ISSN 2532-5604

CSCC Centro di ricerca sulle Scienze Cognitive e la Comunicazione
Cognitive Science and Communication Research Centre

CSCC Working Papers 01/17

Trust behind bars: a longitudinal study of inmates' prosocial preferences

Mario A. Maggioni*, Domenico Rossignoli*, Simona Beretta*, and Sara Balestri*

*DISEIS and CSCC Università Cattolica del Sacro Cuore, Milano Italy

CSCC
Centro di ricerca in Scienze Cognitive e della Comunicazione
Università Cattolica del Sacro Cuore
Via Necchi, 5 - 20123 - Milano ITALY

La pubblicazione nella Collana CSCC Working Papers è soggetta a valutazione da parte di due referees indipendenti attraverso una procedura di valutazione *single blind*.

To be published in the CSCC Working Papers series the article must pass a *single blind* referee procedure by two independent referees.

CSCC Working Papers 01/17

CSCC Centro di ricerca in Scienze Cognitive e della Comunicazione
Università Cattolica del Sacro Cuore
Via Necchi, 5 - 20123 - Milano ITALY

ISSN 2532-5604
CSCC Working Papers
[Online]

Trust behind bars: a longitudinal study of inmates' prosocial preferences

Mario A. Maggioni[*], Domenico Rossignoli[*], Simona Beretta[*], and Sara Balestri[*]

[*]*DISEIS and CSCC Università Cattolica del Sacro Cuore, Milano Italy*

March, 2017

Abstract

The paper presents the results of a Longitudinal Lab-in-the-Field Experiment implemented between September 2015 and July 2016 in two State Prisons in California (USA). A subset of eligible inmates willing to undertake GRIP (Guiding Rage Into Power), an "offender accountability program", were randomly assigned to it. The paper tests whether the participation to this program (used as a treatment in the experiments), based on building strong relationships and mutual help, affects prosocial preferences of participants, with specific reference to trust. The results of a Difference-in-Differences (DID) estimation procedure show that trust significantly increased in GRIP participants compared to the control group. This result is robust to alternative estimation techniques and to the inclusion of an endogenous behavioral measure of altruism.

Corresponding author: mario.maggioni@unicatt.it

Acknowledgements:the authors wish to thank E. Colombo, P.A. David, E. Mullet, B. Paulle, V. Rotondi, L. Sullivan, F. Trombetta and the participants to the CSCC LBR Workshop held in Milan (2017) for comments and observations. The usual *caveats* apply. We thank E. Esposto for research assistantship. Financial support by the Fetzer Institute (Project # 3520.00) and the D3.2 Competitive Research Funds Università Cattolica del Sacro Cuore are gratefully acknowledged.

Keywords:Trust; Trust Game, Lab-in-the-Field Experiment; Longitudinal Study; Prisons

JEL Classification: C91, C92, D03, Z10

1 Introduction

The United States of America notably host the largest correctional population in the world (Walmsley et al., 2013). High incarceration rates entail huge social and economic consequences for inmates and their families and also for the entire society, since prosocial attitudes can seriously decline in prison. In particular, long-term imprisonment leads to the so-called process of "prisonization" (Naderi, 2014), that produces both psychological and behavioral effects[1](Haney et al., 2003), affecting inmates' ability to trust people (McCorkle, 1992) and hampering the possibility of effective resocialization (Zingraff, 1975). These negative individual psychological and behavioral effects hinder the effective reintegration of inmates in local communities and cause a process of social exclusion (Morenoff and Harding, 2014; Wakefield and Uggen, 2010; Lynch and Sabol, 2004) which in turn exacerbates the decrease in their prosocial behavior (Twenge et al., 2007), eventually contributing to recidivism (Figlio et al., 1972; Durose et al., 2015).

Within the US national framework, California hosts the second largest prison population in the country and represents a specially interesting case, since it has embarked, together with New York and New Jersey, on one of the most relevant prison downsizing experiment adopted during the last decade (Newman and Scott, 2011). Facing a U.S. Supreme Court decision, which ordered the State of California to reduce its prison population by a quarter within two years, in 2011 Governor Jerry Brown signed the Public Safety Realignment Act which propelled a number of measures aimed at achieving this target by transferring lower-level offenders from state to county prisons, and supporting a number of evidence-based community corrections programs (Petersilia, 2014, p. 802).

This paper aims at testing whether the participation to an offender accountability program (GRIP[2]), run by the US-based NGO *Insight-Out*, significantly changes prosocial attitudes of violent crime offenders, with a special focus on trust[3]. It is worthwhile stressing that this paper does not directly aim at testing the impact or effectiveness of the program in reaching its core targets[4]; but rather uses behavioral economics games, within an experimental setting, to measure effects of the program on changing inmates' prosocial preferences and attitudes. In particular, we focus on generalized trust, since it is widely acknowledged (Putnam, 1993; Fukuyama, 1995; Knack and Keefer, 1997; Warren, 1999; Guiso et al., 2008; Grimalda and Mittone, 2011) as one of the ingredients of well-functioning societies, efficient and growing economic systems and effective participative institutions.

The way Insight-Out administers the program particularly fits the requirements

[1]These effects are not necessarily pathological, rather normal adaptive processes to abnormal demands, due to the highly insecure and stressful prison's environment. See Haney et al. (2003) for a detailed survey of the psychological processes occurring during long-term imprisonment.

[2]GRIP is the acronym for *Guiding Rage Into Power* an offender accountability program which is currently administered in seven Californian prisons and involves around 500 inmates. See more detail about GRIP in the dedicated section 2.

[3]In a somehow similar experimental setting, Blattman et al. (2017) investigates the "malleability" of a number of noncognitive skills and preferences in criminally engaged adults.

[4]That would be mostly based on evidence of low or no recidivism for inmates "graduated" in the program.

for an experimental setting. The NGO's limited capacity (until 2016 Insight-Out has been able to manage only one GRIP class of about 25 inmates per prison each year) allows the implementation of a Lab-in-the-Field Experiment[5] involving 42 treated (inmates participant to GRIP) and 38 controls (inmates not attending the program), in two Californian State prisons. All 80 inmates were surveyed twice, the first time before the beginning of GRIP and the second time about ten months later (after the end of the program). The experimental protocol, designed to run a panel data Difference-in-Differences analysis, envisaged a set of behavioral games, devised to elicit prosocial behavior (and in particular altruism and trust) and a series of questions based on a set of validated psychological scales of forgiveness and self-forgiveness[6].

The novelty of our contribution is threefold:

1. we implemented a longitudinal study in the framework of prosocial behavioral games, addressing a promising, but still little investigated[7], research question concerning the change of prosocial attitudes and preferences over time, with reference to trust and altruism, experienced by people exposed to a "rehab" program;

2. we devised an "augmented" experimental setting in which we analyse trust by controlling for an endogenously determined characteristic of the inmates, by including, as a covariate in the DID model specification, the outcome of a Dictator Game (as proxy for altruism) in the baseline survey;

3. we applied a behavioral economics set of games to a non-standard marginalized adult population (prison inmates) and are able to show that trust significantly increased after taking part to an offender accountability program.

The paper is organized as follows: section 2 outlines the main features of the GRIP program, highlighting its core targets; section 3 presents the research design and experimental methods; section 4 presents the methodology to estimate the Average effect of the Treatment on the Treated (ATT); section 5 provides results and robustness checks, and section 6 provides a discussion of the main findings and conclusions.

[5]Following the taxonomy proposed by Harrison and List (2004), a Lab-Like Field or Lab-in-the-Field experiment involves participants drawn from the field and asks them "to perform laboratory tasks that are not part of their day-to-day environment."(Viceisza, 2016, p. 836)

[6]The forgiveness-related questions are not analyzed in this paper due to their different nature and scope. Therefore, we have postponed their illustration to a future work.

[7]A recent survey shows that while the stability of risk and time preferences have been extensively studied, social preferences are investigated by only 4 papers, and only two of them address trust and altruism, based on a very limited sample of university students (Chuang and Schechter, 2015, Table 3, p. 154).

2 *Guiding Rage Into Power* (GRIP): outline of the rehabilitation program

Jacques Verduin has been running the GRIP (*Guiding Rage Into Power*) Program in San Quentin Prison (California, USA) since 2011[8], the founder of Insight-Out, an NGO based in the San Francisco's area. GRIP originates from its founder's vision that the lack of relations is the main driver of violence and unlawful behavior[9] and is classified as an "offender accountability program", according to the California Department of Corrections and Rehabilitation (CDCR)[10]. GRIP aims at providing inmates with the skills to undo and prevent violent behavior so to become "agents of change", i.e. "people with skills to defuse conflicts around them"[11]. In particular, the program focuses on the origins of behaviors and habits that are conducive to crime, with the specific purpose of undoing "the characteristic destructive behavioral patterns (...) that lead to transgressions[12]". The program usually spans over an "academic year", roughly ten months long (between September and July) and develops through weekly/fortnightly lessons, each one focused on a specific topic, aiming at (1) stopping violent behavior, (2) cultivating mindfulness, (3) achieving emotional intelligence and (4) understanding victim impact. GRIP targets unobservable behavioral traits, that indirectly affect other-regarding preferences and beliefs. For this reason, the program is particularly suited to the experimental analysis we devised: neither trust, nor altruism are explicitly "taught" during GRIP classes, hence we expect experimental outcomes not to be driven by inmates' adherence to the program content[13]. However, both trust and altruism can be indirectly affected by the program and the implementation of behavioral games is especially devised to elicit them by the observed preferences in a set of task.

GRIP classes are held through a variety of didactic methods, spanning from traditional frontal lessons, to group-works and intervention of external guests[14]. Great emphasis and efforts are spent on creating a strong group identification, so that GRIP participants realize and experience that they are not alone, but part of a community that is pursuing the same objectives[15]. Every cohort of GRIP participants is named "tribe" followed by a number that consists of how many years all the men (new participants as well as co-facilitators) have been incarcerated in any type of correctional facility: from juvenile detention to county jails to state prisons[16]. The inmates are also

[8] His experience with prison programs is much longer: in 1997, he founded the Insight Prison Project (IPP), pioneering innovative in-prison rehabilitation programs designed to create transformational change among prisoners at California's San Quentin State Prison.

[9] A summary of both vision and mission of Insight-Out can found on the NGO's website. See https://goo.gl/uMFgna for further details.

[10] An independent presentation of the program's features can be found in Paulle (2017).

[11] See http://insight-out.org/index.php/programs/grip-program.

[12] *ibidem*.

[13] No reference to trust and altruism/generosity can be found in the final "Pledge", signed by each inmate taking part in GRIP.

[14] For a detailed presentation on how the GRIP program works, refer to: https://vimeo.com/63489782.

[15] A recent analysis of the effectiveness of a somehow similar program (called STYL) in changing a series of noncognitive skills and preferences of criminally engaged men in Liberia can be found in Blattman et al. (2017)

[16] In a given class of around 30 men, the total often climbs higher than 600 years (Paulle, 2017). In this

asked to work on their own or in small groups on specific assignments, in order to keep track of their progresses in achieving the four above mentioned goals. At the end of the course, a Graduation ceremony is held, involving inmates, relatives and the prison's warden. During the Graduation ceremony inmates, or "students" as they are called throughout the program, receive their title and diploma of "peace maker".

3 Experimental design

3.1 Identification strategy

GRIP does not target "specific" kinds of inmates, being virtually open to any inmate who applies. Before starting each program, Insight-Out offers a program orientation day in each prison to illustrate the main features of GRIP to all the potentially interested inmates. After the presentation of GRIP, inmates are asked to signal their interest and take an interview aimed at identifying mental disabilities or other severely impairing conditions. The actual number of selected participants (usually around 25 per prison) depends on both the NGO capacity constraints and prison security rules. This feature of the enrollment process particularly fits the methodological requirements of a pre-/post-treatment research design: the treatment (i.e. the participation to the program) was randomly assigned to inmates that signaled their interest and passed the interview, up to the filling of all available places. The random assignment of the treatment has been implemented according to the following procedure: after the motivational open day, Insight-Out provided us with information about the inmates (identified by an anonymous code) interested in taking part into GRIP. We randomly assigned the resulting pool of inmates to either treatment or control by balancing for ethnic group, to accomplish with CDCR requirements for prison programs and with GRIP's aim.

Henceforth, inmates enrolled in GRIP will be referred to as the "Treated" group (T); inmates who asked to enter the program but were not enrolled because of capacity constraints will be referred to as the "Waiting list" group (W), since, if the program is repeated in the same prison, they will have a chance to attend the program in the subsequent years; and, finally, inmates who attended the orientation day, but did not show up for the interview, as the "No interview" group (N).

The experiment has been actually conducted on 80 inmates of two Californian State Prisons (Avenal and Mule Creek), operated by the CDCR, according to the above mentioned three sample of inmates (T, W, N)[17]. Since the perspective of this research especially focuses on changes in prosocial preferences and attitudes, we devised a longitudinal study by administering the same questionnaires twice to each inmate: the first one in September 2015, before the start of the treatment (i.e. the GRIP program), the second one in July 2016 after the end of the program. The first survey aimed at measuring the initial level of the parameters of interest; the second survey aimed at measuring whether a significant change in the parameters occurred after the treatment.

particular way each cohort obtains its unique group identification, its name, that inmates often use to recall for years after the of the program. See Paulle (2017) for more details.

[17] Given the small size of the W group, compared with the T group, in the paper we use the sum of W + N groups as the main control group in the estimation. However results in which the sole W group is used as control are also presented as a robustness check.

3.2 Incentivized tasks: Trust Game and Dictator Game

In order to assess whether GRIP affects the prosocial preferences and attitudes of inmates, we devised an experimental setting that included two incentivized tasks[18], a Dictator Game and a Trust Game, that were administered separately, after instructions had been read loud and clarification questions have been answered[19].

In the *Dictator Game* (Kahneman et al., 1986) a Proponent is provided with an exogenous endowment (usually a fixed monetary amount), he/she is matched to an anonymous partner who has received no endowment, and his/her choice consists of how to split the endowment between himself and the partner. Within the game theory framework, assuming a Proponent with self-regarding preferences, the Dictator Game has a unique Nash equilibrium in which the Proponent maximizes his/her payoff by keeping all the endowment and sending 0 to the partner. Therefore, deviation from the *selfish* equilibrium solution in the Dictator Game are used to measure empathy, altruism and/or pure generosity (Forsythe et al., 1994; Camerer, 2003; Guala and Mittone, 2010).

In the *Trust Game*, also known as Investment Game (Camerer and Weigelt, 1988; Berg et al., 1995), a Proponent is provided with an exogenous endowment, and he/she is matched to an anonymous partner who has received no endowment. His/her decision now concerns whether and how much of his/her endowment to send to the anonymous partner; the Proponent is also informed that the experimenter will multiply (triple) any amount sent. The Respondent, once has received the total transfer (the amount sent by the Proponent, duly multiplied) is then told to choose if, and how much of the total amount received, to send back to the Proponent. Therefore, the final payoff of the Proponent will be equal to the initial endowment, less the amount sent to the Respondent, plus the amount sent back by the Respondent to the Proponent. This game has a unique sub-game perfect Nash equilibrium in which the Proponent maximizes his/her payoff by keeping all the endowment and sending 0 to the partner: in fact, solving by backward induction, since a *selfish* Respondent has no reason to send back any strictly positive amount, the Proponent maximizes his/her payoff by keeping the entire initial endowment. Sending a positive share of the initial endowment to anonymous partners signals agents' propensity to interact with unknown partners, providing a proxy for generalized trust (Camerer, 2003; Berg et al., 1995; Johnson and Mislin, 2011), that has been defined, as "the deliberate willingness of a decision maker to making himself vulnerable to the actions of another party" (Sutter and Kocher, 2007, p. 365).

Due to the impossibility of dealing with monetary payoffs (due to CDRC rules) we devised an alternative payoff. After extensive consultation with the staff of NGOs working within prisons in the USA, we decided to use dehydrated soups (henceforth: soups) as rewards[20]. Since the experiment deals with non-monetary incentives, we

[18] As already mentioned in section 1, since this research is part of a broader project, the experiment also included a series of questions based on a set of validated psychological scales of forgiveness (Mullet et al., 2004; Chiaramello et al., 2008), and self-forgiveness (Pelucchi et al., 2013) that are not analyzed int this paper.

[19] For details on the experimental procedure see supplementary materials

[20] Dehydrated soups are highly valuable in maximum security prisons as they allow inmates to have a meal in the relative privacy of their cell. Moreover there is anecdotal evidence that these items are stored

devised a control for the "use value" (`soup like`) and the "exchange value" (`soup value`) that inmates attach to soups.

3.3 Procedures

The experiment has been implemented in paper-and-pencil, the first time in September 2015 (before the start of the program) and the second time in July 2016 (after GRIP graduation took place). The experiment has been administered by six students/interviewers purposely recruited in a local Community College[21]. Interviewers were thus independent both from Insight-Out and CDCR, thus minimizing the risk of possible strategic choices on the inmates sides. In both prisons, the same procedure has been applied, as follows.

Inmates were gathered in a room, equipped with tables and chairs, and sat down at an adequate distance from each other. The interviewers read aloud the instructions of each game before administering them, making sure that everyone in the room understood them[22]. Inmates were informed that only one of the games would have been randomly drawn through the toss of a fair plastic coin at the end of the experiment session and rewarded: in this way, each inmates had the incentive to maximize his outcome in both games.

Both behavioral games were played in an anonymous double blind setting. Inmates were randomly assigned a code; Insight-Out staff held records about the matching between individual names and codes, but could not access individual outcome data (i.e. games results); the research team could access individual outcome data, matched with anonymous codes, but could not access individual names.

Both in the Dictator and Trust games inmates were told they would have been randomly matched with anonymous partners. To administer the payment of the payoffs, we devised the following protocol. Once an inmate submitted his paper sheets[23], an interviewer asked him to pick randomly one out of five "reward booklets", randomly taken from a pile. The booklets included the anonymous partner's choices needed to match the inmate's decision in the Trust Game[24]. The inmate was then asked to toss a plastic coin to select the game (either Dictator or Trust) to be rewarded and his choice in the randomly selected game was matched with the partner's choice. Finally, the resulting payoff was paid in soups. On average, inmates earned 5.7 dried soups (modal value 5). On average, each inmate took around twenty-five minutes to complete the experimental session, administration of the rewards included.

and traded with other inmates, thus they can be thought as imperfect substitutes of money in prison. Recent academic research provides further support for our choice: "Inmates are so unhappy with the quality and quantity of prison food that they receive that they have begun relying on ramen noodles — a cheap, durable food product — as a form of money in the underground economy" (Gibson-Light, 2016).

[21] Gavilan College, Gilroy, CA.

[22] Before starting the experiment, inmates were asked to play some trial sessions in order to verify their actual understanding of the instructions.

[23] Including the forgiveness-related questions, not presented in this paper.

[24] These booklets contained the outcome of the choices of University non-academic staff (janitors, wardens, cleaners, cooks, etc.), that were asked to play the Trust Game as Respondent in a previous experiment run by the research team. Therefore, all the choices included in the booklet were generated by real people, and the inmates were aware of this fact.

4 Analysis of the Treatment Effect

4.1 Estimation technique

The experimental framework allowed to perform a Difference-in-Differences (DID) analysis in which the change registered by the T group was compared to the change of W (in order to control for the self-selection related to the individual willingness a/o motivation to join the program) and, more generally, to a wider "control group" obtained by grouping together inmates belonging to N and W groups (N+W). The average effect of the treatment on the treated (ATT) is the effect of GRIP on the inmates taking part into the program.

DID allows to test the ATT in a pre-/post- treatment setting, by controlling for possible confounding factors, including fixed time-invariant individual characteristics. Formally, the effect of a treatment (Treat) on an outcome (Y) can be tested through the model:

$$Y_{it} = \alpha + \beta Treat_{it} + \gamma Post_{it} + \rho(Post * Treat)_{it} + \delta X_i + \epsilon_{it} \tag{1}$$

where the subscripts i and t respectively refer to prisoners and periods (surveys); α represent the constant term; $Post$ is the time dummy, taking value 1 for observations belonging to the second survey and zero otherwise; $Treat$ is the treatment dummy; X_i are individual inmates' characteristics; and ϵ_{it} is the usual error term, while β, γ, ρ and δ are the parameters to be estimated. The ATT effect is estimated by the coefficient ρ. Formally, being g the groups in our sample, namely treated (T), waiting (W), and the wider control group (N+W) and t the two surveys (1 and 2), then ρ is defined as follows[25]:

$$\begin{aligned} \rho = (E[Y_{igt}|g = T, t = 2] - E[Y_{igt}|g = T, t = 1]) \\ -(E[Y_{igt}|g = W, t = 2] - E[Y_{igt}|g = W, t = 1]) \end{aligned} \tag{2}$$

A DID can be estimated in a regression framework by creating dummy variables for g and t. In this way it is possible to estimate a model that fully takes into account a set of possible confounding factors.

4.2 Data and variables

Equation 1 is applied to the analysis of both the outcome of the Dictator Game and the Trust Game. In the Dictator Game, the outcome (i.e. the number of soups sent to the Respondent) is transformed an indicator of the relative endowment (within the range 0-1) that the inmate shares with the anonymous partner. Analogously, the Trust Game yields a discrete outcome expressed in number of soups (between 1 and 10) that has been re-arranged in relative terms, into the 0-1 range.

The benchmark model of our analysis includes control variables related to the following individual characteristics that could affect the inmate's attitude to trust other people:

[25] The same applies to the wider control group (N+W).

- age declared by inmate as of July 2016 (age);
- a dummy variable, coded 1 if inmates was not involved in a stable relationship, i.e. whether he was not married, separated/divorced, engaged or widowed (single);
- a control for personal preferences for soups (soup like);
- a control for the value of soups as means of exchange, independently of individual tastes (soup value).

In an extended specification, we also included dummies for ethnic identity, to control for possible cross-ethnic heterogeneity that could affect the propensity to trust an anonymous partner[26].

Table 1 provides the summary statistics of the individual characteristics of the treated and of the two control groups (N+W and W); Table 2 records mean comparison tests for all three samples. In particular, the last three columns show that no significant differences occur between Treated and control groups on the main covariates, with the sole exception of soup like.

Table 1: Summary of samples characteristics, at baseline

T Group Variable	Obs	Mean	Std. Dev.	Min	Max
age	42	43.86	9.03	22	59
single	42	0.57	0.50	0	1
soup like	41	5.44	3.08	1	10
soup value	41	5.59	3.54	1	10
N+W Group Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
age	38	47.00	9.42	26	63
single	38	0.50	0.51	0	1
soup like	38	7.26	2.83	1	10
soup value	38	6.39	2.95	1	10
W Group Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
age	22	47.59	8.63	30	63
single	22	0.59	0.50	0	1
soup like	22	7.64	2.77	1	10
soup value	22	6.32	2.95	1	10

Furthermore, since the literature (e.g. Camerer, 2003; Johnson and Mislin, 2011) suggests that the outcome of a Trust Game can be actually driven by pure altruism rather than trust, in order to estimate the effect of pure altruism on trust, we provide an augmented version of our benchmark model that also include the outcome of the Dictator Game as measured in the first survey.

Finally, since the experiment is repeated twice, we take into account possible autocorrelation of the error term at the individual level. To tackle this issue, all models are estimated with robust standard errors, clustered at the inmates' level.

[26]Ethnic group identities are self-reported by inmates by choosing among not mutually exclusive categories, hence all included in the model estimation, taken from the US Census official definition of ethnic and racial groups as mandated by the Office of Management and Budget's (OMB) 1997 standards. For further information see https://www.census.gov/quickfacts/meta/long_RHI225215.htm

Table 2: T-tests, by treatment group at baseline

	Sample means			Mean comparison T-tests		
Variable	T	(N+W)	W	T/(N+W)	T/W	W/N
age	43.86	47.00	47.59	−3.143	−3.734	1.403
single	0.57	0.50	0.59	0.071	−0.020	0.216
soup like	5.44	7.26	7.64	−1.824***	−2.197***	0.886
soup value	5.59	6.39	6.32	−0.809	−0.733	−0.182
Ethnic Groups						
White	0.40	0.34	0.41	0.063	−0.004	0.159
Native American/Alaska	0.12	0.13	0.14	−0.013	−0.017	0.011
Hawaiian Native/Pacific	0.10	0.11	0.14	−0.010	−0.041	0.074
Asian	0.10	0.13	0.14	−0.036	−0.041	0.011
Black/African American	0.33	0.34	0.45	−0.009	−0.121	0.267*
Hispanic/Latino	0.48	0.53	0.45	−0.050	0.022	−0.170

*** p<0.01, ** p<0.05, * p<0.10

5 Results and discussion

5.1 Main results

Figure 1 illustrates the distributions of the inmates' choices in the Trust Game, by survey and treatment group. While the wider (N+W) control group in the upper and lower panel report similar distributions, the distribution for the Treated group is clearly different in survey 1 and survey 2, showing that Treated inmates, after the program, actually reported very different answers[27].

Table 3 reports the results of the benchmark models' estimations: model (1) shows that, after the completion of the GRIP Program, the Treated inmates significantly increased the fraction sent to the anonymous partner by 0.184, as shown by the ATT coefficient in Table 3 (corresponding to a 20 per cent increase of the initial amount, given that soups were only available as integer numbers). This finding is robust to the identification of alternative control groups, as shown in model (2): the size of the two coefficients is substantially comparable, since both can be converted into an average increase of about 2 soups [28]. As far as covariates are concerned[29], Trust is negatively associated with age, although the size of the coefficient is very small. Furthermore, `single` is highly significant and negatively associated with Trust, suggesting that the absence of involvement in a stable romantic relationship (either present or past, successful or unsuccessful) signals a less trustful attitude (in the Sutter and Kocher, 2007 sense). The benchmark model's estimations thus support the occurrence of a "trust-increasing" effect of GRIP in participant inmates.

[27] This "at a glance" evidence is supported by a Kruskal-Wallis test's p-values, equals to 0.66 for the Treated and to 0.23 for the Control group. An almost identical result is obtained through a Wilkoxon rank-sum test.

[28] In fact - since dehydrated soups (the goods used in the actual lab-experiment) are indivisible - the two coefficients may be interpreted as identical.

[29] Table 3 does not report further alternative specifications that include also the payoff earned in the first survey: this variable is not significant, implying that the time interval between the two sessions is large enough to cancel out potential "memory effect" of the previous payoff.

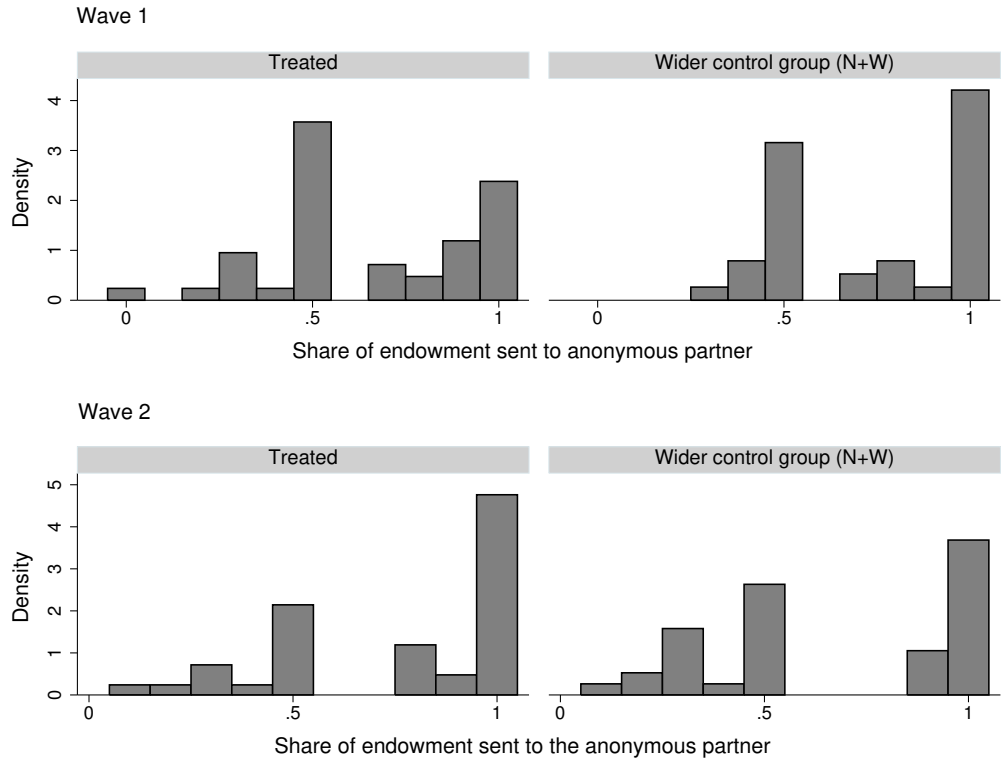


Figure 1: Trust Game: distribution of inmates' choices

Since the observed outcome of a Trust Game can be driven by alternative multiple motivations, such as other-regarding preferences, and *beliefs* over the trustworthiness of the anonymous partner (Sapienza et al., 2013; Fehr, 2009), we estimate the effect of pure altruism on trust (Rabin, 1993) by including the outcome of a Dictator Game. In this way, we devise an "augmented" model in which an endogenously determined attitude of the participants is included as a covariate for trust.

Before moving to the augmented specification of our model, we tested the effect of GRIP on the outcome of the Dictator game through a DID estimation. As models (3) and (4) in Table 3 show, the ATT coefficient is not significant in the benchmark model for the two control groups[30]. Thus, the augmented specification of the DID model presented in Table 4 treats `altruism` as an endogenously determined individual attitude and includes it among the covariates at the value observed at the first survey.

[30]This result is robust for alternative model specifications, including an extended set of covariates to control for ethnicity, and no covariates at all. Furthermore, both the Wilckoxon rank-sum and Kruskal-Wallis tests fail to reject the null hypothesis that the distributions of altruism for Treated and Controls are different across surveys.

Table 3: Benchmark models' results: Trust and Dictator, Difference-in-Differences

Dep. Var:	Trust		Dictator	
Control Group:	(N+W)	W	(N+W)	W
	(1)	(2)	(3)	(4)
ATT	0.184**	0.212**	0.119	0.132
	(0.075)	(0.087)	(0.073)	(0.090)
age	−0.008***	−0.008***	−0.004	−0.004
	(0.002)	(0.003)	(0.003)	(0.003)
single	−0.146***	−0.133**	−0.118**	−0.127**
	(0.049)	(0.059)	(0.047)	(0.053)
prison dummy	0.018	−0.020	−0.088*	−0.150***
	(0.054)	(0.060)	(0.050)	(0.052)
soup value	0.001	0.006	−0.006	0.003
	(0.008)	(0.009)	(0.008)	(0.008)
Observations	158	126	158	126
Adj. R-squared	0.11	0.08	0.07	0.09

Robust standard errors in parentheses, clustered at prisoners' level.
*** p<0.01, ** p<0.05, * p<0.10

Table 4: Augmented Trust results: Difference-in-Differences.

Dep. var.: Trust	including altruism				including altruism & ethnic groups			
Control Group:	(N+W)		W		(N+W)		W	
Model:	OLS	GLM	OLS	GLM	OLS	GLM	OLS	GLM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ATT	0.184**	0.171***	0.212**	0.199***	0.184**	0.172***	0.212**	0.201***
	(0.075)	(0.061)	(0.087)	(0.072)	(0.077)	(0.060)	(0.090)	(0.072)
altruism	0.377***	0.407***	0.375***	0.414***	0.370***	0.395***	0.300**	0.343**
	(0.091)	(0.101)	(0.116)	(0.123)	(0.091)	(0.105)	(0.131)	(0.143)
age	−0.006**	−0.007**	−0.006*	−0.006*	−0.006**	−0.007**	−0.006*	−0.007**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
single	−0.138***	−0.142***	−0.122**	−0.127**	−0.143***	−0.140***	−0.140**	−0.142***
	(0.042)	(0.044)	(0.051)	(0.053)	(0.046)	(0.046)	(0.055)	(0.054)
prison dummy	0.038	0.045	0.024	0.036	0.024	0.027	−0.013	−0.011
	(0.047)	(0.049)	(0.055)	(0.058)	(0.045)	(0.048)	(0.056)	(0.059)
soup value	0.005	0.005	0.006	0.006	0.008	0.009	0.011	0.013
	(0.007)	(0.008)	(0.008)	(0.008)	(0.007)	(0.008)	(0.009)	(0.010)
White					0.043	0.046	0.071	0.069
					(0.052)	(0.059)	(0.062)	(0.065)
Native American/ Alaska					−0.023	−0.028	−0.109	−0.095
					(0.074)	(0.071)	(0.099)	(0.103)
Hawaiian Native/ Pacific					−0.009	−0.009	0.039	0.026
					(0.113)	(0.158)	(0.135)	(0.166)
Asian					−0.020	−0.043	−0.015	−0.033
					(0.087)	(0.105)	(0.137)	(0.133)
Black/ African American					−0.079	−0.074	−0.079	−0.083
					(0.053)	(0.065)	(0.058)	(0.068)
Hispanic/ Latino					0.038	0.038	0.060	0.047
					(0.050)	(0.058)	(0.065)	(0.071)
Observations	158	158	126	126	158	158	126	126
Adj. R-squared	0.21		0.17		0.21		0.18	
McFadden R-Sq.		0.12		0.29		0.13		0.31
AIC		0.974		1.009		1.037		1.086
BIC		−695.2		−516.7		−666.8		−490

Treatment and time dummies included. Clustered robust s.e. in parentheses. *** p<0.01, ** p<0.05, * p<0.10
OLS and GLM for the binomial family (link function: logit): for GLM models, the table reports marginal effects.

In the augmented specification, as shown in Table 4, models (1) to (4) include the same covariates as in Table 3 as well as the endogenous proxy for altruism; models (5) to (8) also include dummy variables for self-reported ethnic identities. For all specifications, the table reports the coefficients for both the wider control group (N+W) and the subsample of motivated inmates (W). Finally, all specifications are estimated both through standard OLS, as common practice in the field, and through GLM for the binomial family with a logit link function. The latter estimation technique has been implemented following the suggestions by Papke and Wooldridge (1996), for bounded dependent variables[31]. ATT, the Average Treatment effect on the Treated, is always positive and significant and it is robust to different model specifications. The coefficient of `altruism` is positive and significant, as expected.

In the extended version of the model reported on columns (5) to (8), none of the group dummies is significant, allowing to exclude effects of ethnic differences on trust[32].

One may raise a concern on whether inmates - despite being told they were interacting with a real person, who had played the same games in a previously administered session, whose answers were recorded in the "reward booklet" - were actually believing to be matched with real persons, rather than simulated ones. Previous behavioral (Bottom et al., 2006) and neuroscience (Sanfey et al., 2003) studies show that people behave differently according to their believes about the nature of partners (real vs. simulated persons). Empirical evidence shows that being convinced of playing against a machine or a simulated person exerts a downward bias on the fraction sent in the Trust Game (Johnson and Mislin, 2011, p. 873).

However this does not hinder our results for a twofold reason: firstly, since our research design entails a DID, any potential downward bias occurring in both surveys is eliminated by the estimation technique; secondly, any residual downward effect would *a fortiori* strengthen our results.

5.2 Robustness checks

Another potential concern for the results shown in Table 4 is related to possible self-selection bias in the control group. As a robustness check we provide a Propensity Score Matching (PSM) procedure for the results. PSM is an estimation technique to evaluate the effectiveness of treatment in observational studies, extensively used since the seminal work by Rosenbaum and Rubin (1983). The aim is to estimate ATT by comparing treated and controls, conditioning on a set of relevant covariates (Caliendo and Kopeinig, 2008; Becker et al., 2002). In order to proceed with the test we ignore

[31] In fact our dependent variable is the fraction of soups sent to the other person and is therefore constrained within the range 0-1. Some researchers addressed this issue by implementing a Tobit model, but Papke and Wooldridge (1996) showed that censored regression techniques do not apply for variables with infeasible values beyond the censoring point. Therefore, GLM models are included in the table as a robustness check.

[32] Burks et al. (2003, p. 196-7) show that "Non-White participants exhibit less trust than whites in a mostly white environment". In our case, however, one should consider that the incarceration rate in the USA is higher for ethnic minorities. In an alternative model specification, not reported here, we run the same models as in columns (1) to (4) including a single ethnic dummy for Non-White inmate: the coefficient is always never statistically significant.

the pre-/post- treatment design and consider only the outcomes of Trust in the second survey, by conditioning them on inmates' individual characteristics. The Propensity Score, i.e. the probability of being part of the Treated group, is therefore calculated by balancing the same covariates that we include in extended model specification (columns 5 to 8 in Table 4)[33]. By imposing "common support" on Treated and Control, the actual matching will consider only Treated and Control with propensity scores within the range of the control group values. In this way, potential outliers in the Treated group are ignored, and the estimated ATT is robust to potential unobserved self-selection mechanisms.

Table 5: Robustness check: Propensity Score Matching, dep. var.: trust

Matching method:	NN† (1)	Radius† (2)	Kernel‡ (3)
ATT	0.188**	0.173**	0.109
	(0.094)	(0.082)	(0.077)
T-stat	2.00	2.10	1.42
Observations	79	79	79
Controls	20	29	37
Treated	41	32	41

Note: common support always requested; balancing property satisfied.
Covariates: altruism, age, single, soup value, ethnic group dummies.
Standard errors (†) or Bootstrap standard errors (‡) in parentheses.
*** p<0.01, ** p<0.05, * p<0.10

Once the propensity score has been calculated[34] and the balancing property successfully tested, the choice of the matching algorithm is of pivotal importance since it affects both the bias and efficiency of the estimated ATT (Caliendo and Kopeinig, 2008). For this reason, Table 5 reports the estimated ATT according to three alternative matching algorithms extensively applied in the PSM literature, namely: 1-to-1 Nearest-Neighbor (NN), Radius Caliper and Kernel[35].

Model (1) and (2), shown in Table 5, report statistically significant ATT coefficients, very close to those estimated in Table 4 (ranging from 0.17 to 0.19) and rely on algorithms that are based on similar procedures, identifying the most suitable observation in the control group for each treated one. These procedures are the most straightforward application of PSM and allow to compare similar individuals. For sake of completeness, column (3) in Table 5 reports the estimation of the ATT when the Kernel technique is applied. The lack of statistical significance in the Kernel model estimation is likely to be driven by the fact that this technique uses all the available information to generate the counterfactual outcome for the treated, thus including also potential "bad matches", i.e. controls with propensity scores very far from the treated (Caliendo and Kopeinig, 2008). Therefore, PSM overall confirms and strengthens our main result,

[33] Due to the limited size of the sample, the prison dummy has been not included since it would have further reduced the number of available propensity scores for the matching. All other covariates are indeed included in the propensity score estimation.

[34] The results of the logit estimation are shown in the Appendix.See Table A2

[35] See Caliendo and Kopeinig (2008) for a survey of pros and cons of different matching algorithms.

supporting the evidence that trust, as measured by the fraction of soups sent to an anonymous partner in the Trust Game, increased in inmates participating to GRIP.

6 Conclusions

The present study tests whether GRIP, a specific offender accountability program for long-term sentenced inmates, implemented in two Californian State Prisons (Avenal and Mule Creek), changes the prosocial behavioral attitudes of inmates. The paper describes the results of a Lab-in-the-Field experiment, based on a longitudinal design spanning over a period of 10 months with a specific focus on trust. The research protocol envisaged the administration of two questionnaires including a set of behavioral situations ("games"), widely used in the experimental and behavioral economics literature, namely the Dictator Game (Kahneman et al., 1986) and the Trust Game (Berg et al., 1995), to 80 inmates, 42 treated (enrolled in the program) and 38 controls. The first questionnaire was administered before the start of the program; the second questionnaire at the end of the program.

A DID estimation procedure shows that trust significantly increased in GRIP participants compared to the control group. This result is robust to alternative estimation techniques and to the inclusion of an endogenous behavioral measure of altruism (measured by a Dictator Game).

The results of the paper support the claim that an offender accountability program, such as GRIP, produces beneficial effects on the inmates' prosocial preferences and attitudes beyond its primary aims. A wider application of these programs can thus be thought as an effective instrument, in the short-term, to re-establish or strengthen prosocial behavior in inmates and, in the long term, to facilitate rehabilitation processes by fostering inmates' reintegration and re-socialization in their communities, thus potentially contributing to the reduction of recidivism.

References

References

Becker, S. O., A. Ichino, et al. (2002). Estimation of average treatment effects based on propensity scores. *The stata journal 2*(4), 358–377.

Berg, J., J. Dickhaut, and K. McCabe (1995). Trust, reciprocity, and social history. *Games and economic behavior 10*(1), 122–142.

Blattman, C., J. C. Jamison, and M. Sheridan (2017). Reducing crime and violence: Experimental evidence from cognitive behavioral therapy in liberia. *The American Economic Review 107*(4), 1165–1206.

Bottom, W. P., J. Holloway, G. J. Miller, A. Mislin, and A. Whitford (2006). Building a pathway to cooperation: Negotiation and social exchange between principal and agent. *Administrative Science Quarterly 51*(1), 29–58.

Burks, S. V., J. P. Carpenter, and E. Verhoogen (2003). Playing both roles in the trust game. *Journal of Economic Behavior & Organization 51*(2), 195–216.

Caliendo, M. and S. Kopeinig (2008). Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys 22*(1), 31–72.

Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton: Princeton University Press.

Camerer, C. and K. Weigelt (1988). Experimental tests of a sequential equilibrium reputation model. *Econometrica 56*(1), 1–36.

Chiaramello, S., M. T. M. Sastre, and E. Mullet (2008). Seeking forgiveness: Factor structure, and relationships with personality and forgivingness. *Personality and Individual Differences 45*(5), 383–388.

Chuang, Y. and L. Schechter (2015). Stability of experimental and survey measures of risk, time, and social preferences: A review and some new results. *Journal of Development Economics 117*, 151–170.

Durose, M. R., H. N. Snyder, and A. D. Cooper (2015). Multistate criminal history patterns of prisoners released in 30 states. Special Report, Bureau of Justice Statistics.

Fehr, E. (2009). On the economics and biology of trust. *Journal of the european economic association 7*(2-3), 235–266.

Figlio, R. M., T. Sellin, and M. E. Wolfgang (1972). *Delinquency in a birth cohort*. Chicago: University of Chicago Press.

Forsythe, R., J. L. Horowitz, N. E. Savin, and M. Sefton (1994). Fairness in simple bargaining experiments. *Games and Economic behavior 6*(3), 347–369.

Fukuyama, F. (1995). *Trust: The social virtues and the creation of prosperity*. New York: Free Press Paperbacks.

Gibson-Light, M. (2016). Must work for food: The politics of nutrition and informal economy in an american prison. Paper presented at the 111th ASA Annual Meeting, Seattle, 20-23 August.

Grimalda, G. and L. Mittone (2011). Generalized trust: An experimental perspective. In *Social Capital, Corporate Social Responsibility, Economic Behaviour and Performance*, pp. 260–281. London: Palgrave Mac-Millan.

Guala, F. and L. Mittone (2010). Paradigmatic experiments: the dictator game. *The Journal of Socio-Economics 39*(5), 578–584.

Guiso, L., P. Sapienza, and L. Zingales (2008). Trusting the stock market. *the Journal of Finance 63*(6), 2557–2600.

Haney, C. et al. (2003). The psychological impact of incarceration: Implications for post-prison adjustment. *Prisoners once removed: The impact of incarceration and reentry on children, families, and communities 33*, 66.

Harrison, G. W. and J. A. List (2004). Field experiments. *Journal of Economic literature 42*(4), 1009–1055.

Johnson, N. D. and A. A. Mislin (2011). Trust games: A meta-analysis. *Journal of Economic Psychology 32*(5), 865–889.

Kahneman, D., J. L. Knetsch, and R. Thaler (1986). Fairness as a constraint on profit seeking: Entitlements in the market. *The American economic review 76*(4), 728–741.

Knack, S. and P. Keefer (1997). Does social capital have an economic payoff? a cross-country investigation. *The Quarterly journal of economics 112*(4), 1251–1288.

Lynch, J. P. and W. J. Sabol (2004). Assessing the effects of mass incarceration on informal social control in communities. *Criminology & Public Policy 3*(2), 267–294.

McCorkle, R. C. (1992). Personal precautions to violence in prison. *Criminal justice and behavior 19*(2), 160–173.

Morenoff, J. D. and D. J. Harding (2014). Incarceration, prisoner reentry, and communities. *Annual review of sociology 40*, 411–429.

Mullet, E., M. Girard, and P. Bakhshi (2004). Conceptualizations of forgiveness. *European psychologist 9*(2), 78–86.

Naderi, N. (2014). *Prisonization*. Blackwell Publishing Ltd.

Newman, W. J. and C. L. Scott (2011). Brown v. plata: prison overcrowding in california. *The journal of the American Academy of Psychiatry and the Law 40*(4), 547–552.

Papke, L. E. and J. M. Wooldridge (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics 11*(6), 619–632.

Paulle, B. (2017). Stumbling on the rehabilitation gold? foucault vs. foucault in san quentin and beyond. *Ethnography 0*(0), 0.

Pelucchi, S., F. G. Paleari, C. Regalia, and F. D. Fincham (2013). Self-forgiveness in romantic relationships: It matters to both of us. *Journal of Family Psychology 27*(4), 541.

Petersilia, J. (2014). California prison downsizing and its impact on local criminal justice systems. *Harvard Law & Policy Review 8*(2).

Putnam, R. D. (1993). *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton: Princeton University Press.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *The American economic review 83*(5), 1281–1302.

Rosenbaum, P. R. and D. B. Rubin (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika 70*(1), 41–55.

Sanfey, A. G., J. K. Rilling, J. A. Aronson, L. E. Nystrom, and J. D. Cohen (2003). The neural basis of economic decision-making in the ultimatum game. *Science 300*(5626), 1755–1758.

Sapienza, P., A. Toldra-Simats, and L. Zingales (2013). Understanding trust. *The Economic Journal 123*(573), 1313–1332.

Sutter, M. and M. G. Kocher (2007). Trust and trustworthiness across different age groups. *Games and Economic Behavior 59*(2), 364–382.

Twenge, J. M., R. F. Baumeister, C. N. DeWall, N. J. Ciarocco, and J. M. Bartels (2007). Social exclusion decreases prosocial behavior. *Journal of personality and social psychology 92*(1), 56.

Viceisza, A. C. (2016). Creating a lab in the field: economics experiments for policy-making. *Journal of Economic Surveys 30*(5), 835–854.

Wakefield, S. and C. Uggen (2010). Incarceration and stratification. *Annual review of sociology 36*, 387–406.

Walmsley, R. et al. (2013). *World prison population list* (tenth ed.). London: International Centre for Prison Studies.

Warren, M. E. (1999). *Democracy and trust*. Cambridge: Cambridge University Press.

Zingraff, M. T. (1975). Prisonization as an inhibitor of effective resocialization. *Criminology 13*(3), 366–388.

A Correlation matrix

Table A1: Pairwise correlation of main variables

	trust	altruism	age	single	soup like
altruism	0.349				
	(0.000)				
age	−0.235	−0.174			
	(0.003)	(0.030)			
single	−0.234	−0.002	−0.114		
	(0.003)	(0.982)	(0.151)		
soup like	0.027	−0.119	0.206	−0.118	
	(0.738)	(0.135)	(0.009)	(0.139)	
soup value	0.076	−0.14	−0.005	-0.202	0.350
	(0.345)	(0.079)	(0.952)	(0.011)	(0.000)

P-values in parentheses.

B Propensity Score estimation

Table A2: Propensity Score estimation: Logit model

Dep. var.:	Treatment
Covariates	(1)
altruism	−0.891
	(1.026)
age	−0.055*
	(0.029)
single	0.015
	(0.513)
soup value	−0.107
	(0.080)
White	0.471
	(0.616)
Native American/Alaska	0.477
	(0.859)
Hawaiian Native/Pacific	0.224
	(1.575)
Asian	−1.225
	(1.366)
Black/African American	0.190
	(0.637)
Hispanic/Latino	−0.140
	(0.573)
Observations	79
Log-Likelihood	−51.24

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.10
Note: Logit estimation, marginal effects reported.

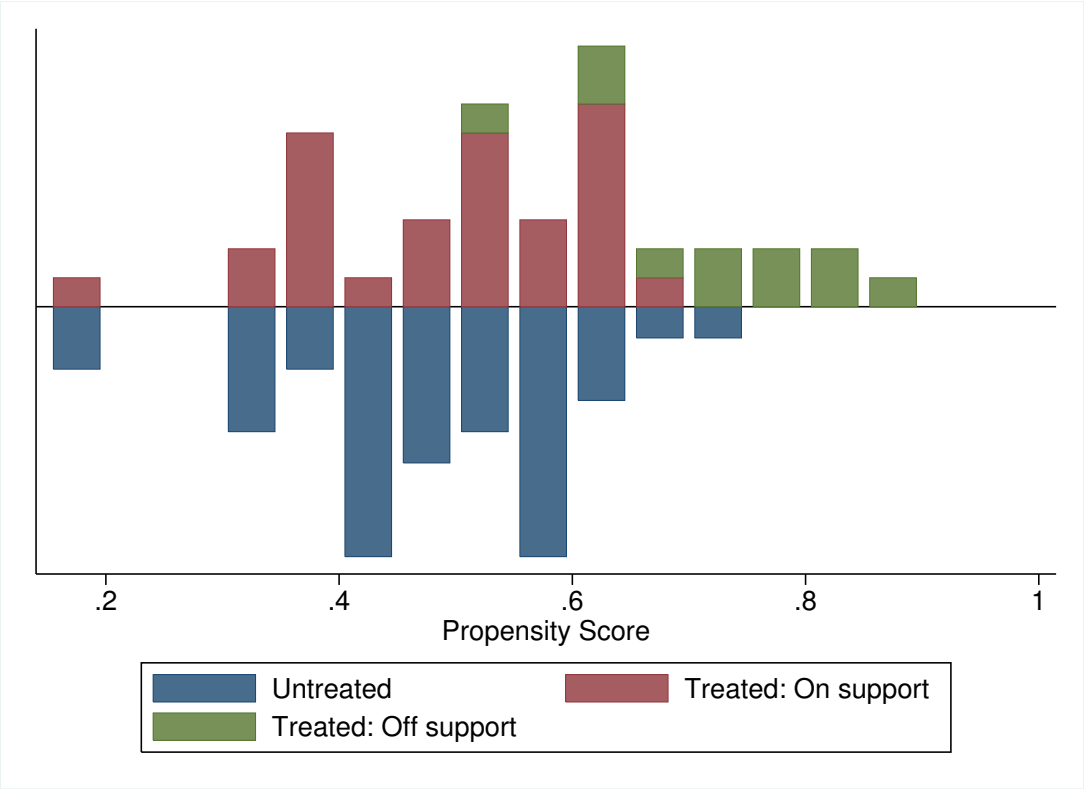


Figure 2: Estimated Propensity Scores: Common Support

C Supplementary materials

C.1 Survey instructions

Before each session of the survey begins, the following instructions are read aloud by the interviewer and handed in to inmates (below horizontal line):

Welcome and thank you for participating in this survey.

Your choices and answers will be treated in a strictly anonymous manner. Your name will not appear in the survey and will not be attached in any way to your answers. The interviewers are students who will not be involved in any further stage of the survey. **Results will be used and published in an anonymous way**.

We are interested in checking some factors that influence the way people make decisions. During the survey you will be asked to answer some questions and asked to make decisions in a number of situations. In one of these situations you must decide individually; in another one, your answer will be matched with that of another person who has already performed this task. This interaction will take place in a completely anonymous way. The identity of the other person will never be revealed, either during or after the end of the survey.

The other person has received these same instructions. During the survey you will be asked to make some choices. The choices you make will give you the opportunity to receive a certain number of dried soups. In the situation involving an individual decision, your choice will directly determine the number of earned dried soups. In the situation where you will be matched with another person, your choice will be combined with the choice made by the other person and will determine the final number of dried soups you receive.

You will answer some questions and make decisions in 2 situations. Each decision will get you a number of dried soups. At the end of the survey, one situation out of the two will be randomly selected (through a coin toss).This will determine how many soups you will ultimately receive. Immediately at the end of the survey, you will receive the number of dried soups earned in the selected situation.

For the participation to the survey, you will get a gift.

During this session you will be asked to:

1. fill in the survey;

2. make your choices in a 6 different situations.

Overall, the session will last approximately twenty minute. We ask you to work alone and in silence.

Thank you for your participation!

C.2 Tutorial Booklet

Starting from the following page we present the tutorial booklet that includes visual instruction of the incentivized tasks (Dictator Game and Trust Game), illustrated to inmates before each experimental session started.

TUTORIAL BOOKLET
SECTION 1

Section 1

In this situation you are interacting with another anonymous person, whose identity will not be revealed, either during or after the end of the survey. The same applies to the other person: he will never know who you are. A different role will be assigned to each of you. **You have been assigned to be Person A.**

You are provided with **10 dried soups**, while Person B is provided with none. You must decide how many dried soups to give to Person B.

Eventually you will obtain the initial amount of dried soups **minus the dried soups you have given** to Person B.

Person B will get the dried soups that you have decided to give to him.

Section 1: INSTRUCTION

You have been assigned to be Person A.

You are endowed with 10 dried soups.
Person B is endowed with none.

A	B
	
10	**0**

1

SECTION 1

Section 1: INSTRUCTION

You have been assigned to be Person A.

You are endowed with 10 dried soups.
Person B is endowed with none.

A

10

You must decide how many soups to give to Person B.

Eventually you will obtain the initial amount of soups less the soups you have given to Person B.

Person B will get the soups that you have decided to give to him.

Section 1: INSTRUCTION

You have been assigned to be Person A.

Remember, you are endowed with 10 soups, B with zero.

A

10

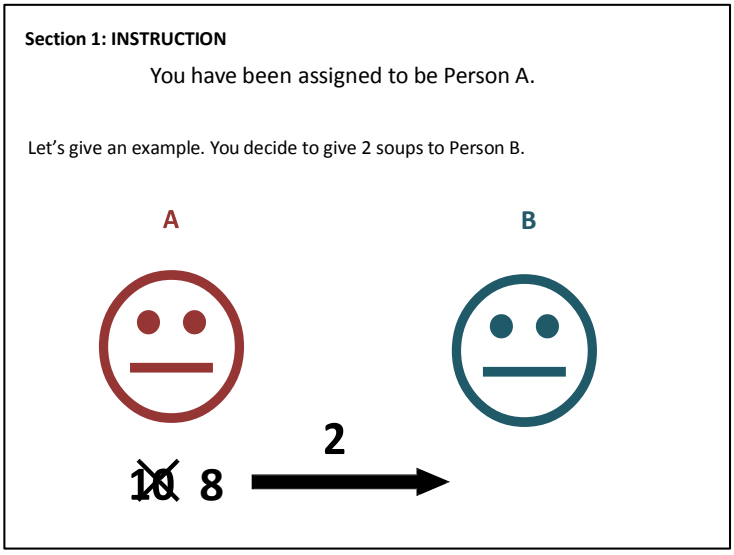
You have to decide if keeping these soups for yourself or giving part or all of them to Person B.

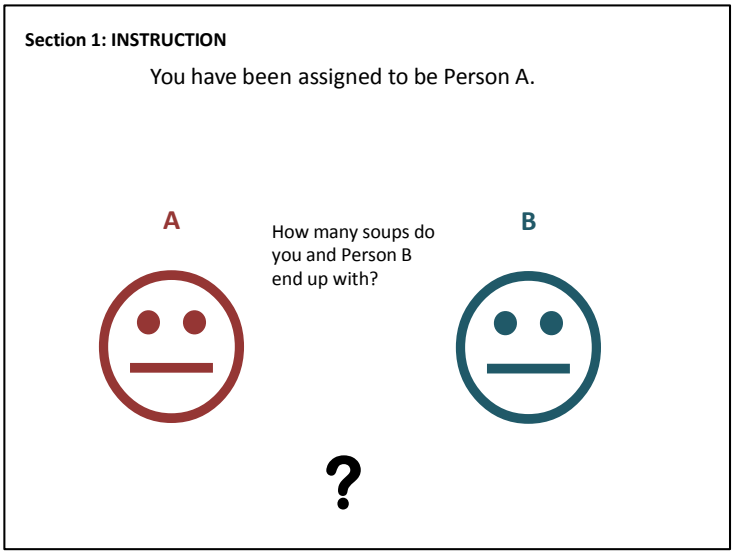
B

?

2

SECTION 1


Section 1: INSTRUCTION
You have been assigned to be Person A.
Let's give an example. You decide to give 2 soups to Person B.
A
B
10
8
2


Section 1: INSTRUCTION
You have been assigned to be Person A.
A
How many soups do you and Person B end up with?
B
?

3

SECTION 1

Section 1: INSTRUCTION

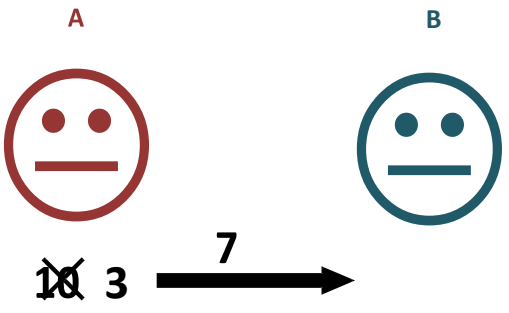
You have been assigned to be Person A.

At the end of this interaction, you retain 8 soups and Person B receives 2.

Section 1: INSTRUCTION

You have been assigned to be Person A.

Let's give another example. You are still endowed with 10 soups.
You decide to give 7 soups to Person B.



4

SECTION 1

Section 1: INSTRUCTION

You have been assigned to be Person A.

A

How many soups do you and Person B end up with?

B

?

Section 1: INSTRUCTION

You have been assigned to be Person A.

At the end of this interaction, you retain 3 soups and Person B receives 7.

A

3

B

7

5

SECTION 2

Section 2

In this situation you are interacting with another anonymous person, whose identity will not be revealed either during or after the end of the survey. The same applies to the other person: he will never know who you are. A different role will be assigned to each of you. **You have been assigned to be Person A.**

You have been provided with **10 dried soups** as initial endowment. You have to decide if keeping these dried soups for yourself or giving part or all of them to Person B.

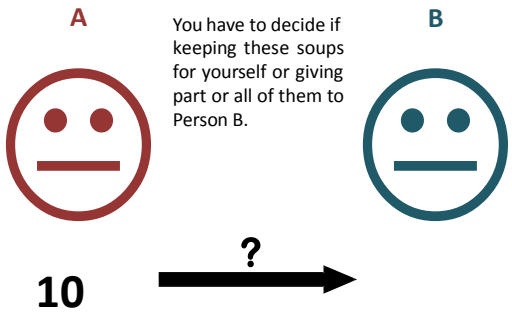
All the dried soups that you choose to give will be tripled; so for every dried soup you decide to give, Person B will receive 3 dried soups. Once the other person has received the dried soups that you have decided to give him (multiplied by three), he will have to decide in turn if and how many dried soups to give back to you. At the end, you will obtain:

- The amount of dried soups equal to 10 (the initial endowment)
- Minus the amount of dried soups you give to him
- Plus the amount of dried soups that he has decided to give you back.

Section 2: INSTRUCTION

You have been assigned to be Person A.

You are endowed with 10 dried soups.



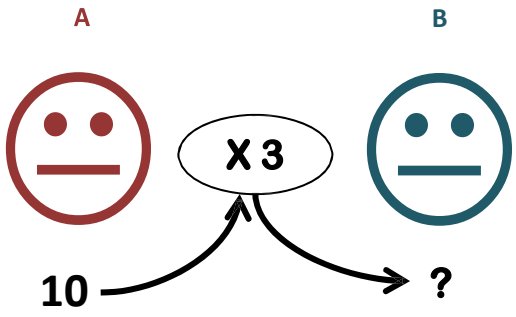
11

SECTION 2

Section 2: INSTRUCTION

You have been assigned to be Person A.

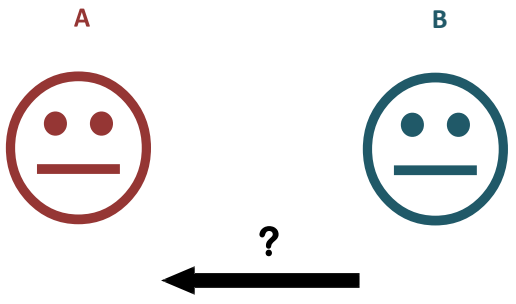
All the soups that you choose to give will be tripled, and so, for every soup you decide to give, Person B will receive 3 soups.



Section 2: INSTRUCTION

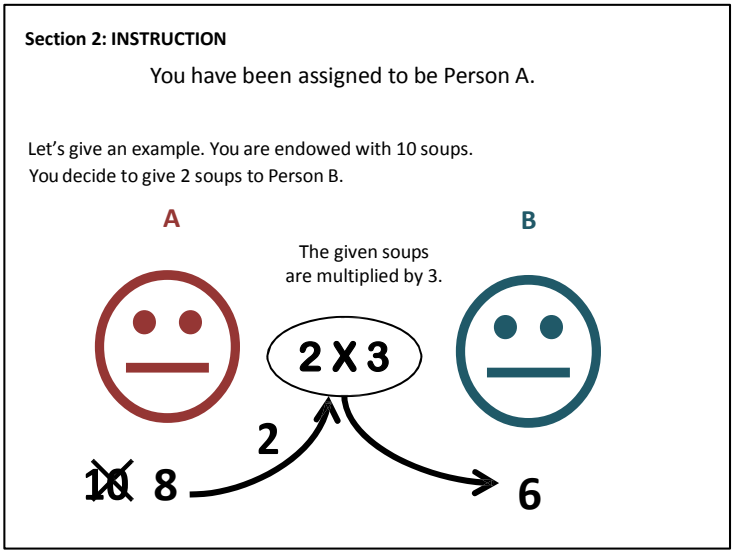
You have been assigned to be Person A.

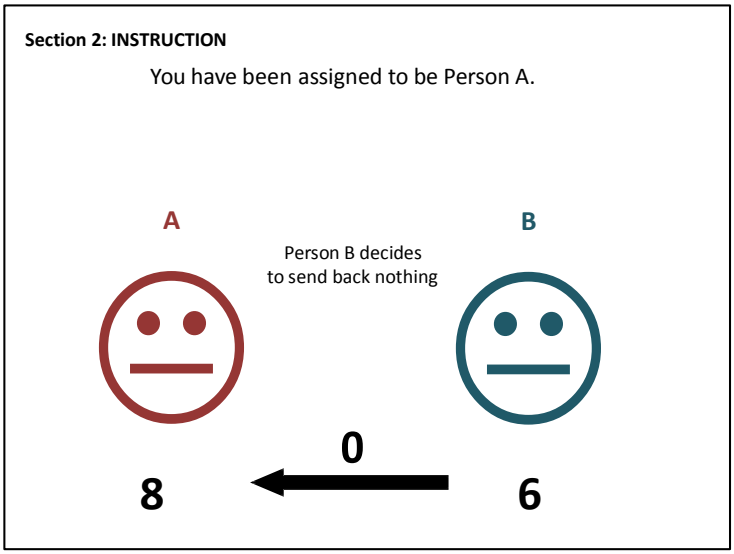
Once the other person has received the soups that you have decided to give him (multiplied by three), he will have to decide in turn if and how many soups giving back to you.



12

SECTION 2


Section 2: INSTRUCTION
You have been assigned to be Person A.
Let's give an example. You are endowed with 10 soups.
You decide to give 2 soups to Person B.
A
B
The given soups
are multiplied by 3.
2 X 3
2
10 8
6

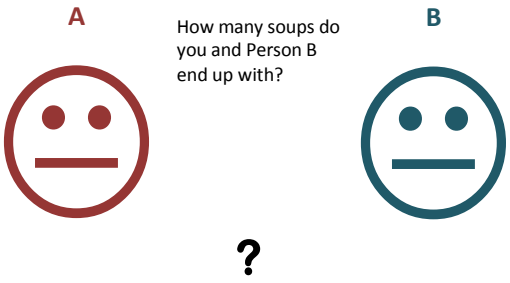

Section 2: INSTRUCTION
You have been assigned to be Person A.
A
B
Person B decides
to send back nothing
0
8
6

13

SECTION 2

Section 2: INSTRUCTION

You have been assigned to be Person A.



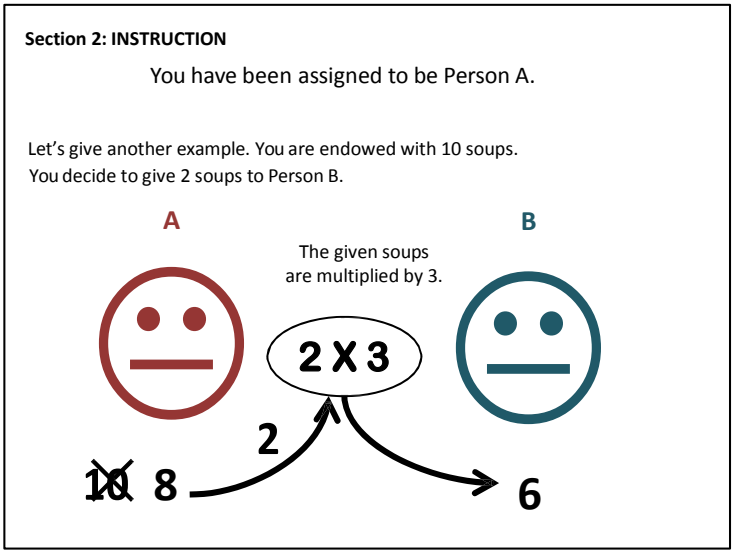
Section 2: INSTRUCTION

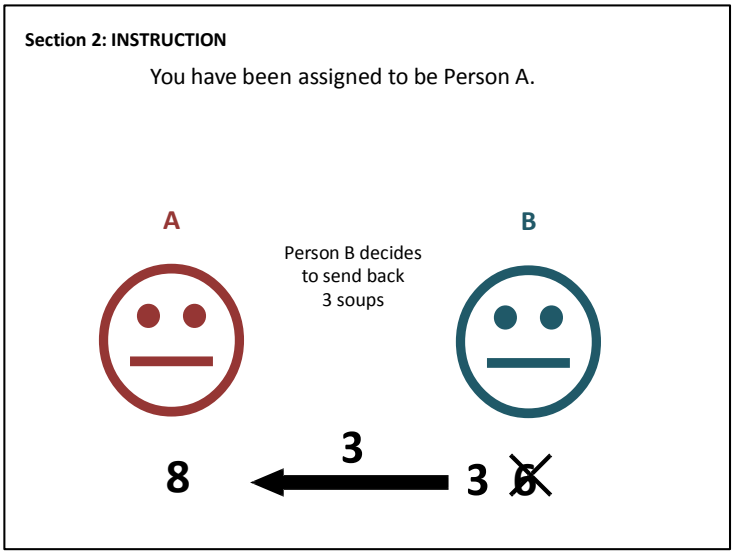
You have been assigned to be Person A.

At the end of this interaction you receive 8 soups and Person B receives 6.

14

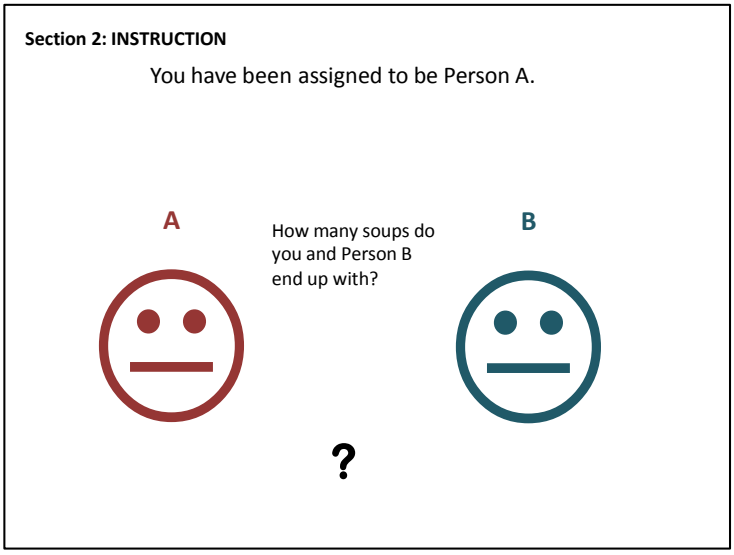
SECTION 2

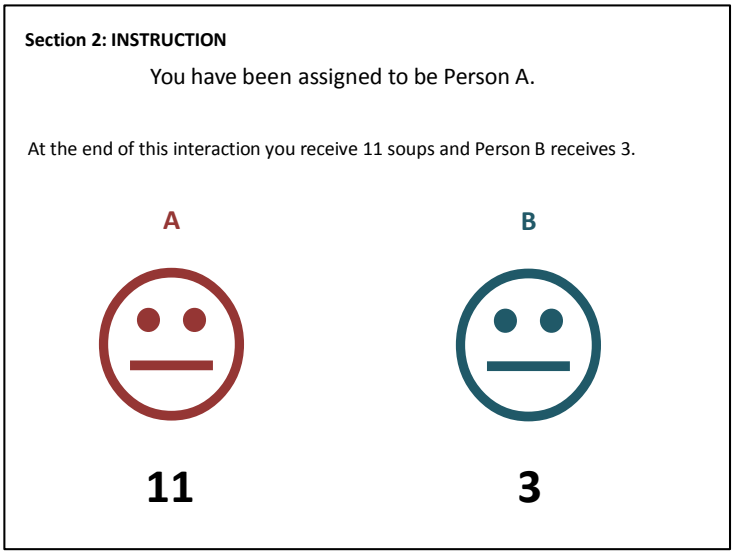

Section 2: INSTRUCTION
You have been assigned to be Person A.
Let's give another example. You are endowed with 10 soups.
You decide to give 2 soups to Person B.
A
B
The given soups
are multiplied by 3.
2 X 3
2
10 8
6


Section 2: INSTRUCTION
You have been assigned to be Person A.
A
B
Person B decides
to send back
3 soups
3
8
3 0

15

SECTION 2


Section 2: INSTRUCTION
You have been assigned to be Person A.
A
How many soups do you and Person B end up with?
B
?


Section 2: INSTRUCTION
You have been assigned to be Person A.
At the end of this interaction you receive 11 soups and Person B receives 3.
A
B
11
3

16

C.3 Dictator Game questionnaire

Section 1

In this situation you are interacting with another anonymous person, whose identity will not be revealed, either during or after the end of the survey. The same applies to the other person: he will never know who you are. A different role will be assigned to each of you. **You have been assigned to be Person A**.You are provided with 10 dried soups, while Person B is provided with none. You must decide how many dried soups to give to Person B. Eventually you will obtain the initial amount of dried soups minus the dried soups you have given to Person B.

Person B will get the dried soups that you have decided to give to him.

How many dried soups do you choose to give to the other person?

Please choose **only one of the following**, by ticking the appropriate circle:

- ◯ 0 dried soups
- ◯ 1 dried soup
- ◯ 2 dried soups
- ◯ 3 dried soups
- ◯ 4 dried soups
- ◯ 5 dried soups
- ◯ 6 dried soups
- ◯ 7 dried soups
- ◯ 8 dried soups
- ◯ 9 dried soups
- ◯ 10 dried soups

C.4 Trust Game questionnaire

Section 2

In this situation you are interacting with another anonymous person, whose identity will not be revealed either during or after the end of the survey. The same applies to the other person: he will never know who you are. A different role will be assigned to each of you. **You have been assigned to be Person A**. You have been provided with 10 dried soups as an initial gift. You have to decide if keeping these dried soups for yourself or giving part or all of them to Person B. All the dried soups that you choose to give will be tripled; so for every dried soup you decide to give, Person B will receive 3 dried soups. Once the other person has received the dried soups that you have decided to give him (multiplied by three), he will have to decide in turn if and how many dried soups to give back to you.

At the end, you will obtain:

– The amount of dried soups equal to 10 (the initial endowment);

– Minus the amount of dried soups you give to him;

– Plus the amount of dried soups that he has decided to give you back.

How many of your 10 dried soups do you decide to give to the other person?

Please choose **only one of the following**, by ticking the appropriate circle:

- ◯ 0 dried soups
- ◯ 1 dried soup
- ◯ 2 dried soups
- ◯ 3 dried soups
- ◯ 4 dried soups
- ◯ 5 dried soups
- ◯ 6 dried soups
- ◯ 7 dried soups
- ◯ 8 dried soups
- ◯ 9 dried soups
- ◯ 10 dried soups

C.5 Questions related to preferences for soups

To obtain information about preferences and value attached to soups by inmates, at the end of the survey the following questions were provided (below horizontal line):

Conclusion

Thanks for participating in this survey.
Before submitting the survey, please answer the following questions:

1. How much do you like dried soups? Please choose the appropriate answer by ticking the appropriate item

(1 = Don't like them at all, 10 = Like them very much)

I don't like them at all									I like them very much
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
○	○	○	○	○	○	○	○	○	○

2. Apart from your personal consumption, how much dried soups can be valuable for you? Please choose the appropriate answer by ticking the appropriate item

(1 = Not valuable at all, 10 = Very valuable)

Soups are not valuable at all									Soups are very valuable
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
○	○	○	○	○	○	○	○	○	○

C.6 Reward Booklet

Starting from the following page we present an example of the reward booklets used in the survey.

Each booklet reports different values, depending on the real choices made by the anonymous partners in a previous survey.

The first page of the booklet include an alphanumeric code relating to the anonymous partner: the answers included in the booklet refer to this specific person. Both situation 1 and situation 2 in the booklet include a self-explanatory table to allow the interviewer to perform the rewarding procedure. The numbers in the table presented in situation 2 vary according to the choices made by each anonymous partner and are therefore different each booklet.

In this supplementary materials section we present a single example.

REWARDS

PERSON B's CHOICES

(NOTE: you have been randomly matched with **person ####)**

TO THE INTERVIEWER, VERY IMPORTANT:

At the end of the survey, the prisoner tosses a plastic coin.

The number marked on the UP side will indicate which section of the survey will be rewarded.

Go to the corresponding page:

- If coin toss yields 1, go to page 1
- If coin toss yields 2, go to page 2

Then follow the instruction to correspond the correct reward!

PAYOFF INSTRUCTION **1**

If coin toss yields **1**

The situation to be rewarded is Section 1.

INSTRUCTION FOR THE INTERVIEWER:

1. Go to **Section 1 of the *GRIP-CSCC Survey*** filled in by the inmate: this section involves Person A's choice only.
2. Identify the prisoner's answer in the **grey column** of the REWARD TABLE below.
3. Read the corresponding reward in the **white column** and circle the reward.
4. Show the outcome to the inmate, in order to make clear to him the way he has been rewarded.

REWARD TABLE

Person A's answer	reward to Person A
0	10
1	9
2	8
3	7
4	6
5	5
6	4
7	3
8	2
9	1
10	0

5. Reward the prisoner with the correct amount of soups.

PAYOFF INSTRUCTION **2**

If coin toss yields **2**

The situation to be rewarded is Section 2.
Person B has made his choice about how many soups sending back to Person A (the prisoner).
The choices of both A and B generate the following REWARD TABLE for Person A (the prisoner):

REWARD TABLE

Soups sent by Person B	0	0	1	2	3	3	4	3	3	3
Person A's answer	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Person A's REWARD	**9**	**8**	**8**	**8**	**8**	**7**	**7**	**5**	**4**	**3**

INSTRUCTION FOR THE INTERVIEWER:

1. Go to **Section 2 of the *GRIP-CSCC Survey*** filled in by the prisoner and read Person A's answer**.**
2. Identify Person A's answer in the **grey row**.
3. The reward is given by the number reported in the corresponding **white cell** right below Person A's answer.
4. Circle the reward cell.
5. Show the outcome to the inmate, in order to make clear to him the way he has been rewarded.
6. Reward the prisoner with the correct amount of soups.

C.7 Plastic coin

Figure 3: Side 1 (head) and side 2 (tail) of the plastic coin used in the rewarding procedure