run in the field or in the lab; α is the local MPCR for each treatment; β is the global MPCRs for each treatment; M is the number of local group members; N is the number of global group members; Group identity refers to the strategy used to manipulate group identity.

bias in favor of the local. This evidence was standard in the MLPGG experiments (Blackwell and McKee, 2003; Chakravarty and Fonseca, 2017; Fellner and Lünser, 2014), up until Gallier et al. (2019), which could not replicate it. However, despite the robustness of this effect across studies, its interpretation is still controversial since the normalized case maintains an imbalance between the two public goods in terms of strategic uncertainty and opportunity cost in the contribution. Chakravarty and Fonseca (2017), for instance, see it as a consequence of the lower degree of strategic uncertainty in cooperation at the local level due to the lower number of players (*size effect*). A similar conclusion is reached by Gallier et al. (2019) who, in reviewing the previous findings, point out that a larger contribution to the local public good in the normalized treatment is not *per se* evidence of parochialism since this may derive from the contribution being responsive to MPCR and irresponsive to group size. The role of strategic uncertainty might also explain why, in Fellner and Lünser (2014), higher returns alone are not sufficient to sustain contribution to the global public good unless they are combined with feedback on the contribution of others.

Another common result in the literature is that, albeit lower, contri-

bution to the local public good persists even when the MPCRs are equal. This result somehow questions the role of efficiency as the sole driver of contribution. Chakravarty and Fonseca (2017) interpret it as a sign that financial considerations do not totally overcome the effect of (local) group social identity. However, the literature has not yet tested whether the contribution to the global public good persists when no financial incentives exist.

Finally, in a standard PGG, Todd L. Cherry and David L. Dickinson (2008) and Bernasconi et al. (2009) show that the addition of an identical public good to the players' choice set leads to an increase in total contribution. More recently, Chakravarty and Fonseca (2017) document the presence of the same categorical crowding-in effect in the context of a MLPGG, with endogenously formed groups, by showing that adding a financially dominated local public good to an already available global one increases total contribution.

3.2.2 Experimental design

The general objective pursued by our pre-registered design is to provide robust evidence of efficiency effects in the MLPGG. Firstly, we investigate the robustness of the leveling up by studying whether the contribution to the global public good always increases whenever its relative efficiency rises. Secondly, we investigate whether such an increase in efficiency produces a marginal crowding in that increases total contribution or induces a substitution with subjects simply shifting their contribution choice between the two public goods.

The review of experimental evidence suggests that results are sensitive to the specific characteristics of the designs. Namely, the variety of strategies adopted to induce group identity might condition the replication of stable tendencies in contribution decisions. Consequently, we opted to avoid any manipulation of group identity in order to minimize its effects on the allocation decisions between the local and the global public good. Accordingly, we provided participants with no group characterizations or feedback on group composition.

Moreover, since the experiment was run online, no other visual reference was available to subjects, thus making it possible to avoid other sources of potential identification. Finally, the decision to implement a one-shot game instead of a repeated one reduces the opportunity for the individuals in the local groups to learn and adopt strategic spillovers across rounds.

In a between-subjects design, we keep α at a fixed value of 0.6 across all treatments, whereas β takes values of 0.15, 0.30, 0.45, and 0.6. Table 4 provides a summary of all the parameters across treatments and, to better clarify the social efficiency of each public good, the value of the total benefit (TB), defined by Gallier et al. (2019) as the individual earnings from a good obtained when every group member makes a one-token contribution to it (i.e., αM and βN respectively).

Table 4: Summary of the treatments parameters.

Treatment	**Local PG**			**Global PG**		
	M	α	TB	N	β	TB
T_0	4	0.6	2.4	-	-	-
T_1	4	0.6	2.4	8	0.15	1.2
T_2	4	0.6	2.4	8	0.3	2.4
T_3	4	0.6	2.4	8	0.45	3.6
T_4	4	0.6	2.4	8	0.6	4.8

In line with Blackwell and McKee (2003), treatments involve only the manipulation of β. Specifically, T_2 and T_4 represent the two commonly implemented special cases. On the one hand, T_2 corresponds to the situation where the returns of the public goods are normalized ($\alpha M = \beta N$), thus sterilizing any efficiency effect due to scale. Consequently, the local good is less costly and hence less risky, given that the individual return from a token contributed to this public good is higher than the return of a token contributed to the global public good.

Conversely, T_4 corresponds to the opposite case in which marginal returns are equal ($\alpha = \beta$). Therefore, the public goods are equally costly,

but the global public good is more efficient because of the scale effect. This feature has two main implications. Firstly, for the individual player, the two public goods are equally risky as the return from the contribution is the same. Secondly, while in T_2 the members of the local group are better off if their fellow member i contributes to the local account rather than to the global one (as $\alpha > \beta$), this is not the case for T_4 (given that α and β are equal). Therefore, contributing to the local public good in T_4 is neither less costly for the contributors nor does it provide higher payoffs for their fellow local group members. Thus, the only difference between the two public goods in T_4 is that contribution to the local public good excludes the members of the other group from the benefit of the public good provision.

Differently, in treatment T_1 – which is a specific novelty of our design – we introduce a global public good that is worse than the local one in all respects. It is more costly – β is lower – and the TB is lower as well. Hence, payoff-wise, there is no incentive to contribute to the global public good, and the decision to contribute may then be motivated by concerns about equity and fairness.

T_3, which is analogous to the treatment used by Fellner and Lünser (2014), is an intermediate case where both the trade-offs of cost and total benefit are present – $\alpha > \beta$ but $\alpha M < \beta N$ – and affect the decision in opposite directions, favoring contribution to the local and to the global public good, respectively. Finally, treatment T_0 is designed to test for the categorical crowding-in effect, given that subjects in this treatment can only contribute to a local public good. Thus, we test our hypothesis by adding a global good to a situation where only the local good is present, and not *viceversa* as in Chakravarty and Fonseca (2017).

3.2.3 Hypotheses

Our design enables us to single out three main hypotheses which address the main efficiency effects investigated in the MLPGG literature.

Hypothesis 1 (levelling up) *The average contribution to the global public good $\bar{C}$ is an increasing function of β; i.e., individuals tend to increase their*

contributions to the global good as its relative efficiency increases.

Hypothesis 2 (substitution effect) *The average contribution to the local public good $\bar{c}$ is a decreasing function of β; i.e., individuals tend to decrease their contributions to the local good as the relative efficiency of the global good increases.*

Hypothesis 3 (marginal crowding in) *The average total contribution $\bar{T}$ is an increasing function of β; i.e., individuals tend to increase their overall contributions as the relative efficiency of the global good increases.*

As for the categorical crowding in, we formulate the following hypothesis.

Hypothesis 4 (categorical crowding in) *Average total contribution $\bar{T}$ increases as a consequence of the addition of a global good per se.*

3.2.4 Implementation

The experiment was implemented using oTree (Chen et al., 2016) and conducted online on the Prolific platform (Palan and Schitter, 2018), which allowed for the recruitment of a socio-demographically varied and well-powered sample with a guarantee of complete anonymity and full randomization. A total of 802 UK nationals participated in two different sessions. 80 subjects participated in the first session (run as a pilot), and the remaining 722 in the second session.[2] Each subject was randomly assigned to one of the treatments and then to a local and a global group. We succeeded in obtaining sub-samples of almost the same size, although some dropouts led to slight imbalances due to the substitution procedure, which randomly assigns new entrants to treatments. Table 5 reports on our sample's size and demographics and shows that the treatment sub-samples were homogeneous in terms of key individual-specific variables confirming that the randomization of individuals across treatments worked successfully.[3] It is also worth noticing that compared to

[2]We aggregated the two sessions because no substantial changes occurred between sessions 1 and 2, and we have chosen out of caution the same time slots and days of the week to launch them.

[3]There is no statistically significant difference across treatments at any level of significance. We performed Kruskal-Wallis tests for the variables: age, income, socioeconomic

experiments in the lab, which are standard in the MLPGG literature, the average age of our participants is notably higher, and the fraction of students is lower, thus making our sample more representative of the actual population.

Table 5: Sample sizes and participants' average characteristics by treatment.

	N	Age	Male	Income	Student	Soc. Status	Educ.	Employed
T_0	164	36.28	0.32	2.59	0.23	5.39	3.68	0.70
T_1	160	35.01	0.31	2.42	0.20	5.31	3.79	0.74
T_2	164	33.89	0.30	2.27	0.26	5.36	3.64	0.70
T_3	160	34.28	0.37	2.59	0.18	5.46	3.72	0.68
T_4	154	34.16	0.30	2.64	0.20	5.32	3.65	0.76

Notes: Education is coded as: 1 "no formal qualifications", 2 "secondary education", 3 "high school diploma", 4 "undergraduate degree", 5 "graduate degree", 6 "doctorate degree". Personal income is coded as: 1 "less than 10k", 2 "10–20k", 3 "20–30k", 4 "30–40k", 5 "40-50k", 6 "50-60k", 7 "60-70k", 8 "80-90k", 9 "greater than 90k". Socioeconomic status refers to participants' self-reported place on a ladder representing society from 1 to 10.

After going through the instructions, subjects faced the decision on the main task, i.e., how to allocate their endowment between their personal account, the local public good, and the global public good. After the decision task, participants answered questions to measure their empirical expectations, personal normative beliefs, and normative expectations (Bicchieri and Chavez, 2010; Bicchieri and Xiao, 2009).[4] At the end of the experimental questionnaire, subjects replied to three control questions and a 3-item Cognitive Reflection Test in the standard version proposed by Frederick (2005), followed by subjects' elicitation of their social and risk preferences using questions AF.1.2, AF.2.1, AF.3.2, AF.4.3, AF.5.1. and AF.6. from Falk, A. Becker, et al. (2018).

Each participant was endowed with 10 Points and advised in the instructions that they would be converted into pounds at the end of the

status, and education, and Fisher's tests for the dichotomous variables: gender, student status, and employment status.

[4]The effect of norms in shaping contributions is a growing topic in the PGG literature e.g., Bašić and Verrina, 2021; Catola, D'Alessandro, et al., 2021; Engel and Kurschilgen, 2020; Kandul and Lanz, 2021; Otten et al., 2021.

experiment at the given rate of 1 Point = £ 0.025. Overall, the average payment was £ 1.13 (out of which £ 0.50 were show-up fees).

3.3 Results

We start our analysis by providing a visual representation of the average outcome choices, i.e., contributions, per treatment. Figure 4 depicts the average contribution to the local public good, Figure 5 the average contribution to the global public good, and Figure 6 the average of the sum of the two types of contribution. Local and global contributions are always positive across treatments and show opposite trends as β increases. In contrast, total contributions appear to be stable between T_1 and T_4, but lower in T_0.

These general trends are only partially confirmed by non-parametric tests of the differences between consecutive treatments. The difference in the global contribution is shown to be statistically significant only in the comparison between T_2 and T_3 (*Wilcoxon rank-sum* tests, $T_1 - T_2$, $p = 0.0502$; $T_2 - T_3$, $p = 0.0003$; $T_3 - T_4$, $p = 0.3700$). A similar result holds for the contribution to the local public good. Indeed, the decrease in contribution is only statistically significant when moving up from T_2 to T_3 (*Wilcoxon rank-sum* tests, $T_1 - T_2$, $p = 0.6124$; $T_2 - T_3$, $p = 0.0020$; $T_3 - T_4$, $p = 0.2135$). However, we must note that comparisons between non-consecutive treatments always provide statistically significant differences for contributions both to local and global public goods.

Also, the non-parametric tests confirm that there is no significant increase in total contribution as β increases from T_1 to T_4 (*Wilcoxon rank-sum* tests, $T_1 - T_2$, $p = 0.1974$; $T_2 - T_3$, $p = 0.1237$; $T_3 - T_4$, $p = 0.4479$). In contrast, when only a local good is present, the total contribution is lower than in all the other treatments (*Wilcoxon rank-sum* tests, $p < 0.001$ for each comparison between T_0 and other treatments). It is worth underlining the statistical significance of the comparison between T_0 and T_1, as it shows that the addition of an inefficient public good is enough to increase total contributions.

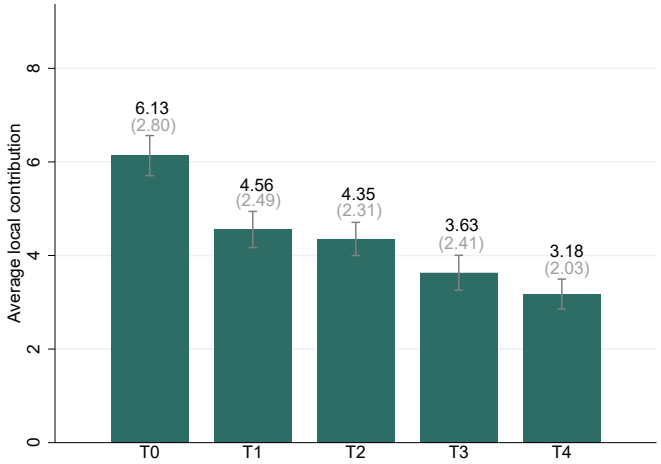

Figure 4: Average contribution (and standard deviation) to the local public good by treatment. Standard deviations in parentheses. Confidence intervals at the 95% level.

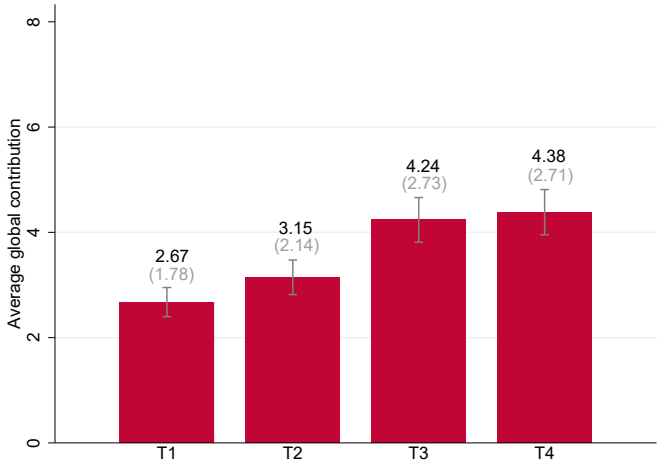

Figure 5: Average contribution (and standard deviation) to the global public good by treatment. Standard deviations in parentheses. Confidence intervals at the 95% level.

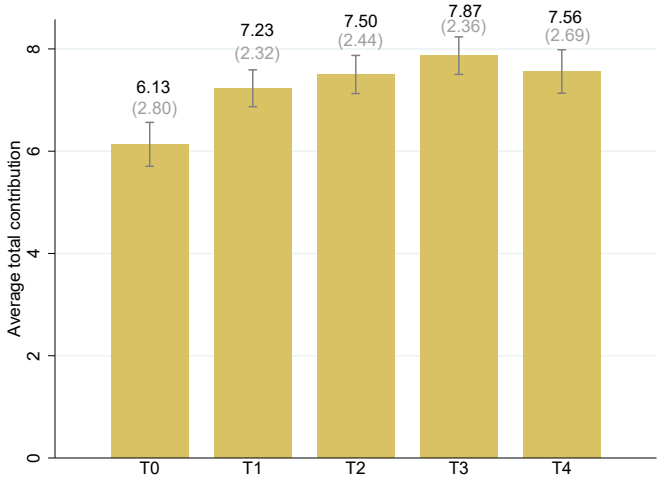

Figure 6: Average total contribution (and standard deviation) by treatment. Standard deviations in parentheses. Confidence intervals at the 95% level.

3.3.1 Contribution to the local and global public goods

In this subsection, we focus our analysis on the contributions to each public good and leave the study of the total contribution to the following subsection. Accordingly, we exclude the observations of T_0 from this analysis, given that subjects in that treatment do not face the decision on whether (and how much) to contribute to the local or the global good since there is no global public good in T_0.

To test our hypotheses, we perform a set of OLS regressions using β – i.e., the MPCR of the global public good – as the main regressor to estimate the average effect of changes in efficiency on the local and global contributions, respectively (see Table 6). We chose the OLS for comparability with the main studies in the literature (see Blackwell and McKee, 2003; Gallier et al., 2019), however, applying Tobit models provides consistent results (see Table B.4 in Appendix B). Control variables include socio-demographic information collected through Prolific (age, gender, income, socioeconomic status, education, employment status, and stu-

dent status) and a set of individual-specific characteristics about preferences (altruism, patience, risk, trust, negative and positive reciprocity) collected in the post-task questionnaire. We also include a variable to measure the response time in the task, a score variable for correct answers in the Cognitive Reflection Test, and a measure of the performance in three comprehension questions. Given that the task, although simple, entails computational difficulties, we include the individual comprehension score as a control variable, thus allowing for some degree of miscalculation.[5]

Table 6: OLS regressions examining the local and global contribution choices in the MLPGG.

	(1) Local contribution	(2) Global contribution	(3) Local contribution	(4) Global contribution
β	-3.245***	4.158***	-3.206***	4.189***
	(0.541)	(0.550)	(0.609)	(0.605)
Constant	5.147***	2.052***	2.961***	0.639
	(0.231)	(0.197)	(0.755)	(0.782)
Controls	✗	✗	✓	✓
Observations	638	638	525	525
R^2	0.052	0.079	0.105	0.147

Notes: Columns (1)-(2) show the results from regressions without controls. Columns (3)-(4) show the results from regressions that include control variables. Robust standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 6 shows that, on average, the effect of β is positive for the contribution to the global good (leveling-up effect) and negative for the contribution to the local good (substitution effect), thus leading to our first two results.

Result 1 (levelling up) *Contribution to the global public good on average increases as β increases.*

[5]When we consider sub-samples based on the number of correct answers, we find no difference in the results, except for the categorical crowding-in effect. See Table B.7 and Table B.8 in Appendix B.

Result 2 (substitution effect) *Contribution to the local public good on average decreases as β increases.*

Therefore, Result 1 and Result 2 indicate that individuals tend to substitute their contribution to the local public good with that to the global good as the relative efficiency of the latter increases. However, the robustness of these results might appear in contrast with the non-parametric tests on the differences in contribution between consecutive treatments presented above. To further investigate this potential limitation, we run an OLS analysis employing treatment dummy variables rather than regressor β. The results confirm that leveling up and substitution do not always occur between successive steps across our treatments. However, we need to consider that the differences in the relative efficiency between consecutive treatments are very small, potentially reducing their impact on changes in contributions. Indeed, when bigger jumps are considered – i.e. comparisons between non-consecutive treatments – the differences in contributions both to the local and global public good are always statistically significant. Therefore, we cannot, in principle, exclude that the lack of statistical significance is due to a lack of power.

To further analyze the relationships between Result 1 and Result 2, we check, within each treatment, which public good receives the higher average contribution. Figure 7 shows the difference between the average contribution to the local good and the average contribution to the global good by treatment (i.e., $\bar{c} - \bar{C}$). This difference is positive in T_1 and T_2, whereas in T_3 and T_4 it would appear to be negative. Indeed, in both T_1 and T_2 the average contribution to the local good is significantly higher than the average contribution to the global good (*Wilcoxon signed-rank* tests, $p < 0.001$ in both cases), while the opposite is true for T_3 and T_4, even though this difference is statistically significant only in T_4 (*Wilcoxon signed-rank* tests, $p = 0.1670$ for T_3; $p = 0.0007$ for T_4).
Therefore, we observe that subjects contribute more to the local public good unless the global one yields a higher total benefit. This analysis is connected to the debate concerning the interpretation of the treatment where the total benefits are equal (i.e., T_2). The existing literature finds the same positive difference as in our T_2, with the exception of Gallier et

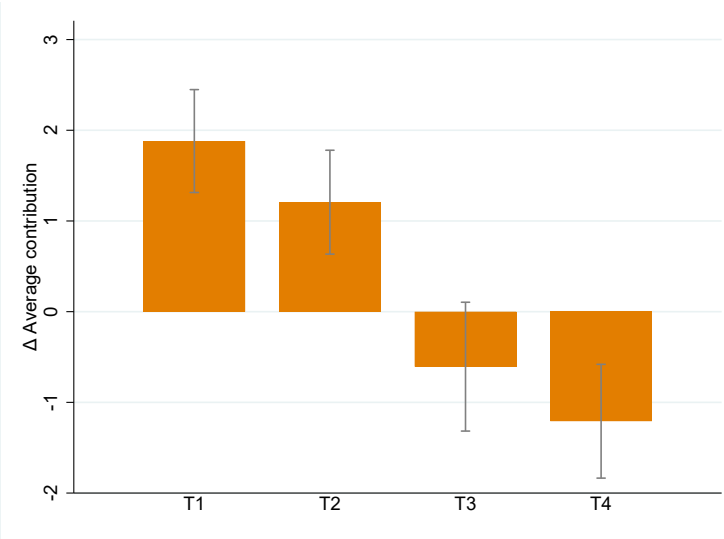

Figure 7: Difference between average contributions to the local and global goods per treatment. Confidence intervals at the 95% level.

al. (2019). Even if the interpretation of this result, which relies on strategic risk and *size effect* (as proposed by Chakravarty and Fonseca, 2017), seems more suitable in our case, our design does not allow us to exclude that, indeed, in-group bias plays a role. Finally, the trade-off between opportunity cost and potential returns in T_3 may explain why our result differs from the literature. Indeed, while Fellner and Lünser (2014) obtain an average contribution to the global public good that is significantly higher than the average contribution to the local public good, we find that this difference is not statistically different from zero.

3.3.2 Total contribution

Result 1 and Result 2, while questioning the possibility of an increase in the total contribution, cannot rule it out. Indeed, the presence of a reallocation of resources between the local and global public goods does not exclude the possibility of an overall increment in the total amount con-

tributed. To investigate this possibility, we again use regressor β, representing the MPCR of the global public good – which we impute to 0 for T_0 –, and estimate its impact on total contribution to test for the existence of a marginal crowding-in effect. Differently from the analysis in Table 6, however, we add a distinct regressor, G, to identify, if present, a categorical crowding-in effect. G is a dummy variable that is equal to 1 if there is a global public good (hence, for observations in T_1, T_2, T_3, and T_4) and 0 otherwise (hence, for observations in T_0).

Table 7 reports on the results of the regression on the total contribution of regressors G and β (Column 1), with the inclusion of control variables (Column 2). We can derive our third and fourth results from this analysis.

Result 3 (marginal crowding in) *There is no statistically significant evidence of a marginal crowding-in effect.*

Result 4 (categorical crowding in) *The introduction of an additional global public good produces per se a statistically significant increase in total contribution.*

Table 7: OLS regressions examining total contributions in the MLPGG.

	(1) Total contribution	(2) Total contribution
G	1.065***	1.160***
	(0.319)	(0.334)
β	0.914	0.745
	(0.594)	(0.642)
Constant	6.134***	2.168**
	(0.218)	(0.762)
Controls	✗	✓
Observations	802	658
R^2	0.051	0.164

Notes: Column (1) shows the results of the baseline specification. Column (2) shows the results of the regression that includes control variables. In T_0, we impute the value of 0 to β. Robust standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

While introducing an additional global good increases the overall level of contributions, the marginal increase in efficiency is completely ineffective in increasing total contribution. In fact, changes in relative efficiency have only redistributive effects and do not induce subjects to increase their overall contribution.

Our analysis of marginal crowding in does not consider the fact that the total efficiency of overall contribution varies across treatments. To provide further detail on total contribution, we compute an index of relative efficiency (REI) as the ratio between the actual generated public good per treatment and the maximum attainable level per treatment, that is:

$$REI = \frac{\bar{c} \cdot TB_c + \bar{C} \cdot TB_C}{10 \cdot \max\{TB_c, TB_C\}}, \tag{3.2}$$

where TB_c is the total benefit of the local public good, and TB_C is the total benefit of the global public good (see Table 4). Results are shown in Table 8.[6] By construction, the value of the index in T_0 and T_2 is equal to $1/10$ of the total contribution (as the total benefits cancel out), while it is lower for all other treatments (as $TB_c \neq TB_C$ and the contribution to both public goods is always positive). Therefore, for any given level of the total contribution, T_2 produces the highest relative efficiency because, in terms of efficiency, the two goods are perfect substitutes. As long as players contribute, it does not matter how they allocate their resources since there are no "wrong choices".

The sharp decline in the REI in T_2, T_3, and T_4 is caused specifically by the combination of a lack of marginal crowding in and the persistence of the contribution to the local public good. In other words, as the difference in total benefits between the global and the local public goods increases, subjects throw away the opportunity for a greater total benefit by keeping on contributing to the local good and, at the same time, by not increasing their total contribution. The same reasoning applies to the difference between T_0 and T_1. Subjects choose to partially contribute to the inefficient public good, thus obtaining a total benefit lower than

[6] Non-parametric tests show that these values differ significantly by treatment (*Kruskal–Wallis* test, $p < 0.001$; *Wilcoxon rank-sum* tests for pairwise comparisons, all p's < 0.001).

the maximum attainable level even though the total contribution in T_1 is higher than in T_0.

Table 8: Relative Efficiency Index (REI) per treatment.

	T_0	T_1	T_2	T_3	T_4
REI	61.3%	58.93%	75.0%	66.6%	59.7%

Finally, it is worth mentioning that the presence of the categorical crowding-in effect is the only result that does not hold in the restricted analysis where we select only those participants who perform well in the comprehension questions. Indeed, introducing a relatively inefficient additional public good does not produce a statistically significant increase in total contributions for this category of people, even though their contribution to both public goods remains significantly positive.

3.4 Discussion and conclusions

In an online multilevel public goods experiment, we investigated the effects of changing the MPCR of the global public good on contribution decisions. The general objective was to systematize the evidence and interpretations provided in the literature while adding, at the same time, new insights on some aspects which have either been neglected or not well-understood. In particular, we aimed to shed light on whether and to what extent increasing the MPCR of the global public good induces the leveling up of contribution to the global good and, if this was the case, whether and to what extent this effect is accompanied by a decrease of contribution to the local good – i.e., by the substitution effect – or by an increase in total contributions – i.e., by the marginal crowding-in effect. Moreover, by adding a control treatment where only the local public good is provided, we were able to measure the effect of the mere addition of a global public good *per se* – i.e., the categorical crowding-in effect.

Table 9 summarises the evidence collected in previous studies for each of the effects analyzed. We briefly discuss them in the summary

of our main findings.

Table 9: Summary of the main results in the MLPGG literature (including our study).

Paper	Leveling up	Substitution	Marginal crowd. in	Categorical crowd. in
Gallier et al. (2019)	✓	✓	✗	–
Chakravarty, Fonseca (2017)	✓	✓	✗	✓
Fellner, Lünser (2014)	✓	✓	✓	–
Blackwell and McKee (2003)	✓	✗	✓	–
Our Study	✓	✓	✗	✓

Notes: ✓= the effect is found, ✗= not found, – = not investigated.

The leveling-up effect is the most robust evidence in the literature, as it has been repeatedly replicated, including in the recent papers by Chakravarty and Fonseca (2017) and Gallier et al. (2019). We confirm this effect with our Result 1 and provide a generalization by extending the analysis to a series of efficiency increases of the global good, which allowed for an estimation of the average linear effect and many more pairwise comparisons than those usually referred to in the standard literature.

The evidence concerning the substitution effect and the marginal crowding-in effect is much more mixed. For Blackwell and McKee (2003) there is no substitution from the local to the global, but only an increase in the total contribution; Fellner and Lünser (2014) find that both the effects are jointly active following the rise in the productivity of the global good; only Chakravarty and Fonseca (2017) and Gallier et al. (2019) find that substitution cancels out any increase in total contribution. This latter finding is consistent with our Result 2 and Result 3 as we also observe that as the efficiency of the global public good increases, the leveling-up is financed out of a complete substitution of the contribution to the local public good, thus leaving total contribution unchanged. However, our design offers more robust evidence for both the decrease in the contribution to the local public good and the stability of total contribution, which

we test at several levels of relative and absolute efficiency. Notably, the decision to sterilize the group identity condition – usually manipulated in the standard multilevel design (one exception being the baseline condition in Gallier et al. (ibid.)) – might have contributed to clearing these results.

With Result 4, we confirm the findings of Todd L. Cherry and David L. Dickinson (2008) for the standard PGG, who show that adding the possibility to contribute to a larger number of public goods brings about a rise in total contribution. Moreover, we produce a new piece of evidence in the context of the MLPGG design. Differently from Chakravarty and Fonseca (2017) who add a local public good to a baseline condition with only a global good, we added a global good to the local good in the baseline. However, we do find the same positive effect on total contribution.

Finally, by looking at the within-treatments analyses, we also confirm several standard results in the literature and provide some novel insights. Firstly, the circumstance that subjects contribute more to the local good until the global good has a higher total benefit confirms a common finding in the MLPGG literature (Blackwell and McKee, 2003; Chakravarty and Fonseca, 2017; Fellner and Lünser, 2014). While we believe that in our anonymous and one-shot setting, this result is likely to be explained by the lower strategic uncertainty of the local public good, this interpretation cannot be considered the only plausible one unless a disentanglement of the individual propensity to reduce strategic risk (by opting for the public good where fewer players are involved) is implemented by design. Secondly, we focused on two treatments in which one of the public goods is financially dominated by the other. In the case of T_4, it is the local public good that is (weakly) dominated; in accordance with the literature, we find that subjects keep contributing to the local public good despite the lack of incentives. This suggests that allocation criteria other than individual payoff maximization are at stake. On the other hand, with our T_1, we provide a new test of a condition where the global public good is both riskier and less productive. In this case, the positive and significant contribution to the global public good cannot be justified on the grounds of individual payoff and depends on a specific

willingness to contribute to a public good that benefits all players, such as, for example, fairness concerns or inequity aversion. In this sense, it is worth noting that the contribution to the global public good in T_1 and the local public good in T_4 remain significantly positive even for the restricted sample of those with a relatively higher comprehension of the task.

However, the explanation of this kind of decision in terms of some preferences that do not respond to individual or group utility maximization is beyond the scope of our design and is left for further research. Likewise, additional investigation of the motivations that explain the stability of total contribution is required. It might be the case that a heuristic imposing a stable diversification between one's private account and the total contribution is at stake. However, the validity and robustness of this hypothesis require testing with a dedicated design (e.g., by comparing T_0 with a multilevel setting where more than one public good is added). Moreover, this hypothesis does not apply to the leveling up and substitution effects since, under such invariant automatic heuristics, changes in efficiency could not affect contribution decisions.

One limitation of this study pertains to the implementation of a one-shot game instead of a repeated one. This decision was deliberately made to minimize the opportunity for individuals within local groups to learn and strategically influence outcomes across multiple rounds. Moreover, employing a one-shot game allowed for not limiting the potential to use individual data as independent observations. However, it is worth considering potential follow-up studies involving repeated interactions to establish a stronger connection between our findings and real-world phenomena. In reality, individuals often encounter a trade-off between making local or global contributions, as exemplified in the environmental domain. Expanding the application of the MLPGG framework could involve analyzing the cooperation of different countries within international climate agreements or exploring environmental conservation and protection initiatives (e.g., forests, water, biodiversity) where stakeholders at various levels—such as local communities, government agencies, and international organizations—are involved.

Chapter 4

A cross-country experiment on the multilevel public goods game

This essay is based on the paper: M. Catola, P. Guarnieri, V. Pizziol, C. Rapallini "Measuring the attitude towards a European public budget: A cross-country experiment." Dipartimento di Economia e Management (DEM), University of Pisa, Pisa, Italy, 2023.

4.1 Introduction

The pandemic crisis and the war in Ukraine represent an unprecedented challenge for the European Union (EU) towards greater cohesion of policies, particularly of political economies, to counterbalance unfavorable shocks. Up to March 2020, the European fiscal policy was guaranteed – with doubtful success – by fiscal rules (i.e., the Stability and Growth Pact), while the European budget was not used as a fiscal policy instrument (e.g., Caselli and Wingender, 2021; De Grauwe and Ji, 2019). Despite being improperly referred to as its own resources, Member States' contributions have always been the source of revenue for the European budget, and European-level taxes have not been directly levied on cit-

izens (M. Bordignon and Scabrosetti, 2016). At present, the debate is focused on the need to revise the Stability and Growth Pact, which is *de facto* suspended, without modifying the European Treaties because of the long and politically challenging process that the latter would require (e.g., Blanchard et al., 2021; Maduro et al., 2021). There is sizable support for the view that the new fiscal constraints must be flanked with a European fiscal capacity (i.e., common resources) that should be activated in specific contingencies or for the realization of common projects that are exceptional in nature (e.g., in the energy sector (Romanelli et al., 2022)). At the same time, there is a slight possibility of a reform allowing European institutions the power to tax, given that this would require support from the European Parliament, Member States, and European citizens.

In this regard, assessing the attitude of European citizens towards a direct contribution to the European budget appears relevant. However, this assessment is difficult mainly because the acceptability of a fiscal policy depends on the perceived return that subjects expect from the use of the revenues (Maestre-Andrés, Drews, and Bergh, 2019; Maestre-Andrés, Drews, Savin, et al., 2021), which is currently uncertain from the perspective of EU citizens. Indeed, no Eurobarometer survey – the standard tool used by the European Commission to assess the attitudes of citizens towards EU institutions and policies – has directly addressed this acceptability issue, nor has it been discussed in the micro or behavioral literature. To fill this gap, we perform an incentivized online experiment to measure EU citizens' willingness to contribute to the European public budget. We frame it as a public good provision problem to capture the impact of perspective returns on this propensity. Specifically, we use a multilevel public goods game (MLPGG) that makes experimental subjects face a trade-off between contributing to a European public budget or to a national public budget.

In the MLPGG, subjects are assigned to a local group and asked how much of their private endowment they would like to contribute to the public good of their local group or to the public good of a global group that contains other local groups in addition to their local group (Blackwell and McKee, 2003; Buchan, Brewer, et al., 2011; Buchan, Grimalda,

et al., 2009; Chakravarty and Fonseca, 2017; Fellner and Lünser, 2014; Gallier et al., 2019). In our experiment, the decision is framed as one concerning the alternative between contributing to the (local) national public budget or to the (global) European public budget. We assign subjects to local groups based on their country of residence. We selected six EU Member States (Italy, Germany, France, The Netherlands, Poland, and Portugal) following a geopolitical criterion representative of the different positions in terms of macroeconomic policies within the EU.

The MLPGG allows us to investigate two main effects: the first is connected to group identity[1], while the second is the impact on contribution decisions of the relative efficiency of the local and global public goods. Regarding the former, the literature highlights that when group identity is primed in the local groups, it drives some degree of in-group favoritism that motivates contributions to the local group. Priming group identity is attained through different kinds of manipulations but typically involves the way in which the local groups are formed and the minimal identity approach (Blackwell and McKee, 2003; Chakravarty and Fonseca, 2017; Fellner and Lünser, 2014; Gallier et al., 2019). In this study, we prime group identity by revealing to the subjects that they are assigned to local groups composed only of individuals residing in the same country as they do. Accordingly, a stronger sense of belonging to the local group (with respect to that activated by minimal identity manipulation) could be driven by the actual different citizenship of subjects and the related cultural, institutional, and political differences.

Regarding the relative efficiency of the local and the global public good, the standard treatments in the MLPGG experiments investigate to what extent increasing the marginal per-capita return (MPCR) of the global public good (while keeping the MPCR of the local public good constant) a) increases the contribution to the global public good (*leveling up* effect) b) decreases contribution to the local public good (*substitution* effect) and/or increases the total contribution (the sum of the contributions to the local and the global public goods) (*marginal crowding in*). While the positive effect of increased efficiency on willingness to con-

[1]For a review on the group identity literature, see for instance Bronchal, 2023.

tribute is an established result in the standard PGG (Chaudhuri, 2011; Ledyard, 1995), this effect is more controversial in the case of a strategically more complex game such as the MLPGG. Indeed, the results offered by the literature are mixed and sensitive to the magnitude of efficiency changes (see Catola, D'Alessandro, et al., 2023, for a detailed discussion on the differences in results and experimental designs). Following this line of research, we set up two different treatments to measure whether there are differences in how citizens from the selected EU countries respond to an increase in the efficiency of a European public budget. This change in efficiency can be thought of as the return to citizens of a public expenditure potentially funded by the European public budget, especially in those sectors that address transnational challenges such as the environment and the energy policy, defense, and public health in the face of pandemic events. In this sense, investigating decision-making in the context of the MLPGG suggests useful insights to improve the ability of European institutions to overcome particularism and guarantee cohesion by sustaining citizens' welfare as a return of a direct European fiscal policy.

Using a public good design to investigate support for institutions is not new in the experimental literature (Barrett and Dannenberg, 2017; Battaglini et al., 2020; Gallier, 2020), but to the best of our knowledge, no study has addressed propensities towards strengthening the European budget by means of direct taxation. A tax game that is usually applied to identify drivers of compliance/evasion to a given tax (Bazart and Bonein, 2014; Coricelli et al., 2010; Górecki and Letki, 2021; Spicer and L. A. Becker, 1980; Spicer and Hero, 1985) does not seem suited to our purpose since it deals with the response to exogenously imposed fiscal pressure and not with voluntary (economic) contribution to an institution that is new and holds a spending power in return. In contrast, the MLPGG design links the propensity to contribute to a public institution to the sense of belonging to it in addition to its efficiency in distributing returns. In this regard, this study is closely related to Buchan, Brewer, et al., 2011; Buchan, Grimalda, et al., 2009, who use the MLPGG to study the effects of globalization on the willingness to contribute to national

versus international public goods and to Gallier et al., 2019, who assess the willingness to pay for local and regional public goods among German living in two different regions. However, two main features distinguish our design from these studies. First, national identity is not only used to prime group identity in local groups but to frame the whole decision context since it relates to a potential sense of belonging to European society. Second, by framing the decision as an alternative between two different public budgets, subjects are confronted with two labels that may represent the actual institutions to which they act as citizens, thus adding realism to the decision at stake. In the same realistic vein, after the experimental task, subjects completed a questionnaire aimed at collecting information to test if the most recent crisis calling for an EU response affects EU citizens' propensity as measured in our MLPGG.

4.2 Methods

4.2.1 Experimental design

In the main task of our experiment, we ask participants to play a one-shot linear MLPGG. This game is characterized by a nested structure where two or more local groups are part of a higher-level global group. Figure 8 depicts the specific configuration we employ in our experiment.

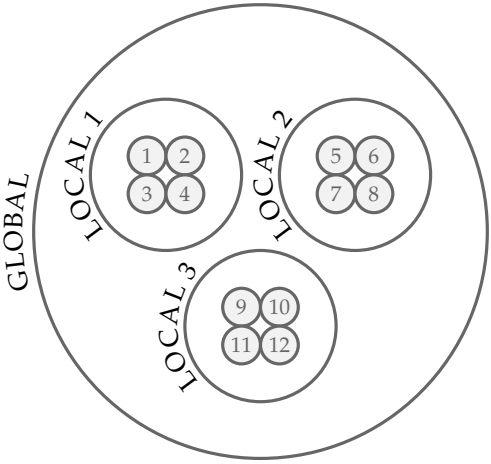

Figure 8: Configuration of the MLPGG in this study.

Participants are randomly matched in local groups of $M = 4$ and, at the same time, in global groups of $N = 12$. Thus, each of the global groups is composed of 3 local groups. Each individual i receives an endowment e_i, which he/she can keep for herself in the private account, contribute to the local public good provided at the local-group level, or contribute to the global public good provided at the global-group level. We set each endowment e_i equal to 10 points. Any amount c_i contributed to the local public good is multiplied by a local-specific factor and divided equally among the 4 local group members. We refer to this ratio as α, the local MPCR. Any amount C_i contributed to the global public good is multiplied by a global-specific factor and divided equally among the 12 global group members. We refer to this ratio as β, the global MPCR.[2]

Given the game structure, the payoff that each player i receives by playing the game is equal to:

$$\pi_i = e_i - c_i - C_i + \alpha \sum_{j=1}^{M} c_j + \beta \sum_{k=1}^{N} C_k. \tag{4.1}$$

[2]It is worth noting that $(1 - \alpha)$ and $(1 - \beta)$ then represent the actual costs that player i incurs by contributing 1 point to the local and to the global public goods, respectively.

In our experiment, we set $\alpha = 0.6$, while the value of β is treatment specific:

(i) in treatment Low, we set $\beta = \alpha/3 = 0.2$;

(ii) in treatment $High$, we set $\beta = \alpha = 0.6$.

These two treatments are most commonly used in the literature to investigate whether and to what extent participants react to variations in the relative efficiency of the two nested public goods. We measure efficiency in terms of total benefit (TB), which, following Gallier et al., 2019, is defined as the individual earnings obtained from a public good when every group members make a 1-point contribution to it (i.e., αM and βN, respectively).

Table 10 provides a full summary of the relevant parameters for each treatment.

Table 10: Summary of the treatments parameters.

Treatment	**Local PG**			**Global PG**		
	M	α	TB	N	β	TB
Low	4	0.6	2.4	12	0.2	2.4
High	4	0.6	2.4	12	0.6	7.2

In the Low treatment, the TBs of the two nested goods are equalized ($\alpha M = \beta N$), thus sterilizing efficiency effects due to scale. Indeed, the local public good is less costly and less risky than the global one since the individual return from 1 point contributing to it is higher than the return of 1 point contributing to the global public good. Thus, in the Low treatment, players have only a weak incentive to contribute to the global public good.

The $High$ treatment corresponds to the case where the MPCRs of the two goods are equal, i.e., $\alpha = \beta$. Here, the two public goods are equally costly, but the global public good is more efficient due to scale effects. This, in turn, means that for each player i, the two goods are equally risky, as the return from contributing is the same in both cases. Additionally, while in the Low treatment, the local group members are better

off if their fellow member i contributes to the local public good rather than to the global one ($\alpha > \beta$), this is not the case for $High$ ($\alpha = \beta$). Hence, contributing to the local public good in $High$ is neither less costly for contributors nor does it provide higher payoffs for their fellow local group members. Thus, the only monetary difference due to contributing to the local public good in $High$ vs contributing to it in Low, is that of excluding the members of the other two local groups from the benefits of the public good provision.

In conclusion, the implementation of these two treatments provides a straightforward way to test the impact of efficiency on contribution decisions as, from a game-theoretical point of view, in each treatment, one good is better than the other (the local good is better than the global in the *Low* treatment, and *viceversa* in the *High* treatment) given that any strategical trade-off is sterilized.

4.2.2 Groups formation

To address our main research questions, we rely on a between-subjects design to expose each subject from each country to only one of the two efficiency-related treatments. Each participant is randomly matched with other 3 participants of the same country of residence to form local groups and also with 8 other participants from two other local groups, each composed of residents from one of the other 5 EU countries, to form the global group. Therefore, each global group is formed by 3 local groups, each being homogeneous in terms of the country of residence.

Participants are informed about the matching protocol; thus, they are aware that their group was homogeneous with respect to the country of residence and that the other groups were formed of participants from other countries. However, participants do not have any other information about the specific countries involved other than that they also belong to the EU.

We opted to frame the experiment to enhance the connection to the real world, help understand the environment, and reduce confusion (Alekseev et al., 2017). The public goods of the MLPGG were referred to as the

participants' Country Public Budget and the EU Public Budget. Therefore, the combination of the information provided to players and the framing of the task allows us to capture the willingness of players to contribute to either a group of their fellow citizens or three groups of generic EU citizens.

For the sample selection, while in principle it could have been possible to recruit participants from each country in the EU, for most EU countries, there was a limited sample of registered subjects on the platform that we used to run the experiment. Therefore, (as in Buchan, Brewer, et al., 2011; Buchan, Grimalda, et al., 2009) we rely on a sample of countries that was selected by combining the availability of subjects on the platform with a geopolitical criterion. We include Italy, Germany, and France since they are all founding countries and the three largest economies in the EU. Moreover, they represent different positions in terms of macroeconomic policies within the EU. The Netherlands is one of the so-called *Frugal Four*, a block of northern countries, including also Denmark, Sweden, Austria, and, lately, Finland (historically, the strongest advocates for austerity programs within the EU). Poland is a member of the *Visegrád Group*, a group of 4 countries in Eastern Europe (Czech Republic, Hungary, Poland, and Slovakia) that joined the EU in 2004 and have disagreed with other EU countries on several topics in the last decade. Finally, Portugal is one of the so-called *PIGS*, a group of Southern European countries characterized by high public debt that has come under strong economic and political pressure since the 2008 economic crisis. In terms of governmental structure, it is worth mentioning that Germany is the only federal state in our sample, which is a rather rare case within the EU (the only other cases being Austria, Belgium, and, to a certain degree, Spain). Finally, France, Germany, Italy, and The Netherlands are net contributors to the EU budget, while Poland and Portugal are net receivers.

Concerning this selection criterion, it is worth noting that we rely on governmental positions (at least up to March 2020), even if we investigate the willingness to contribute to the European budget with potential "own resources" and not to a "derivative budget" financed with member

states' contributions as it is currently. In other words, our experimental design set up a framework similar to that advanced by the fiscal federalism literature (Ambrosanio and Massimo Bordignon, 2015), according to which a political body has its own resources if these revenues are levied directly from taxpayers and accrue directly to the budget of the entity, without being determined by decisions taken by some other political bodies. Different from "tax shares" own resources are also usually accompanied by some autonomy (e.g., the possibility of varying the tax rate), although not necessarily by the right to impose the tax or to determine its characteristics.

At the same time, the criterion of the governmental position appeared the most appropriate for framing our decision problem. If we consider that contributions to public budgets are likely to be affected by evaluations about how to spend those budgets, political opinions about fiscal policies and public investments in the EU were, in principle, expected to correlate with decisions in our sample. However, the reliability of this criterion rests on the assumption that governments' positions are representative of the population's opinions in the selected countries. This holds only under the assumption that voting systems can ensure effective and updated political representation in modern democracies. Despite its limitations, this assumption appeared valid for the purposes of our study.

4.2.3 The post-experimental questionnaire

The post-experimental questionnaire includes three sets of questions to assess if the participant has an immigrant background, his/her feelings of belonging to the country of residence, to Europe, and his/her (positive or negative) feelings toward the EU, as well as whether these feelings changed following the most recent dramatic events, e.g., the COVID-19 pandemic crisis and the war in Ukraine. Based on the answers to these questions, we define the control variables of our estimation strategy. The numbered list of questions is available in section C.2 of Appendix C. All the answers are on a 5-point scale unless otherwise specified.

The questionnaire begins with three preliminary questions to assess the possible immigration background of participants. First, we ask about the participant's birth country (Q1) to verify if he/she is a first-generation immigrant. Participants born in the country of residence are considered not to have an immigration background, even if they can be second-generation immigrants. Then, we ask first-generation immigrants how old they were when they moved to the country of residence (Q2) to control for the recency of their immigration. Finally, we ask about the country of birth of the participant's parents (Q3 and Q4) to control the parents belonging to an EU country. In sum, our working hypothesis is that participants' decisions to contribute to the national and EU budgets can be altered by having recently immigrated to an EU country. To assess their feelings towards the country of residence and towards Europe, we ask participants how strongly they identify themselves with the country (e.g., how strongly they feel Italian if Italy is the country of residence) and how strongly they feel they are an EU citizen (Q5 and Q6, respectively). Then, we ask for a personal judgment on the EU image (Q7). For the COVID-19 questions, we take inspiration from one of the multinational surveys delving into European citizens' attitudes and opinions over the course of the crisis commissioned by the European Parliament and conducted at the end of April 2020 (European Parliament, 2020). We ask participants' opinions about the benefit for their country of being part of the EU before the pandemic (Q8), if they are satisfied with the solidarity between the EU member states in fighting the pandemic (Q9), and if their opinion about the benefits of being part of the EU changed after the pandemic (Q10). Concerning the war in Ukraine, the main aim is to control participants' propensity to contribute to national and EU defense and whether the war has affected this. National defense is one of the clearest examples of a public good, and common defense has always been one of the open issues in the European agenda since its foundation in the 1950s. However, it is not granted that every individual looks favorably upon national defense expenditures, as someone may think that not having an army and being neutral makes the country safer than otherwise having an army. To control for this attitude, we first ask participants to assess,

on a scale from 0 to 10, how much they agree that higher military spending increases the level of safety (Q11). Then, we ask whether, after the beginning of the war, they were in favor of higher military expenses in their country (Q12) and whether they were in favor of financing a European army before the beginning of the war (Q13) and after the beginning of the war (Q14).

4.2.4 Procedures

The experiment, which was preregistered (AsPredicted number: #89021) and approved by the Ethical Committee of the University of Florence (Italy), was programmed in oTree (Chen et al., 2016) and conducted online between the 19th and 20th May 2022. The participants were recruited from the EU adult population of the six selected countries through the Prolific platform (Palan and Schitter, 2018). An overall sample of 1,200 subjects living in the EU (i.e., 600 participants per efficiency treatment, equally distributed between the selected countries) was recruited to participate in the experiment. Recruitment was based on the country of residence rather than the country of nationality. We considered this criterion more representative of the individual sense of citizenship since civil rights, such as the right to vote and to stand as a candidate in elections to the European Parliament (Article 22(1) TFEU, 2008), are given to residents of the Member State. The sample size was determined by an *a-priori* power analysis expecting a small effect size (Cohen's $d = 0.35$) with alpha=0.05 and power 0.80 for a two-tailed t-test for a between-subjects design.

Before starting the experiment, subjects were asked to confirm their current country of residence.[3] Then, participants had the opportunity to choose whether to complete the experiment in English or switch to their national language. Before facing the task, subjects had to answer some control questions to test their comprehension of the decision at stake. The experiment did not start until the participants had answered all the

[3]Out of the 1203 participants joining the study on Prolific, 3 declared not to live anymore in the country of residence for which they were recruited. We granted them a fixed participation fee without making them proceed with the experiment.

questions correctly.

The payoffs were expressed in points that were converted to pounds at the rate of 1 point = £ 0.025 at the end of the experiment. Over all the treatments, mean earnings amounted to £ 1.53 (including a £ 0.50 fixed participation fee), and the experiment took, on average, 7 minutes to complete. The average earnings in the experiment corresponded to a £ 13 hourly compensation, and thus, they were perfectly in line with the salary of a student assistant in the EU (namely, approximately € 15). Additionally, by keeping the game monetary reward much greater than the fixed participation fee, we ensured that the payoffs of the task were salient.

4.3 Sample characteristics

4.3.1 Demographics

Table 11 reports, separately for each efficiency treatment, summary statistics of demographic characteristics of our sample. The last column reports p-values from either Kruskal–Wallis tests for continuous variables or Fisher's exact tests for dummy variables.

Overall, the average age is approximately 29 years old, there is an almost perfect split between females and males, and 16.50% of participants were not born in the same country where they currently reside. Approximately 47% are students. Our sample is, on average, well-educated: 33.91% hold a high school diploma (or equivalent), 25.58% an undergraduate degree, and 35.33% (at least) a graduate degree. Based on the participants' self-reported measure, our sample is, on average, in a middle socioeconomic status in all treatments. Finally, it is clear that, on average, our sample is younger, better educated, and has a higher share of students than the average population in each country. While this could represent a limitation for the representativeness of our results, it is also worth mentioning that this sample is more diverse than the samples usually employed in laboratory experiments, which is one of the advantages of running an online experiment.

Table 11: Means (and standard deviations) of participants' characteristics per treatment.

	Low	**High**	**p-value**
Age	28.60 (8.99)	28.39 (8.61)	0.606
Female	0.51 (0.50)	0.49 (0.50)	0.729
Student	0.45 (0.50)	0.49 (0.50)	0.183
Socioeconomic status	5.55 (1.52)	5.56 (1.46)	0.883
Secondary education	0.33 (0.47)	0.35 (0.48)	0.428
Undergraduate degree	0.26 (0.44)	0.25 (0.43)	0.791
Graduate and Post-graduate	0.36 (0.48)	0.35 (0.48)	0.763
Migrant	0.16 (0.36)	0.17 (0.38)	0.485
Observations	604	596	

Notes: Age is the age of the participant at the time of the study. *Female* is a dummy variable that equals 1 if the participant is female. *Student* is a dummy variable that equals 1 if the participant is a student. *Migrant* is a dummy variable that equals 1 if the participant was not born in the country of residence. *Socioeconomic status*measures the self-reported place occupied by the participant on a ladder representing the society that goes from 1 to 10. *Secondary education* is a dummy variable that equals 1 if the participant holds a high school diploma or equivalent. *Undergraduate degree education* is a dummy variable that equals 1 if the participant holds an undergraduate degree. *Graduate and post-graduate* is a dummy variable that equals 1 if the participant holds a graduate or doctorate degree. *p-values* refer to Kruskal–Wallis tests for continuous variables and Fisher's exact tests for dummy variables.

Table 12: Means (and standard deviations) of participants' characteristics per country.

	IT	DE	FR	NL	PL	PT	p-value
Age	28.91 (8.93)	29.9 (9.35)	29.93 (9.66)	27.86 (7.47)	26.49 (8.42)	27.91 (8.41)	0.001
Female	0.50 (0.50)	0.50 (0.50)	0.50 (0.50)	0.50 (0.50)	0.51 (0.50)	0.50 (0.50)	1.000
Student	0.50 (0.50)	0.47 (0.50)	0.35 (0.48)	0.44 (0.50)	0.56 (0.50)	0.51 (0.50)	0.001
Socioec. status	5.73 (1.44)	5.61 (1.52)	5.49 (1.51)	5.75 (1.68)	5.24 (1.43)	5.51 (1.30)	0.003
Secondary educ.	0.45 (0.50)	0.36 (0.48)	0.17 (0.38)	0.33 (0.47)	0.47 (0.50)	0.26 (0.44)	0.001
Undergrad. degree	0.20 (0.40)	0.27 (0.45)	0.17 (0.38)	0.38 (0.49)	0.23 (0.43)	0.28 (0.45)	0.001
Grad. and Post-grad.	0.32 (0.47)	0.29 (0.45)	0.61 (0.49)	0.27 (0.45)	0.21 (0.41)	0.42 (0.49)	0.001
Migrant	0.07 (0.25)	0.29 (0.45)	0.27 (0.44)	0.30 (0.46)	0.01 (0.07)	0.06 (0.25)	0.001
Observations	200	200	200	200	200	200	

Notes: p-values refer to Kruskal–Wallis tests for continuous variables and Fisher's exact tests for dummy variables.

While descriptive statistics do not present statistically significant differences when comparing treatments, this is not the case when we compare countries. This is not surprising given that there are actual socio-demographic differences across our selected countries. Moreover, it is not possible to recruit stratified samples through Prolific, but we were able to at least impose balanced samples with respect to gender. Table 12 presents the descriptive statistics divided by country in the same fashion as Table 11.

It is interesting to note that participants from Germany and France have a higher average age, but for France, this is explained by a sample with a relatively small share of students and a substantially higher share of highly educated participants (approximately 78% of participants hold a university degree, with a remarkable 61% holding masters degree or

higher). It is also worth mentioning how the distribution of immigrants in the sample is largely uneven. First-generation immigrants comprise one-third of the samples of Germany, France, and The Netherlands, but comprise a fairly small share of the samples of Italy, Portugal, and especially Poland.

Furthermore, we control whether the randomization in the treatment allocation worked well within countries. Our tests reject the hypothesis of any statistically significant differences between demographics in the treatment subsamples for each country (results of the tests can be found in Table C.3)

4.3.2 The post-experimental questionnaire

We now turn to the answers collected through the post-experimental questionnaire. The following figures present the average answers to each question by country (descriptive statistics by country and the statistical tests can be found in Table C.4, Table C.5, Table C.6 in Appendix C).

Figure 9 depicts the average answers to the questions assessing feelings towards own country and the EU.

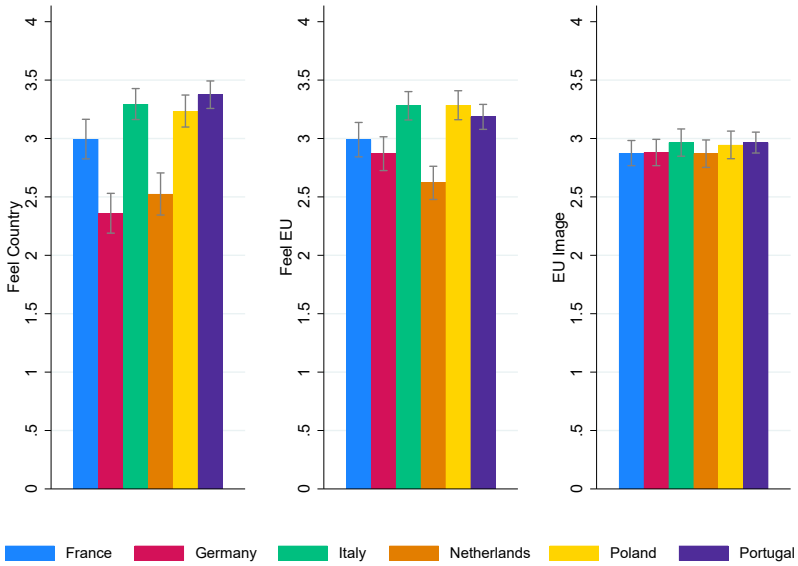

Figure 9: Mean answers to Feel questions by country. Confidence intervals at the 95% level.

KW tests for *Feel EU* and *Feel Country* find significant differences across countries, while no differences are found for *Image EU*. The pairwise comparisons between each country show that the differences in *Feeling EU* are driven by weaker feelings of belonging to the EU among Dutch residents compared to all others, except for Germany, whose citizens also show a weaker feeling of belonging to the EU compared to Italy and Poland. Similarly, for *Feeling Country*, German and Dutch residents show a weaker feeling of belonging to their own countries compared to all others.

Moreover, as shown in Figure 10, countries display significant differences in the answers to the COVID-related questions. More specifically, Polish residents feel that their country has benefited from being a member of the EU more than the French, German, Dutch, and Portuguese residents, and the Dutch and French residents also reported lower benefits

compared to Portuguese and Italian residents. Additionally, Italians and Portuguese display higher levels of satisfaction regarding the solidarity between the EU Member States in fighting COVID-19 compared to the Dutch and Germans, and Portuguese also compared to the French and the Polish. These answers reflect the type of event at stake. The COVID-19 pandemic has been a huge symmetric exogenous shock for the euro area and the world, but with asymmetric impacts across countries both because of the timing of the spread of the virus and of the differences in underlying economic structures. Accordingly, starting in 2020, the European Commission adopted measures to support national economies (i.e., SURE and NGEU) that are differentiated across countries. Italy was the first country to experience the pandemic, which resulted in a highly severe impact in terms of lives, and thus was one of the first recipients of European support.[4]

Finally, Figure 11 plots the average answers to the questions concerning the war in Ukraine. We do find some cross-country variability in the answers to the questions. Particularly, Italian and German residents are less convinced that increasing public expenditures on national defense makes them safer than Polish and Dutch residents, and for Italians, this also holds in comparison with Portuguese residents. The Polish also hold a stronger positive belief about military spending compared to the French. For the *National Army*, the Polish agree that their country should increase its public expenditures on the national army after the war's outbreak, more than any other country in our sample. Italians show the lowest level of adherence to that statement compared to all other countries, except for the French (whose answers to this question are not significantly different from those of the Italians). Much less variation emerges when looking at the answers to the two questions on an EU army, with Germany displaying the lowest levels of agreement to the necessity of an EU army financed by the EU budget, both before and after the Russian-Ukrainian war.

[4]In 2021, Italy received slightly less than one-third of the entire SURE funding, while the second recipient is Spain, which received almost one-fourth. For the NGEU program, Italy is expected to receive the equivalent of 11 percent of its GDP, while France and Germany will receive the equivalent of 1.5 and 1 percent of GDP, respectively.

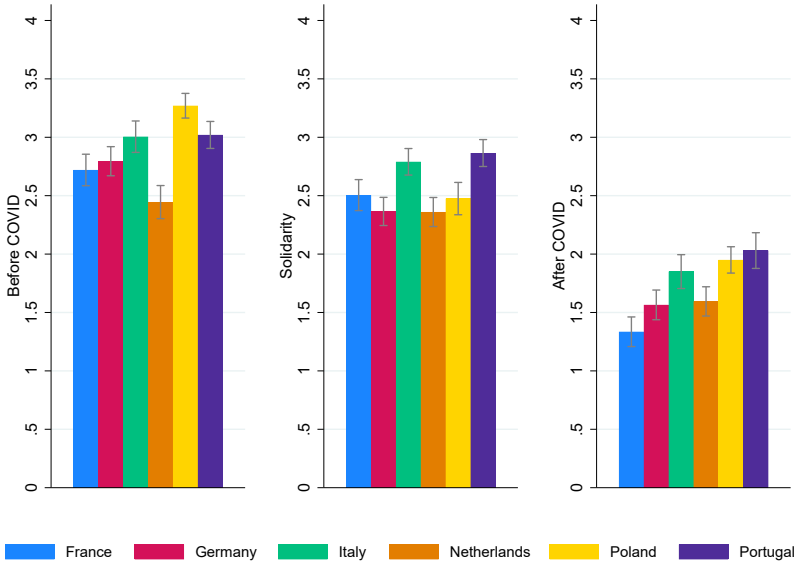

Figure 10: Mean answers to COVID-19 questions by country. Confidence intervals at the 95% level.

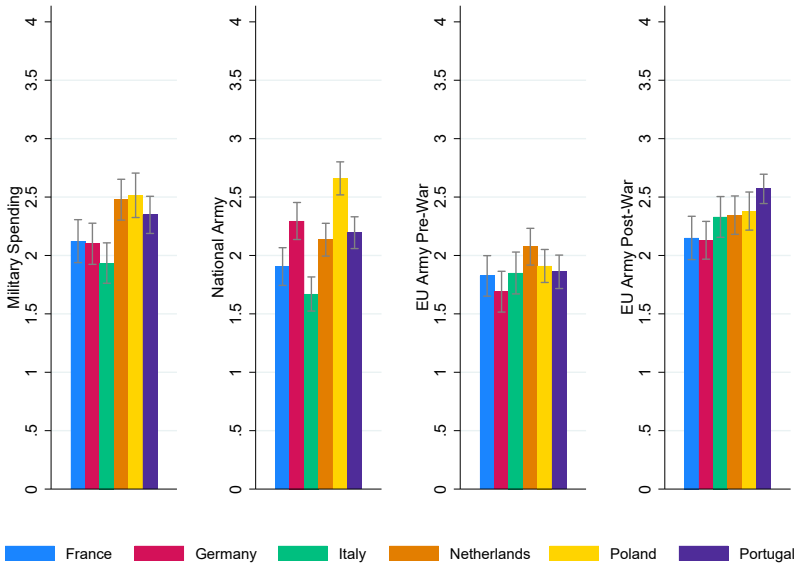

Figure 11: Mean answers to war questions by country. The Military Spending question is standardized to vary between 0 and 5 for graphical comparability. Confidence intervals at the 95% level.

4.4 Results

In this section, we present our results. We first display some descriptive and nonparametric analyses of the contributing behavior in all countries. We then investigate the presence of efficiency-related effects by making use of regressions, which allow us to control for heterogeneity in participants' demographic characteristics and individual preferences and beliefs. Finally, as an additional analysis, we restrict the studied sample to only those subjects displaying a migration background.

4.4.1 Contributing behavior across countries

Mean contributions to the Country Budget are 37.30% of the initial endowment (41.90% in the Low treatment and 32.50% in the High treatment), and mean contributions to the EU Budget are 38.50% of the initial endowment (32.40% in the Low treatment, and 44.70% in the High treatment). The first noteworthy fact documented is, hence, that, over all countries, the mean total contribution (i.e., the sum of contributions to the Country and EU Budgets) is, out of 10 points, approximately 7.43 in the Low treatment and 7.72 in the High treatment. This finding shows that contribution levels are higher compared to other most recent online one-shot PGGs that report contributions amounting to 60% of the initial endowment (Berg et al., 2020; Bilancini, Boncinelli, Nardi, et al., 2023; Catola, D'Alessandro, et al., 2021; Isler, Gächter, et al., 2021), but are in line with recent one-shot MLPGGs where average total contributions in the game are approximately 75% of the endowment (Catola, D'Alessandro, et al., 2023; Gallier et al., 2019).

Although this cross-study comparison can only be qualitative in its nature, it can suggest that the mere addition of a global public good (in our case, the EU one) compared to a situation where only a local one is provided (in our case, the country one) can positively impact total contributions (*categorical crowding-in effect*). This evidence aligns with that found by Todd L. Cherry and David L. Dickinson, 2008, which shows that adding the possibility to contribute to a larger number of public goods results in greater total contributions, and by Chakravarty and Fon-

seca, 2017 and Catola, D'Alessandro, et al., 2023, which obtain the same result in an MLPGG context.

In Figure 12, Figure 13, Figure 14, we provide mean contributions by country and treatment for each of the three variables of interest.[5] We test whether the decisions in the MLPGG from different countries come from the same distribution in both efficiency treatments. In the *High* treatment, KW tests do not reject the null hypothesis that contributions to the Country Budget (χ^2=8.959, p=0.1107), contributions to the EU Budget (χ^2=3.624, p=0.6047), and the Total budget (χ^2=3.910, p=0.5624, respectively) come from the same distribution for all the countries considered. This holds for contributions to the EU Budget (χ^2=1.334, p=0.9314) and Total contribution (χ^2=7.576, p=0.1812) also in the *Low* treatment, while in this condition the only statistically significant difference appears in contributions to the Country Budget (χ^2=11.433, p=0.0434). To further investigate this evidence, we run a set of pairwise comparisons using Wilcoxon rank-sum tests. They indicate that this result is driven by lower contributions performed by German participants to their Country Budget compared to the others. However, after applying Bonferroni corrections, no difference remained statistically significant. This analysis leads to our first result.

Result 1 *Contributions to the Country and EU Budgets, and Total Contribution, at each efficiency level, are not significantly different across countries.*

4.4.2 Efficiency-related effects

We now turn to investigating the efficiency-related effects. Looking at the mean values, it appears that total contributions do not vary between the *Low* and the *High* treatment, suggesting the *marginal crowing-in effect* is not at stake. On the other hand, the average contributions to the EU budget in each country seem relatively higher in the *High* treatment compared to *Low* while contributions to the country budget seem to decrease when switching from *Low* to *High*. This reading allows for hypothesizing

[5]Related details about exact mean values and standard deviations can be found in Table C.1 and Table C.2 of Appendix C.

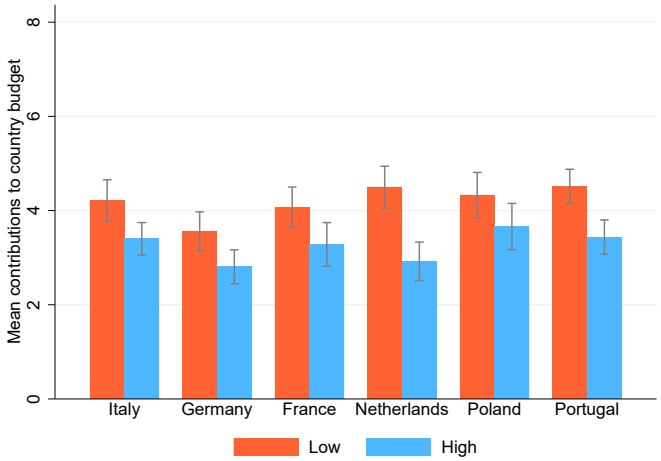

Figure 12: Mean contributions to the Country Budget by country and treatment. Confidence intervals at the 95% level.

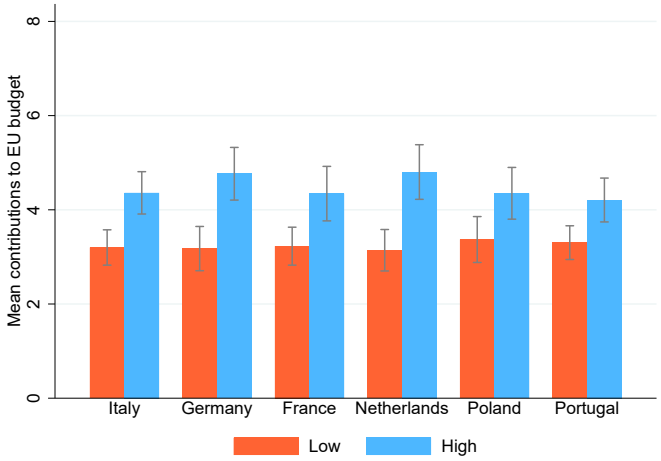

Figure 13: Mean contributions to the EU Budget by country and treatment. Confidence intervals at the 95% level.

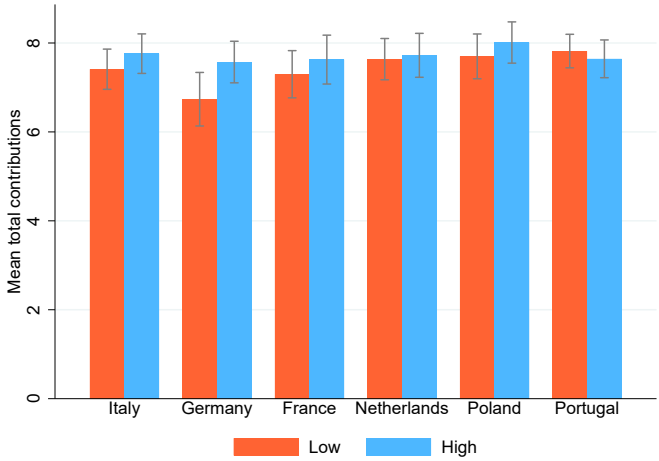

Figure 14: Mean Total contribution by country and treatment. Confidence intervals at the 95% level.

the presence of both *leveling-up* and *substitution* effects while ruling out the marginal crowding-in effect. We test these hypotheses through OLS regressions.

With the regressions displayed in Table 13, we aim to estimate the impact of the efficiency manipulation on the contribution to the Country Budget, the EU Budget, and the Total contribution. Accordingly, our main independent variable is the dummy variable High, which is equal to 1 if the observation is from the *High* treatment, and 0 otherwise. We also include country dummies to control for country-fixed effects, as well as their interactions with the treatment dummy (Columns 1-3). Additionally, we include demographics and answers to the post-experimental questionnaire as control variables in Columns 4-6.

The positive and significant coefficient of High in (2) and (5) indicates that there is robust evidence of a leveling-up effect. Indeed, subjects are responsive to efficiency concerns since their contribution to the EU Budget is higher when its relative efficiency is higher We also find robust evidence of a substitution effect given the negative and significant coeffi-

Table 13: OLS regressions examining contribution choices to the Country Budget (Columns 1, 4), the EU Budget (Columns 2, 5), and the sum of the two (Columns 3, 6) in the MLPGG.

	(1) Country	(2) EU	(3) Total	(4) Country	(5) EU	(6) Total
High	-1.574***	1.661***	0.086	-1.532***	1.766***	0.233
	(0.309)	(0.371)	(0.345)	(0.309)	(0.361)	(0.314)
DE	-0.936**	0.035	-0.901*	-0.849**	-0.117	-0.966*
	(0.310)	(0.328)	(0.387)	(0.316)	(0.334)	(0.379)
FR	-0.426	0.086	-0.339	-0.466	0.019	-0.447
	(0.316)	(0.304)	(0.359)	(0.328)	(0.315)	(0.359)
IT	-0.285	0.059	-0.226	-0.272	-0.300	-0.572
	(0.321)	(0.295)	(0.330)	(0.341)	(0.299)	(0.337)
PL	-0.165	0.228	0.063	-0.214	-0.192	-0.406
	(0.334)	(0.335)	(0.349)	(0.350)	(0.346)	(0.349)
PT	0.020	0.162	0.182	0.024	-0.130	-0.106
	(0.293)	(0.289)	(0.305)	(0.302)	(0.296)	(0.310)
High × DE	0.821*	-0.072	0.750	0.765	-0.239	0.527
	(0.417)	(0.525)	(0.520)	(0.416)	(0.519)	(0.496)
High × FR	0.788	-0.545	0.243	0.827	-0.621	0.206
	(0.446)	(0.517)	(0.521)	(0.445)	(0.507)	(0.499)
High × IT	0.764	-0.501	0.264	0.711	-0.631	0.080
	(0.421)	(0.476)	(0.473)	(0.423)	(0.458)	(0.448)
High × PT	0.904	-0.679	0.225	0.855	-0.834	0.021
	(0.467)	(0.527)	(0.491)	(0.468)	(0.518)	(0.468)
High × PT	0.495	-0.756	-0.261	0.528	-0.879	-0.351
	(0.405)	(0.477)	(0.451)	(0.404)	(0.475)	(0.435)
Age				-0.000	0.003	0.003
				(0.009)	(0.010)	(0.009)
Female				0.228	-0.098	0.130
				(0.127)	(0.145)	(0.150)
Student				-0.185	0.153	-0.032
				(0.148)	(0.168)	(0.173)
Socioeconomic Status				-0.022	0.102*	0.081
				(0.044)	(0.051)	(0.051)
Education				-0.151*	-0.012	-0.164*
				(0.068)	(0.075)	(0.075)
Migrant				0.176	-0.826***	-0.650**
				(0.207)	(0.228)	(0.247)
Feel Country				0.250***	-0.188*	0.062
				(0.072)	(0.078)	(0.077)
Feel EU				-0.013	0.279**	0.265**
				(0.083)	(0.093)	(0.101)
EU Image				0.230*	-0.001	0.229
				(0.106)	(0.126)	(0.129)
Before COVID				-0.081	0.285***	0.204*
				(0.076)	(0.086)	(0.095)
Solidarity				-0.129	0.003	-0.126
				(0.080)	(0.089)	(0.090)
After COVID				-0.063	-0.074	-0.137
				(0.066)	(0.073)	(0.076)
Military Spending				0.006	-0.076*	-0.071
				(0.035)	(0.035)	(0.038)
National Army				0.038	0.062	0.100
				(0.081)	(0.087)	(0.091)
EU Army Pre-War				0.061	0.041	0.102
				(0.066)	(0.079)	(0.082)
EU Army Post-War				-0.084	0.136	0.051
				(0.075)	(0.086)	(0.088)
Constant	4.495***	3.141***	7.636***	4.417***	1.690**	6.107***
	(0.228)	(0.224)	(0.237)	(0.587)	(0.602)	(0.655)
Observations	1200	1200	1200	1200	1200	1200
R^2	0.066	0.064	0.016	0.094	0.121	0.088

Notes: Baseline category for treatment dummies is *Low*. Baseline category for country dummies is NL (=1 when observation is from The Netherlands, and 0 otherwise). Robust standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

cients of the treatment variable in the regressions about Country-budget contributions (Columns 1 and 4). Therefore, when the relative efficiency of the Country Budget is lower, subjects contribute less to it. Finally, if we consider the Total contribution variable, the effect of the treatment is not statistically significant, thus suggesting that the leveling-up and the substitution effects balance out, leaving overall contributions unchanged.

We can also capture cross-country differences in both their attitudes towards the diverse types of contributions as well as in their reactions to efficiency changes, given the presence of countries' dummies and their interaction with the treatment dummy. The coefficients of the countries' dummies "FR", "IT", "PL", "PT" in all columns are always insignificant, confirming that the contributions in these countries do not statistically differ from those in The Netherlands (the reference category). The coefficients of these country dummies are also similar (i.e., not statistically different) in magnitude (see, as shown by the post-estimation equality of coefficient tests reported in Table C.7 in Appendix C). There seems to be a difference, instead, when comparing Germans with the Dutch in the regressions explaining Country-budget contributions and Total contributions: the "DE" coefficient is negative and statistically significant in (1), (3), (4), and (6). Furthermore, we can note that the interactions between the High variable and countries' dummies are also always insignificant, apart from that of "High × DE" in (1), whose significance, however, disappears once including control variables in (4). Given the post-estimation equality of coefficient tests also for the other interaction terms, we can safely disregard any significant differences across countries as regards efficiency-driven effects.

Overall, we can state the following results:

Result 2 *Contributions to the EU Budget increase on average as its relative efficiency increases, in all countries.*

Result 3 *Contributions to the Country Budget decrease on average as its relative efficiency decreases, in all countries.*

Result 4 *There is no statistically significant evidence of an increase in total contribution due to an increase in the relative efficiency of the EU budget in all countries.*

These three results are in line with most of the MLPGG literature (Catola, D'Alessandro, et al., 2023; Fellner and Lünser, 2014; Gallier et al., 2019). However, when looking at the coefficient of our control variables (Columns 4-6) some further considerations concerning the conditions and the possible drivers of contribution decisions can be advanced. The first consideration concerns the status of being a migrant which on average drives subjects in such conditions to contribute less to the European budget and to decrease their total contribution. The second consideration regards the significance of the variables measuring the feeling of belonging towards the country or European community, i.e., *Feel Country* and *Feel EU*. As one would have expected, feeling more attached to one's own country leads subjects to increase their contribution to the Country budget (to the detriment of contribution to the EU budget), while feeling more attached to Europe leads them to contribute relatively more to the European budget and also to increase their total contribution. Overall, these considerations point out the relevance of factors connected to one's sense of identity.

Lastly, in our data, the positive attitude toward the European budget seems to be grounded on beliefs that precede the most recent events, especially COVID (Solidarity and After COVID variables do not display significant coefficients).

4.4.3 Additional analysis on migrants

In this section, we rely on subjects' answers to our post-experimental questionnaire to further investigate how their sense of identity affects their contribution decisions. We consider, in particular, the status of being a migrant as a relevant variable worth investigating for additional analysis.

The status of being a migrant represents a strong identity trait that significantly affects decisions. However, this effect could vary depend-

ing on the country of origin. Accordingly, we further develop our analysis by testing whether moving from a country that belongs to the EU or not affects migrants' contribution decisions. We, therefore, consider a dummy variable, *Migrant EU*, that takes value 1 if the country of origin of the migrant belongs to the EU, and 0 otherwise.

Table 14 reports the results of an OLS regression where only the migrants are included. We include – in addition to all the regressors of our main analysis – the variable *Age of Moving* obtained from question Q2 (replacing *Age*). Indeed, the age of moving to the host country could affect the feelings of identity connected to the status of being a migrant. Moreover, we exclude Poland from this analysis since there is only one migrant in the entire subsample. The results show that migrants who come from another EU country tend to contribute less to the Country Budget compared to migrants who come from a country outside the EU. This is not unexpected since these subjects could maintain stronger ties with their native country: it may be easier for them to move back to their countries (due to lighter regulations and travel expenses) and (consequently) the decision concerning their permanence in the host country could be felt as less definitive. These reasons can potentially explain why this group is less willing than the other group to contribute to a budget that benefits only subjects from their host country. In the same fashion, one could expect this group to be more willing to contribute to the EU Budget since such a contribution would benefit also participants from their native country. However, this is not the case, as there is no statistically significant difference in the contribution behavior toward the EU Budget between the two groups. Finally, it is worth noting how migrants react to the change in the relative efficiency of the European public good by showing only the substitution effect (and not the leveling up). In other words, subjects in the High treatment contribute to the Country Budget less than subjects in the Low treatment; however, they do not contribute more to the EU Budget.

Table 14: OLS regressions examining the contribution decisions of the subsample of migrants to the Country Budget, to the EU Budget, and the sum of contributions to both budgets.

	(1) Country	(2) EU	(3) Total
High	-1.179***	0.611	-0.568
	(0.293)	(0.346)	(0.405)
DE	-0.538	0.467	-0.071
	(0.403)	(0.448)	(0.539)
FR	-0.760	0.839	0.079
	(0.422)	(0.505)	(0.589)
IT	0.991	0.181	1.171
	(0.525)	(0.595)	(0.736)
PT	0.061	1.594	1.654
	(0.686)	(0.920)	(0.917)
Migrant EU	-0.793*	0.383	-0.410
	(0.375)	(0.490)	(0.513)
Age of moving	-0.012	0.025	0.013
	(0.022)	(0.026)	(0.026)
Female	0.342	-0.338	0.005
	(0.310)	(0.367)	(0.423)
Student	-0.145	0.416	0.272
	(0.303)	(0.380)	(0.426)
Socioeconomic Status	-0.071	0.005	-0.066
	(0.107)	(0.139)	(0.146)
Education	-0.304	-0.091	-0.395
	(0.163)	(0.211)	(0.212)
Feel Country	-0.146	-0.135	-0.281
	(0.143)	(0.193)	(0.188)
Feel EU	0.408*	0.254	0.662**
	(0.175)	(0.188)	(0.222)
EU Image	0.211	0.168	0.379
	(0.203)	(0.278)	(0.308)
Before COVID	0.151	0.390	0.541
	(0.190)	(0.219)	(0.284)
Solidarity	-0.294	-0.215	-0.510
	(0.223)	(0.233)	(0.259)
After COVID	0.059	-0.308	-0.249
	(0.173)	(0.198)	(0.226)
Military Spending	-0.043	0.002	-0.040
	(0.072)	(0.088)	(0.114)
National Army	0.041	0.216	0.257
	(0.199)	(0.227)	(0.253)
EU Army Pre-war	0.158	-0.211	-0.053
	(0.211)	(0.211)	(0.247)
EU Army Post-War	-0.251	0.259	0.008
	(0.214)	(0.223)	(0.249)
Constant	5.382***	0.812	6.194***
	(1.321)	(1.489)	(1.491)
Observations	194	194	194
R^2	0.218	0.184	0.205

Notes: Baseline category for treatment dummies is *Low*. Baseline category for country dummies is NL. Robust standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

4.5 Discussion and conclusions

In this paper, we investigated European citizens' willingness to financially sustain a European public budget compared to the public budget of the country in which they live. For this purpose, we relied on an online multilevel public good game involving a sample of 1,200 participants from six EU member states. We implemented two treatments that differ with respect to the relative efficiency of the public good representing the European public budget. Specifically, its efficiency is increased across treatments while the efficiency of the country public good remains constant. By applying this design, we are able to address two main research questions: a) To what extent do contribution decisions to the two public budgets differ across countries? and b) To what extent do reactions to the increase in the efficiency of the European public budget differ across countries?

We do find evidence of a sustained willingness to contribute to the European public budget and a positive response to the increase in its efficiency (*leveling up effect*) – which is, however, not accompanied by an increase in the total contribution (*marginal crowding in*) but by a decrease in the contribution to the country public budget (*substitution effect*). This evidence lets us make a preliminary and provisional point to address the current debate about the opportunity to introduce increasingly stable financial resources to the European budget in the form of direct taxation rather than the current reliance on transfers from the member states. Overall, European citizens in our sample show that they would support a European institution that is strengthened in its budget capacity, especially if this increased budget capacity translates into higher returns to EU citizens. However, the relevance of this general result must be discussed by referring to two main potential limitations of our work.

The first limitation is apparent in the lack of evidence for differences across countries for both our research questions. This lack of evidence could, in principle, reveal a limited power of our analysis to actually grasp such differences rather than the fact that these differences are not at stake. However, it must be noted that our analysis confirms the main

findings in the literature for each of the considered countries, i.e., the positive contribution to both public goods, the leveling-up effect, and the substitution effect. These results seem to confirm the reliability of our analysis to the extent that they can be considered a genuine robust replication of standard phenomena, with no exceptions across our country samples. However, if this is the case, we can claim to obtain an actual lack of differences in the propensities of citizens of the selected countries, who appear equally motivated in their support towards an (efficient) European public budget.

The second limitation relates to the external validity of our experiment, which appears constrained by our procedure of selection of countries. In the paper, we provided both a clarification of the technical need that made us select countries and a justification of our geopolitical criterion of selection. We must acknowledge that the possibility of inferring actual support for contribution to a European public budget by European citizens is conditioned by the fact that our country samples only partially represent the institutional, cultural, and socio-political diversity within the EU. However, the homogeneity of our results across countries can again be referred to as a basis for a reasonable generalization. Indeed, our selection of Member States embraces quite a large variability at the level of country-level characteristics, and, notwithstanding, citizens express quite an identical contribution behavior. Thus, it appears not too risky to infer that such a behavior can be considered representative of the overall European population.

The third limitation relates to the variability of individuals' characteristics within our overall sample and specifically across countries. While the imbalance in the individual characteristics detected may reflect real-world sources of heterogeneity across countries, they partially limit the interpretation and validity of the results reported in the study as regards the cross-country comparisons.

Lastly, future research can entail considering an experimental treatment in which subjects are informed about the nationality of the individuals included in the other groups. This idea builds on the assumption that stereotyped mental and social contrapositions do exist across Eu-

ropean countries. Thus, revealing the nationality of participants in the other groups may serve to enforce the sense of group identity by levying on the home-country bias and to bring new evidence on the potential relevance of these stereotyped contrapositions in a controlled experimental setting.

Chapter 5

Cooperation under the threat of severe adverse events

This essay is based on the paper: E. Bilancini, L. Boncinelli, C. Nardi, V. Pizziol "Cooperation is unaffected by the threat of severe adverse events in Public Goods Games." OSF Preprint, 2023.

5.1 Introduction

While most of the existing studies on the PGG have focused on deterministic situations, actual decisions about how much to contribute to public goods are made in situations entailing some form of environmental risk. By the term "environmental risk", we intend the existence of an exogenous stochastic process that can generate adverse events that negatively affect individuals' payoff. Environmental risk has accompanied a vast part of human history (e.g., climate change, production shocks, technological change, floods, and earthquakes). So, understanding if and to what extent environmental risk may affect cooperative behaviors seems natural and relevant, given the importance of cooperation for humankind's success and development.

In this paper, we try to understand the role of a specific form of environmental risk in an experimental setting. In particular, we consider the case where the individual marginal return to cooperation is small and, in addition, there is a low probability that an adverse event will occur, which has a considerable negative impact on individuals' payoff independently of individuals' behavior. We focus on this case because, on the one hand, this is a widespread situation for social dilemmas involving cooperation, and, on the other hand, it is the simplest and most basic setting for studying the role of environmental risk. One may want to consider cases where risk depends on individuals' behavior (e.g., the public good is a defense against the adverse event) or where adverse events are very likely (e.g., the gains from the public good are structurally very volatile). However, both these characteristics could have additional effects besides those of the kind of environmental risks we study here, presumably blurring the interpretation of results. Also, one may want to consider the situation where individual return to cooperation is substantial (e.g., the public good is very local or the group is small). Still, besides being possibly less relevant for actual social dilemmas, this case would lead to a relatively too small expected negative payoff generated by the small-chance adverse event, potentially diluting effects.

There is reliable evidence showing that environmental risk can affect cooperation in the linear PGG (e.g., Fischbacher, Schudy, et al., 2014; Gangadharan and Nemes, 2009; Levati, Morone, and Fiore, 2009). However, only a few papers in this line of research have compared the role of risk correlation across individuals (Corazzini and Sugden, 2011; Théroude and Zylbersztejn, 2020; Vesely, Wengström, et al., 2017), finding mixed results and leaving the scope for further investigations. Moreover, none of these papers focuses on low-probability events. This feature may be relevant to understanding the evolution of cooperative behaviors in areas with a threat of natural disasters, social emergencies, and targeted sacrifices. In particular, one may wonder whether cooperation might be more likely when a village is subject to the risk of floods, random kidnapping by bandits, or necessary sacrifice by one of its members.

In our experimental setting, we consider a one-shot linear Public Goods

Game with groups of 40 members. We introduce stochastic adverse events that induce three different risk correlations across individuals: independent risk (each individual has a 2.5% probability of experiencing the adverse event), perfectly positively correlated risk (there is a 2.5% probability that all group members experience the adverse event), and perfectly negatively correlated risk (1 member out of 40 is randomly selected and experiences the adverse event for sure). This latter type of risk has led us to work with relatively large groups, allowing adverse events to occur with low probability. To the best of our knowledge, no experimental study has explored this setting.

More specifically, we run an incentivized online experiment with between-subject conditions: i) a negative event independently affecting a different number of group members, depending on a random draw at the individual level (*Independent Risk treatment*); ii) a negative event that strikes either all or nobody in the group, which depends on the realization of a random draw happening at the group level (*Positively Correlated Risk treatment*); iii) a negative event that hits only one member with certainty, depending on a random draw at the group level (*Negatively Correlated Risk treatment*).[1] We compare these conditions to a *Control treatment* with deterministic payoffs in the absence of environmental risks. While it is always socially optimal to contribute, the incentive to free-ride is significant and constant across all conditions. This is especially true considering that we consider a one-shot framework.

Differently from other papers investigating the effect of shared versus idiosyncratic risks (e.g., Corazzini and Sugden, 2011; Zhang, 2019), our one-shot experimental design permits us to sterilize the impact of potential confounding factors, such as self-insurance or risk-sharing considerations, as well as learning effects. Given our focus on large groups, our experiment would be hard to implement in a laboratory setting, which is the main reason why we opted for an online setting. In turn, an online setting makes it hard to run repeated games due to asynchronicity

[1] Freundt and Lange (2021) introduce the concept of a negatively correlated risk in the PGG but, very differently from our design, apply it to the riskiness of internal and external returns.

and frequent drop-outs. This would especially hold for our case, where the group is quite large, and learning effects require a high number of repetitions due to small probabilities. Like most other papers, we also study these differences in the absence of payoff-driven concerns because payoffs are equivalent in expectation across all conditions.[2]

We establish that there is no significant difference in cooperation levels across the four conditions. So, the presence of a slight chance of severe adverse events does not affect cooperation, and risk correlation across individuals – positive or negative – does not appear to play any role. These findings support the generalizability of previous results based on the deterministic PGG workhorse. Likewise, they also support standard choice models that rely on expected utility theory with other-regarding preferences, excluding specific effects of low probabilities or risk correlations. Therefore, we are in line with Théroude and Zylbersztejn (2020)'s findings and extend them to a negatively correlated risk and a low probability of substantial losses.

5.2 Related literature

The present paper is generally connected to the experimental economic literature that studies the effects of uncertainty on the provision of public goods. Scholars have been employing many different ways to introduce uncertainty in the PGG. For instance, David L Dickinson (1998) does so through the possibility of *ex-post* exclusion from the public good's benefits. Others induce uncertainty by allowing for the production (or enhancement) of the joint investment's benefits only if the total amount of contributions overcomes a target level with a variant of the PGG known as the threshold or step-level PGG (e.g., Gueth et al., 2015; Sonnemans et al., 1998).[3]

[2]This is a standard approach used in other social dilemmas as well e.g., Xiao and Kunreuther, 2016.

[3]When failure to reach the target entails a chance to lose funds, the game is also known under the name of "collective-risk social dilemma" (e.g., Brown and Kroll, 2017; Dannenberg et al., 2015; Milinski, Röhl, et al., 2011; Milinski, Sommerfeld, et al., 2008; Tavoni et al., 2011). This framework has been extensively used to model environmental dilemmas

Some other studies induce uncertainty in the PGG by seeding risk via a lottery-style MPCR. Levati, Morone, and Fiore (2009) is the first study to combine risk preferences with voluntary contributions in this setting. They show that introducing risk on the MPCR, which is randomly selected for all group members, decreases contributions and that risk aversion has a strong negative effect on them. Levati and Morone (2013) find that this result cannot be extended to the case where the minimum value of the stochastic MPCR still allows for efficiency gains, even when probabilities are unknown. Also Artinger et al. (2012) and Todd L Cherry et al. (2015) study cooperation in a linear PGG with risky MPCRs, finding that cooperation in the risky settings compared to deterministic ones is lower (Artinger et al., 2012; Todd L Cherry et al., 2015) or comparable when the negative event's probability is very low (Artinger et al., 2012). Very differently from our design, however, in these papers, the payoff of the public good is the only at risk, so the private account represents a safe investment.

Lastly, within this same branch of literature, only a few recent papers vary, as we do, the level at which the environmental risk arises, i.e., whether at the individual or the group level, namely, Théroude and Zylbersztejn (2020) and Vesely, Wengström, et al. (2017). Despite the fact that we do not have stochastic MPCRs, our work closely relates to these papers precisely because of the risk correlation's treatments. Table 15 provides a summary of the differences in terms of parameters, treatments, and results in these studies, as well as in Zhang (2019) and Corazzini and Sugden (2011), who also manipulates risk correlations but pick as a negative event the risk of losing each period's payoff. These studies share some characteristics (which are not put in the table) that, instead, deviate from our design: they all employ lab experiments with students and have small group sizes—groups are made up of 3 or 4 members. In Zhang (2019)'s repeated PGG, the probability of experiencing the adverse event—that is, the loss of all the payoff in a period—is negatively

related to the fight against climate change. A typical finding of this branch of literature is that groups fail to cooperate when they perceive a low probability for a catastrophic event to occur, while the perception of a likely catastrophe fosters cooperation.

related to the payoffs from the game in that period. What is found is that cooperation is higher in the presence of risk at the group level than at the individual one. Corazzini and Sugden (2011), similarly, design a negative event consisting in the loss of the period's payoff; however, with a fixed probability of 67%. They find that contributions are higher when the risk is positively correlated across group members, while they are almost the same in the independent fate and the rival fate treatments—a treatment where there is only one "winner" out of the group and everyone else loses out their payoff. This condition involves asymmetric outcomes, as in our *Negatively Correlated Risk treatment*, but in our case, there is only one "loser" out of the group. Théroude and Zylbersztejn (2020) keep the risk, which is embodied in the stochasticity of the MPCR, to be wholly exogenous and compares the risky treatments also to a control treatment with deterministic payoffs. No statistically significant and systematic effect of risk on the patterns of cooperation is found across all conditions, neither in the one-shot nor in the repeated version of the game. Likewise, Vesely, Wengström, et al. (2017) compares these same three conditions in a setting of risky MPCRs, where only a repeated version of the game is present. They instead find that risk stimulates cooperation, with a higher effect when risk is at the individual rather than at the group level.

Overall, these results provide mixed evidence and leave space for further investigation on the role of risk correlation across individuals. Also, these papers never focus on a very low probability of the adverse event, as we do by keeping it constant to a value as low as 2.5%.

Since we have groups of 40 members, our work also relates to the literature on PGGs with big group sizes. Contrary to the intuition that cooperation should be more attainable in smaller groups, some studies find that larger groups cooperate moderately or significantly more than smaller ones, concluding that group size positively affects cooperative behavior (e.g., Barcelo and Capraro, 2015; Diederich et al., 2016; Isaac and Walker, 1988; Isaac, Walker, and Williams, 1994; Nosenzo et al., 2015). Although we do not manipulate group size, we bring new evidence on PGGs characterized by a low MPCR and a high number of

Table 15: Summary of experimental designs and results employed to manipulate the role of risk correlations.

Paper	Type of interaction	Treatments	α, no risk	α low, α high	Loss	Probability	Results
TZ (2020)	One-shot, Repeated (10 periods)	Baseline, Heterogeneous Risk, Homogeneous Risk	0.4	0.3, 0.5	-	0.5	Heterogeneous Risk no effect. Homogeneous Risk positive effect only in early rounds.
Z (2019)	Repeated (20 periods)	Independent Risk, Common Risk	0.4	-	Payoff of the period	Positively related to risk level & negatively to payoffs	Common Risk positive effect compared to Independent Risk.
VW (2017)	Repeated (20 periods)	No Risk, Independent Risk, Correlated Risk	0.5	0, 2	-	0.75	Independent Risk and Correlated Risk positive effect compared to No Risk. Independent Risk strongest effect.
CS (2011)	Repeated (25 periods)	Common Fate, Independent Fate, Rival Fate	0.5	-	Payoff of the period	0.67	Common Fate positive effect compared to Independent Fate and Rival Fate.

Notes: TZ (2020) is Théroude and Zylbersztejn (2020), Z (2019) is Zhang (2019), VW (2017) is Vesely, Wengström, et al. (2017), CS (2011) is Corazzini and Sugden (2011). α is the MPCR in the non-risky conditions in TZ (2020) and VW (2017), and always in Z (2019) and CS (2011). α *low,* α *high* are the two MPCRs in the risky conditions in TZ (2020) and VW (2017). *Loss* is the loss type in Z (2019). *Probability* indicates the probability of the negative event.

members, enhancing the connection to real-world scenarios where public goods naturally provide small marginal returns in big communities.

5.3 Methods

5.3.1 Experimental design

The main task of our experiment is a linear PGG. Participants are randomly matched in large groups of $N = 40$ and interact only once. Each individual $i \in N$ receives an endowment e_i which he can either keep for himself (private account) or contribute to a public good. Any contribution to the public good is multiplied by 2 and divided equally among the members of the group, implying that the MPCR is 0.05.

To investigate whether and to what extent different types of environmental risk influence cooperation, contribution decisions in the PGG are collected under four treatments.

(i) In the *Control* treatment (C), participants play the standard deterministic (i.e., risk-free) PGG.

(ii) In the *Independent Risk* treatment (IR), participants face the risk of being hit by an exogenous adverse event. The adverse event – which takes the form of a lump-sum loss λ – happens with probability p. In each group, the participants' chances of being hit by the adverse event are independent, meaning that none, some, or all group members can be hit.

(iii) In the *Positively Correlated Risk* treatment (PCR) participants face the same probability p of being hit by the adverse event (loss λ) as in IR. However, contrary to IR, the participants' chances of being hit by the adverse event are positively correlated, meaning that none or all group members can be hit.

(iv) In the *Negatively Correlated Risk* treatment (NCR), participants face, once again, the same probability p of being hit by the adverse event (loss λ). Their chances of being hit by the adverse event are now negatively correlated, meaning that only one randomly selected group member can be hit.

In the risk-involving treatments (IR, PCR, and NCR), the adverse event is realized after the game choices are made. The probability of the adverse event, p, is the same across all three treatments, and it is set to be equal to $1/N$ (i.e., $1/40$ or 2.5%). When a participant is hit by a negative event, a loss λ of 40 Points is deducted from his earnings.[4] To ensure that the risk-involving treatments are equivalent to the standard public goods game (treatment C) in terms of expected payoffs, we set the endowment in IR, PCR, and NCR equal to 60 Points and the endowment in C to 59 Points.[5]

Furthermore, to avoid negative payoffs in case of adverse events, participants' contributions in all treatments are restricted to integer numbers between 0 and 20 Points, i.e., $c_i \in \{0, 1, 2, \ldots, 20\}$.

It is worth mentioning that the risk of an exogenous adverse event does not change the incentive structure of the PGG. Given that all treatments apply to a one-shot setting, a rational and selfish participant has the incentive to be a free-rider and to contribute nothing ($c_i = 0$), whereas a full contribution ($c_i = 20$) represents the social optimum.

5.3.2 Procedures

The experiment – preregistered (AsPredicted number: #85704) and approved by the Joint Ethical Committee of Scuola Normale Superiore and Scuola Superiore Sant'Anna (Italy) – was programmed in oTree (Chen et al., 2016) and conducted online between the end of January and the beginning of March 2022. The participants were recruited through Prolific (Palan and Schitter, 2018) among the US adult population. Upon entering the study, they were asked to provide informed consent and to read the instructions. The instructions contained a simple attention check to ensure that participants were reading them carefully.[6] Before starting the

[4] The exchange rate between Points and Pounds is set at 10 Points = £ 0.20 for all participants.

[5] The difference in endowments between the risk-involving treatments and the risk-free treatment is equal to the expected loss (i.e., 1 Point). Such a small difference is very unlikely to produce endowment effects.

[6] As stated in the preregistration, only subjects who did not fail the attention check were allowed to participate in the experiment.

experiment, subjects had to answer some control questions testing their comprehension of the decision task. The experiment did not start until the participants had answered all the questions correctly. We can, therefore, safely assume that they understood the game.

After making their game choices and before receiving any feedback, participants had to report their (first-order) beliefs about others' contributions. Beliefs were elicited by asking each participant to guess the average contribution of the group members. We gave participants a financial incentive to report their beliefs accurately. We paid them 10 Points if they estimated the actual contribution of others correctly ($+/-0.5$ Points) and nothing otherwise. Incentives in the belief task were kept small relative to incentives in the PGG to avoid hedging (Blanco et al., 2010). Participants were unaware of the subsequent belief elicitation task when making their game decisions. This avoids any influence of beliefs on game decisions. Notwithstanding the extensive body of literature devoted to the question of how beliefs should be elicited (before or after choices), this is not a settled issue (Charness et al., 2021). We preferred asking first about choices because these are our most important data. Upon completion of the belief elicitation task, participants filled out a post-experimental questionnaire asking them about their risk tolerance and their general preferences (positive reciprocity, altruism, and trust).[7] The risk tolerance was measured with a non-incentivized question from the German Socio-Economic Panel asking participants to rate their willingness to take risks in general on an 11-point scale ranging from 0 (not at all willing to take risks) to 10 (very willing to take risks). The behavioral validity of this survey risk measure has been confirmed by Dohmen et al. (2011). Positive reciprocity, altruism, and trust were elicited with questions from the Global Preference Survey (Falk, A. Becker, et al., 2018). More specifically, they were respectively measured by asking partici-

[7]The post-experimental questionnaire did not include questions on the participants' demographic characteristics – namely, age, gender, and student status – as this information can be retrieved from Prolific. The questionnaire also elicited loss aversion using the lottery choice task proposed by Gächter et al. (2021). Yet, given the pitfalls of this task in settings (like ours) in which the stakes can no longer be considered small, in the rest of the paper, we overlook such measure.

pants to self-assess their willingness i) to return a favor, ii) to give to good causes without expecting anything in return, and iii) to assume that people have only the best intentions. The three answers had to be provided on a scale from 0 to 10, where a higher rating indicated a higher willingness to act in the described way.

The post-experimental questionnaire also included two mathematical questions testing the participants' literacy about probability. These questions were intended to measure both basic knowledge of probabilities and the so-called 'conjunction fallacy,' which occurs when it is assumed that the conjunction of two events is more – rather than less – likely to occur than one of the events alone.[8] A math score was then constructed as the sum of correct answers, ranging from 0 to 2.

We used a between-subjects design, i.e., each subject was exposed to only one of the four treatments. Averaging over all treatments, mean earnings amounted to £ 2.18 (inclusive of a £ 0.75 fixed participation fee) and participants took about 10 minutes to complete the experiment. The incentives in the experiment were thus substantial and perfectly resembled the hourly compensation usually provided in lab experiments (namely, £ 13).

5.3.3 Participants

Overall, 1280 subjects participated in the experiment, i.e., 320 participants (8 groups) per treatment. The sample size was determined using an *a-priori* power analysis for a t-test with a mean contribution in the control treatment equal to 14,[9] a power of 0.80, an alpha of 0.05, and an alleged effect size of 0.275. We aimed at having an effect size between

[8]The questions read: "Two fair six-sided dice are rolled. What is the probability that their sum is exactly equal to 2? a) 1/3, b) 1/6, c) 1/18, d) 1/36" and "Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. Which of the following statements is more probable? a) Linda is a bank teller, b) Linda is a bank teller and is active in the feminist movement". The latter question is due to Kahneman and Tversky (1982).

[9]This is a conservative expectation: in an online, standard PGG experiment conducted on MTurk, with a group size of 4 and a MPCR of 0.4, Arechar et al. (2018b) reported an average contribution of 15 out of 20.

Table 16: Means (and standard deviations) of participants' characteristics and preferences.

	C	IR	PCR	NCR
Age	28.00	29.31	29.40	29.33
	(10.13)	(11.35	(11.49)	(12.39)
Female	0.62	0.63	0.63	0.63
	(0.48)	(0.48)	(0.48)	(0.48)
Student	0.29	0.28	0.24	0.28
	(0.45)	(0.45)	(0.43)	(0.45)
Experienced	0.30	0.31	0.32	0.31
	(0.46)	(0.46)	(0.47)	(0.46)
Risk tolerance	5.05	5.06	5.50	5.42
	(2.22)	(2.38)	(2.30)	(2.17)
Positive reciprocity	8.71	8.86	8.76	8.83
	(1.32)	(1.25)	(1.30)	(1.27)
Altruism	7.59	7.63	7.64	7.72
	(2.03)	(1.99)	(1.96)	(2.02)
Trust	5.22	5.09	5.25	5.31
	(2.34)	(2.48)	(2.26)	(2.38)
Math score	1.52	1.43	1.47	1.49
	(0.58)	(0.60)	(0.59)	(0.56)
Observations	320	320	320	320

0.2 and 0.3 because we wanted to improve on the previous related work significantly, while at the same time excluding economically irrelevant effects.

Table 16 reports summary statistics of our sample's demographic characteristics and individual preferences, divided by treatment. Overall, the average age is around 29, and about two-thirds of the participants are female. Approximately thirty percent of the participants are students, and about the same percentage are experienced Prolific users (i.e., have completed at least 150 studies). Based on the participants' responses to the SOEP question, our sample is, on average, risk-neutral in all treatments. Finally, our sample is well-balanced in terms of general preferences (positive reciprocity, altruism, and trust) and probability literacy, which is measured by the math score.[10]

[10]According to a series of χ^2 tests, we find no differences in gender, student status, and experience in using Prolific across treatments (p-values equal 0.996, 0.528, and 0.964, respectively). Similarly, a series of Kruskal-Wallis tests do not reveal any differences in

5.4 The role of uncertainty

Following the so-called *perceived target of the threat* principle outlined by Weisel and Zultan (2016), one could expect that when individuals perceive their group to be under threat, they tend to act for the group's good and contribute more. In contrast, they tend to act more selfishly and withhold their contributions when they perceive the threat to be personally upon themselves. However, in a context where uncertainties cannot be reduced by cooperation, the risk might not play an influential role (for instance, null effects are found in Björk et al., 2016). It is not easy to advance specific hypotheses in this regard. *A priori*, it is unclear whether inducing different types of environmental risks, affecting the whole community, part of it, or only one member, can overcome or boost the free-riding problem and to what extent. We believe that the first step is to document if and to what extent cooperative behavior is affected.

For the above reasons, we just test the following two null hypotheses:

Hypothesis 1: *No difference exists in contribution levels between the control and any of the risky experimental conditions.*

Hypothesis 2: *No difference exists in contribution levels between any pair of risky experimental conditions.*

The answers to these hypotheses are given in Subsection 5.5.2 (Result 1 and Result 2, respectively) by testing the significance of the different conditions in Tobit regressions run on the experimental data collected.

5.5 Results

In this section, we present our results. We first display some descriptive and non-parametric analyses. We then investigate the presence of

age, positive reciprocity, altruism, trust, and math score (p-values equal 0.689, 0.376, 0.824, 0.864, and 0.306, respectively). Although the risk tolerance seems to vary across treatments (Kruskal-Wallis test, p-value $= 0.025$), this variation becomes statistically insignificant applying the Bonferroni correction for multiple testing.

treatment effects by making use of regressions, which allow us to control for heterogeneity in participants' demographic characteristics and individual preferences. Finally, we briefly report on the elicited first-order beliefs.

5.5.1 Descriptive and non-parametric analyses

Figure 15 depicts, separately for each treatment, the mean contributions to the public good. Its visual inspection reveals two noteworthy features. First, the Control treatment replicates the most recent findings in online, one-shot PGGs (Berg et al., 2020; Catola, D'Alessandro, et al., 2021; Isler, Gächter, et al., 2021): the mean contributions are equal to 11.78, or, alternatively, 59% of the points are available for the allocation decision. Remarkably, contributions to the C treatment are substantial, even though – compared to the previous studies – we implement a larger group size ($N = 40$) and a much smaller marginal per capita return ($MPCR = 0.05$).

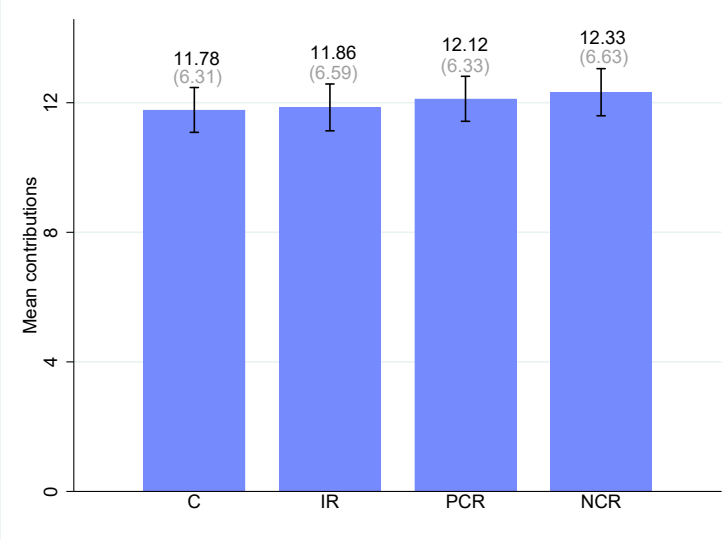

Figure 15: Mean contributions by treatment. Standard deviations in parentheses. Confidence intervals at the 95% level.

The second fact documented through Figure 15 is that the mean contributions in the risk-involving treatments are slightly higher than in C, especially in the PCR and NCR treatments. Yet, the differences are not statistically significant, either when simultaneously comparing all treatments (Kruskal-Wallis test, p-value equal 0.5254) or when implementing pairwise comparisons between treatments (Wilcoxon rank-sum tests, all p-values > 0.1653).[11] The lack of treatment effects is further confirmed by looking at the distributions of contributions across treatments, which are displayed in Figure 16, Figure 17, Figure 18, Figure 19. The figure shows that the game-theoretic prediction of universal free riding, based on general opportunism, is clearly rejected in all treatments: the proportions of free-riders are stable across treatments and are as low as 7.5% in C and IR, 6.5% in PCR and 8.5% in NCR. Moreover, the contributions are bimodal (at 10 and 20 Points) in all treatments, with a higher proportion of people contributing 10 or 20 in the risk-involving treatments than in the Control. Although there seems to be some variation in the fraction of half and full contributors between the risk-free and the risk-involving treatments, a series of Epps-Singleton tests do not reject the null hypothesis of equal distributions across treatments (all p-values > 0.0545).

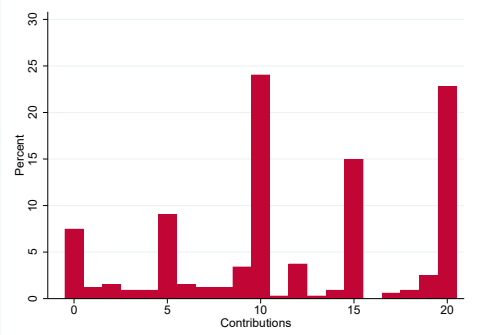

Figure 16: Distribution of contribution choices in the Control treatment (C).

[11] All p-values in the paper are two-tailed.

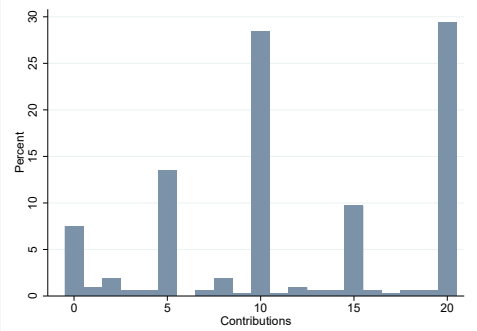

Figure 17: Distribution of contribution choices in the Independent Risk treatment (IR).

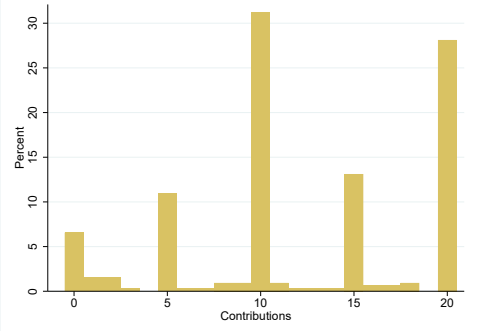

Figure 18: Distribution of contribution choices in the Positively Correlated Risk treatment (PCR).

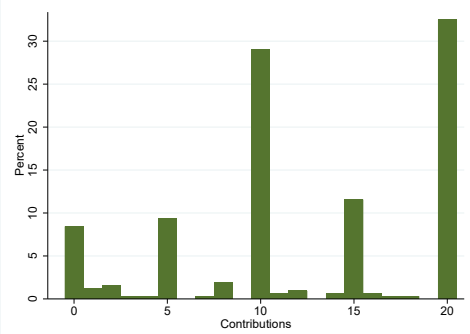

Figure 19: Distribution of contribution choices in the Negatively Correlated Risk treatment (NCR).

5.5.2 Treatment effects on contributions

Table 17 shows the results of Tobit models aimed at examining the contribution choices in the PGG, which are bounded between 0 and 20. The coefficients of the treatment dummies – "IR", "PCR" and "NCR" – in column (1) are positive and insignificant, confirming that the contributions in the risk-involving treatments do not statistically differ from those in the Control treatment (the reference category). The coefficients of the treatment dummies are also similar (i.e., not statistically different) in magnitude (see the post-estimation equality of coefficient tests reported at the bottom of the table). This holds true even if we add controls for participants' demographics and preferences as well as for the time spent on the decision page (see column (2)).[12] Among the added control variables, "Age", "Risk tolerance", "Positive reciprocity", "Altruism", and "Trust" positively and significantly impact contributions. More specifically, contributions are found to increase with age. This evidence is consistent with psychological research reporting that older adults value contributions to the public good more than younger ones (Freund and

[12]The effect of different types of environmental risks on contributions remains null even if double-hurdle regressions, which allow to separately consider the decision to contribute (extensive margin) and the decision of how much to contribute (intensive margin), are used. Results are available in Table D.2 of Appendix D.

Blanchard-Fields, 2014). A higher willingness to take risks – as measured by the SOEP question – is associated with a higher propensity to contribute (which is not surprising since the participants receive a lower payoff if their group members do not contribute anything) and participants with a higher positive reciprocity disposition are more inclined to contribute. Finally, as one would intuitively expect, more altruistic participants and those who exhibit higher levels of trust in others tend to contribute more.

In conclusion, we state the following two results:

Result 1: *Keeping the expected payoff constant for given contribution levels, the mere addition of environmental risk – taking the form of an exogenous low chance of a substantial negative shock – does not produce appreciable changes in contribution decisions.*

Result 2: *Different risk correlations (zero, positive, negative) of the environmental shock do not appreciably affect contribution decisions.*

Given these null results, we deem it important to discuss the statistical power related to our sample size. We substantially improved the statistical power of our analyses with respect to the previous literature. For instance, Théroude and Zylbersztejn, 2020 also report a null result, but with a sample size of around 70 subjects per treatment. With an *ex ante* Cohen's d equal to 0.275, we would have been able to detect statistically significant differences in contributions between treatments if and only if these differences had been at least equal to 1.78 Points. Clearly, the observed effect sizes are much smaller. Hence, we can confidently conclude that our experimental treatments have no economically meaningful impact on the contribution decisions.

5.5.3 Beliefs

Figure 20 plots the mean values of elicited first-order expectations about others' behavior, divided by treatment. On average, participants expect

Table 17: Tobit regressions examining the contribution choices in the PGG.

	(1)	(2)
IR	0.475	0.282
	(0.769)	(0.718)
PCR	0.732	0.137
	(0.752)	(0.713)
NCR	1.138	0.525
	(0.786)	(0.740)
Age		0.096***
		(0.029)
Female		-1.133
		(0.661)
Student		-0.890
		(0.579)
Experienced		-0.832
		(0.650)
Risk tolerance		0.847***
		(0.134)
Pos. Reciprocity		0.602*
		(0.236)
Altruism		0.498**
		(0.161)
Trust		0.295*
		(0.125)
Math score		-0.070
		(0.470)
Log(Time)		-0.302
		(0.514)
Constant	12.694***	-2.667
	(0.523)	(2.732)
Tests of coefficients (p-values)		
IR vs. PCR	0.744	0.8448
IR vs. NCR	0.417	0.7523
PCR vs. NCR	0.612	0.6118
Observations	1280	1280
Pseudo R^2	0.001	0.022

Notes: Robust standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

the group members to contribute about half of the available points, and this is stable across treatments.

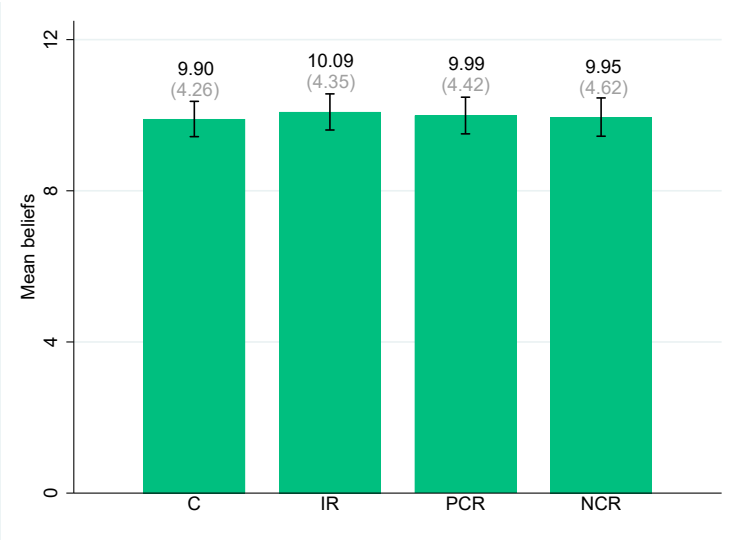

Figure 20: Mean beliefs by treatment. Standard deviations in parentheses. Confidence intervals at the 95% level.

The participants' beliefs are strongly and positively correlated with their own behavior in the PGG (Pearson's correlation coefficients are equal to 0.4854, 0.5814, 0.5311, and 0.5206 in C, IR, PCR, and NCR, respectively; all coefficients are statistically significant at the 0.1% level). This finding can be interpreted as a signal of compliance with social norms. Indeed, in many contexts, social norms can help explain why individuals behave prosocially at a cost for themselves (e.g., Bicchieri, Dimant, et al., 2022). Alternatively, it could reflect the so-called *false consensus effect* (Ross et al., 1977), suggesting that participants who are more prone to contribute have more optimistic beliefs about others' behavior.

As for the accuracy of beliefs, we find that only a small fraction of subjects – i.e., less than 10% – perfectly predicts the actual average contribution of the group members (±0.5). The mean difference between beliefs and others' actual contributions is always negative and ranges from -1.77

(SD = 4.45) in IR to -2.37 (SD = 4.78) in NCR. Hence, participants underestimate the degree of others' prosocial behavior in all treatments. This is in line with recent findings for linear PGG games played online (e.g., Bilancini, Boncinelli, and Celadin, 2022; Catola, D'Alessandro, et al., 2021), while for laboratory experiments it has often been found the opposite (e.g., Fehr, Hoff, et al., 2008; Kocher et al., 2015). It is not straightforward to rationalize such mixed evidence.

5.6 Discussion and conclusions

A large body of experimental evidence reports that people typically cooperate in the PGG, even in one-shot anonymous interactions. Most studies focus on the case with no environmental risk. In this paper, we add to the literature by documenting that this tendency is fundamentally preserved in the presence of a low probability of an adverse event having a considerable negative impact on individuals' payoff independently of individuals' behavior. More specifically, we document that cooperation levels are considerable (about 60% of resources available for contribution) even though the marginal return of contributing is as little as 0.05 and, interestingly, these cooperative levels are in line with what is found in other online one-shot PGGs employing small group sizes with much larger individual returns.

Most importantly, from our experimental findings, we can conclude that the mere addition of environmental risk does not change cooperative behaviors with respect to deterministic scenarios. Additionally, we find that the nature of environmental risk – i.e., whether it is independent, positively, or negatively correlated across individuals – does not appreciably affect cooperation levels.

Our results can be considered in the light of decision theories in uncertain environments. For instance, following Tversky and Kahneman (1992) and Prelec (1998), one could expect that people tend to overweight low probabilities when dealing with described probabilities in scenarios entailing some risk, like ours. However, the actual effect on the behavior of such over-weighting depends crucially on the expected value of the

negative shock and individuals' risk attitudes. In our experimental setting, expected payoffs conditional on group members' contributions are identical across all treatments. Moreover, there is only a 2.5% chance that the final payoff is reduced by about 2/3 of the initial endowment. So, under the assumption of risk neutrality, the expected value of the negative shock is little and should not affect behavior even if over-weighting is strong. Our results are consistent with this prediction. In general, although individuals might not follow the expected utility theory (see, e.g., Starmer, 2000), it needs not show up in our data, provided that risk attitudes are not too far from risk neutrality, as it seems to be the case with our experimental subjects.

Furthermore, one may consider the role of other-regarding preferences, such as altruism (Anderson et al., 1998; Andreoni and Miller, 2002) or inequity aversion (Bolton and Ockenfels, 2000; Fehr and Schmidt, 1999; Fischbacher, Schudy, et al., 2014). In principle, one might expect that such other-regarding preferences affect behavior depending on the presence of risk and the type of risk correlation since the realization of the adverse event will not affect group members in the same way. However, given the additive nature of the stochastic component in our setting and its small expected value in absolute terms, the expected welfare changes in a large group are quite diluted. So, even substantial altruism or strong inequity aversion are not expected to affect behavior across treatments, in line with what we observe.

We stress that our experimental data improve, in terms of statistical power and detectable effect size, upon previous work (Théroude and Zylbersztejn, 2020). Hence, the lack of treatment effects suggests that the nature of environmental risk – i.e., whether it is independent, positively, or negatively correlated across individuals – is not a primary source of behavioral effects, at least as a single source of variation as we tested in our experiment. It remains to explore whether this neutrality survives in different settings with an endogenous risk, endogenous group membership and size, or adverse event mitigation.

Indeed, in our study, we focus on a kind of environmental risk where contributing to the public good does not affect the probability of the ad-

verse event or the size of its effects upon realization. Thus, we leave out the relation between investments in the PGG and the negative environmental shock. A different research line can investigate this aspect, along the lines of David L Dickinson (1998). Also, it seems interesting to inquire about the reactions to the realization of a disaster by looking at the *ex-post*, rather than *ex-ante*, cooperative behavior. Further research could also investigate the role of conditional cooperators (Fischbacher, Gächter, et al., 2001) to check whether there are differences in the behavior of such player types that do not mirror the average behavior.

Finally, it is crucial to acknowledge the limitations inherent in this study. First, the study's reliance on an online experimental design may limit the generalizability of the findings to real-world, offline interactions. Hence, the findings may not fully capture the intricate dynamics and complexities that occur in face-to-face interactions or different contexts. Furthermore, it is important to note that the study primarily focuses on a one-shot game, meaning that participants made decisions without the benefit of repeated interactions or opportunities for learning and experience. We cannot exclude that this scope of the experiment might have hindered the emergence of significant effects. Lastly, given the absence of any discernible treatment effects, it is worth considering whether the null result stems from the inherent nature of risk itself or perhaps from the nuanced manner in which risk was perceived within the experiment. It is plausible that the way in which risk was presented or understood by participants mitigated any potential impact on the observed cooperation choices. To conclude, these limitations should be carefully considered when interpreting the study's findings, considering their broader implications, and when contemplating future research in this literature.

Chapter 6

Conclusion

In this dissertation, I presented four papers focused on studying cooperative behaviors with Experimental and Behavioral Economics methods in the context of public goods provision. They investigate different aspects of human cooperation in the public good game (PGG) and its variant, the multilevel public goods game (MLPGG).

The first paper, presented in Chapter 2, reports results from an experiment designed to examine the comparative role of personal and social norms on individuals' decisions in a PGG scenario. While also contributing methodologically to the literature on norms elicitation, this paper shows that personal norms, compared to social ones, have a more prominent role in cooperative decision-making, at least in online one-shot anonymous interactions. The results thus provide further support to the recent, but rapidly growing, stream of literature that emphasizes the importance of personal norms in decision-making (Bašić and Verrina, 2021; Capraro and Perc, 2021; Capraro and Rand, 2018).

The second paper, discussed in Chapter 3, reports results from an experiment designed to examine the effects of efficiency changes on individuals' decisions in a MLPGG scenario without priming group identity and with a neutral framing of the task. The findings reveal that people positively react to increases in the upper-level public good efficiency but at the expense of contributions to the lower-level public good. More-

over, we document that the simple addition of a public good (in our case, the global public good) in the decision set is enough to increase overall contributions. Our findings, hence, clear up the previous mixed evidence found in the MLPGG literature (Blackwell and McKee, 2003; Chakravarty and Fonseca, 2017; Fellner and Lünser, 2014; Gallier et al., 2019).

The third paper, reported in Chapter 4, also builds on an experiment based on a MLPGG framework. It explores the willingness of EU citizens from different countries to contribute to the local public good, referred to as the country budget, versus a potential European Union (EU) public budget (i.e., the global public good) and how they react to changes in the relative efficiency of the two nested goods. This work, hence, connects not only to the MLPGG literature on efficiency-driven effects but also to that on national identity (Buchan, Brewer, et al., 2011; Buchan, Grimalda, et al., 2009). The results show no significant differences in contributing behaviors across different EU countries. Also, they indicate that the efficiency-driven findings established in the previous essay are replicated under the new different conditions, specifically with subjects from different EU countries and primed to perceive a sense of national identity (at the local level) and European identity (at the global level), as well as with a framed task that evokes real-world institutions (i.e., public budgets). This indicates that the efficiency-driven findings established in the previous essay are likely to be generalized for larger applications.

The fourth and final paper, presented in Chapter 5, experimentally examines the presence and type of environmental risk on individuals' decisions in a PGG scenario. More specifically, it investigates whether the presence of a risk, which is modeled as a slight chance of severe losses, and whether its nature, i.e., if independent, negatively, or positively correlated across individuals of the group, impacts cooperative behavior. The findings indicate that the presence or nature of our environmental risk does not affect cooperation, thus supporting and expanding Théroude and Zylbersztejn, 2020's conclusions. Hence, in our experiment, standard choice models based on expected utility theory (with other-regarding preferences) can perfectly explain behaviors.

Overall, this dissertation aims to contribute to our understanding of human cooperative behavior in the context of public goods provision, which has important implications not only for scholars willing to uncover the unknown about human behaviors but also for policymakers willing to design effective policies or institutions aimed at achieving optimal outcomes for society as a whole. Indeed, the findings of this dissertation provide several important insights as well as depict possible avenues for future research.

The results of the first essay suggest interventions aimed at giving salience to personal norms - for instance, by simply asking people to stop and think about their personal norm in a given situation - to switch people prior from social norms (i.e., what they expect others to do/believe) to their personal norm. This intervention would help change priors when they are based on a guess of what would other people think is the right thing to do (which might lead to an underestimation of the likelihood of socially efficient outcomes) or what other group members would do (which, again, would lead to inefficiently lower contributions). Moreover, future research could investigate whether the same documented patterns hold in scenarios entailing repeated decisions where anonymity is loosened, and social proximity is augmented.

The results of the second and third essays show that individuals tend to increase their contributions to the global (European) public good as its relative efficiency increases while decreasing their average contributions to the local (national) public good. One implication of this stream of research is that policymakers should consider the potential trade-offs between investing in local versus global public goods. While investing in local public goods may be important for meeting the needs of specific communities, investing in global public goods may have broader benefits for society as a whole and be more efficient, especially in sectors that deal with transnational problems like pollution or health in times of pandemics. Another implication is that policymakers should design policies and institutions that are capable of creating mechanisms for coordinating contributions across different levels of government or between different countries. Further research can be carried out to study the ef-

fects of integration vs. segregation on group-transcendent prosociality using other paradigms of group identity (Bronchal, 2023), such as ethnicity (e.g., Chuah et al., 2014), religion (e.g., Isler, Yilmaz, et al., 2021), or political ideology (e.g., Romano et al., 2021).

The findings of the fourth essay suggest that the presence of environmental risk and its correlation among group members do not significantly affect cooperative behavior. This has important implications for policymakers who aim to encourage contributions to public goods provision in environments with potential risks of adverse events entailing low probabilities and huge damages, such as catastrophic natural disasters. Our null results could indicate that policymakers should focus on promoting a sense of shared community membership that goes beyond a possible "common fate" narrative in order to encourage cooperation. However, it is important to note that our study investigated *ex-ante* behavior prior to any realization of adverse events. A different research approach could investigate the reactions to the realization of a disaster (*ex-post* behavior) and shed light on human cooperative behavior after facing catastrophes such as earthquakes, tsunamis, or wars. Lastly, to bring new evidence, a new experiment could be designed to compare the behavior of anonymous groups with no information about fellow members, homogeneous groups where all members share the same identity, and heterogeneous groups where members have mixed identities, with the intent to provide insights into the impact of group identity on cooperative behavior in risky public good game scenarios.

Appendix A

Supplementary materials for Chapter 2

A.1 Additional tables

Table A.1: Participants' average characteristics by experiment.

	Age	**Male**	**Student**	**Soc. Status**	**Education**	**Employed**
Experiment 1	36.38 (12.74)	0.33 (0.47)	0.23 (0.42)	5.40 (1.54)	3.69 (1.00)	0.71 (0.46)
Experiment 2	35.45 (14.64)	0.34 (0.48)	0.30 (0.46)	5.58 (1.38)	3.50 (0.96)	0.71 (0.46)
p-value	0.252	0.893	0.194	0.399	0.104	1.000

Notes: P-values refer to the results of Wilcoxon rank-sum tests for the variables Age, Socioeconomic status, and Education, and of Fisher's tests for the variables Male, Student status, and Employment status.

Table A.2: Tobit regressions examining the contribution choices in the PGG with controls.

	(1)	(2)	(3)	(4)
PN	0.828***	1.114***		0.944***
	(0.0986)	(0.104)		(0.109)
EE	0.423**		1.098***	0.591***
	(0.145)		(0.159)	(0.139)
NE		-0.0969	0.128	-0.336*
		(0.130)	(0.144)	(0.140)
Altruism	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Patience	0.234*	0.234*	0.435**	0.253**
	(0.095)	(0.107)	(0.141)	(0.090)
Risk tolerance	0.011	0.017	-0.003	-0.050
	(0.072)	(0.078)	(0.116)	(0.066)
Trust	0.072	0.141	-0.026	0.063
	(0.071)	(0.073)	(0.093)	(0.071)
Neg. Reciprocity	-0.038	-0.022	-0.073	-0.034
	(0.058)	(0.065)	(0.079)	(0.055)
Pos. Reciprocity	-0.017	-0.014	0.050	-0.038
	(0.111)	(0.114)	(0.138)	(0.108)
Comprehension	0.037	0.022	0.087	0.039
	(0.148)	(0.153)	(0.215)	(0.137)
CRT	0.144	0.077	0.469*	0.224
	(0.136)	(0.139)	(0.197)	(0.135)
Age	0.000	-0.006	0.009	-0.010
	(0.013)	(0.014)	(0.021)	(0.014)
Male	-0.198	-0.225	-0.354	-0.333
	(0.357)	(0.384)	(0.518)	(0.366)
Stud. Status	-0.574	-0.658	-0.088	-0.511
	(0.435)	(0.498)	(0.553)	(0.438)
Socioeco. Status	-0.085	-0.077	-0.118	-0.027
	(0.089)	(0.098)	(0.124)	(0.091)
Education	0.266	0.334	0.171	0.285
	(0.170)	(0.171)	(0.208)	(0.166)
Constant	-3.409**	-2.919*	-3.757**	-3.048**
	(1.120)	(1.201)	(1.386)	(1.097)
Observations	158	158	158	158

Notes: The dependent variable is the contribution. Robust standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.3: Tobit regressions examining different types of norms with the pooled data from the two experiments.

	(1) PN	(2) EE	(3) NE
Play	-0.603	-0.488	-0.220
	(0.548)	(0.370)	(0.393)
Constant	6.975***	5.581***	5.408***
	(0.479)	(0.325)	(0.330)
Observations	263	263	263

Notes: Play is a dummy variable indicating the experiment subjects participated in. It takes value 1 if an observation comes from Experiment 1 and 0 if it comes from Experiment 2. Robust standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.4: Tobit regressions examining the contribution choices in the PGG with controls and interaction terms with Male.

	Contribution
PN	0.812***
	(0.133)
Male	0.314
	(0.926)
Male×PN	0.333
	(0.205)
EE	0.736***
	(0.165)
Male×EE	-0.286
	(0.305)
NE	-0.254
	(0.132)
Male×NE	-0.250
	(0.268)
Stud. Status	-0.408
	(0.415)
Education	0.260
	(0.154)
Socioeco. Status	-0.005
	(0.095)
Age	-0.008
	(0.013)
Altruism	0.001
	(0.001)
Patience	0.249**
	(0.088)
Risk	-0.057
	(0.063)
Trust	0.090
	(0.076)
Neg. Reciprocity	-0.049
	(0.058)
Pos. Reciprocity	-0.019
	(0.100)
Comprehension	0.046
	(0.138)
CRT	0.240
	(0.128)
Constant	-3.623**
	(1.233)
Observations	158

Notes: Robust standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.5: Tobit regressions examining the contribution choices in the PGG with controls and interaction terms with Education.

	Contribution
PN	1.282***
	(0.282)
3.Education	-3.665*
	(1.693)
4.Education	-3.168
	(1.755)
5.Education	-3.350
	(1.955)
6.Education	-3.591
	(2.717)
3.Education×PN	-0.285
	(0.310)
4.Education×PN	-0.330
	(0.309)
5.Education×PN	-0.659
	(0.382)
6.Education×PN	-1.012
	(0.960)
EE	-0.547
	(0.414)
3.Education×EE	1.019*
	(0.472)
4.Education×EE	1.220**
	(0.454)
5.Education×EE	2.043***
	(0.551)
6.Education×EE	0.653
	(0.650)
NE	-0.317
	(0.237)
3.Education×NE	0.100
	(0.267)
4.Education×NE	-0.023
	(0.263)
5.Education×NE	-0.478
	(0.337)
6.Education×NE	1.484
	(1.005)
Male	-0.128
	(0.302)
Stud. Status	-0.275
	(0.377)
Socioeco. Status	-0.040
	(0.087)
Age	-0.010
	(0.013)
Altruism	0.000
	(0.001)
Patience	0.301**
	(0.092)
Risk	-0.082
	(0.071)
Trust	0.074
	(0.071)
Neg. Reciprocity	-0.071
	(0.054)
Pos. Reciprocity	0.065
	(0.089)
Comprehension	-0.147
	(0.122)
CRT	0.208
	(0.126)
Constant	0.950
	(1.809)
Observations	158

Notes: Robust standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.6: Tobit regressions examining the contribution choices in the PGG with controls and interaction terms with Student Status.

	Contribution
PN	0.959***
	(0.106)
Stud. Status	-1.559
	(1.240)
Stud. Status×PN	-0.037
	(0.295)
EE	0.507**
	(0.168)
Stud. Status×EE	0.270
	(0.271)
NE	-0.327*
	(0.135)
Stud. Status×NE	-0.016
	(0.317)
Male	-0.348
	(0.358)
Education	0.267
	(0.152)
Socioeco. Status	-0.022
	(0.092)
Age	-0.011
	(0.014)
Altruism	0.001
	(0.001)
Patience	0.268**
	(0.090)
Risk tolerance	-0.060
	(0.064)
Trust	0.062
	(0.070)
Neg. Reciprocity	-0.029
	(0.056)
Pos. Reciprocity	-0.037
	(0.106)
Comprehension	0.029
	(0.133)
CRT	0.225
	(0.131)
Constant	-2.763*
	(1.098)
Observations	158

Notes: Robust standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.7: Tobit regressions examining the contribution choices in the PGG with controls and interaction terms with Socioeconomic Status.

	Contribution
PN	1.415***
	(0.278)
2.Socioeco. Status	-2.612
	(2.158)
3.Socioeco. Status	6.254***
	(1.600)
4.Socioeco. Status	5.304***
	(1.502)
5.Socioeco. Status	3.789**
	(1.166)
6.Socioeco. Status	2.443
	(1.236)
7.Socioeco. Status	3.938**
	(1.229)
8.Socioeco. Status	2.050
	(1.653)
9.Socioeco. Status	-4.164***
	(1.007)
2.Socioeco. Status×PN	1.558***
	(0.391)
3.Socioeco. Status×PN	-0.225
	(0.398)
4.Socioeco. Status×PN	-0.242
	(0.303)
5.Socioeco. Status×PN	-0.510
	(0.279)
6.Socioeco. Status×PN	-0.749*
	(0.323)
7.Socioeco. Status×PN	-0.303
	(0.312)
8.Socioeco. Status×PN	-0.860**
	(0.297)
EE	1.521***
	(0.317)
2.Socioeco. Status×EE	-0.000
	(0.592)
3.Socioeco. Status×EE	-1.285*
	(0.556)
4.Socioeco. Status×EE	-1.916***
	(0.420)
5.Socioeco. Status×EE	-0.729
	(0.372)
6.Socioeco. Status×EE	-0.541
	(0.398)
7.Socioeco. Status×EE	-0.950*
	(0.402)
NE	-0.524*
	(0.207)
2.Socioeco. Status×NE	-1.417**
	(0.425)
3.Socioeco. Status×NE	-0.147
	(0.528)
4.Socioeco. Status×NE	0.627
	(0.323)
5.Socioeco. Status×NE	0.031
	(0.277)
6.Socioeco. Status×NE	0.450
	(0.254)
7.Socioeco. Status×NE	0.131
	(0.410)
Controls	✓
Constant	-6.867***
	(1.603)
Observations	158

Notes: Robust standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

A.2 Experimental instructions (Experiment 1)

The Task

In this study, you will be randomly assigned to a group of 4 participants. You are given 10 points and have to decide whether to contribute to the common pool of the group. Contributions can be any integer from 0 to 10. You keep the remaining points. The other participants face the same decision.

> The **group common pool** yields the following returns: the contributions of the 4 participants are added up and the total is multiplied by a factor of 2.4. The resulting amount is equally split among the 4 participants.

Your payoff equals your earnings from the group common pool plus the amount you keep for yourself. The final conversion will be as follows: 40 points correspond to 1 GBP.

YOUR DECISION

Please decide how to distribute your 10 points among the two options.

Your contribution to the group:

What you keep for yourself:

Remind: The total amount contributed to the group common pool will be multiplied by 2.4 and divided by 4.

YOUR OPINION

Consider the decision task you faced. In your opinion, how much the other members of your group contribute to the group common pool? Please indicate in the box below what you believe was the average contribution of the other members of your group to the group common pool. You will receive additional 4 points for each correct answer. An answer is considered to be correct if it is less than 0.50 close to the true value.

Consider again the decision task you faced. How much do you believe a member of your group ought to contribute to the group common pool? Please indicate in the box below what you believe a member of your group ought to contribute to the group common pool.

In your opinion, how have the other members of your group answered to the previous question? Please indicate in the box below what you believe was the average answer by the other members of your group to the previous question. You will receive additional 4 points for each correct answer. An answer is considered to be correct if it is less than 0.50 close to the true value.

Control questions

Please answer the following questions.

QUESTION 1: What are your total earnings if all members of the group (you included) contribute 10 to their group?
. . . points
QUESTION 2: What level of your contribution to the group earns the highest payoff for you personally if all others contribute 0 to the group?
. . . points
QUESTION 3: What level of your contribution to the group earns the highest payoff for you personally if all others contribute 10 to the group?
. . . points

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?
. . . cents
If it takes 5 minutes for five machines to make five widgets, how long would it take for 100 machines to make 100 widgets?
. . . 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?
. . . days

We now ask for your willingness to act in a certain way in different areas. Please indicate your answer on a scale from 0 to 10, where 0 means you are "completely unwilling to do so" and a 10 means you are "very willing to do so".

- **How willing are you to give up something that is beneficial for you today in order to benefit more from that in the future?**

  ◦ 0 ◦ 1 ◦ 2 ◦ 3 ◦ 4 ◦ 5 ◦ 6 ◦ 7 ◦ 8 ◦ 9 ◦ 10

- **How willing are you to take risks?**

  ◦ 0 ◦ 1 ◦ 2 ◦ 3 ◦ 4 ◦ 5 ◦ 6 ◦ 7 ◦ 8 ◦ 9 ◦ 10

- **How willing are you to punish someone who treats others unfairly, even if there may be costs for you?**

  ◦ 0 ◦ 1 ◦ 2 ◦ 3 ◦ 4 ◦ 5 ◦ 6 ◦ 7 ◦ 8 ◦ 9 ◦ 10

How well does the following statement describe you as a person? Please indicate your answer on a scale from 0 to 10. A 0 means "does not describe me at all" and a 10 means "describes me perfectly".

- **I assume that people have only the best intentions.**

  ◦ 0 ◦ 1 ◦ 2 ◦ 3 ◦ 4 ◦ 5 ◦ 6 ◦ 7 ◦ 8 ◦ 9 ◦ 10

Please now imagine yourself in the following situations and think about what you would do.

- **Today you unexpectedly received 1,000 £. How much of this amount would you donate to a good cause?**

  ○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

- **You are in an area you are not familiar with, and you realize you lost your way. You ask a stranger for directions. The stranger offers to take you to your destination. Helping you costs the stranger about 20 £ in total. However, the stranger says he or she does not want any money from you. You have six presents with you. The cheapest present costs 5 £, the most expensive one costs 30 £. Do you give one of the presents to the stranger as a "thank-you" gift? If so, which present do you give to the stranger?**

  ○ None ○ The one worth 5 £ ○ The one worth 10 £ ○ The one worth 15 £ ○ The one worth 20 £ ○ The one worth 25 £ ○ The one worth 30 £

A.3 Experimental instructions (Experiment 2)

Instructions

In this brief survey, we will ask you some opinions about the decision task in an experiment that took place on July 27th 2020. By answering, you will have the chance to win some bonus payments. The conversion rate for these payments is 40 points = £ 1. The survey is structured as follows:

- In the next slide, we will show you the exact text of the decision that participants faced;
- In the following one, you will be asked to express your opinions;
- Then, a few questions about your attitudes will conclude the survey.

The decision in the experiment of July 27:

"In this study, you will be randomly assigned to a group of 4 participants. You are given 10 points and have to decide whether to contribute to the common pool of the group. Contributions can be any integer from 0 to 10. You keep the remaining points. The other participants face the same decision.

> The **group common pool** yields the following returns: the contributions of the 4 participants are added up and the total is multiplied by a factor of 2.4. The resulting amount is equally split among the 4 participants.

Your payoff equals your earnings from the group common pool plus the amount you keep for yourself. The final conversion will be as follows: 40 points correspond to 1 GBP."

YOUR OPINION

Remind: The total amount contributed to the group common pool was multiplied by 2.4 and divided by 4.

Consider the decision task the participants faced in the experiment. In your opinion, how much did the group members contribute to the group common pool? Please indicate in the box below what you believe was the average contribution of the group members to the group common pool. You will receive additional 4 points for each correct answer. An answer is considered to be correct if it is less than 0.50 close to the true value.

Consider again the decision task the participants faced in the experiment. How much do you believe a member of the group ought to contribute to the group common pool? Please indicate in the box below what you believe a member of the group ought to contribute to the group common pool.

In your opinion, how did the participants in the experiment answer to the previous question? Please indicate in the box below what you believe was the average answer by the participants in the experiment to the previous question. You will receive additional 4 points for each correct answer. An answer is considered to be correct if it is less than 0.50 close to the true value.

We now ask for your willingness to act in a certain way in different areas. Please indicate your answer on a scale from 0 to 10, where 0 means you are "completely unwilling to do so" and a 10 means you are "very willing to do so".

- **How willing are you to give up something that is beneficial for you today in order to benefit more from that in the future?**

  ○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

- **How willing are you to take risks?**

  ○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

- **How willing are you to punish someone who treats others unfairly, even if there may be costs for you?**

  ○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

How well does the following statement describe you as a person? Please indicate your answer on a scale from 0 to 10. A 0 means "does not describe me at all" and a 10 means "describes me perfectly".

- **I assume that people have only the best intentions.**

  ○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

Please now imagine yourself in the following situations and think about what you would do.

- **Today you unexpectedly received 1,000 £. How much of this amount would you donate to a good cause?**

  ○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

- **You are in an area you are not familiar with, and you realize you lost your way. You ask a stranger for directions. The stranger offers to take you to your destination. Helping you costs the stranger about 20 £ in total. However, the stranger says he or she does not want any money from you. You have six presents with you. The cheapest present costs 5 £, the most expensive one costs 30 £. Do you give one of the presents to the stranger as a "thank-you" gift? If so, which present do you give to the stranger?**

○ None ○ The one worth 5 £ ○ The one worth 10 £ ○ The one worth 15 £ ○ The one worth 20 £ ○ The one worth 25 £ ○ The one worth 30 £

Appendix B

Supplementary materials for Chapter 3

B.1 Additional tables

Table B.1: Number of subjects, mean (and standard deviation) of the contributions to local and global public goods, and total contributions by treatment.

Treatment	Subjects	Local contribution	Global contribution	Total contribution
T_0	164	6.13 (2.80)	-	6.13 (2.80)
T_1	160	4.56 (2.49)	2.67 (1.78)	7.23 (2.32)
T_2	164	4.35 (2.31)	3.15 (2.14)	7.50 (2.44)
T_3	160	3.63 (2.41)	4.24 (2.73)	7.87 (2.36)
T_4	154	3.18 (2.03)	4.38 (2.71)	7.56 (2.69)

Table B.2: P-value results of the Wilcoxon rank-sum test over the full sample for each pairwise comparison.

	Local contribution		**Global contribution**		**Total contribution**	
	Δ	p	Δ	p	Δ	p
T_0 vs. T_1	1.57	0.0000	–	–	-1.1	0.0000
T_0 vs. T_2	1.78	0.0000	–	–	-1.37	0.0000
T_0 vs. T_3	2.50	0.0000	–	–	-1.74	0.0000
T_0 vs. T_4	2.95	0.0000	–	–	-1.43	0.0000
T_1 vs. T_2	0.21	0.6124	-0.48	0.0502	-0.27	0.1974
T_1 vs. T_3	0.93	0.0004	-1.57	0.0000	-0.64	0.0026
T_1 vs. T_4	1.38	0.0000	-1.71	0.0000	-0.33	0.0386
T_2 vs. T_3	0.72	0.0020	-1.09	0.0003	-0.37	0.1237
T_2 vs. T_4	1.17	0.0000	-1.23	0.0000	-0.06	0.4859
T_3 vs. T_4	0.45	0.2135	-0.14	0.3700	0.31	0.4579

Notes: Δ corresponds to the difference between the mean values.

Table B.3: OLS regressions examining the contributions in the MLPGG with the inclusion of controls.

	(1) Local contribution	(2) Global contribution	(3) Total contribution
G			1.160***
			(0.334)
β	-3.206***	4.189***	0.745
	(0.609)	(0.605)	(0.642)
Altruism	0.000	-0.000	0.000
	(0.001)	(0.001)	(0.001)
Patience	0.064	0.014	0.193**
	(0.072)	(0.073)	(0.070)
Risk tolerance	-0.094	0.132*	0.024
	(0.062)	(0.060)	(0.054)
Trust	0.087	0.174***	0.235***
	(0.051)	(0.049)	(0.048)
Neg. Reciprocity	0.031	-0.027	-0.013
	(0.042)	(0.043)	(0.043)
Pos. Reciprocity	0.034	0.070	0.120*
	(0.059)	(0.064)	(0.060)
Time	0.000	-0.001	-0.001
	(0.003)	(0.003)	(0.003)
Comprehension	0.093	-0.056	0.069
	(0.125)	(0.131)	(0.121)
CRT	0.054	0.137	0.222*
	(0.093)	(0.096)	(0.088)
Age	0.008	0.011	0.011
	(0.009)	(0.009)	(0.010)
Gender	-0.146	0.109	-0.180
	(0.237)	(0.249)	(0.229)
Personal income	-0.134	-0.123	-0.174*
	(0.083)	(0.084)	(0.085)
Student status	0.555	-0.293	0.169
	(0.292)	(0.282)	(0.277)
Socioeco. status	0.022	-0.051	-0.054
	(0.067)	(0.075)	(0.064)
Education	0.241*	0.006	0.188
	(0.106)	(0.109)	(0.111)
Employment status	0.344	-0.442	0.094
	(0.254)	(0.269)	(0.245)
Constant	2.961***	0.639	2.168**
	(0.755)	(0.782)	(0.762)
T_0	✗	✗	✓
Observations	525	525	658
R^2	0.105	0.147	0.164

Notes: Columns (1)-(2) report the results over the restricted sample, which excludes observations from T_0. The regressors are β, a set of individual-specific characteristics about preferences, time spent on the task page, a score variable for correct answers to the control questions, a score variable for correct answers in the Cognitive Reflection Test, and socio-demographic characteristics. In Column (3), the same regressors are used, with the addition of the dummy G, which is equal to 1 if there is a global public good (i.e., in T_1, T_2, T_3, and T_4) and 0 otherwise (i.e., in T_0). Robust standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table B.4: Tobit models examining the contribution choices in the MLPGG.

	(1) Local contribution	(2) Global contribution	(3) Total contribution	(4) Total contribution
β	-3.862***	4.782***	1.597	3.205***
	(0.666)	(0.690)	(0.885)	(0.652)
Constant	5.284***	1.723***	7.622***	6.920***
	(0.276)	(0.252)	(0.345)	(0.227)
T_0	✗	✗	✗	✓
Observations	638	638	638	802
Pseudo R^2	0.012	0.016	0.001	0.008

Notes: In Columns (1)-(3) the sample is restricted excluding observations from T_0. Robust standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table B.5: Seemingly Unrelated Regressions (SUR) examining contribution choices in the MLPGG.

	(1) Local contribution	(2) Global contribution	(3) Total contribution	(4) Total contribution
β	-3.245***	4.158***	0.914	2.359***
	(0.550)	(0.564)	(0.583)	(0.425)
Constant	5.147***	2.052***	7.199***	6.553***
	(0.224)	(0.230)	(0.238)	(0.155)
T_0	✗	✗	✗	✓
Observations	638	638	638	802
Pseudo R^2	0.052	0.079	0.004	0.037

Notes: In Columns (1)-(3) the sample is restricted excluding observations from T_0. Robust standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table B.6: OLS regressions examining contribution choices in the MLPGG with treatment dummy variables.

	(1) Local contribution	(2) Global contribution	(3) Total contribution	(4) Local contribution	(5) Global contribution	(6) Total contribution
T_0	1.780***		-1.366***	1.829***		-1.433***
	(0.284)		(0.290)	(0.296)		(0.300)
T_1	0.203	-0.471*	-0.269	0.325	-0.598*	-0.264
	(0.267)	(0.219)	(0.264)	(0.282)	(0.241)	(0.274)
T_3	-0.722**	1.091***	0.369	-0.674*	0.931**	0.273
	(0.262)	(0.273)	(0.267)	(0.286)	(0.300)	(0.280)
T_4	-1.178***	1.237***	0.058	-1.184***	1.188***	0.016
	(0.244)	(0.275)	(0.288)	(0.273)	(0.296)	(0.313)
Constant	4.354***	3.146***	7.500***	1.565*	1.907*	3.642***
	(0.181)	(0.167)	(0.190)	(0.660)	(0.768)	(0.773)
Controls	✗	✗	✗	✓	✓	✓
Observations	802	638	802	658	525	658
R^2	0.149	0.085	0.054	0.210	0.150	0.166

Notes: Columns (1)-(3) report coefficients of models that only include treatment dummies, where the omitted (baseline) category is T_2. Columns (4)-(6) report coefficients from regressions containing also control variables. Robust standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table B.7: OLS regressions examining contribution choices in the MLPGG with a restricted sample of people who gave at least two correct answers out of the three control questions.

	(1) Local contribution	(2) Global contribution	(3) Total contribution	(4) Total contribution
β	-7.355***	5.195***	-2.160	0.680
	(1.231)	(1.277)	(1.466)	(1.005)
Constant	6.767***	1.673***	8.440***	7.179***
	(0.532)	(0.437)	(0.535)	(0.320)
T_0	✗	✗	✗	✓
Observations	182	182	182	251
R^2	0.159	0.076	0.013	0.002

Notes: The sample includes only those subjects who obtained a score at least equal to 2/3 in the control questions. Columns (1)-(3) report the results excluding observations from T_0. Robust standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table B.8: OLS regressions examining contribution choices in the MLPGG with a restricted sample of people who gave three correct answers out of three control questions.

	(1) Local contribution	(2) Global contribution	(3) Total contribution	(4) Total contribution
β	-11.425***	8.925***	-2.501	2.345
	(1.872)	(2.304)	(2.447)	(2.042)
Constant	8.615***	0.211	8.826***	6.676***
	(0.777)	(0.727)	(0.775)	(0.658)
T_0	✗	✗	✗	✓
Observations	58	58	58	79
R^2	0.338	0.208	0.017	0.020

Notes: The sample includes only those subjects who obtained a score equal to 3/3 in the control questions. Columns (1)-(3) report the results excluding observations from T_0. Robust standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

B.2 Experimental instructions

Instructions

In this study, you will be randomly assigned to a group of 4 participants. Your group will be randomly matched with another group of the same size. You are given 10 points and have to decide whether to contribute to a local common pool (the common pool of your group) and a global common pool (the common pool of both your group and the other group). Contributions can be any integer from 0 to 10. You keep the remaining points. The other participants face the same decision.
The local and the global common pools yield the following returns.

- Your **Local common pool**: The contributions of the 4 participants are added up and the total is multiplied by a factor of 2.4. The resulting amount is equally split among the 4 participants.
- **Global common pool**: The contributions of the 8 participants are added up and the total is multiplied by a factor of 1.2. The resulting amount is equally split among the 8 participants.

Your payoff equals your earnings from the local common pool, plus your earnings from the global common pool, plus the amount you keep for yourself. The final conversion will be as follows: 40 points correspond to 1 GBP.

YOUR DECISION

Please decide how to distribute your 10 points among the three options.

Your contribution to the local common pool:

Your contribution to the global common pool:

What you keep for yourself:

Remind: The total amount contributed to the local common pool will be multiplied by 2.4 and divided by 4; The total amount contributed to the global common pool will be multiplied by 1.2 and divided by 8.

In your opinion, how much the other members of your group contribute to the local common pool and to the global common pool? Please indicate in the boxes below what you believe was the average contribution of the other members of your group to the local common pool and to the global common pool. You will receive additional 4 points for each correct answer. An answer is considered to be correct if it is less than 0.50 close to the true value.

Local common pool:

Global common pool:

Consider again the decision task you faced. How much do you believe a member of your group ought to contribute to the local common pool and to the global common pool? Please indicate in the boxes below what you believe a member of your group ought to contribute to the local common pool and to the global common pool.

Local common pool:

Global common pool:

In your opinion, how have the other members of your group answered to the previous question? Please indicate in the boxes below what you believe was the average answer by the other members of your group to the previous question. You will receive additional 4 points for each correct answer. An answer is considered to be correct if it is less than 0.50 close to the true value.

Local common pool:

Global common pool:

Control questions

Please answer the following questions.

QUESTION 1: What are your total earnings if all (you included) contribute 10 to their group?

. . . points

QUESTION 2: What level of your contribution to the group earns the highest payoff for you personally if all others contribute 0 to the group?

. . . points

QUESTION 3: What level of your contribution to the group earns the highest payoff for you personally if all others contribute 10 to the group?

. . . points

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?

. . . cents

If it takes 5 minutes for five machines to make five widgets, how long would it take for 100 machines to make 100 widgets?

. . . 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?

. . . days

We now ask for your willingness to act in a certain way in different areas. Please indicate your answer on a scale from 0 to 10, where 0 means you are "completely unwilling to do so" and a 10 means you are "very willing to do so".

- **How willing are you to give up something that is beneficial for you today in order to benefit more from that in the future?**

  ○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

- **How willing are you to take risks?**

  ○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

- **How willing are you to punish someone who treats others unfairly, even if there may be costs for you?**

  ○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

How well does the following statement describe you as a person? Please indicate your answer on a scale from 0 to 10. A 0 means "does not describe me at all" and a 10 means "describes me perfectly".

- **I assume that people have only the best intentions.**

  ○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

Please now imagine yourself in the following situations and think about what you would do.

- **Today you unexpectedly received 1,000 £. How much of this amount would you donate to a good cause?**

  ○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

- **You are in an area you are not familiar with, and you realize you lost your way. You ask a stranger for directions. The stranger offers to take you to your destination. Helping you costs the stranger about 20 £ in total. However, the stranger says he or she does not want any money from you. You have six presents with you. The cheapest present costs 5 £, the most expensive one costs 30 £. Do you give one of the presents to the stranger as a "thank-you" gift? If so, which present do you give to the stranger?**

  ○ None ○ The one worth 5 £ ○ The one worth 10 £ ○ The one worth 15 £ ○ The one worth 20 £ ○ The one worth 25 £ ○ The one worth 30 £

Appendix C

Supplementary materials for Chapter 4

C.1 Additional tables

Table C.1: Mean (and standard deviation) of contribution choices by treatment.

	Country Budget	EU Budget	Total contribution
Low	4.19 (2.22)	3.24 (2.17)	7.43 (2.54)
High	3.25 (2.09)	4.47 (2.70)	7.72 (2.41)
Total	3.73 (2.21)	3.85 (2.53)	7.58 (2.48)

Table C.2: Mean (and standard deviation) of contribution choices by treatment and country.

	Country		**EU**		**Total**	
	Low	High	Low	High	Low	High
Italy	4.21 (2.26)	3.40 (1.76)	3.20 (1.91)	4.36 (2.29)	7.41 (2.30)	7.76 (2.26)
Germany	3.56 (2.13)	2.81 (1.81)	3.18 (2.42)	4.76 (2.82)	6.73 (3.10)	7.57 (2.36)
France	4.07 (2.20)	3.28 (2.34)	3.23 (2.07)	4.34 (2.93)	7.30 (2.72)	7.63 (2.79)
The Netherlands	4.50 (2.26)	2.92 (2.10)	3.14 (2.23)	4.80 (2.97)	7.63 (2.35)	7.72 (2.53)
Poland	4.33 (2.48)	3.66 (2.47)	3.37 (2.52)	4.35 (2.76)	7.70 (2.60)	8.01 (2.33)
Portugal	4.51 (1.83)	3.44 (1.87)	3.30 (1.81)	4.21 (2.38)	7.82 (1.91)	7.64 (2.17)

Table C.3: Mean (and standard deviation) of participants' characteristics by country and treatment.

		Age	Female	Stud.	Socioec. Status	Sec. Educ.	Under-grad	Grad. & Postgrad.	Migrant
FR	Low	30.00	0.55	0.32	5.38	0.15	0.20	0.60	0.25
		(10.23)	(0.50)	(0.47)	(1.46)	(0.36)	(0.40)	(0.49)	(0.43)
	High	29.83	0.45	0.37	5.60	0.19	0.14	0.62	0.28
		(9.08)	(0.50)	(0.49)	(1.55)	(0.39)	(0.35)	(0.49)	(0.45)
	p-value	0.944	0.157	0.554	0.246	0.580	0.348	0.886	0.632
DE	Low	30.45	0.46	0.43	5.73	0.36	0.26	0.27	0.33
		(9.50)	(0.50)	(0.50)	(1.56)	(0.48)	(0.44)	(0.44)	(0.47)
	High	29.32	0.54	0.50	5.47	0.37	0.28	0.31	0.23
		(9.19)	(0.50)	(0.50)	(1.48)	(0.48)	(0.44)	(0.46)	(0.43)
	p-value	0.283	0.322	0.395	0.195	1.000	0.875	0.643	0.158
IT	Low	28.70	0.50	0.54	5.71	0.43	0.20	0.35	0.08
		(9.25)	(0.50)	(0.50)	(1.50)	(0.50)	(0.40)	(0.48)	(0.27)
	High	29.12	0.50	0.45	5.74	0.46	0.21	0.28	0.06
		(8.64)	(0.50)	(0.50)	(1.37)	(0.50)	(0.40)	(0.45)	(0.23)
	p-value	0.540	1.000	0.258	0.758	0.776	1.000	0.361	0.783
NL	Low	27.89	0.46	0.43	5.81	0.32	0.37	0.27	0.23
		(8.06)	(0.50)	(0.50)	(1.66)	(0.47)	(0.49)	(0.45)	(0.42)
	High	27.83	0.53	0.45	5.68	0.33	0.39	0.27	0.37
		(6.86)	(0.50)	(0.50)	(1.71)	(0.47)	(0.50)	(0.44)	(0.48)
	p-value	0.832	0.396	0.888	0.653	1.000	0.885	1.000	0.045
PL	Low	26.96	0.54	0.53	5.10	0.43	0.25	0.24	0
		(8.43)	(0.50)	(0.50)	(1.51)	(0.50)	(0.44)	(0.43)	(0)
	High	25.98	0.45	0.60	5.38	0.51	0.22	0.19	0.01
		(8.42)	(0.50)	(0.49)	(1.33)	(0.50)	(0.41)	(0.39)	(0.10)
	p-value	0.177	0.258	0.394	0.240	0.321	0.618	0.390	0.485
PT	Low	27.59	0.52	0.43	5.54	0.26	0.27	0.40	0.05
		(7.89)	(0.50)	(0.50)	(1.36)	(0.44)	(0.45)	(0.49)	(0.22)
	High	28.22	0.50	0.57	5.48	0.26	0.28	0.44	0.79
		(8.92)	(0.50)	(0.50)	(1.25)	(0.44)	(0.45)	(0.50)	(0.27)
	p-value	0.943	0.779	0.066	0.691	1.000	1.000	0.670	0.568

Notes: *Age* is the age of the participant at the time of the study. *Female* is a dummy variable that equals 1 if the participant is female. *Student* is a dummy variable that equals 1 if the participant is student. *Migrant* is a dummy variable that equals 1 if the participant was not born in the country of residence. *Socioeconomic status* measures the self-reported place occupied by the participant on a ladder representing society that goes from 1 to 10. *Secondary education* is a dummy variable that equals 1 if the participant holds a high school diploma or equivalent. *Undergraduate* is a dummy variable that equals 1 if the participant holds an undergraduate degree. *Graduate and post-graduate* is a dummy variable that equals 1 if the participant holds a graduate or doctorate degree.

Table C.4: Mean (and standard deviation) of the answers to the post-experimental questionnaire by country.

	France	**Germany**	**Italy**	**Netherlands**	**Poland**	**Portugal**
Feel Country	3.00	2.36	3.29	2.52	3.23	3.38
	(1.22)	(1.23)	(0.96)	(1.30)	(0.99)	(0.85)
Feel EU	2.99	2.87	3.28	2.62	3.29	3.19
	(1.06)	(1.04)	(0.87)	(1.03)	(0.89)	(0.77)
EU Image	2.88	2.88	2.96	2.87	2.94	2.96
	(0.78)	(0.81)	(0.84)	(0.85)	(0.85)	(0.64)
Before COVID	2.72	2.79	3.00	2.44	3.27	3.02
	(0.97)	(0.90)	(0.97)	(1.02)	(0.76)	(0.83)
Solidarity	2.50	2.37	2.79	2.36	2.48	2.87
	(0.96)	(0.87)	(0.82)	(0.90)	(1.00)	(0.83)
After COVID	1.33	1.56	1.85	1.59	1.95	2.03
	(0.91)	(0.92)	(1.05)	(0.90)	(0.81)	(1.10)
Military Spending	4.25	4.20	3.87	4.96	5.03	4.70
	(2.65)	(2.54)	(2.49)	(2.52)	(2.74)	(2.29)
National Army	1.91	2.29	1.67	2.13	2.66	2.19
	(1.16)	(1.14)	(1.05)	(1.01)	(1.02)	(0.98)
EU Army Pre-War	1.82	1.69	1.85	2.08	1.91	1.86
	(1.25)	(1.26)	(1.29)	(1.13)	(1.02)	(1.03)
EU Army Post-War	2.15	2.13	2.33	2.35	2.38	2.57
	(1.34)	(1.17)	(1.26)	(1.18)	(1.18)	(0.91)

Table C.5: Kruskal–Wallis tests on the answers to the post-experimental questionnaire by country.

Variable	χ^2	**p**
Feel EU	78.968	< 0.001
Feel Country	136.374	< 0.001
Image EU	4.830	0.4370
Before COVID	96.000	0.001
Solidarity	67.127	< 0.001
After COVID	69.788	< 0.001
Military Spending	32.916	< 0.001
National Army	950148	< 0.001
EU Army Pre-War	12.691	0.0265
EU Army Post-War	16.396	0.0058

Table C.6: Wilcoxon rank-sum tests results of pairwise comparisons of answers to the post-experimental questionnaire between countries.

Country		Country	z	p-value
		Feel EU		
Germany	vs.	Italy	-4.350	$p < 0.001$
Germany	vs.	Poland	-4.474	$p < 0.001$
The Netherlands	vs.	Italy	7.148	$p < 0.001$
The Netherlands	vs.	France	4.039	p=0.002
The Netherlands	vs.	Poland	-7.213	$p < 0.001$
The Netherlands	vs.	Portugal	-5.906	$p < 0.001$
		Feel Country		
Germany	vs.	Italy	-8.215	$p < 0.001$
Germany	vs.	France	5.622	$p < 0.001$
Germany	vs.	Poland	-8.882	$p < 0.001$
Germany	vs.	Portugal	-4.474	$p < 0.001$
The Netherlands	vs.	Italy	6.557	$p < 0.001$
The Netherlands	vs.	France	4.131	$p < 0.001$
The Netherlands	vs.	Poland	-5.992	$p < 0.001$
The Netherlands	vs.	Portugal	-7.161	$p < 0.001$
		Before COVID		
Poland	vs.	France	-5.998	$p < 0.001$
Poland	vs.	Germany	-5.596	$p < 0.001$
Poland	vs.	The Netherlands	-8.337	$p < 0.001$
Poland	vs.	Portugal	3.245	$p < 0.001$
Portugal	vs.	France	-3.222	p=0.020
Portugal	vs.	The Netherlands	-6.054	$p < 0.001$
Italy	vs.	The Netherlands	5.912	$p < 0.001$
Italy	vs.	France	-3.359	p=0.012
		Solidarity		
Italy	vs.	Germany	-5.033	$p < 0.001$
Italy	vs.	The Netherlands	4.920	$p < 0.001$
Portugal	vs.	Germany	-6.468	$p < 0.001$
Portugal	vs.	France	-4.226	$p < 0.001$
Portugal	vs.	The Netherlands	-6.309	$p < 0.001$
Portugal	vs.	Poland	-4.304	$p < 0.001$
		Military Spending		
Germany	vs.	Poland	-3.066	p=0.033
Germany	vs.	The Netherlands	-3.005	p=0.044
Italy	vs.	Poland	-4.214	$p < 0.001$
Italy	vs.	Portugal	-3.471	p=0.008
Italy	vs.	The Netherlands	-4.191	$p < 0.001$
Poland	vs.	France	-3.057	p=0.033
		National Army		
Italy	vs.	Germany	5.399	$p < 0.001$
Italy	vs.	Poland	-8.861	$p < 0.001$
Italy	vs.	Portugal	-4.874	p=0.008
Italy	vs.	The Netherlands	-4.258	$p < 0.001$
Germany	vs.	France	-3.387	p=0.011
Germany	vs.	Poland	-3.333	p=0.014
Poland	vs.	France	-6.563	$p < 0.001$
Poland	vs.	The Netherlands	-5.333	$p < 0.001$
Poland	vs.	Portugal	4.888	$p < 0.001$
		EU Army Pre-War		
Germany	vs.	The Netherlands	-3.279	p=0.015
		EU Army Post-War		
Germany	vs.	Portugal	-3.960	p=0.002

Table C.7: Post-estimation tests on the equality of coefficients.

Specification	(1) Country	(2) EU	(3) Total	(4) Country	(5) EU	(6) Total
DE vs. IT	0.0354	0.9388	0.0790	0.5563	0.2738	0.7357
DE vs. FR	0.0933	0.8710	0.1700	0.0786	0.5664	0.3289
DE vs. PL	0.0171	0.5771	0.0161	0.8679	0.7327	0.6376
DE vs. PT	0.0007	0.6743	0.0028	0.3172	0.5152	0.1355
IT vs. FR	0.6551	0.9213	0.7507	0.2313	0.6789	0.2097
IT vs. PL	0.7186	0.5900	0.4023	0.4631	0.5268	0.9152
IT vs. PT	0.2956	0.6967	0.1742	0.1008	0.6090	0.3299
FR vs. PL	0.4273	0.6613	0.2812	0.0602	0.8346	0.1632
FR vs. PT	0.1194	0.7840	0.1166	0.0027	0.9687	0.0210
PL vs. PT	0.5457	0.8307	0.7103	0.4357	0.8387	0.3395
High×DE vs. High×IT	0.8861	0.3688	0.3363	0.7874	0.9823	0.7999
High×DE vs. High×FR	0.9368	0.3607	0.3570	0.8922	0.3989	0.3704
High×DE vs. High×PL	0.8541	0.2502	0.3151	0.7495	0.6604	0.8990
High×DE vs. High×PT	0.3935	0.1524	0.0372	0.6371	0.5534	0.3304
High×IT vs. High×FR	0.9564	0.9245	0.9673	0.8840	0.4577	0.5548
High×IT vs. High×PL	0.7576	0.7094	0.9351	0.9530	0.6742	0.7186
High×IT vs. High×PT	0.4877	0.5462	0.2269	0.4652	0.5832	0.2505
High×FR vs. High×PL	0.8070	0.7961	0.9726	0.8410	0.2498	0.3233
High×FR vs. High×PT	0.4798	0.6523	0.2993	0.5323	0.1788	0.0711
High×PL vs. High×PT	0.3498	0.8729	0.2841	0.4508	0.9251	0.4146

Table C.8: OLS regressions examining the contribution decisions of the subsample of migrants to the Country Budget, to the EU Budget, and the sum of contributions to both budgets.

	(1) Country	(2) EU	(3) Total
High	-1.179***	0.611	-0.568
	(0.293)	(0.346)	(0.405)
DE	-0.538	0.467	-0.071
	(0.403)	(0.448)	(0.539)
FR	-0.760	0.839	0.079
	(0.422)	(0.505)	(0.589)
IT	0.991	0.181	1.171
	(0.525)	(0.595)	(0.736)
PT	0.061	1.594	1.654
	(0.686)	(0.920)	(0.917)
Migrant EU	-0.793*	0.383	-0.410
	(0.375)	(0.490)	(0.513)
Age of moving	-0.012	0.025	0.013
	(0.022)	(0.026)	(0.026)
Female	0.342	-0.338	0.005
	(0.310)	(0.367)	(0.423)
Student	-0.145	0.416	0.272
	(0.303)	(0.380)	(0.426)
Socioeconomic Status	-0.071	0.005	-0.066
	(0.107)	(0.139)	(0.146)
Education	-0.304	-0.091	-0.395
	(0.163)	(0.211)	(0.212)
Feel Country	-0.146	-0.135	-0.281
	(0.143)	(0.193)	(0.188)
Feel EU	0.408*	0.254	0.662**
	(0.175)	(0.188)	(0.222)
EU Image	0.211	0.168	0.379
	(0.203)	(0.278)	(0.308)
Before COVID	0.151	0.390	0.541
	(0.190)	(0.219)	(0.284)
Solidarity	-0.294	-0.215	-0.510
	(0.223)	(0.233)	(0.259)
After COVID	0.059	-0.308	-0.249
	(0.173)	(0.198)	(0.226)
Military Spending	-0.043	0.002	-0.040
	(0.072)	(0.088)	(0.114)
National Army	0.041	0.216	0.257
	(0.199)	(0.227)	(0.253)
EU Army Pre-war	0.158	-0.211	-0.053
	(0.211)	(0.211)	(0.247)
EU Army Post-War	-0.251	0.259	0.008
	(0.214)	(0.223)	(0.249)
Constant	5.382***	0.812	6.194***
	(1.321)	(1.489)	(1.491)
Observations	194	194	194
R^2	0.218	0.184	0.205

Notes: Baseline category for treatment dummies is *Low*. Baseline category for country dummies is NL. Robust standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

C.2 List of variables from the post-experimental questionnaire

All the questions included in the post-experimental questionnaire are reported below with the corresponding label (e.g., Q1) and variable name (in parenthesis). Besides Q1-Q4 and Q11, all answers are based on a 0-5 scale.

Migration

Q 1 (Migrant) *Were you born in (country of residence)?*

Q 2 (Age of Migration) *How old were you when you moved to (country of residence)?*

Q 3 (Mother Country) *In which country was you mother born?*

Q 4 (Father Country) *In which country was you father born?*

Feelings

Q 5 (Feel Country) *How strongly do you feel (country of residence)?*

Q 6 (Feel EU) *How strongly do you feel an EU citizen?*

Q 7 (EU Image) *In general, does the EU conjure up for you a very positive, fairly positive, neutral, fairly negative or very negative image?*

COVID-19

Q 8 (Before COVID) *Before Coronavirus pandemic, would you say that (country of residence) has on balance benefited from being a member of the EU?*

Q 9 (Solidarity) *How satisfied are you with the solidarity between the EU Member States in fighting the Coronavirus pandemic?*

Q 10 (After COVID) *Has your opinion on the benefits for (country of residence) from being a member of the EU changed after the Coronavirus pandemic?*

War in Ukraine

Q 11 (Military Spending) *Each person has no choice but to consume the service of the national defence. For those who believe increasing public expenditures on national defence makes them safer, an increase in these expenditures is positive. Others think additional expenditures on armies only lead to arms races and decrease national security. Such individuals value additional public expenditures on national defence negatively. On a scale from 0 to 10, how much do you consider belonging to the first group?*

Q 12 (National Army) *After the beginning of the Russian-Ukrainian war, do you think your country (of residence) should increase its public expenditures on the army?*

Q 13 (EU Army Pre-War) *Before the Russian-Ukrainian war, have you ever thought that the EU should have an army financed with the EU budget?*

Q 14 (EU Army Post-War) *After the Russian-Ukrainian war, do you think the EU should get an army and finance it with an EU budget?*

C.3 Experimental instructions

This appendix reports the English instructions we used for the Low treatment with Italian residents. The instructions for the $High$ treatment and other countries were adapted accordingly and are available upon request.

You have been selected to take part in this study since you declared on Prolific.co that you are an Italian resident.

Are you still an Italian resident?

○ Yes ○ No

Do you prefer to read the following instructions in Italian or in English?

○ Italian ○ English

Instructions 1/2

In this study, you will be firstly asked to make a decision. Depending on your decision and on the decisions made by other participants, who face the same decision, you will have the opportunity to get some bonus payments.

After this decision, you will be asked to fill in a short questionnaire. You will receive any bonus payment only after the questionnaire is completed.

All amounts will be expressed in Points rather than pound sterling. The exchange rate is 10 Points = 0.25.

Instructions 2/2

You are randomly assigned to a group of 4 including you and your fellow citizens.

Your group is randomly matched with other two groups of the same size, making up an overall set of 12 participants. Each of these two groups is composed of people belonging to the same country selected from a group of 5 European Union (EU) countries members.

...

...

You are given 10 Points and have to decide how much to contribute to your country public budget (the fund of the group with your fellow citizens) and to the EU public budget (the fund of both your group and the other two groups).

- Your **country public budget** yields the following return: the contributions of the 4 participants are added up and the total is multiplied by 2.4. The resulting amount is equally split among the 4 participants.
- The **EU public budget** yields the following return: the contributions of the 12 participants are added up and the total is multiplied by 2.4. The resulting amount is equally split among the 12 participants.

You keep the Points you do not wish to contribute to the two public budgets. Consequently, your bonus payments equal your earnings from your country budget, plus your earnings from the EU public budget, plus the amount you keep for yourself.

Control questions

Please answer the following questions. You will be allowed to go on, only after you correctly respond to both of them.

QUESTION 1: How much do you need to contribute to your country public

budget/the EU public budget to earn the highest payoff for you personally if all others contribute 0 to your country public budget/the EU public budget?

○ 10 ○ 0 ○ 5

QUESTION 2: How much do you need to contribute to your country public budget/the EU public budget to allow your fellow citizens/all participants to earn the highest payoff if all them contribute 10 to your country public budget/the EU public budget?

○ 10 ○ 0 ○ 5

YOUR DECISION

Please decide how to distribute your 10 Points among the three options (please enter an integer number from 0 to 10, i.e. 0, 1, 2, . . . , 9, 10).

Your contribution to your country public budget:

Your contribution to the European Union public budget:

What you keep for yourself:

Remind: The total amount contributed to your country public budget will be multiplied by 2.4 and divided by 4; The total amount contributed to the European Union public budget will be multiplied by 2.4 and divided by 12.

And now, just a few questions about you and your opinions. There are no wrong or correct answers. Please answer with honesty.

- Were you born in Italy?
  ○ Yes ○ No
- How old were you when you moved to Italy? ☐
  [if "No" to previous question]
- In which country was your mother born? ☐
- In which country was your father born? ☐

- **How strongly do you feel Italian?**
  ○ Not at all strongly ○ Not very strongly ○ Neutral ○ Fairly strongly ○ Very strongly
- **How strongly do you feel an EU citizen?**
  ○ Not at all strongly ○ Not very strongly ○ Neutral ○ Fairly strongly ○ Very strongly
- **In general, does the EU conjure up for you a very positive, fairly positive, neutral, fairly negative or very negative image?**
  ○ Very negative ○ Fairly negative ○ Neutral ○ Fairly positive ○ Very positive

---

- **Before Coronavirus pandemic, would you say that Italy has on balance benefited from being a member of the EU?**
  ○ Strongly agree ○ Agree ○ Neither agree nor disagree ○ Disagree ○ Strongly disagree

- **How satisfied are you with the solidarity between the EU Member States in fighting the Coronavirus pandemic?**
  ○ Very satisfied ○ Fairly satisfied ○ Not very satisfied ○ Not at all satisfied ○ Don't know

- **Has your opinion on the benefits for Italy from being a member of the EU changed after the Coronavirus pandemic?**
  ○ Strongly agree ○ Agree ○ Neither agree nor disagree ○ Disagree ○ Strongly disagree

---

- **Each person has no choice but to consume the service of the national defense. For those who believe increasing public expenditures on national defense makes them safer, an increase in these expenditures is positive. Others think additional expenditures on armies only lead to arms races and decrease national security. Such individuals value additional public expenditures on national defense negatively.**
  On a scale from 0 to 10, how much do you consider belonging to the first group?

  ○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

- **After the beginning of the Russian-Ukrainian war, do you think your country should increase its public expenditures on the army?**
  ○ Strongly agree ○ Agree ○ Neither agree nor disagree ○ Disagree ○ Strongly disagree

- **Before the Russian-Ukrainian war, have you ever thought that the EU should have an army financed with the EU budget?**

○ Strongly agree ○ Agree ○ Neither agree nor disagree ○ Disagree ○ Strongly disagree

- **After the Russian-Ukrainian war, do you think the EU should get an army and finance it with an EU budget?**
  ○ Strongly agree ○ Agree ○ Neither agree nor disagree ○ Disagree ○ Strongly disagree

---

Appendix D

Supplementary materials for Chapter 5

D.1 Additional tables

Table D.1: Wilcoxon rank-sum tests results over the contribution variable.

	z	p-value
C vs. PCR	-0.861	0.3894
C vs. IR	-0.277	0.7821
C vs. NCR	-1.387	0.1653
PCR vs. IR	0.575	0.5654
PCR vs. NCR	-0.534	0.5936
IR vs. NCR	-1.052	0.2930

Table D.2: Hurdle model examining contribution choices in the PGG.

	(1)	(2)
Logit		
IR	0.000	0.065
	(0.300)	(0.311)
PCR	-0.144	0.018
	(0.310)	(0.327)
NCR	0.128	0.301
	(0.293)	(0.310)
Constant	-2.512***	2.658*
	(0.212)	(1.073)
Poisson		
IR	0.007	0.001
	(0.037)	(0.035)
PCR	0.019	-0.003
	(0.036)	(0.034)
NCR	0.056	0.037
	(0.035)	(0.034)
Constant	2.544***	2.134***
	(0.025)	(0.128)
Controls	✗	✓
Observations	1280	1280

Notes: Robust standard errors are in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

D.2 Simulations on best-response analysis

We perform a series of numerical calculations using Python code based on a best-response analysis. In these calculations, we explore expected utility models with the following social preferences: (i) inequality aversion (IA) in the model suggested by Fischbacher, Schudy, et al., 2014 and based on the Equity, Reciprocity, and Competition (ERC) framework by Bolton and Ockenfels, 2000; (ii) IA in the model proposed by Fehr and Schmidt, 1999; (iii) altruism as in the model put forward by Anderson et al., 1998 and Andreoni and Miller, 2002. We use the following general form of utility function:

$$U_i = \alpha_U(\pi_i)^{1/\gamma} + \beta_U(sign)(Function)^{1/\delta} \tag{D.1}$$

where $sign \in \{+, -\}$ is the sign of the Social Preference and $Function$ provides the functional specification of the Social Preference core, including its specific parameters as proposed in the original contributions:

- α_U and β_U are specific coefficients for, respectively, the payoff and the social preference functions;
- γ and δ express the risk preference of agent i with respect to the relevant part of the utility (i.e., his own payoff or his social preferences): risk averse when > 1, risk-neutral when $= 1$, and risk seeker when < 1.

We consider agent i best replying to the other agents, considering the remaining agents $(N-1)$ to be homogeneous with respect to their contribution levels, each contributing $\bar{c}_{-i}$. We also assume that agent i anticipates all the possible realized distributions of the losses within the group induced by the bad luck, denoting such distributions with $D_1, \ldots, D_K$ along with their respective probabilities denoted $Pr(D_1), \ldots, Pr(D_K)$. The resulting expected utility of agent i that we used for our calculations is given, for choosing c_i given $\bar{c}_{-i}$, by:

$$EU_i(c_i \mid \bar{c}_{-i}) = \sum_{k=1}^{K} Pr(D_k) U_i(c_i \mid \bar{c}_{-i}, D_k) \quad . \tag{D.2}$$

In the following table, we provide a summary of the main results of the numerical calculations, which are reported graphically at the end of this section, based on the best response functions for each treatment. In most cases, we find no appreciable difference between treatments, and when present, differences are very small, if not negligible. Moreover, these slight differences do not seem to be systematic.

Table D.3: Summary of simulations' results from best-response analyses.

#	SP	Conditions	C	PCR	NCR	IR
1	IA-ERC	$\nu < 250000$	0	=	=	=
2	IA-ERC	$35000 < \nu < 400000 + RA$	45°	=	$\uparrow +1$	=
3	IA-ERC	$250000 < \nu < 10000000$	0 + 45°	$\sim =$	=	=
4	IA-ERC	$\nu > 550000 + RA$	45°	=	$\uparrow +1$	$\uparrow +1$
5	Altruism	g=0, $\alpha < 0.49$	0	=	=	=
6	Altruism	g=0, $\alpha < 0.49 + RA$	0	=	=	=
7	Altruism	g=0, $\alpha >= 0.49$	20	=	=	=
8	IA-FS	$\alpha = 0.94, \beta = 0.94$	0	=	=	=
9	IA-FS	$\alpha = 0.96, \beta = 0.96$	45°	=	$\downarrow 0$	$\downarrow 0$
10	IA-FS	$\alpha = 0.96, \beta = 0.96 + RA$	45°	=	=	=

Notes: Column *SP* indicates the Social Preference considered. Column *Conditions* summarizes the relevant parameter utilized (with *RA* if risk aversion is present, i.e., $\gamma = \delta > 1$). Column *C* reports the trend of the best response function for the *Control* treatment. 0 indicates that the function lies on the x axis. 20 indicates that it is a flat line in correspondence of a contribution of 20. 45° indicates that it coincides with the bisector. Columns *PCR*, *NCR*, *IR* report the trend of the best response functions for the respective treatment in comparison to *C* results. $=$ indicates identical functions. $\downarrow 0$ indicates that in the considered treatment, the function detaches from the *C* and is a flat line in correspondence of the x axis. $\uparrow +1$ indicates that it detaches upward from *C* by about 1 point.

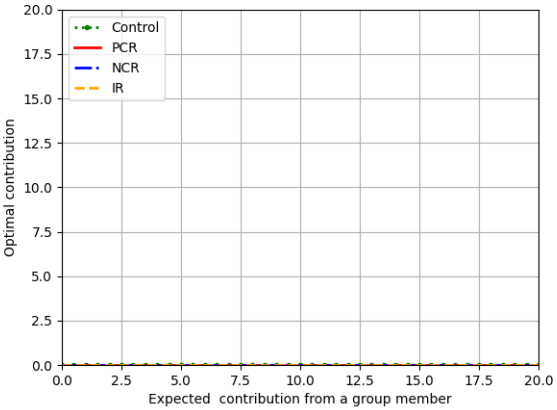

Figure D.1: Best-response analysis, simulation #1 results.

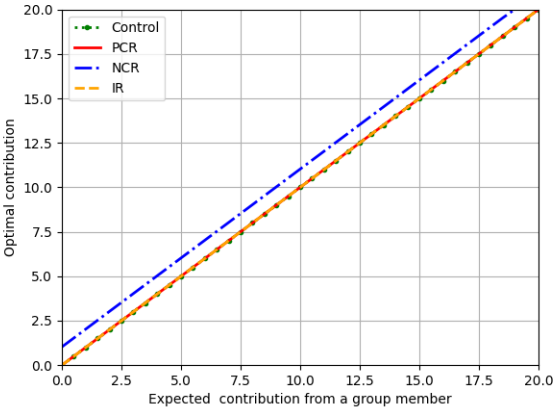

Figure D.2: Best-response analysis, simulation #2 results.

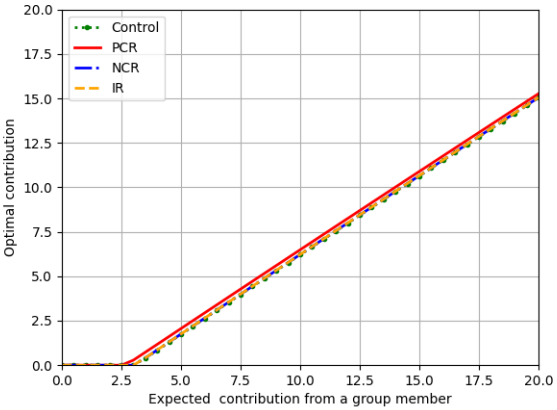

Figure D.3: Best-response analysis, simulation #3 results.

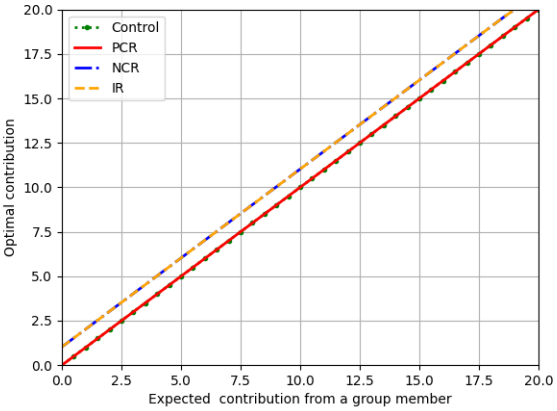

Figure D.4: Best-response analysis, simulation #4 results.

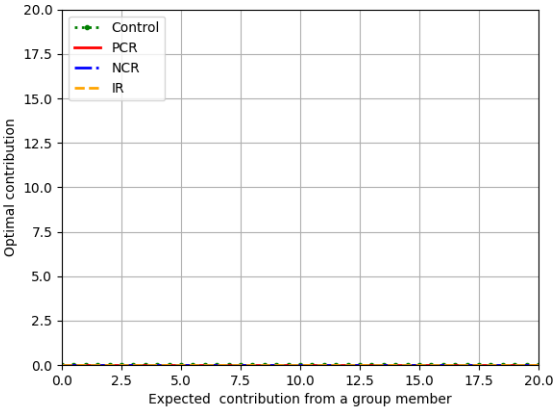

Figure D.5: Best-response analysis, simulation #5 results.

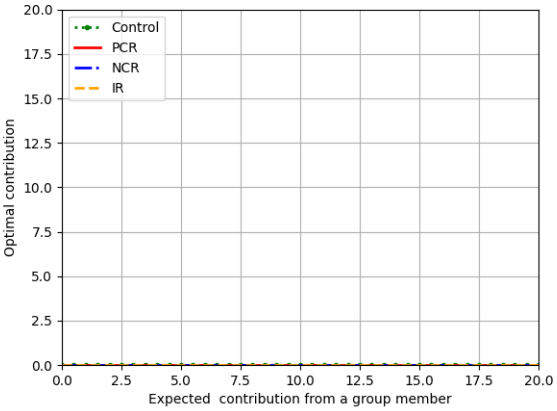

Figure D.6: Best-response analysis, simulation #6 results.

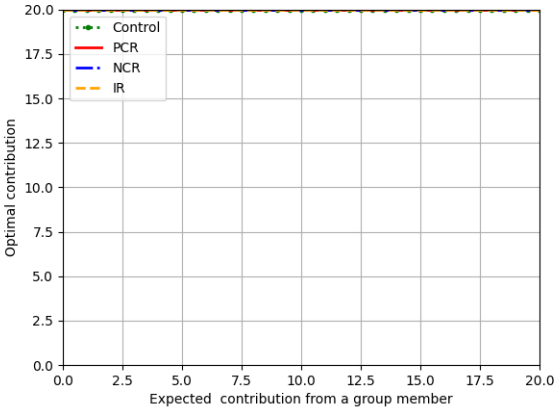

Figure D.7: Best-response analysis, simulation #7 results.

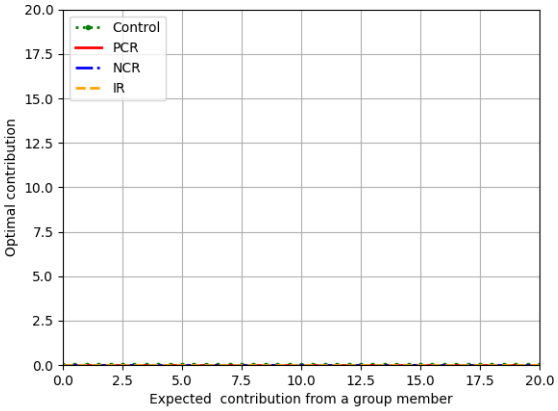

Figure D.8: Best-response analysis, simulation #8 results.

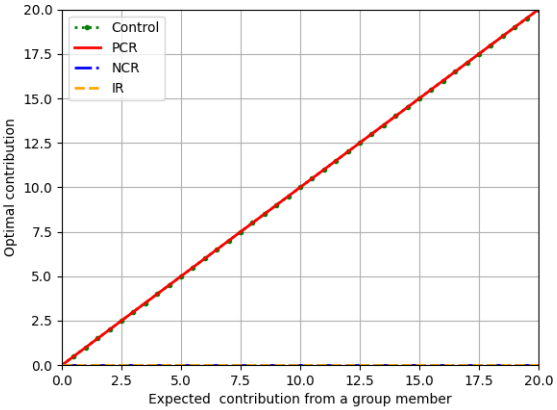

Figure D.9: Best-response analysis, simulation #9 results.

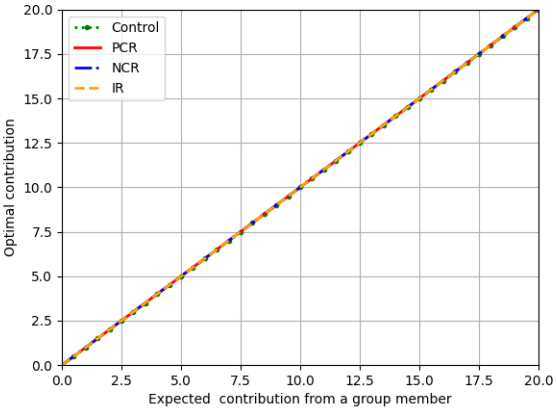

Figure D.10: Best-response analysis, simulation #10 results.

D.3 Experimental instructions

Group formation and exchange rate

In this study, you will be placed in a group of 40 people. The group will be randomly formed. Nobody will ever learn the identity of the other members of the group. In this study all amounts will be expressed in Points rather than pounds. The exchange rate is 10 Points = £ 0.20.

Decisions

You (as well as the other members of your group) will be endowed with 60/59 Points. You have to decide how many of the Points that you have you want to contribute to a project that yields Points for you as well as for the other group members. More specifically, the sum of contributions that you and your group members make to the project is multiplied by 2 (return from the contribution in the public project), and then divided by 40 (number of members in the group). Your contribution can be any integer number between 0 and 20 Points (i.e., 0, 1, ..., 20). The Points that you do not contribute you keep (they are your own and yield income just for you).

Your earnings

Your earnings are calculated as the sum of:

a) "Points from the project" = sum of contributions to the project made by you and your group members, multiplied by (2/40 =) 0.05;

b) "Points that you keep" = 60/59 minus your contribution to the project.

...

...
The calculation of the other group members' earnings will be completely similar.
When you have finished reading the above Instructions, please press 1 on your keyboard instead of clicking NEXT.

[Independent Risk treatment:]

Risk of negative event on each member of the group

There is the risk that 40 Points are deducted from the earnings calculated above. To determine whether to deduct the 40 Points, the computer will randomly select an integer number between 1 and 40 (i.e., 1, 2, ..., 40). If the selected number is equal to 1, the 40 Points will be deducted from the earnings; if the selected number is between 2 and 40, the earnings will remain unchanged. The computer will select a number for EACH member of the group. Consequently, the 40 Points will be deducted from the earnings of none, some, or all members of the group.

[Positively Correlated Risk treatment:]

Risk of negative event on all members of the group

There is the risk that 40 Points are deducted from the earnings calculated above. To determine whether to deduct the 40 Points, the computer will randomly select an integer number between 1 and 40 (i.e., 1, 2, ...,40). If the selected number is equal to 1, the 40 Points will be deducted from the earnings; if the selected number is between 2 and 40, the earnings will remain unchanged. The computer will select a number for ALL members of the group. Consequently, the 40 Points will be deducted from the earnings of none or all members of the group.

[Negatively Correlated Risk treatment:]

Risk of negative event on one member of the group

There is the risk that 40 Points are deducted from the earnings calculated above. The 40 Points will be deducted from the earnings of ONE member of the group. This member will be randomly selected by the computer from the 40 people in the group.

The following examples and control questions should help you test your understanding of the decision task. Once you have answered all questions correctly, the task will start.

A copy of the Instructions is reported at the bottom of the page.

- **EXAMPLE 1:** Suppose that you contribute 0 Points to the project and that the other 39 group members contribute 20 Points, then the sum of contributions is (0 + 39 × 20 =) 780 Points and the individual earnings from the project amount to (0.05 × 780 =) 39 Points. Furthermore, you keep for yourself (60/59 − 0 =) 60/59 Points. It follows that, if the negative event does not hit you, your total earnings are (39 + 60/59 =) 99/98 Points.
  - How many Points do you earn from the project? ☐
  - How many Points do you keep? ☐
  - How many Points do you earn in total if the negative event does not hit you ? ☐

- **EXAMPLE 2:** Suppose that you and the other 39 group members contribute 10 Points to the project, then the sum of contributions is (40 × 10 =) 400 Points and the individual earnings from the project amount to (0.05 × 400 =) 20 Points. Furthermore, you keep for yourself (60/59 − 10 =) 50/49 Points. It follows that, if the negative event does not hit you , your total earnings are (20 + 50/49 =) 70/69 Points.
  - How many Points do you earn from the project? ☐
  - How many Points do you keep? ☐
  - How many Points do you earn in total if the negative event does not hit you ? ☐

- **EXAMPLE 3:** Suppose that you contribute 20 Points to the project and that the other 39 group members contribute 0 Points, then the sum of contributions is (20 + 39 × 0 =) 20 Points and the individual earnings from the project amount to (0.05 × 20 =) 1 Point. Furthermore, you keep for yourself (60/59 − 20 =) 40/39 Points. It follows that, if the negative event does not hit you , your total earnings are (1 + 40/39 =) 41/40 Points.
  - How many Points do you earn from the project? ☐
  - How many Points do you keep? ☐
  - How many Points do you earn in total if the negative event does not hit you ? ☐

- **The negative event (i.e., deduction of 40 Points):** *[Only in the treatments with risk]*
  - ○ can independently affect each member of the group.
  - ○ surely affects none or all members of the group.
  - ○ surely affects only one member of the group.

- can be avoided if all members of the group contribute to the public project.
- can be avoided if all members of the group do not contribute to the public project.

YOUR DECISION

Please choose how many of the 60/59 Points that you have been endowed with
you want to contribute to the project.

Recall that you can contribute any integer number between 0 and 20 Points (i.e., 0, 1, ..., 20).

GUESS OTHERS' CHOICES

We now ask you to guess the average contribution of your group members. You can earn an extra amount of money depending on how close your estimate is to the actual average contribution of the other group members. If your estimate is exactly right or not more than 0.5 Points away from the actual average contribution, you will earn 10 Points. Otherwise, you will earn 0 Points.

In your opinion, what is the average contribution of your group members? You can insert any number (with two digits) between 0 and 20.

A SHORT QUESTIONNAIRE ABOUT YOU

We kindly ask you to answer some questions about yourself.

Most of the questions are descriptive,
and your responses are completely confidential.

Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please indicate your answer on a scale from 0 to 10, where 0 means "unwilling to take risks" and 10 means "fully prepared to take risks".
○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

How well do the following statements describe you as a person? Please indicate your answer on a scale from 0 to 10, where 0 means "does not describe me at al" and a 10 means "describes me perfectly".

When someone does me a favor I am willing to return it.
○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

I assume that people have only the best intentions.
○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

We now ask for your willingness to act in a certain way in a specific area. Please indicate your answer on a scale from 0 to 10, where 0 means "completely unwilling to do so" and 10 means "very willing to do so".

How willing are you to give to good causes without expecting anything in return?
○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

We now ask you to make 6 different decisions. Each decision implies a choice between two options:

- OPTION A gives you a 50% chance to win 6 and a 50% chance to lose an amount x, and
- OPTION B gives you nothing with certainty.

Please make your 6 decisions, choosing each time your preferred option.

	Option A	Option B	Decision
1	50% chance to win 6, 50% chance to lose 2	0 for sure	A ○ ○ B
2	50% chance to win 6, 50% chance to lose 3	0 for sure	A ○ ○ B
3	50% chance to win 6, 50% chance to lose 4	0 for sure	A ○ ○ B
4	50% chance to win 6, 50% chance to lose 5	0 for sure	A ○ ○ B
5	50% chance to win 6, 50% chance to lose 6	0 for sure	A ○ ○ B
1	50% chance to win 6, 50% chance to lose 7	0 for sure	A ○ ○ B

We now ask you to answer these last two questions.

- Two fair six-sided dice are rolled. What is the probability that their sum is exactly equal to 2?
  ○ 1/3

- 1/6
- 1/18
- 1/36

- Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. Which is more probable?
  - Linda is a bank teller.
  - Linda is a bank teller and is active in the feminist movement.

Bibliography

Abraham, Diya Elizabeth, Luca Corazzini, Miloš Fišar, and Tommaso Reggiani (2021). "Delegation and overhead aversion with multiple threshold public goods". In: *MUNI ECON Working Paper*.

Alekseev, Aleksandr, Gary Charness, and Uri Gneezy (2017). "Experimental methods: When and why contextual instructions are important". In: *Journal of Economic Behavior & Organization* 134, pp. 48–59.

Ambrosanio, Flavia and Massimo Bordignon (2015). "Normative versus positive theories of revenue assignments in federations". In: *Handbook of Multilevel Finance*. Edward Elgar Publishing, pp. 231–263.

Anderson, Simon P, Jacob K Goeree, and Charles A Holt (1998). "A theoretical analysis of altruism and decision error in public goods games". In: *Journal of Public Economics* 70.2, pp. 297–323.

Andreoni, James and John Miller (2002). "Giving according to GARP: An experimental test of the consistency of preferences for altruism". In: *Econometrica* 70.2, pp. 737–753.

Arechar, Antonio A, Simon Gächter, and Lucas Molleman (2018a). "Conducting interactive experiments online". In: *Experimental Economics* 21, pp. 99–131.

— (2018b). "Conducting interactive experiments online". In: *Experimental Economics* 21.1, pp. 99–131.

Artinger, Florian, Nadine Fleischhut, M Vittoria Levati, and Jeffrey R Stevens (2012). "Cooperation in a risky environment: Decisions from experience in a stochastic social dilemma". In: *Jena Economic Research Papers*.

Barcelo, Hélene and Valerio Capraro (2015). "Group size effect on cooperation in one-shot social dilemmas". In: *Scientific Reports* 5.1, pp. 1–8.

Barrett, Scott and Astrid Dannenberg (2017). "Tipping versus cooperating to supply a public good". In: *Journal of the European Economic Association* 15.4, pp. 910–941.

Bašić, Zvonimir and Eugenio Verrina (2021). "Personal norms—and not only social norms—shape economic behavior". In: *MPI Collective Goods Discussion Paper* 2020/25.

Battaglini, Marco, Salvatore Nunnari, and Thomas R Palfrey (2020). "The political economy of public debt: A laboratory study". In: *Journal of the European Economic Association* 18.4, pp. 1969–2012.

Bazart, Cécile and Aurélie Bonein (2014). "Reciprocal relationships in tax compliance decisions". In: *Journal of Economic Psychology* 40, pp. 83–102.

Beekman, Gonne, Stephen L. Cheung, and Ian Levely (Dec. 2017). "The effect of conflict history on cooperation within and between groups: Evidence from a laboratory experiment". In: *Journal of Economic Psychology* 63, pp. 168–183.

Berg, Pieter van den, Peter Dewitte, Ine Aertgeerts, and Tom Wenseleers (2020). "How the incentive to contribute affects contributions in the one-shot public goods game". In: *Scientific Reports* 10.1, pp. 1–5.

Bergstrom, Theodore, Lawrence Blume, and Hal Varian (1986). "On the private provision of public goods". In: *Journal of Public Economics* 29.1, pp. 25–49.

Bernasconi, Michele, Luca Corazzini, Sebastian Kube, and Michel André Maréchal (July 2009). "Two are better than one!. Individuals' contributions to "unpacked" public goods". In: *Economics Letters* 104.1, pp. 31–33.

Bicchieri, Cristina and Alex Chavez (2010). "Behaving as expected: Public information and fairness norms". In: *Journal of Behavioral Decision Making* 23.2, pp. 161–178.

Bicchieri, Cristina and Eugen Dimant (2019). "Nudging with care: The risks and benefits of social information". In: *Public Choice*, pp. 1–22.

Bicchieri, Cristina, Eugen Dimant, Simon Gächter, and Daniele Nosenzo (2022). "Social proximity and the erosion of norm compliance". In: *Games and Economic Behavior* 132, pp. 59–72.

Bicchieri, Cristina and Erte Xiao (2009). "Do the right thing: But only if others do so". In: *Journal of Behavioral Decision Making* 22.2, pp. 191–208.

Bilancini, Ennio, Leonardo Boncinelli, and Tatiana Celadin (2022). "Social value orientation and conditional cooperation in the online one-shot

public goods game". In: *Journal of Economic Behavior & Organization* 200, pp. 243–272.
Bilancini, Ennio, Leonardo Boncinelli, Roberto Di Paolo, Dario Menicagli, Veronica Pizziol, Emiliano Ricciardi, and Francesco Serti (2022). "Prosocial behavior in emergencies: Evidence from blood donors recruitment and retention during the COVID-19 pandemic". In: *Social Science & Medicine* 314, p. 115438.
Bilancini, Ennio, Leonardo Boncinelli, Chiara Nardi, and Veronica Pizziol (2023). "Cooperation is unaffected by the threat of severe adverse events in Public Goods Games". In: *OSF Preprints*.
Birnbaum, Michael H (2004). "Human research and data collection via the Internet". In: *Annu. Rev. Psychol.* 55, pp. 803–832.
Biziou-van-Pol, Laura, Jana Haenen, Arianna Novaro, Andrés Occhipinti Liberman, and Valerio Capraro (2015). "Does telling white lies signal pro-social preferences?" In: *Judgment and Decision Making* 10, pp. 538–548.
Björk, Lisa, Martin Kocher, Peter Martinsson, and Pham Nam Khanh (2016). "Cooperation under risk and ambiguity". In: *University of Gothenburg Working Paper.*
Blackwell, Calvin and Michael McKee (2003). "Only for my own neighborhood? Preferences and voluntary provision of local and global public goods". In: *Journal of Economic Behavior and Organization* 52.1, pp. 115–131.
Blanchard, Olivier, Alvaro Leandro, and Jeromin Zettelmeyer (2021). "Redesigning EU fiscal rules: From rules to standards". In: *Economic Policy* 36.106, pp. 195–236.
Blanco, Mariana, Dirk Engelmann, Alexander K Koch, and Hans-Theo Normann (2010). "Belief elicitation in experiments: is there a hedging problem?" In: *Experimental Economics* 13.4, pp. 412–438.
Bolton, Gary E and Axel Ockenfels (2000). "ERC: A theory of equity, reciprocity, and competition". In: *American Economic Review* 90.1, pp. 166–193.
Bordignon, M. and S. Scabrosetti (2016). "The Political Economy of Financing the EU budget". In: *The Future of EU-finances*. Ed. by Thiess Büttner and Michael Thöne. Mohr Siebeck: Mohr Siebeck. Chap. 5, pp. 63–90.
Bowles, Samuel (2016). *The moral economy: Why good incentives are no substitute for good citizens*. Yale University Press.

Bowles, Samuel and Sandra Polania-Reyes (2012). "Economic incentives and social preferences: substitutes or complements?" In: *Journal of Economic Literature* 50.2, pp. 368–425.

Bronchal, Adrià (2023). "Better the devil you know: The effects of group identity uncertainty on coordination efficiency". In: *Journal of Economic Behavior & Organization* 214, pp. 634–656.

Brown, Thomas C and Stephan Kroll (2017). "Avoiding an uncertain catastrophe: climate change mitigation under risk and wealth heterogeneity". In: *Climatic Change* 141.2, pp. 155–166.

Buchan, Nancy R., Marilynn B. Brewer, Gianluca Grimalda, Rick K. Wilson, Enrique Fatas, and Margaret Foddy (2011). "Global Social Identity and Global Cooperation". In: *Psychological Science* 22.6, pp. 821–828.

Buchan, Nancy R., Gianluca Grimalda, Rick Wilson, Marilynn Brewer, Enrique Fatas, and Margaret Foddy (2009). "Globalization and human cooperation". In: *Proceedings of the National Academy of Sciences* 106.11, pp. 4138–4142.

Buso, Irene Maria et al. (2021). "Lab-like findings from online experiments". In: *Journal of the Economic Science Association* 7.2, pp. 184–193.

Capraro, Valerio (2019). "The dual-process approach to human sociality: A review". In: *arXiv preprint arXiv:1906.09948*.

Capraro, Valerio and Matjaž Perc (2021). "Mathematical foundations of moral preferences". In: *Journal of the Royal Society Interface* 18.175.

Capraro, Valerio and David G. Rand (2018). "Do the right thing: Experimental evidence that preferences for moral behavior, rather than equity or efficiency per se, drive human prosociality". In: *Judgment and Decision Making* 13.1, pp. 99–111.

Caselli, Francesca and Philippe Wingender (2021). "Heterogeneous effects of fiscal rules: The Maastricht fiscal criterion and the counterfactual distribution of government deficits". In: *European Economic Review* 136, p. 103748.

Catola, Marco, Simone D'Alessandro, Pietro Guarnieri, and Veronica Pizziol (2021). "Personal and social norms in a multilevel public goods experiment". In: *Discussion Papers del Dipartimento di Economia e Management – Università di Pisa* 272.

— (2023). "Multilevel public goods game: Levelling up, substitution and crowding-in effects". In: *Journal of Economic Psychology* 97, p. 102626.

Catola, Marco, Simone D'Alessandro, Pietro Guarnieri, and Veronica Pizziol (2021). "Personal norms in the online public good game". In: *Economics Letters* 207, p. 110024.

Chakravarty, Surajeet and Miguel A Fonseca (2017). "Discrimination Via Exclusion: An Experiment On group Identity and Club Goods". In: *Journal of Public Economic Theory* 19.1, pp. 244–263.

Charness, Gary, Uri Gneezy, and Vlastimil Rasocha (2021). "Experimental methods: Eliciting beliefs". In: *Journal of Economic Behavior & Organization* 189, pp. 234–256.

Chaudhuri, Ananish (2011). "Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature". In: *Experimental Economics* 14.1, pp. 47–83.

Chen, Daniel L., Martin Schonger, and Chris Wickens (2016). "oTree-An open-source platform for laboratory, online, and field experiments". In: *Journal of Behavioral and Experimental Finance* 9, pp. 88–97.

Cherry, Todd L, E Lance Howe, and James J Murphy (2015). "Sharing as risk pooling in a social dilemma experiment". In: *Ecology and Society* 20.1.

Cherry, Todd L. and David L. Dickinson (2008). "Voluntary contributions with multiple public goods". In: *Environmental Economics, Experimental Methods*. Ed. by Todd L. Cherry, S. Kroll, and J. F. Shorgen. London: Routledge, pp. 184–193.

Chuah, Swee-Hoon, Robert Hoffmann, Bala Ramasamy, and Jonathan HW Tan (2014). "Religion, ethnicity and cooperation: An experimental study". In: *Journal of Economic Psychology* 45, pp. 33–43.

Corazzini, Luca, Christopher Cotton, and Tommaso Reggiani (2020). "Delegation and coordination with multiple threshold public goods: experimental evidence". In: *Experimental Economics* 23.4, pp. 1030–1068.

Corazzini, Luca, Christopher Cotton, and Paola Valbonesi (2015). "Donor coordination in project funding: Evidence from a threshold public goods experiment". In: *Journal of Public Economics* 128, pp. 16–29.

Corazzini, Luca and Matteo M Marini (2022). "Focal points in multiple threshold public goods games: A single-project meta-analysis". In: *MUNI ECON Working Paper*.

Corazzini, Luca and Robert Sugden (2011). "Common fate, game harmony and contributions to public goods: experimental evidence". In: *International Review of Economics* 58, pp. 43–52.

Coricelli, Giorgio, Mateus Joffily, Claude Montmarquette, and Marie Claire Villeval (2010). "Cheating, emotions, and rationality: an experiment on tax evasion". In: *Experimental Economics* 13.2, pp. 226–247.

Dannenberg, Astrid, Andreas Löschel, Gabriele Paolacci, Christiane Reif, and Alessandro Tavoni (2015). "On the provision of public goods

with probabilistic and ambiguous thresholds". In: *Environmental and Resource Economics* 61.3, pp. 365–383.
Dawes, Robyn M (1980). "Social dilemmas". In: *Annual Review of Psychology* 31.1, pp. 169–193.
De Grauwe, Paul and Yuemei Ji (2019). "Time to change budgetary priorities in the Eurozone". In: *Intereconomics* 54.5, pp. 285–290.
Di Paolo, Roberto and Veronica Pizziol (2023). "Gamification and Sustainable Water Use: The Case of the BLUTUBE Educational Program". In: *Simulation & Gaming*, p. 10468781231181652.
Dickinson, David L (1998). "The voluntary contributions mechanism with uncertain group payoffs". In: *Journal of Economic Behavior & Organization* 35.4, pp. 517–533.
Diederich, Johannes, Timo Goeschl, and Israel Waichman (2016). "Group size and the (in) efficiency of pure public good provision". In: *European Economic Review* 85, pp. 272–287.
Dohmen, T., D. Huffman, J. Schupp, A. Falk, U. Sunde, and G. G. Wagner (2011). "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences". In: *Journal of the European Economic Association* 9.3, pp. 522–550.
Engel, Christoph and Michael Kurschilgen (2020). "The Fragility of a Nudge: the power of self-set norms to contain a social dilemma". In: *Journal of Economic Psychology* 81, p. 102293.
Engel, Christoph and Bettina Rockenbach (2011). "We are not alone: the impact of externalities on public good provision". In: *MPI Collective Goods Preprint* 2009.29.
Eriksson, Kimmo, Pontus Strimling, Per A Andersson, and Torun Lindholm (2017). "Costly punishment in the ultimatum game evokes moral concern, in particular when framed as payoff reduction". In: *Journal of Experimental Social Psychology* 69, pp. 59–64.
European Parliament (2020). *Uncertainty/EU/Hope: Public opinion in times of COVID-19*. Tech. rep. European Parliament, p. 83.
Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde (2018). "Global Evidence on Economic Preferences". In: *The Quarterly Journal of Economics* 133.4, pp. 1645–1692.
Falk, Armin, Urs Fischbacher, and Simon Gächter (2013). "Living in two neighborhoods-social interaction effects in the laboratory". In: *Economic Inquiry* 51.1, pp. 563–578.
Fehr, Ernst, Karla Hoff, and Mayuresh Kshetramade (2008). "Spite and development". In: *American Economic Review* 98.2, pp. 494–99.

Fehr, Ernst and Klaus M Schmidt (1999). "A theory of fairness, competition, and cooperation". In: *The Quarterly Journal of Economics* 114.3, pp. 817–868.
Fellner, Gerlinde and Gabriele K. Lünser (2014). "Cooperation in local and global groups". In: *Journal of Economic Behavior and Organization* 108, pp. 364–373.
Fiala, Lenka and Sigrid Suetens (2017). "Transparency and cooperation in repeated dilemma games: a meta study". In: *Experimental Economics* 20, pp. 755–771.
Fischbacher, Urs, Simon Gächter, and Ernst Fehr (2001). "Are people conditionally cooperative? Evidence from a public goods experiment". In: *Economics Letters* 71.3, pp. 397–404.
Fischbacher, Urs, Simeon Schudy, and Sabrina Teyssier (2014). "Heterogeneous reactions to heterogeneity in returns from public goods". In: *Social Choice and Welfare* 43.1, pp. 195–217.
Frederick, Shane (2005). "Cognitive reflection and decision making". In: *Journal of Economic Perspectives* 19.4, pp. 25–42.
Freund, Alexandra M and Fredda Blanchard-Fields (2014). "Age-related differences in altruism across adulthood: making personal financial gain versus contributing to the public good." In: *Developmental Psychology* 50.4, p. 1125.
Freundt, Jana and Andreas Lange (2021). "On the voluntary provision of public goods under risk". In: *Journal of Behavioral and Experimental Economics*, p. 101727.
Gächter, Simon, Eric J Johnson, and Andreas Herrmann (2021). "Individual-level loss aversion in riskless and risky choices". In: *Theory and Decision*, pp. 1–26.
Gallier, Carlo (2020). "Democracy and compliance in public goods games". In: *European Economic Review* 121, p. 103346.
Gallier, Carlo, Timo Goeschl, Martin Kesternich, Johannes Lohse, Christiane Reif, and Daniel Römer (2019). "Leveling up? An inter-neighborhood experiment on parochialism and the efficiency of multi-level public goods provision". In: *Journal of Economic Behavior and Organization* 164, pp. 500–517.
Gangadharan, Lata and Veronika Nemes (2009). "Experimental analysis of risk and uncertainty in provisioning private and public goods". In: *Economic Inquiry* 47.1, pp. 146–164.
Goette, Lorenz and Egon Tripodi (2020). "Does positive feedback of social impact motivate prosocial behavior? A field experiment with blood donors". In: *Journal of Economic Behavior & Organization* 175, pp. 1–8.

Górecki, Maciej A and Natalia Letki (2021). "Social norms moderate the effect of tax system on tax evasion: Evidence from a large-scale survey experiment". In: *Journal of Business Ethics* 172.4, pp. 727–746.
Granulo, Armin, Rudolf Kerschreiter, and Martin G Kocher (2023). "Cooperation and confusion in public goods games: confusion cannot explain contribution patterns". In: *Journal of the Economic Science Association*, pp. 1–11.
Guala, Francesco (2005). *The methodology of experimental economics*. Cambridge University Press.
Gueth, Werner, M Vittoria Levati, and Ivan Soraperra (2015). "Common and private signals in public goods games with a point of no return". In: *Resource and Energy Economics* 41, pp. 164–184.
Güth, Werner and Lauri Sääksvuori (2012). "Provision of multilevel public goods by positive externalities: Experimental evidence". In: *The B.E. Journal of Economic Analysis and Policy* 12.1.
Herrmann, Benedikt, Christian Thöni, and Simon Gächter (2008). "Antisocial punishment across societies". In: *Science* 319.5868, pp. 1362–1367.
Horton, John J, David G Rand, and Richard J Zeckhauser (2011). "The online laboratory: Conducting experiments in a real labor market". In: *Experimental Economics* 14, pp. 399–425.
Iajya, Victor, Nicola Lacetera, Mario Macis, and Robert Slonim (2013). "The effects of information, social and financial incentives on voluntary undirected blood donations: Evidence from a field experiment in Argentina". In: *Social Science & Medicine* 98, pp. 214–223.
Isaac, R Mark and James M Walker (1988). "Group size effects in public goods provision: The voluntary contributions mechanism". In: *The Quarterly Journal of Economics* 103.1, pp. 179–199.
Isaac, R Mark, James M Walker, and Arlington W Williams (1994). "Group size and the voluntary provision of public goods: Experimental evidence utilizing large groups". In: *Journal of Public Economics* 54.1, pp. 1–36.
Isler, Ozan, Simon Gächter, A John Maule, and Chris Starmer (2021). "Contextualised strong reciprocity explains selfless cooperation despite selfish intuitions and weak social heuristics". In: *Scientific Reports* 11.1, pp. 1–17.
Isler, Ozan, Onurcan Yilmaz, and A John Maule (2021). "Religion, parochialism and intuitive cooperation". In: *Nature Human Behaviour* 5.4, pp. 512–521.

Kahneman, Daniel and Amos Tversky (1982). "On the study of statistical intuitions". In: *Cognition* 11.2, pp. 123–141.
Kandul, Serhiy and Bruno Lanz (2021). "Public good provision, in-group cooperation and out-group descriptive norms: A lab experiment". In: *Journal of Economic Psychology* 85, p. 102382.
Kerr, Norbert L (1983). "Motivation losses in small groups: A social dilemma analysis." In: *Journal of Personality and Social Psychology* 45.4, p. 819.
Kocher, Martin G, Peter Martinsson, Dominik Matzat, and Conny Wollbrant (2015). "The role of beliefs, trust, and risk in contributions to a public good". In: *Journal of Economic Psychology* 51, pp. 236–244.
Krupka, Erin L and Roberto A Weber (2013). "Identifying social norms using coordination games: Why does dictator game sharing vary?" In: *Journal of the European Economic Association* 11.3, pp. 495–524.
Lange, Andreas, Jan Schmitz, and Claudia Schwirplies (2022). "Inequality, role reversal and cooperation in multiple group membership settings". In: *Experimental Economics* 25.1, pp. 68–110.
Ledyard, John O. (1995). "Public Goods: A Survey of Experimental Research". In: *The Handbook of Experimental Economics*. Ed. by John H. Kagel and Alvin E. Roth. Princeton: Princeton University Press. Chap. 2, pp. 111–194.
Levati, M Vittoria and Andrea Morone (2013). "Voluntary contributions with risky and uncertain marginal returns: the importance of the parameter values". In: *Journal of Public Economic Theory* 15.5, pp. 736–744.
Levati, M Vittoria, Andrea Morone, and Annamaria Fiore (2009). "Voluntary contributions with imperfect information: An experimental study". In: *Public Choice* 138.1-2, pp. 199–216.
Lorko, Matej, Maroš Servátka, Robert Slonim, and Michal Ďuriník (2023). *Intertemporal Coordination in Volunteer Markets*. Tech. rep.
Maduro, Miguel Poiares, Philippe J Martin, Jean-Claude Piris, Jean Pisani-Ferry, Lucrezia Reichlin, Armin Steinbach, and Beatrice Weder (2021). *Revisiting the EU framework: Economic necessities and legal options*. Centre for Economic Policy Research.
Maestre-Andrés, Sara, Stefan Drews, and Jeroen van den Bergh (2019). "Perceived fairness and public acceptability of carbon pricing: a review of the literature". In: *Climate Policy* 19.9, pp. 1186–1204.
Maestre-Andrés, Sara, Stefan Drews, Ivan Savin, and Jeroen van den Bergh (2021). "Carbon tax acceptability with information provision and mixed revenue uses". In: *Nature Communications* 12.1, pp. 1–10.

Martinangeli, Andrea F.M. (2021). "Do what (you think) the rich will do: Inequality and belief heterogeneity in public good provision". In: *Journal of Economic Psychology* 83, p. 102364.
McCarter, Matthew W., Anya Samek, and Roman M. Sheremeta (Dec. 2014). "Divided Loyalists or Conditional Cooperators? Creating Consensus About Cooperation in Multiple Simultaneous Social Dilemmas". In: *Group & Organization Management* 39.6, pp. 744–771.
Milinski, Manfred, Torsten Röhl, and Jochem Marotzke (2011). "Cooperative interaction of rich and poor can be catalyzed by intermediate climate targets". In: *Climatic Change* 109.3, pp. 807–814.
Milinski, Manfred, Ralf D Sommerfeld, Hans-Jürgen Krambeck, Floyd A Reed, and Jochem Marotzke (2008). "The collective-risk social dilemma and the prevention of simulated dangerous climate change". In: *Proceedings of the National Academy of Sciences* 105.7, pp. 2291–2294.
Nosenzo, Daniele, Simone Quercia, and Martin Sefton (2015). "Cooperation in small groups: the effect of group size". In: *Experimental Economics* 18.1, pp. 4–14.
Nowak, Martin A (2006). "Five rules for the evolution of cooperation". In: *Science* 314.5805, pp. 1560–1563.
Ordaz-Cuevas, Jesús Erubiel and Joss Sánchez-Pérez (2023). "Cooperation in problems of common-pool resources". In: *Journal of Dynamics and Games* 10.4, pp. 390–402.
Otten, Kasper, Vincent Buskens, Wojtek Przepiorka, and Naomi Ellemers (2021). "Cooperation between newcomers and incumbents: The role of normative disagreements". In: *Journal of Economic Psychology* 87, p. 102448.
Palan, Stefan and Christian Schitter (Mar. 2018). "Prolific.ac—A subject pool for online experiments". In: *Journal of Behavioral and Experimental Finance* 17, pp. 22–27.
Pizziol, Veronica, Xhiselda Demaj, Roberto Di Paolo, and Valerio Capraro (2023). "Political ideology and generosity around the globe". In: *Proceedings of the National Academy of Sciences* 120.15, e2219676120.
Prelec, Drazen (1998). "The probability weighting function". In: *Econometrica*, pp. 497–527.
Romanelli, Marzia, Pietro Tommasino, and Emilio Vadala' (2022). *The future of European fiscal governance: a comprehensive approach*. Questioni di Economia e Finanza (Occasional Papers) 691. Bank of Italy, Economic Research and International Relations Area.
Romano, Angelo, Matthias Sutter, James H Liu, and Daniel Balliet (2021). "Political ideology, cooperation and national parochialism across 42

nations". In: *Philosophical Transactions of the Royal Society B* 376.1822, p. 20200146.
Ross, Lee, David Greene, and Pamela House (1977). "The "false consensus effect": An egocentric bias in social perception and attribution processes". In: *Journal of Experimental Social Psychology* 13.3, pp. 279–301.
Samuelson, Charles D and David M Messick (1986). "Inequities in access to and use of shared resources in social dilemmas." In: *Journal of Personality and Social Psychology* 51.5, p. 960.
Sonnemans, Joep, Arthur Schram, and Theo Offerman (1998). "Public good provision and public bad prevention: The effect of framing". In: *Journal of Economic Behavior & Organization* 34.1, pp. 143–161.
Spicer, Michael W and Lee A Becker (1980). "Fiscal inequity and tax evasion: An experimental approach". In: *National Tax Journal* 33.2, pp. 171–175.
Spicer, Michael W and Rodney E Hero (1985). "Tax evasion and heuristics: A research note". In: *Journal of Public Economics* 26.2, pp. 263–267.
Starmer, Chris (2000). "Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk". In: *Journal of Economic Literature* 38.2, pp. 332–382.
Tajfel, H. (Apr. 1974). "Social identity and intergroup behaviour". In: *Social Science Information* 13.2, pp. 65–93.
Tajfel, Henri (1970). "Experiments in intergroup discrimination". In: *Scientific American* 223.5, pp. 96–103.
— (1982). "Social psychology of intergroup relations". In: *Annual Review of Psychology* 33.1, pp. 1–39.
Tavoni, Alessandro, Astrid Dannenberg, Giorgos Kallis, and Andreas Löschel (2011). "Inequality, communication, and the avoidance of disastrous climate change in a public goods game". In: *Proceedings of the National Academy of Sciences* 108.29, pp. 11825–11829.
TFEU (2008). "Treaty on the Functioning of the European Union". In: *Official Journal of the European Union. C* 115, p. 49.
Théroude, Vincent and Adam Zylbersztejn (2020). "Cooperation in a risky world". In: *Journal of Public Economic Theory* 22.2, pp. 388–407.
Thöni, Christian and Stefan Volk (2018). "Conditional cooperation: Review and refinement". In: *Economics Letters* 171, pp. 37–40.
Tversky, Amos and Daniel Kahneman (1992). "Advances in prospect theory: Cumulative representation of uncertainty". In: *Journal of Risk and Uncertainty* 5.4, pp. 297–323.

Vesely, Stepan, Erik Wengström, et al. (2017). "Risk and cooperation: Experimental evidence from stochastic public good games". In: *Lund University Working Paper*.

Weisel, Ori and Ro'i Zultan (2016). "Social motives in intergroup conflict: Group identity and perceived target of threat". In: *European Economic Review* 90, pp. 122–133.

Xiao, Erte and Howard Kunreuther (2016). "Punishment and cooperation in stochastic social dilemmas". In: *Journal of Conflict Resolution* 60.4, pp. 670–693.

Zelmer, Jennifer (2003). "Linear public goods experiments: A meta-analysis". In: *Experimental Economics* 6.3, pp. 299–310.

Zhang, Huanren (2019). "Common fate motivates cooperation: The influence of risks on contributions to public goods". In: *Journal of Economic Psychology* 70, pp. 12–21.