The influence of norms and norm deviance on punishment behaviour in an asymmetric public good game.

Stijn Beekhuis, 6011705
Supervisor: Kasper Otten
Second assessor: Dr. Anne-Rigt Poortman
Sociology, Utrecht , The Netherlands
14-06-2020

Abstract

Humanity has punished undesirable behaviour for centuries. Punishment is a mechanism which enforces cooperation in society. Previous research shows that without punishment, cooperation levels drop in social situations. However, actors themselves need to punish when there is no central authority to enforce laws. One potential reason to punish for individuals is self-interest, an individual punishes to achieve maximum payoff in the situation next. Another reason is that an individual supports a social norm and punishes norm deviations because norm deviations endanger the group. This paper researches what drives people to punish in social dilemmas without a central authority. The possible effects of the norms equality (equal payoff) and efficiency (equal contribution) are examined, and in what way these norms causes people to punish. An asymmetric public good game is experimentally conducted to measure the possible effects of norms and deviation from these norms. Multiple regression analysis is used to examine the data. The results show that norms have an effect on punishment, and norm deviation increases this effect. This research contributes to the sociological literature by examining the effect of norms on punishment behaviour in social dilemmas, in future research the effect of other norms than equality and efficiency could be examined.

Keywords: Punishment; asymmetric public good game; contribution norms; norm deviations

Utrecht University

Introduction

The first-ever punishment, according to Christian history, was the punishment of Adam and Eve by God. Adam and Eve were allowed to eat every kind of food in the garden of Eden, but the fruit from the tree of knowledge of good and evil was forbidden. God warned them for the consequences of eating the forbidden fruit and the first rule for humankind was made: Do not eat from the tree that yields knowledge of good and evil. Adam and Eve were seduced and they broke the rule. Therefore, they were expelled by God from the garden of Eden as punishment (Osman, 2000).

Rules and punishment have always been a part of human life. From the supposed punishment of God until the daily sanctions in today's life. Adam and Eve had a central authority, God, who made and executed rules and punishments. However, in current society there are social situations without a central authority, but behavioural rules are still applied. For example, waiting for your turn in a queue is a social rule. Everybody stays in line and expects this behaviour from other people involved. If somebody should jump the queue, he will most likely be addressed for his wrong behaviour. These rules are norms. According to Liefbroer & Billari (2010) norms describe what should be done, what is allowed to be done or what should not be done in a social situation. A norm is shared by a group of actors, this can be a whole country or a small group, like a family. The final part of the definition of norms contains punishment. If an actor fails to follow a norm, there will often be a sanction. (Liefbroer & Billari, 2010). A social sanction of being addressed for jumping the queue is an example, in this case shame serves as the punishment. To avoid punishment, people will normally conform to the norm. Norms can provide a way to prevent conflict and provide cooperation between actors in social situations. Cooperation is the collaboration of two or more actors who cooperatively achieve the best collective outcome (Reuben & Riedl, 2013).

In evolution, the cooperation of individuals ensures the persistence of certain genes or whole species. Cooperation to survive works in nature. Most of the time cooperation is between kin. The human race is an exception, people do not tend to exclusively limit cooperative behaviour with family members. Humans cooperate with non-kin more than any other species (Axelrod & Hamilton, 1981). A possible explanation is the level of intelligence of humans. Humans process complex information to decide the best of all actions. With the understanding of more complex information, social dilemmas and game theory become more important. People can deduce long-term consequences because of their intelligence, long-

term advantages for cooperating are considered. Therefore, cooperating with others is often the best choice to achieve goals or avoid sanctions (Axelrod & Hamilton, 1981).

Despite the high levels of cooperation among humans, norm deviations are still tempting and frequent. These deviating actors profit from the cooperation between others but do no invest themselves. This kind of behaviour is called free-riding. The cooperation in a public good and the risk of deviant behaviour like free-riding is of importance for the current society. A public good is a good which is non-rival and non-excludable, there is no rivalry because actors cannot be excluded from use and the good is non-excludable because use does not decrease availability for other actors. People can contribute to a public good to increase the collective value, or they take advantage of others by free-riding to collect a higher relative payoff. Actors face a social dilemma when they take face a public good because of possible contribution loss. Do they cooperate by contributing to the public good or will they free-ride on the contribution of others (Carpenter & Matthews, 2004). A well-known example of a current social dilemma is climate change. Each country benefits when greenhouse gas emissions are reduced in total, but the costs to reduce emissions are high. There is a potential free-riding risk if countries do not invest in reducing emissions but still profit from other countries investments (Hasson, 2010).

This paper contributes to the literature by examining punishment behaviour in an asymmetric public good game. The public good game describes a social dilemma where actors are able to contribute to a public good. As above explained, every actor gains benefit from the public good, regardless if they contribute or not. If an actor chooses self-interest and free-rides, the collective output will be less in total. The individual gain for the free-rider will be higher. When an actor contributes to the public good, the total collective payoff will be higher but his/her individual payoff will be lower because they lose a part of their contribution. A conflict arises between the two options, partly because of uncertainty what other actors will do. Thus, The individual gain decreases by contributing to the public good the collective payoff increases (Nikiforakis, 2008**)**.

Several articles are written about the symmetric public good game, where every actor benefits the same from investing in the public good (Muehlbacher & Kircler, 2009). Actors are allowed to punish in a public good game, which is costly. If somebody fails to contribute to the public good, people can reduce the earnings of the free-rider. It will cost the punisher a certain amount of earnings. This leads to a social dilemma for an actor; should he/she punish or not? Multiple outcomes of this so-called second-order dilemma will be examined in this paper. The dominant strategy is to never contribute and never to punish because both

contributing and punishing are costly. A rational player will always choose this dominant strategy according to the rational choice theory (Anderson, 2000**)**. However, the results of earlier research show frequent contributions and frequent use of punishment. Frequent use of punishment does increase the level of cooperation over time. Being punished is for free-riders costly, so they will contribute to the next interaction. Thus, the cooperation level will increase when punishment is possible in repeated public good games (Tan, 2008).

Actors are unequal in an asymmetric public good game. There are different asymmetric public good games, in real-life social dilemmas people are also in multiple ways unequal. When there are differences in endowments, some actors can contribute more than others. A difference in benefits or returns from the public good is when some actors receive relatively more from the public good, this asymmetric public good game will be used in this paper. Actors are different in return rate from the public good, high return actors receive relatively more from the public good than low return actors. The difference in return rate affects mechanisms behind the solution of the second-order dilemma as will be explained later (Reuben & Riedl, 2013). Furthermore, this paper will focus on contribution norms and how this alters punishment behaviour. Earlier research about the symmetrical public good game is less realistic in real life situations because individual actors often differ in their relative benefit from the public good. For example; a truck driver benefits more from a good infrastructure than someone who works at home and drives less (Tan, 2008). Actors have different views of fairness principles in asymmetric public games, multiple outcomes can be seen as fair by different players. Contributing the same is different than having the same payoff, while both situations can be seen as fair. People might punish differently because of this difference. Therefore, the research question for this paper is:

In what way do contribution norms influence punishment behaviour in asymmetric social dilemmas?

Multiple mechanisms are used to conduct this research. Two different kinds of models for how humans act are used, the homo economicus and the homo sociologicus. The homo sociologicus will act according to a norm, while the homo economicus acts for maximum individual gain. Furthermore, the fairness principles which are discussed in the paper of Reuben & Riedl (2009) will be a central factor in this paper, by dividing contribution norms in two fairness principles: equality and efficiency. The theory section will explicate more about the mechanisms behind an asymmetric public game. To investigate the influence of

different contribution norms and punishment behaviour, data is collected through an experiment where a public good game is played and punishing for free-riding behaviour is possible. Behavioural norms are measured in the experiment, to connect the norms with actual behaviour. This experiment will be further explained in the method section.

Theory

This section provides a theoretical background in norms and punishment. Firstly, the second-order dilemma is further explained. Secondly, two behavioural insights in the economy and the sociology are described. These models are important for the thoughts behind two basic strategies in a social dilemma, conform to norms or free-ride. Thirdly, fairness principles are clarified. Different thoughts about what is fair influences norms about contributions and punishment behaviour of actors. Lastly, the importance of punishment is explained and the hypotheses are lined up.

Second-order dilemma of punishment

As earlier stated, actors who have access to a public good face a social dilemma. Free-riding to enjoy the benefits without contribution could be tempting instead of contributing to the public good. This social dilemma is also known as the first-order dilemma (Okada, 2008). A way to prevent free-riding because of the first-order dilemma is to introduce a form of sanction, which motivates actors to conform to a norm. However, sanctioning without a central authority must be done by actors who support the public good. The possible social costs or risks of sanctioning by the punisher is another social dilemma which arises because of the first order dilemma. The consideration of punishing free-riders is therefore a second-order dilemma (Decker, Stiehler & Strobel, 2003). The earlier example of someone jumping a queue comes to mind, someone in the queue must point out the free-rider that staying in line is better for everyone. The punisher risks a confrontation with the actor who jumps the queue, thus addressing the hustled free-rider is costly. An important characteristic of the outcome of the second-order dilemma is the possible repeatability of the social situation. People could free-ride and not punish when they participate in a public social dilemma which occurs only one time without long-term consequences, this is the dominant strategy to achieve maximum surplus when people behave rational (Decker, Stiehler & Strobel, 2003; Carpenter & Matthews, 2004). However, in a repeated social situation with the same actors, people will learn and built a reputation which influences the first and second-order dilemma. To achieve

maximum earnings on a collective or individual level, actors must estimate if costly punishing would increase cooperation and thus surplus in future rounds of the social dilemma. Even when an actor wants to maximize his or her individual payoff, the dilemma of punishing is still there in repeated interactions. (Decker, Stiehler & Strobel, 2003; Carpenter & Matthews, 2004).

Homo economicus and homo sociologicus

The emergence, problems and use of norms in society have been researched by many social sciences like psychology, economy, sociology and anthropology. This paper focuses on the combination of the use of norms in two disciplines; sociology and economy. These disciplines are suited for systematic model making concerning norms and their influence on cooperation. Explaining the rules in society is for these two sciences of importance because both sciences seek to predict and expound society. An economist researches rational behaviour and the consequences in market situations. Sociology on the other hand examines bounded rationality in non-economic situations. Both types of research examine the behaviour of individuals and consequences at the macro-level. Norms are of importance for these sciences because individuals often follow norms that clarify and influence behaviour in society. Therefore, norms could help explain and possibly predict certain behaviour in social situations (Bendor & Swistak, 2001).

In some economic views, an individual is seen as a homo economicus. The homo economicus has a few characteristics which simplify human behaviour, he/she acts according to the rational choice theory. Characteristics concerning social exchanges are especially of importance for social dilemmas. The homo economicus is fully self-interested, fully rational and decisions are outcome based. The homo economicus wants to achieve a maximum profit in every social or business situation, but he/she will consider actions which could infer losses in future transactions. Thus, individuals which would be part of this Model take their reputation and possible consequences of selfish behaviour into account (Gintis, 2000). Applying the homo economicus as a model for every individual helps to predict behaviour with relatively simple mathematical models. However, real-life individuals generally do not conform to this model. People do act charitably or not fully rational. One could argue that individuals only act rationally charitable to improve their reputation which would increase their personal outcomes over time (Gintis, 2000). One example could be of conforming yourself to a norm in the neighbourhood to clean your driveaway of snow in the winter. By

doing so, your neighbours could tend to clean their driveaway sooner, which would help you to be mobile. In addition, the reputation of this certain individual will grow in the neighbourhood which would enlarge the chance of increased future help from neighbours. This strategy to increase the contribution by expecting something in return is called the 'tit-for-tat' strategy.

The homo economicus model explains certain behaviour very well, but according to Gintis (2000), there are many social situations where actors show different behaviour. The most important factor why the model fails is that people interact cooperatively and prosocial in strategic interactions, where the economic model predicts free-riding. The homo economicus has a similar consideration for punishing, in a one-time social dilemma he/she would never punish costly. However, in a repeated social interaction providing a stable earning could be the best rational way to maximize individual payoff. Punishing, even costly, could increase earnings over time because free-riders will contribute in the next round to prevent another punishment. This mechanism increases the payoff for the punisher over time. Thus, costly punishing could be the best rational choice to increase earnings in repeated social dilemmas according to the homo economicus (Anderson, 2000).

In sociology, people are traditionally seen as homo sociologicus. The homo sociologicus is derived from the sociological theory where people do not act fully rational and are not fully self-regarding (Abell, 1991). According to the sociological theory, behaviour is determined by group influence. Behaviour which serves the group or society is most likely preferred by a group. Group influence is most noticeable through norms. Individuals will follow these norms for rewards, or to avoid social punishment (Boyd & Richerson, 2001). According to the sociological theory people will almost always follow behavioural norms to serve to collective good under the threat of punishment. The common strategy in social dilemmas is always to conform to the norm, despite the possibilities of not maximizing the individual payoff. The example of cleaning the driveaway in a neighbourhood is applicable. According to the sociology theory, an individual will conform to the norm and clean his driveaway because of group influence. He/she wants to be part of the group and wants to avoid sanctions for deviant behaviour. Following the homo sociologicus characteristics, an actor would punish for the same reason as he/she would cooperate to the norm. Maintaining the norm is the main reason to punish for the homo sociologicus (Anderson, 2000).

What is fair

A public good with heterogenetic returns has two important fairness principles; efficiency and equality. These two principles stand for equal payoffs (equality) and equal contributions (efficiency). Equality and efficiency conflict with each other in a social dilemma with heterogenetic returns. The equality norm requires that every actor ends with the same payoff, despite differences in contributions. Reciprocity of outputs (i.e., payoffs) could ensure solidarity between actors to act according to the equality norm in the next rounds. The efficiency norm requires that every actor contributes the same, differences in payoff are not of importance. Reciprocity of inputs (i.e., contributions) could ensure solidarity between actors to act according to the efficiency norm in future interactions (Bardsley & Sausgruber, 2005; Carpenter & Matthews, 2004; Konow, 2003). A difference in return rates make it impossible to achieve both principles at the same time because equal contributions concludes in unequal payoffs and equal payoffs are only possible with unequal contributions. Without heterogenetic returns, equal contributions and equal payoffs would be the same, but in many real-life situations, there is heterogeneity (Reuben & Riedl, 2013). For example, a student house needs a new washing machine and everyone contributes equally. If one inhabitant washes at his or her parents' house, he/she has no payoff for the new washing machine but did incur the costs for it. This example shows the use of the efficiency norm, where the requirement of equal contributions is met but there are no equal payoffs.

Importance of punishment

The use of punishment in a social dilemma to prevent or discourage free-riding behaviour is widely known in science (Masclet & Villeval, 2008; Nikiforakis, 2008; Tan, 2008). A social situation without a central authority would, at first sight, be sensitive for free-riders. Rational actors would behave selfishly without laws or rules, which would lead to a low or non-existent level of cooperation. Famous research such as that of Hobbes (1651) argues that without an authority one actor would always free-ride, and therefore verbal agreements are useless to divide a public good. Costly punishment would also not be done since a second-order dilemma has the same characteristics as a first-order dilemma. However, Ostrom, Walker & Gardner (1992) concluded that costly punishment was far more common than expected in social dilemmas with a public good as a resource, which led to a higher cooperation level over time. Free-riders are directly disciplined by the potential of punishment to cooperate because being punished leads to a lower individual payoff (Fehr &

Gächter, 2000). With punishment, actors have a way to implement their ideas about fairness. If one of the actors fails to contribute according to certain standards, a different actor is able to show his or her dismay in the form of punishment. Norms like equal contribution or equal payoff would arise in a social dilemma of a public good where punishment is possible. As earlier stated, equal contribution and equal payoff in homogeneous public good games are the same outcome of the dilemma, but these two norms differ in heterogeneous public good games. The equality and efficiency norms arise in a heterogeneous public good game where punishment is possible and individual gain is not the only factor of importance (Fischbacher, Schudy & Teyssier, 2014).

The sociological theory suggests that people contribute and punish according to the norms they support and maintaining the norm to serve the group. The actors who support the equality norm will punish deviations from equal payoff. While others who support the efficiency norm will punish deviation from equal contributions. The mechanism of the sociological theory does not indicate which of the two norms actors will support, but that actors support norms and that norm deviation from any supported norm will be punished. Equality supporters punish group members with a higher payoff than them, while efficiency supporters punish group members with a lower contribution than them. Free-riders are seen different by the different norm supporters. Thus, low contributors are free-riders according to efficiency supporters, while big earners are free-riders according to the equality norm. Therefore, there is a difference between punishment behaviour, even when all actors are seen as homo sociologicus. The fairness principles do not matter assuming actors behave like the homo economicus, who support their self-interest. Contributing and punishing are according to the rational choice theory mechanisms to maximize individual payoff. High return players profit more from the public good and they lose more payoff when group members free-ride. Therefore, high return players would punish free-riders more than low return players because high return players lose more payoff.

Applied to the public good game

This section will apply behavioural theories and fairness principles to a repeated asymmetric public good game with costly punishment. The benefit returns of the particular public good game which will be used in this paper are heterogeneous. Some actors benefit relatively more when they contribute to the public good than others. Following the homo sociologicus, people act according to the fairness principles, because the group determines how an

individual must act (Abell, 1991). Following the fairness principles should be ones goal, so the difference in relative and absolute payoff between actors on an individual level is not a factor for the fairness principles (Carpenter & Matthews 2004). Actors who support equal payoffs view deviance from equality as free-riding, while actors who support equal contributions view deviance from efficiency as free-riding. Both kinds of actors will punish free-riders to maintain norms. The following hypotheses arise:

Hypothesis 1: *The more a person supports efficiency over equality, the more he/she will punish deviations from efficient contributions.*

Hypothesis 2: *The more a person supporters equality over efficiency, the more he/she will punish deviations from equal payoffs.*

Following the homo economicus perspective, high return players lose more benefits on an individual level when the public good stays low because of free-riders. When this is the case, high return players lose more payoff. Assuming players could show characteristics of the homo economicus and intend a maximum individual payoff (Gintis, 2000), high return players would lose more individual gain from free-riders. Therefore punishment behaviour of high return players will be stricter. The next hypothesis arises:

Hypothesis 3: *High return players will punish free-riders harsher than low return players in an asymmetric public good game*

Methods and data

The data is collected by executing a computerized heterogeneous public good game experiment among students from Utrecht University. The program is made with z-Tree software (Fischbacher, 2007). The experiment was conducted in the Experimental Laboratory for Sociology and Economics (ELSE) at Utrecht University in October and November 2019. 192 participants were divided into 64 groups of three people, an internet recruitment system (ORSEE) was used to recruit participants (Greiner, 2015). There were 8 sessions with each 24 participants, a session lasted approximately 75 minutes. Participants were paid on average 15 euros, quantity dependent on the results in the game where the minimum payment was 5 euro and the maximum 22 euros. There were 127 female participants (66%), 62 male (33%)

and 3 other. Participants were mainly students at the Utrecht University, 87 were Dutch and 105 from other countries. They could not communicate with each other in any way during the experiment apart from the anonymous computer interactions. The experiment was explained in written instructions. The repeated public good game is played in two parts of 10 rounds, every participant played 20 rounds in total. The method will be further explored in the next part. Firstly, the experiment is explained in more detail. Secondly, the used variables are operationalized. Lastly, the descriptive statistics are shown and the analysis is explained.

Experimental design

Before the students participate, their individual normative view on contribution fairness to a public good was measured. This is done by using a hypothetical public good game, where an appropriate contribution for hypothetical players is asked to the participant. When participants believe that high return players and low return players should contribute the same, efficiency is supported as normative view. When participants support that high- and low return players should earn the same, the normative view is equality. Groups of three are made to play the public good game, there are groups of players with similar normative views and groups of players with dissimilar normative views. The effect of these conditions will not be examined, this is already done in a different paper (Otten, Buskens & Ellemers, 2020). The rest of the experiment procedure is the same for both conditions. After 10 rounds, a rematch of the groups is done where one person of each group joins a different group. Groups who were sorted on similar views become dissimilar, and vice versa. The normative views are measured three times; before the game, after the first 10 rounds and at the end of the game. This procedure is done to clarify the effect of a newcomer on cooperation. Only the data from the first 10 periods will be studied because the effect of a newcomer is not of interest to the current paper.

The experimental design is comparable with that of Fehr & Gächter (2000). However, there are a few differences which will be elucidated. All individuals *i* start with the same endowment *E*. The endowment for each player is 20 points. Every round of the public good game has two stages, a contribution and a review stage. In the contribution stage each individual is able to contribute an amount of their endowment to the public good, c_i, where c_i ϵ *{*0, 1, …, *E}*. The unused endowment is kept for the individual. The sum of contributions to the public good will be multiplied by the multiplication factor (M > 1) and divided among the group members. Every point which is contributed per individual is worth more than one point

for the group as collective. The multiplication factor is 1.75 is this experiment. For an individual, one contributed point is worth less than one point ($m_i < 1$). This paper researches a heterogenous public good game, the return rates from the public good differ between participants in groups. The groups consist of three persons, where one player received relatively more profit than the other two. The high return player receives 0.75 points for each point in the public good, the low return players each 0.5. Therefore, it is more profitable to contribute nothing for every individual. However, for the group as a collective it is more profitable to contribute fully.

In the second stage, the review stage, participants are able to punish other players in the form of deduction points $p_{ij} \in \{0, 1, \ldots, \max(p_{ij})\}$ to each group member *j*. Every given deduction point costs 1 point for the punisher, but reduces the assigned group member's points by 3 (δ). Thus, punishment is costly. The individuals cannot see who gave them deduction points, to prevent revenge punishment**.** This is of importance to ensure that punishment behaviour is only influenced by contributions, without other motives. After the review stage, the next round starts with the contribution stage. The payoff for each participant after one round is:

$$\pi_i = E - c_i + m_i \sum_j c_j - \sum_{j \neq i} p_{ij} - \delta \sum_{j \neq i} p_{ji}$$

Operationalization

Dependent variable

Punishment behaviour is measured through the number of deduction points which participants give to group members in the review stage. The more deduction points an actor decides to give to an alter, the more strictly he/she punishes. The variable *pun_to_alter* shows the number of punishment points the participant allocates to a certain group member. *pun_to_alter* is a ratio variable, where 0 is the minimum deduction points given and 10 is the maximum. 1 punishment point costs the punisher 1 point of their individual return and reduces the return of the alter with 3 return points.

Independent variables

The predictor variables consist of two aspects; which norm participants follow, efficiency or equality, and how much group members deviate from these norms. To create a variable which

measures how much people follow the efficiency norm or the equality norm, the variable *norm_scale* is created. Participants have given their view on how much high return and low return players should contribute, their view on contributions from low return players is subtracted from their view on contributions from high return players. Thus, *Norm_scale* gives the difference between how much participants believe a high return player must contribute and what a low return player must contribute. If a participant believes a low return player should contribute less than a high return player, he/she supports the norm of equality. To ensure equal payoffs, the low return player should contribute less than the high return player. This would lead to a relative high score on *norm_scale*. A participant who supports fully equality scores a 10 on *norm_scale*. If a participant believes a low return and high return player should contribute approximately the same, he supports the norm of efficiency and *norm_scale* has a relative low score. A participants who supports fully efficiency has a score of 0 on *norm_scale.* Thus, a higher score on *norm_scale* indicates equality support. A lower score on *norm_scale* indicates efficiency support. A negative score indicates that low return participants should contribute more than high return actors, actors who supported this opinion were filtered because it was not supported by one of the used norms.

To test hypothesis 1; *The more a person supports efficiency over equality, the more he/she will punish deviations from efficient contributions*, the variable *efficiency_deviation* is created. The contribution of the alter is subtracted from the contribution of the ego which gives the deviation behaviour form the alter of efficiency. These contributions should be the same according to the efficiency norm, so a higher score on *efficiency_deviation* shows that the efficiency norm is not followed by the alter. A negative difference is also possible, in this case participants contributed less than their group members. These negative scores are excluded from the data with the norm support filter. A negative score implies that a participant does not adhere to their own norm. Therefore, the negative scores are not useful in the predictor variable in this research.

To test hypothesis 2; *The more a person supports equality over efficiency, the more he/she will punish deviations from equal outcomes,* the variable *equality_deviation* is created. The earnings from the game for the participant is subtracted from the output from the game of the alter. *Equality_deviation* shows the difference between the output from the ego and the output from the alter. The earnings should be the same according to the equality norm. A higher score on *equality_deviation* indicates that the alter has more payoff than the ego, the alter did not act according the equality norm in this case. *Equality_deviation* will be low when the equality norm is supported. A negative score means that the ego has a bigger payoff

than a group member. Negative scores are excluded from the analysis with the norm support filter, for the same reason why negative scores from *efficiency_deviation* are filtered.

To test hypothesis 3; *High return players will punish free-riders harsher than low return players in an asymmetric public good game*, the variabele *high_return* is used as predictor variable. *High_return* stands for the return rate of each individual participant in their group. A score of 0 stands for a return rate of 0.5 from the points in the public good, a score of 1 stands for a return rate of 0.75 from the points in the public good.

Control variables

There will be three control variables added to the regression analysis. One of the control variables is *Age*, to show a possible influence of age on punishment. Age is measured in years with a minimum of 18 and a maximum of 68 years. A different control variable is *Politics*, this variable contains information about the political orientation of a participant. The political orientation could influence how people punish and is therefore added in the regression analysis. *Politics* has a range of 1 to 9 wherein a lower score matches a more left political view and a higher score matches a right political view. *Condition* is the last control variable; this variable consists if a player has group members who are in normative agreement or not. Normative agreement (0) means that the participant is placed with group members who share the same contribution view, normative disagreement (1) means that the participant is placed with group members who differ in contribution view. As earlier stated, the influence of *condition* is examined in a different paper (Otten, Buskens & Ellemers, 2020), and is therefore not a predictor in this paper.

Table 1: Descriptive statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Punishment points allocated to alter (*pun_to_alter*)	3840	0	10	.360	1.159
Efficiency – equality scale (*norm_scale)*	3840	-20	10	3.841	4.611
Efficiency deviation alter (*efficiency_deviation)*	3840	-20	20	0	5.927
Equality deviation alter (*equality_deviation)*	3840	-27	27	0	8.874
Return rate ego (*High_return)*	3840	0	1	.333	-
Age	3840	18	68	24.080	6.446
Political view	3840	0	8	3.109	2.065
Condition group	3840	0	1	.500	-

Table 1 shows the descriptive statistics of the used variables. All variables are measured 3840 times, 192 participants played 10 rounds and were able to punish 2 people every round. The dependent variable *punt_to_alter* has a range of 0 to 10 punishment points allocated. The average is .36, which indicates that relatively few people punish or people punish lightly. *Norm_scale* has a range from -20 to 10, with an average of 3.841. A negative score indicates that participants believe a low return player should contribute more than a high return player, people with a score of 10 believe high return players should contribute 10 contribution points more than low return players. Many observations of *norm_scale* are between 0 and 10, a negative observation indicates that a participant does not support one of the norms. The ego believes that he/she should contribute less than the group members when *norm_scale* has a negative observation. The *efficiency_deviation* has a range from -20 to 20, a negative score indicates that an alter contributed more than an ego. A higher positive score means that there is a greater difference between the contribution of the ego and the alter. A higher score stands for more efficiency deviation because of the contribution difference. The mean of this variable is 0 because all participants are seen as ego and as alter. The same applies to *equality_deviation*, but with a range from -27 to 27 and the equality norm. *High_return* and *Condition* are fixed dichotomous variables, therefore their means are

perfectly .33 and .5. Only one-third of the alters receive a high return (1) and the remainder receives a low return (0). Half of the participants play in a normative agreeable group, which gives *condition* a mean of .5 The ranges of the variables, *Age* (mean = 24.08) , *Politics* (mean = 4.11) and *Condition* (mean = .5), are already discussed.

Analysis

To test the hypotheses, a multiple regression analysis is performed in SPSS version 25. The data is nested, observations are not fully independent of each other. A multiple regression is not the best method to analyse the data. Lack of time and experience is the reason why a multilevel analysis is not performed. As earlier mentioned, the negative scores of *norm_scale, equality_deviation* and *efficiency_deviation* were filtered form the analysis to remove participants who do not follow one of the norms. Leaning towards a certain norm could alter punishment behaviour, players who support the efficiency norm could punish more strictly than players who support equality. Equality supporters would accept lower contributions from low return players than efficiency supporters. However, this paper examines the deviation effect from norms on punishment and not the difference in punishing between norms. The effect of norms themselves will not be extensively looked into.

To show an interaction effect of group members who violate a norm, two interaction terms are made. Firstly, *norm_scale* and *efficiency_deviation* are used to create *norm_efficiency_deviation.* The interaction should have a **negative** influence on the effect of norms on punishment. The expectation is that a person that leans more towards equality than efficiency is less likely to punish deviations from efficiency (see figure 1). The structure of *norm_scale* shows that a lower score stands for an increasement for the efficiency norm. Deviation from efficiency and the score on the equality support leads to a negative interaction effect on punishment. However, if an alter deviates from a norm, the effect of this particular norm on punishment will be generally **positively influenced** following the sociological theory (see figure 2). This means that a negative interaction effect between efficiency and efficiency deviation is most likely positive, but the results are negative because of the structure of *norm_scale*. Thus, efficiency support and efficiency deviation will most likely enlarge punishment behaviour, supporting hypothesis 1. Secondly, the interaction term *norm_equality_deviation* is created by using *norm_scale* and *equality_deviation.* The expectation is that the interaction has a **positive** effect on the effect of norms on punishment, supporting hypothesis 2. Deviation triggers punishment to protect future collective payoffs

according to the sociological theory, this enlarges when there is a specific deviation from a supported norm such as equality (see figure 2).

To test hypothesis 3, a multiple regression analysis is performed with given punishment points (*pun_to_alter)* as the dependent variable. High return players punish more strictly according to the hypothesis because they lose more payoff when people free-ride, so a positive effect is expected. The independent variables *age*, *politics* and *condition* are used in every regression analysis to control for effects on punishment.

Figure 1: interaction effect equality

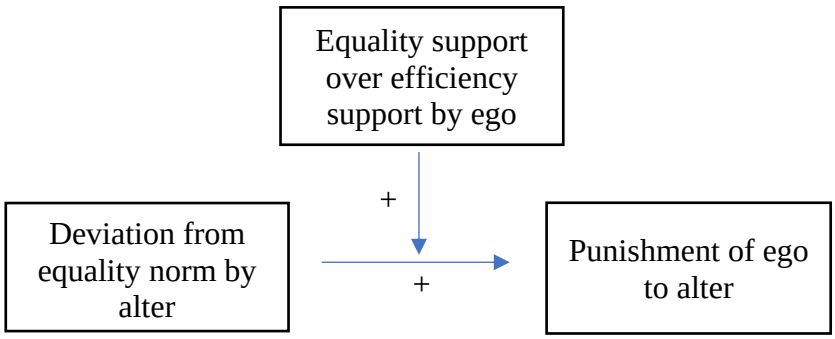
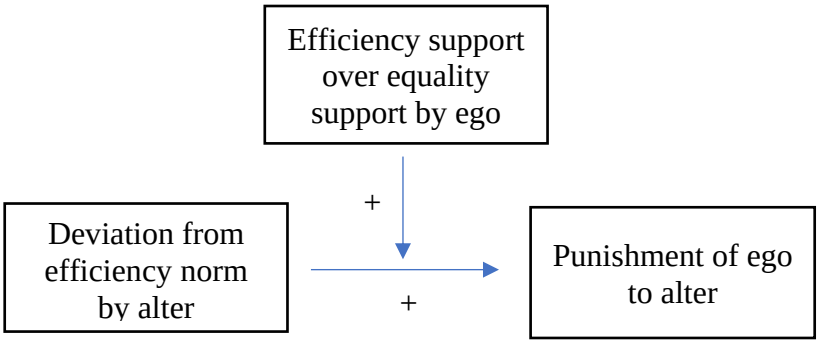


Figure 2: interaction effect efficiency



Results

Hypotheses 1 and 2

The effect of norms and deviation from norms with the support filter is shown in Model 1 and 2. As earlier explained, negative scores are removed because they do not comply with following a norm, and can therefore not be used as a measure for norms or deviation of norms. Model 3 and 4 show the same results but without the support filter, this is done to perform a robustness check between Model 1 and 2 on the one hand and Model 3 and 4 on the other. Model 1 and 3 do not have control variables, these are included in Model 2 and 4. All models are included with two interaction effects between *norm_scale* and *efficiency_deviation* and *norm_scale* and *equality_deviation*. Hypothesis 1: *The more a person supports efficiency over equality, the more he/she will punish deviations from efficient contributions* and hypothesis 2*: The more a person supports equality over efficiency, the more he/she will punish deviations from equal outcomes* are examined in these four models. Table 4 concentrates on hypotheses 3; *High return players will punish free-riders harsher than low return players in an asymmetric public good game.* In Table 4, Model 5 show the results without control variables, Model 6 show the results with control variables included. Regression assumptions have been checked and were met.

The main models for this paper are shown in Table 2. Model 1 is used to the see the influence of control variables on the predictor coefficients ($R2 = .255$, $F(5, 1671 = 115.778$, $p < .001$). Model 2 has an explained variance of 26.6% which is slightly better than the explained variance of Model 1 (25.5%). Model 2 is used to draw conclusions for hypotheses 1 and 2 because of the included control variables ($R2 = .266$, $F(3, 1668 = 75.396)$, $p < .001$). Norm scale has an insignificant positive effect on given punishment points ($B = .007$, $t = .584$, $p = .559$). To illustrate this result, an example with two participants: A participant who scores a zero on *norm_scale* supports efficiency and a participant who scores a 10 supports equality. The equality supporter punishes 10 multiplied by .007 times harsher than the efficiency supporter, this corresponds to .070 times more punishment points given. The punishment range is from 0 to 10, an increasement of .007 is negligible and insignificant. Adding control variables did not importantly alter the results. Deviation from equality has a positive effect on given punishment points ($B = .023$, $t = 3.531$, $p < .001$), participants punish alters who deviate from the equality norm with 10 payoff points will give .230 (.023*10) more punish points than to an alter who does not deviate from the equality norm. Efficiency deviation has a positive effect on given punishment points ($B = .198$, $t = 19.780$, $p < .001$).

Participants punish alters who deviate from the efficiency norm with 10 contribution points with 1.980 (.198 * 10) more punish points than an alter who does not deviate from the efficiency norm. Deviation from efficiency is thus punished harsher (B = .198, t = 19.780, p < .001) than deviation from equality (B = .023, t = 3.531, p < .001). A difference in contribution is seen as more reason to punish than a difference in payoff.

The results of hypotheses 1 and 2 are both shown in Table 2, Model 2. The norm efficiency deviation interaction has a small negative effect on punishment (B = -.018, t = -11.061, p < .001). These results support hypothesis 1. A lower norm score indicates support for the efficient norm and a higher score indicates support for the equality norm. The negative interaction means that persons who support efficiency (score lower on *norm_scale*) punish deviations from efficiency more strictly. For example, two participants are different in the norm they support. Participant A has a score of 0 on *norm_scale* and thus supports efficiency, participant B has a score of 10 on *norm_scale* and thus supports equality. Both participants have group members who deviate from efficiency, the group member contributes 10 points less than both participants. All participants have the same scores for the other variables. The results show that Participant A would give 1.980 - .066 = 1.914 punishment points because the interaction between *norm_scale* and *norm_efficiency_deviation* returns 0 when *norm_scale* is 0. Participant B would give .070 + 1.980 - 1.800 -.066 = .184 punishment points, because the score of 10 on *norm_scale* gives a positive effect (.007 * 10 = .070) but the interaction is a negative effect (10 * 10 * -.018 = -1.800). The equality supporter punishes the efficiency deviation less than the efficiency supporter. The constant is -.066, which is present for both participants. These results show that the punishment of norm deviations depends on a participants own norm. Hypothesis 2 is also confirmed, the interaction term is significant (B = .003, t = 2.125, p = .034). A positive effect indicates the more a participant supports equality, the more he/she would punish equality deviation behaviour from a group member. The interaction effect of *norm_equality_deviation* however is smaller than the interaction effect of *norm_efficiency_deviation*. This indicates that participants who support equality punish equality deviations less harsh than participants who support efficiency and punish efficiency deviations.

Table 2: Model 1 and 2 with norm support filter

	Model 1		Model 2	
	B	SE	B	SE
Norm scale	.015	.011	.007	.011
Efficiency deviation	.200***	.010	.198***	.010
Equality deviation	.023***	.007	.023***	.007
Norm efficiency deviation	-.018***	.002	-.018***	.002
Norm equality deviation	.002	.001	.003*	.001
Age in years	-	-	.011*	.005
Condition experiment	-	-	-.224***	.060
Political orientation	-	-	-.031	.016
(Constant)	-.065	.064	-.066	.190
Adjusted R2	.255		.266	
F change	115.778		6.296	
N	1677		1677	

****p <.001, ** p <.01, *p <.05*

Model 3 in Table 3 is a regression analysis without filter and control variables. Model 3 and 4 function as a robustness check for the analysis, adding the norm support filter did not unexpectedly change the results. Therefore, Model 1 and 2 are considered as leading results for the hypotheses. Model 3 explains for 9% the variances in punishment ($R2 = .090$, $F(5, 3834) = 77.101$, $p < .001$). Model 4 has an explained variance of roughly 10% ($R2 = 103$, $F(3, 3831) = 55.951$, $p < .001$).

The positive effect of *norm_scale* in Model 2 ($B = .007$, $t = .584$, $p = .559$) changes to a negative effect in Model 4 ($B = -.025$, $t = -6.157$, $p < .001$). Deviation from the efficiency norm ($B = .055$, $t = 12.271$, $p < .001$) and deviation from the equality norm ($B = .016$, $t = 5.839$, $p < .001$) are still positive, just like in Model 2. The interaction effect of norm scale with efficiency deviation in Model 4 ($B = -.004$, -6.055, $p < .001$) has also a negative direction just like in Model 2. The interaction of norm scale with equality deviation ($B = .000$, $t = .231$, $p = .817$) decreased to a nonsignificant effect. The robustness check shows that people who do not act according to their own norm punishes efficiency deviance stricter

because the effect of *norm_scale* changes to a negative effect. Apparently, contribution differences are more reason to punish than pay differences when actors who do not follow their norm are included. An important difference in the robustness check is that *norm_equality_deviation* does not affect punishment, this would mean that hypothesis 2 would not be confirmed if the norm support filter was not used. Thus, the deviation of norms has a less distinct effect on punishing in situations where people do not follow their own norms.

Table 3: Model 3 and 4 without filters

	Model 3		Model 4	
	B	SE	B	SE
Norm scale	-.026***	.004	-.025***	.004
Efficiency deviation	.054***	.005	.055***	.005
Equality deviation	.016***	.003	.016***	.003
Norm efficiency deviation	-.004***	.001	-.004***	.001
Norm equality deviation	.000	.001	.000	.001
age in years	-	-	.016***	.003
Condition experiment	-	-	-.140**	.036
political orientation	-	-	.021*	.009
(Constant)	.457***	.023	.050	.095
Adjusted R2	.090		.103	
F change	77.101		18.902	
N	3840		3840	

****p <.001, ** p <.01, *p <.05*

To clarify the interaction results in Model 2, figure 3 and 4 are specified. Model 2 shows a negative effect of the interaction term *norm_efficiency_deviation*, while figure 3 shows a positive effect. As earlier explained, the negative effect arises because of the structure of *norm_scale*. Wherein a lower score indicates a higher support for the efficiency norm. Therefore, the effect of -.018 in Model 2 can be interpreted as a positive interaction effect when taking the viewpoint on efficiency supporters. For every step a participant supports more efficiency than equality, the effect of deviation from efficiency on punishment increases with .018. Figure 4 shows the enlarged effect of support for the equality norm by ego on the positive effect from deviation from equality on punishment by ego. For every step a participant supports equality, the effect of deviation from equality on punishment increases with .003.

Figure 3: *interaction effect efficiency (Model 2)*

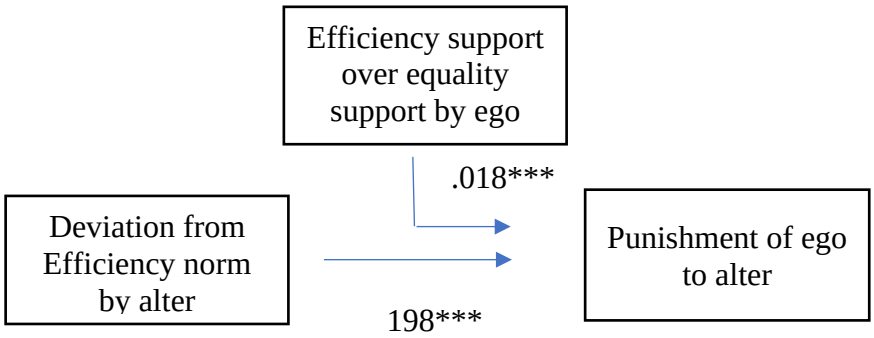
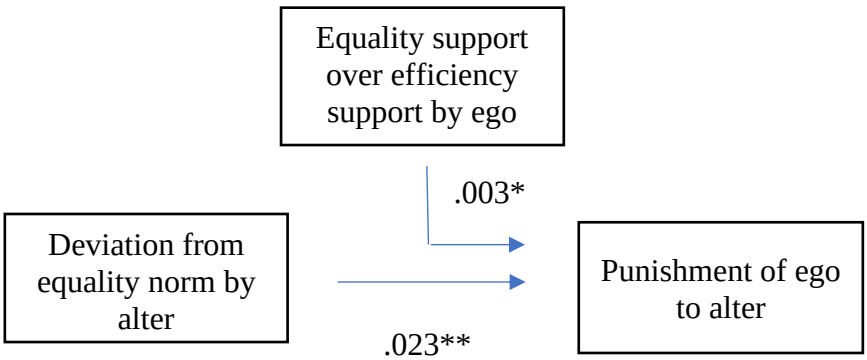


Figure 4: *interaction effect equality (Model 2)*



Hypothesis 3

To test hypothesis 3: *High return players will punish free-riders harsher than low return players in an asymmetric public good game,* a multiple regression analysis is performed. Model 5 in Table 4 shows the analysis without control variables (R2 = .001, F(1, 3838 = 6.262), p =.012). Three control variables are added in Model 6. The control variables do not change the coefficients extremely and the explained variance of the Model increases (R2 = .012, F(3, 3835 = 12.732) , p < .001) Therefore, Model 6 is used to test hypotheses 3. The return rate of a player has a negative effect on his or her punishment behaviour (B = -.092, t = -2.333, p =.020). Thus, a high return player punishes less harsh than a low return player. This is a small effect and in contrast with hypothesis 3. Therefore, hypothesis 3 is not confirmed.

Table 4, Model 5 without control variables and Model 6 with control variables.

	Model 5		Model 6	
	B	SE	B	SE
Return	-.099**	.040	-.092**	.039
Age	-	-	.016***	.003
Condition	-	-	-.124**	.037
Political orientation	-	-	.012	.009
(Constant)	.391***	.023	.014***	.084
Adjusted R2	.001		.012	
F change	6.262**		14.866***	
N	3840		3840	

****p <.001, ** p <.01, *p <.05*

Conclusion and discussion

This paper researches the influence of contribution norms on punishment behaviour. Punishment is centuries old, and it increases cooperation over time in social dilemmas (Reuben & Riedl, 2013). To find a way to promote cooperation, it is of importance to know what convinces people to punish free-riders. People use punishment for different reasons according to earlier literature. This paper examines the most important of these mechanisms by using an experimental design. Therefore, it is relevant to see which mechanisms are responsible for punishment to ensure cooperation in social dilemmas. The research question was: "In what way do contribution norms influence punishment behaviour in asymmetric social dilemmas?" Following the homo sociologicus in combination with the fairness principles, participants would punish deviation from supported norms to maintain these contribution norms. Two of the fairness principles are used to determine two norms; efficiency and equality. According to the theory, people who support the efficiency norm in a public good game will punish deviations from equal contributions harsher than people who follow the equality norm. In contrast, people who support the equality norm in a public good game punish deviations from equal payoffs harsher than people who follow the efficiency norm. Participants who supported one of the used norms in this paper did act according to the fairness principles and the homo sociologicus values. The strengthening interaction effect indicate that participants acted to ensure a collective fair payoff and truly punished according to their norm, whether equal payoffs or equal contributions were seen as fair. Deviations from norms endanger the collective. Therefore, an increased norm support increases the effect of deviations on punishment.

Additionally, to check if not the homo sociologicus but the homo economicus was at play, an alternative hypothesis was tested. High return players would lose more payoff than low return players in a repeated asymmetric public good game from free-riders, therefore high return players would punish more strictly following the homo economicus / rational choice theory. Results showed, however, that high return participants punished slightly less harsh than low return participants. The homo economicus values wherein participants act fully rational to maximize individual profit is therefore not confirmed. People tend to support norms and the collective outcome more than they prefer maximal individual gain, according to this research.

An experimental design is fairly unique in sociological researches, where much survey research is used. The experimental design of this paper has multiple advantages in

comparison with a survey design, but there are also disadvantages to take into account. The experimental approach ensures the causal impacts of norms on punishment by controlling other influences. Experiments allow for a better measure of behaviour, because people act according to real life situations in a controlled environment (Falk & Heckman, 2009). The intern validation is high in a research with an experimental design, results are valid for the researched group. However, surveys have a higher statistical power than experiments because of larger and better representative data, it gives a survey design a better extern validation than an experimental design. Results from a survey research is better generalisable to the population. An experimental design is preferred above survey research when a specific causal mechanism is researched, this research focuses on the causal mechanism behind punishment. An experimental design was therefore preferred in this research, but the results could lack extern validation and thus generalisation. Both kind of researches complement each other and are important to research behaviour (Falk & Heckman, 2009)

This paper contributes to the sociological literature by researching the effect of norms on punishment in an asymmetric public good game, where most prior literature has researched the symmetric public good game. Earlier research from Reuben & Riedl (2013) suggested that norms motivate punishment behaviour in heterogeneous public good games. The effect of norms was not shown, because Reuben & Riedl (2013) did not measure the norms of participants in their experiment. This paper differs to Reuben & Riedl (2013) by explicitly measuring norms and subsequently showing that norms affect punishment behaviour. Additionally, the different norms which arise in different heterogeneous public good games show that the divers social dilemmas in the real world can provoke different norms. Heterogeneity shows a more realistic reproduction of social dilemmas. The significant effect of both norms shows the importance of norms in social situations. In social dilemmas, norms can provide a stable mechanism which leads to cooperation and thus the best collective outcome. More research to examine the effect of different social norms is necessary to understand and regulate social dilemmas.

An important fairness principle named equity is not named in this paper. According to Konow, Saijo & Akai (2009) equity refers to the accountability principle. This principle states that rewards are fair when it is in proportion with individual contributions. The higher people contribute, the more reward they deserve. People who are less able to contribute will most likely not see the equity principle as fair. Being able to contribute less will result in less earnings, which most people try to avoid. This fairness principle is not tested in this paper, the experimental design did not permit equity because return rates were fixed for high and

low return players. The measurement of norms in this paper is limited to efficiency and equality, participants who would support equity in different circumstances are classified as one of these norms. However, the equity norm is not possible in the experimental design where an asymmetric public good game is used. A different experimental design where return rates rises with an increased contribution would be suitable to measure equity.

Next to maximizing individual gain and supporting a norm, there are other mechanisms behind punishment behaviour. Counter punishment for example is the act where actor A is being punished by actor B because actor A punished actor B before. Nikiforakis (2008) researched counter punishment, the threat of counter-punishment could lead to less willingness to punish. A decrease in punishment endangers the cooperation levels in a public good game. Therefore, counter-punishment is an important factor to take into account in following research. Revenge, but also striving to maximal individual gain could lead to the choice to counter-punish because using fear of punishment prevents future sanctions for free-riding. Free-riders could maximize their individual payoff by using counter-punishment to sanction actors who punished them. The effects of counter-punishment are only of importance when actors know who punished them, this information was not accessible for the participants in the experiment. Inclusion of counter-punishment could give a better view on the effects of punishment on cooperation.

The relation between norms and punishment is analysed without taking the different return rates into consideration. There could be a difference in the researched relation between high return players who punish low return players, low return players who punish high return players and low return players who punish low return players. However, a filter to analyse the differences of the researched relation between return rates was unreliable because of high multicollinearity scores between the predictor variables. For further research, the difference in punishment between high return and low return participants could be researched. The effect of a high return on punishment is examined in hypothesis 3, but this result is based on the rational choice theory and not on the influence on chosen norms.

Punishment is a means to achieve cooperation. However, it is important for further scientific research and policy implications that, next to punishment, rewarding is an important mechanism which could improve cooperation. Choi & Ahn (2013) found that rewarding showed comparable results on achieving cooperation in a public good dilemma, but that there were far more research results on punishing behaviour and its effect. By improving cooperation in society, like in companies, a better and more functional work atmosphere could arise. This will include better work results. Punishing does matter to achieve improved

cooperation, but rewarding could just as likely be used. Further research to the effect of norms on rewarding behaviour in a public good game and its effect on cooperation would be justified.

The results of this paper can be used in real life policy decision to achieve the best outcome in social dilemmas. For example, industries which produce much pollution to make their products are sensitive for buyers who consider climate problems when they choose a product. However, buyers can have different norms about environmental measures to reduce pollution. Industries which produce a lot of pollution could reduce it to the same level as industries with less pollution, or they could take the same amount of action to reduce the environmental damage as industries which produce less pollution. Using comparable measures will make payoff, in this case the pollution level, not the same between the industries. These industries will be supported by consumers who follow the efficiency norm, while consumers who follow the equality norm would support industries with the same low pollution levels. Which norm consumers support is of importance for industries to choose their environmental policy, to prevent people from not buying a product which is the punishment in this example.

As a final conclusion, this paper shows that norms, and not personal gain, determine punishment behaviour in asymmetric public good games. Earlier research did not show a clear conclusion of punishment motives. The results show that norms are a clear motive for punishment. Therefore, our results suggest norms are the main reason how cooperation, through punishment, arises in social dilemmas. Cooperation levels can increase when policy makers know which norms are at play, acting according to these norms prevents punishment and results in a higher collective outcome.

Literature

Abell, P. (1991). Homo sociologicus: Do we need him/her?. *Sociological Theory*, *9*(2), 195-198.

Anderson, E. (2000). Beyond homo economicus: New developments in theories of social norms. *Philosophy & Public Affairs*, *29*(2), 170-200.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, *211*(4489), 1390-1396.

Bardsley, N., & Sausgruber, R. (2005). Conformity and reciprocity in public good provision. *Journal of Economic Psychology*, *26*(5), 664-681.

Bendor, J., & Swistak, P. (2001). The evolution of norms. *American Journal of Sociology*, *106*(6), 1493-1545.

Boyd, R., & Richerson, P. J. (2001). Norms and bounded rationality. *Bounded rationality: The adaptive toolbox*, 281-296.

Carpenter, J. P., & Matthews, P. H. (2004). Why punish? Social reciprocity and the enforcement of prosocial norms. *Journal of evolutionary economics*, *14*(4), 407-429.

Choi, J. K., & Ahn, T. K. (2013). Strategic reward and altruistic punishment support cooperation in a public goods game experiment. *Journal of Economic Psychology*, *35*, 17-30.

Decker, T., Stiehler, A., & Strobel, M. (2003). A comparison of punishment rules in repeated public good games: an experimental study. *Journal of Conflict Resolution*, *47*(6), 751-772.

Falk, A., & Heckman, J. J. (2009). Lab experiments are a major source of knowledge in the social sciences. *science*, *326*(5952), 535-538.

Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, *90*(4), 980-994.

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature, 415*(6868), 137–140. https://doi.org/10.1038/415137a

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental economics*, *10*(2), 171-178.

Fischbacher, U., Schudy, S., & Teyssier, S. (2014). Heterogeneous reactions to heterogeneity in returns from public goods. *Social Choice and Welfare, 43*(1), 195-217.

Gintis, H. (2000). Beyond Homo economicus: evidence from experimental economics. *Ecological economics*, *35*(3), 311-322.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association*, *1*(1), 114-125.

Hasson, R., Löfgren, Å., & Visser, M. (2010). Climate change in a public goods game: investment decision in mitigation versus adaptation. *Ecological Economics*, *70*(2), 331-338.

Hobbes, T. (1651). Leviathan, or The Matter, Forme & Power of a Common-Wealth Ecclesiasticall and Civil.

Konow, J. (2003). Which is the fairest one of all? A positive analysis of justice theories. *Journal of economic literature*, *41*(4), 1188-1239.

Konow, J., Saijo, T., & Akai, K. (2009). *Morals and mores: Experimental evidence on equity and equality*. Working paper, Loyola Marymount University, 2009.

Liefbroer, A. C., & Billari, F. C. (2010). Bringing norms back in: A theoretical and empirical discussion of their importance for understanding demographic behaviour. *Population, space and place*, *16*(4), 287-305.

Masclet, D., & Villeval, M. C. (2008). Punishment, inequality, and welfare: a public good experiment. *Social Choice and Welfare, 31*(3), 475-502.

Muehlbacher, S., & Kirchler, E. (2009). Origin of endowments in public good games: The impact of effort on contributions. *Journal of Neuroscience, Psychology, and Economics*, *2*(1), 59.

Nikiforakis, N. (2008). Punishment and counter-punishment in public good games: Can we really govern ourselves?. *Journal of Public Economics*, *92*(1-2), 91-112.

Okada, A. (2008). The second-order dilemma of public goods and capital accumulation. *Public Choice*, *135*(3-4), 165-182

Osman, M. P. (2000). The Adam and Eve story as exemplar of an early-life variant of the Oedipus complex. *Journal of the American Psychoanalytic Association*, *48*(4), 1295-1325.

Ostrom, E., Walker, J., & Gardner, R. (1992). Covenants with and without a sword: Self-governance is possible. *American political science Review*, *86*(2), 404-417.

Otten, K., Buskens, V., Ellemers, N., & Przepiorka, W. (2020). Agree to disagree: Public good provision despite normative disagreement in heterogeneous groups. Unpublished manuscript.

Rawls, J. (2009). *A theory of justice*. Harvard university press.

Reuben, E., & Riedl, A. (2013). Enforcement of contribution norms in public good games with heterogeneous populations. *Games and Economic Behavior*, *77*(1), 122-137.

Tan, F. (2008). Punishment in a linear public good game with productivity heterogeneity. *De Economist*, *156*(3), 269-293.