ELSEVIER

Journal of Economic Behavior & Organization
Vol. 61 (2006) 181–198

JOURNAL OF
Economic Behavior
& Organization

www.elsevier.com/locate/econbase

Income distribution preferences and regulatory change in social dilemmas

Laurie T. Johnson [a,*], E. Elisabet Rutström [b], J. Gregory George [c]

[a] *Economics Department, University of Denver, 2000 E. Asbury, Denver, CO 80208, United States*
[b] *Department of Economics, College of Business Administration and the Institute for Simulation and Training, University of Central Florida, Orlando, FL 32816, United States*
[c] *Division of Business and Economics, Macon State College, Macon, GA 31206, United States*

Received 12 November 2004; received in revised form 2 February 2005; accepted 8 March 2005
Available online 30 May 2006

Abstract

We present results from an experiment where we elicit preferences over regulatory policies of social dilemmas for small groups. These policy choices differ only in income distribution and are made after a common group experience of an unregulated social dilemma game. We model two policies: a traditional grandfathering allocation of pollution permits and an egalitarian public trust fund. We find a sizeable fraction of our participants favor the public trust, indicating that social preferences are not sufficiently expressed during the interactive social dilemma game.

JEL classification: D01 (microeconomic behavior: underlying principles); D31 (personal income, wealth, and their distributions); D63 (equity, justice, inequality, and other normative criteria and measurement); D64 (altruism); Z13 (social norms and social capital, social networks).

Keywords: Income inequality; Social preferences; Social dilemma; Regulatory change

1. Introduction

Understanding distributional preferences is important when assessing the social welfare consequences of economic policies. Such preferences are receiving some attention in policy analysis, including those of social dilemma situations (IPCC, 2001), although it is not clear whether this

* Corresponding author. Tel.: +1 720 270 5241; fax: +1 303 871 2605.
E-mail addresses: lauriejo@du.edu (L.T. Johnson), erutstrom@bus.ucf.edu (E.E. Rutström), ggeorge@mail.maconstate.edu (J.G. George).

0167-2681/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jebo.2005.03.012

attention merely reflects the self-interest of political actors or if it also reflects a concern for income to others. For example, some fishing communities confronted with new regulatory schemes using transferable permits experience contentious disputes over the distribution of quotas (Kaufmann et al., 1999, chapter 5). Further evidence of the potential importance of distributional concerns can be found by looking at solutions to distributional conflicts that communities devise. In irrigation and fishery commons, for example, appropriation rules stipulating equal usage are common, and sometimes lotteries have been used to assign rights to the most productive fishery spots (Schlager, 1994). In irrigation systems, users are often given equal time slots to extract water (Tang, 1994). When access rights in fishing commons are violated, they have sometimes been met with negative reciprocal acts resulting in property damage and, in some cases, even lives, with serious distributional consequences (Schlager). In all of these cases the externality generated affects only the group of users, implying that expressions of self-interest and distributional preferences are confounded.

In other social dilemmas, such as pollution emissions, the externality has broader societal effects, perhaps making it harder to reach voluntary equitable agreements. Tradable pollution permits have become a popular regulatory approach to pollution emissions, generally leading to both environmental and economic efficiency (OECD, 2004; Isaac and Holt, 1999; Cason and Plott, 1996), but the allocation of pollution permits has also generated controversy, leading to intensive lobbying by interested parties (Tietenberg and Johnstone, 2004). These controversies often revolve around distributional issues. Permit allocations generally have been made to incumbent firms, often according to some historic performance and needs, using so called grandfathering allocation schemes (Harrison, 2004; Ellerman, 2004). An alternative to grandfathering is to auction off permits, thereby transferring the rents from permit holders to the government conducting the auction. One example of using auctions to assign permits is in Alaska, where concerns for the allocation of rents motivated the establishment of the Alaska Permanent Fund,[1] and in particular its Permanent Fund Dividend program. The State of Alaska holds annual permit auctions for drilling rights and invests 25 percent of its auction revenues in a Trust Fund. The state then distributes a share of Fund investment earnings to every qualified Alaska resident each year.[2]

Following a trading approach, transferable permits have been proposed as a solution to excessive carbon emissions. Not surprisingly, the distributional consequences of such permits have generated debate at the national and international levels (Najam et al., 2003; Bovenberg and Goulder, 2001). While the Climate Stewardship Act proposed in 2003 includes some concerns regarding the allocation of permit scarcity rents, other proposals, such as Sky Trust,[3] go further and model a trust fund similar to that of the Alaska Permanent Fund. According to this proposal, carbon emission permits would be sold to companies and the income distributed to US citizens in the form of equal dividends. The US Congressional Budget Office (2000) evaluated such a redistribution mechanism as one of several ways in which the government can distribute revenues from permit sales.

Policy proposals such as Sky Trust are based on the premise that preferences over the distribution of the scarcity rents of the licenses exist and that voters (and interest groups) care about the way in which the policy solution distributes these licenses. Recent experimental work on what

[1] Alaska Permanent Fund Corporation: http://www.apfc.org.
[2] State of Alaska, Department of Revenue, Permanent Fund Dividend Division: http://www.pfd.state.ak.us.
[3] US Sky Trust Inc.: http://www.usskytrust.org.

have been termed social preferences offers many reasons to expect that this may be so. In a number of experiments, subjects have been found to be motivated not only by self-interest, but also by a concern for payoffs to others (Charness and Rabin, 2002). For example, social preferences may be based on unconditional distributional preferences (Bolton and Ockenfels, 2000; Fehr and Schmidt, 1999), or on reciprocal motivations (Hoffman et al., 1994; Rabin, 1993; Charness and Rabin, 2002).

We hypothesize that support for programs such as Sky Trust depend upon more than self-interest, including distributional and reciprocal preferences. Demands for public fund dividend payments across all citizens imply that egalitarian outcomes are important. Reciprocity motivations may also be relevant if a grandfathering distribution is perceived to benefit agents who were the strongest contributors to the pollution problem in the past. Both motivations are consistent with the social preference literature, as well as with findings in the field where voluntary enforcement in common pool resource environments often involve both distributional considerations and reciprocity (Ostrom et al., 1994).

Our focus is on individual preferences over the distributional implications of alternative social institutions or policies, and we use the term social preferences in this context to refer to preferences in that domain. We design an experiment to test for the presence of social preferences during a regulatory change in a social dilemma situation. Since political pressure in favor of trust funds may reflect either self-serving interests or fairness perceptions, we test for the presence of social preferences by removing any self-serving motivation from the policy choice. An important aspect of our experiment is that the initial income positions from which redistribution choices are made are endogenously generated through a common history of interactions rather than exogenously imposed by the experimenter. Since our interest is in assessing social preferences in situations where individuals face a regulatory change and therefore have a common history, it is important to recognize that social preferences in this context may reflect perceptions of the fairness of initial income positions. Hoffman and Spitzer (1985) and Hoffman et al. argue that the procedures used for assigning initial income positions in experiments may affect distributional choices; we expect initial positions to be particularly relevant in situations of regulatory change, such as in our experiment.

Participants in our experiment first play an unregulated multi-round social dilemma game, followed by a task where they express distributional preferences over two regulatory solutions. Both of the solutions implement the socially optimal level of earnings, but differ in their distributional consequences. One distributes earnings in proportion to previous income shares, modeling a grandfathering allocation, and the other distributes earnings equally, like a trust fund approach. The design controls for a number of motivations that may influence regulatory choices in the field. These include preferences for efficiency, self-serving payoff increases, strategic uncertainty, and the distribution of income across self and others. This leaves motivations based on the distribution of income that excludes one's own earnings and motivations to reward and punish based on perceptions of fairness during the unregulated past. Such motivations are difficult to observe without the benefit of experimental controls.

We find a sizeable fraction of our participants in favor of the public trust and a pattern of choices that correlate with individual and group cooperation levels during the unregulated history. This pattern allows us to infer that agents have social preferences that include considerations for income distribution and reciprocity. Our findings also suggest that subjects cannot fully express these preferences during the unregulated game, leading to a residual demand for redistribution and reciprocity as offered by the trust fund solution.

Table 1
Per period activity table for stage one (payoffs in cents)

Activity choice	Private earnings	Social cost	Social welfare[a]	SW per person[a]
a	255	10	1170	195
b	265	10	1230	205
c	275	10	1290	215
d	285	10	1350	225
e	290	12	1308	218
f	295	13	1302	217
g	307	18	1194	199
h	326	25	1056	176
i	337	27	1050	175
j	348	29	1044	174
k	365	38	822	137
l	378	49	504	84
m	387	57	270	45
n	390	59	216	36
o	393	61	162	27
p	394	62	132	22
q	395	64	66	11
r	396	66	0	0

Only the first three columns were shown to subjects in the experiment. The activity choices labeled "a" through "r" are coded as "1" to "18" in Fig. 1.

[a] Social welfare (SW) is defined as private earnings minus social cost for the agent's activity choice, minus the sum of the social costs of the other five group participants. In this table SW is calculated based on every player making the same activity choice. The social optimum is activity choice "d".

2. Experimental design

Our experiment consists of two stages. The first stage is a simple ten period pollution game in groups of six with fixed, anonymous matchings. In every period of the pollution game each player chooses from an identical list of discrete activity levels that are increasing in private earnings and external costs. Table 1 displays the activities and the accompanying costs and earnings. Between each period the decisions and resulting earnings of each of the six players in a group are revealed to everyone in the group, identified only by their subject ID. The central purpose of the first stage is to create a common economic history in which all players can identify who made the high and low polluting choices. In a training phase subjects are clearly told that activity level "d," if chosen by every member of the group, will result in the highest group earnings. This way we ensure that subjects are aware of the social dilemma.[4]

The symmetry between the players is an important feature of the pollution game. This symmetry is defined over revenues and costs such that resulting differences in earnings are generated solely by the action choices of the players rather than by initial conditions. Moreover, this is common knowledge to all players, allowing us to avoid confounds that would be generated by fairness perceptions of exogenously allocated initial positions.

[4] Subjects are aware of the existence of a stage two when they participate in the first stage. They are told that they will remain in the same group throughout the experiment but that the second stage has a task that is not similar to the one in stage one.

The second stage of the experiment consists of the introduction of two policy solutions to the pollution problem, both constraining activities so that the social optimum is imposed. The solutions differ only in the implied distribution of income. Subjects are given a default solution that reflects a standard grandfathering of pollution permits, where an allocation of permits is based on demonstrated needs during past production. In this case the rents of the permits are distributed in proportion to past profits. Subjects are then offered the opportunity to exchange this default for an alternative solution that reflects a public trust fund where the rents are distributed equally across all players. We elicit willingness to pay for the trust fund using the incentive compatible Becker–DeGroot–Marschak (BDM) lottery mechanism (Becker et al., 1964). Subjects submit a bid that is compared to a randomly selected value from a commonly known uniform distribution. Bids are restricted to be in percentages of stage one earnings, with a lowest bid of 1 percent and a highest bid of 100 percent.[5] We use a bingo cage that contains 100 balls numbered 1–100 to select the random value. If the subject's bid is higher than (or equal to) this randomly drawn value, the bid becomes binding. Incentive compatibility arises because subjects pay the randomly drawn value rather than the bid, removing any incentives to bid below their true value.

Table 2 illustrates a typical distribution of choices with numbers taken from two of the experimental groups as an illustration.[6] The first column shows the earnings position a subject attained in stage one, and the second his share of total stage one group income, from highest to lowest. The third column shows corresponding earnings under the default grandfathering scheme, and the last column shows earnings under the alternative trust scheme. Both of the schemes have the same aggregate group income for all groups ($135), therefore eliminating efficiency as a motivation for choice.[7] This income is set equal to 10 periods of optimal play in the stage one game so that incentives in stages one and two are commensurate. The trust scheme gives the same earnings to all group members, whereas the grandfathering gives incomes that are proportional to the incomes earned during stage one.[8]

An essential feature of our design is that an individual's policy choice in stage two does not affect his or her own earnings in this stage, only the earnings of the other five group members. This is similar to Engelmann and Strobel and to the multi-person experiments in Charness and Rabin. We elicit choices from all subjects and select one of the six group members at random at the end of the experiment as the one whose decision will be imposed on the group. The random dictator receives a fixed payment equal to what each group member would get if the trust scheme were selected. This guarantees that the trust scheme creates an equal income distribution in stage

[5] We restrict bids to be in percentages of stage one earnings in order to avoid several problems. First, we do not want to create house money effects by giving subjects an endowment for bidding (see Rutström, 1998) and Clark (2002). We also do not want to allow the bid to affect the distribution of earnings in stage two, therefore ruling out stage two income as a source for paying for the bid. This leaves only the stage one earnings for bidding. Second, we want to avoid one well-known problem with the BDM, namely that the probability of winning gets quite small if we allow a very high upper value for the random distribution. Using percentages as bids gives us a natural range of values for the random draw. Finally, we want to use the same mechanism for all subjects when selecting the random value, and we want this mechanism to be public. This makes the experiment shorter (even with public random draws a session lasted 2 hours), and also guarantees that subjects know that we are treating everyone equally.

[6] All subjects in the group saw the same distribution table. On the same screen we included text reminding them that if they were selected to make the choice, their own payoffs would be $22.50 no matter what they chose.

[7] Observations in Engelmann and Strobel (2004) imply that preferences over efficiency can otherwise confound distributive preferences. Charness and Rabin also report many cases where subjects trade off efficiency and redistribution.

[8] Column 3 is simply column 2 multiplied by $135.00.

Table 2
Typical distribution options in stage two

Income position of group participant[a]	Participant's share of stage one group earnings (%)	Default cost solution payments[b] ($)	Alternative cost solution payments[c] ($)
Group 3 (session 3)			
1	23.2	31.29	22.50
2	22.5	30.32	22.50
3	21.5	29.00	22.50
4	17.0	22.89	22.50
5	13.9	18.74	22.50
6	2.0	2.76	22.50
Total	100	135.00[d]	135.00[d]
Group 4 (session 9)			
1	21.3	28.76	22.50
2	21.2	28.62	22.50
3	15.2	20.52	22.50
4	14.4	19.44	22.50
5	14.3	19.31	22.50
6	13.6	18.36	22.50
Total	100	135.00[d]	135.00[d]

[a] This column was not displayed on the subject screen.
[b] Default solution in experiment models the unequal, grandfathering permit distribution.
[c] Alternative solution in experiment models the egalitarian trust fund.
[d] All groups have the same total income in stage 2 ($135.00). The particular distribution that applies to the default grandfathering solution differs across groups and depends on the distribution of earnings after completion of stage 1.

two and that earnings for the random dictator will be the same regardless of which solution he chooses. With this design, revealed preferences reflect neither concerns over stage two group efficiency nor own income comparisons across the two solutions.

We construct the financial incentives in stage one so that we generate a clear temptation to defect but also motivations to cooperate, while keeping the overall incentives in this stage commensurate with those in stage two. When everyone chooses the social optimum, earning $2.25 per period, the temptation to defect to the Nash equilibrium would be 57 cents per period; when everyone chooses the Nash equilibrium, the increased earnings from complete cooperation would be over $2.00. The language used in the experiment was based on terminology such as "activity level," "cost on others," and "cost solution," with no reference to production, consumption, or pollution.

Before playing the first stage externality game, subjects go through instructions and a test of their understanding. In the second stage, the use of the BDM mechanism and the random dictator selection are illustrated in a training phase where participants are paid in Hershey's Chocolate Kisses. Subjects are also given a standard demographic questionnaire, plus a survey with a set of personality and attitudinal questions.[9] These are only included as a control for individual heterogeneity and are given a peripheral role in the analysis. We include a random money endowment to each participant at the beginning of stage one to test for income effects that may confound the other influences. The experiment is programmed and conducted with the software Ztree (Fischbacher, 1999).

[9] These personality questions are taken from a common psychology instrument referred to as The Big Five Personality Factors. John and Srivastava (1999) show how to construct the instrument.

3. Social preferences

The literature on social preferences and distributive justice offers a number of alternative theories for explaining choices of the kind implied by the egalitarian nature of the trust fund. Some of these theories focus entirely on the explanatory power of purely distributional preferences, such as the difference aversion modeled by Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) or altruism (Andreoni and Miller, 2002). Similarly, preferences derived from theories of distributive justice such as maximin or egalitarian preferences are also purely distributional. Charness and Rabin report the presence of maximin type preferences, where players act to maximize the payoffs to the player with the minimum earnings. Any of these preferences could reflect caring about the distributional pattern that is either dependent or independent of one's own income position.

Other theories focus primarily on preferences that are conditional on actions or intentions of other players such as reciprocal preferences (Hoffman et al., 1994; Rabin, 1993; Dufwenberg and Kirchsteiger, 2004). They depend on the intentions revealed through past choices or beliefs about future intentions and actions. More general social preference models that include both distributional and reciprocal preferences include Falk and Fischbacher (2006), Ahn et al. (2001), Charness and Rabin (2002), Cox (2004), and Cox et al. (2004). The experiments in these latter models were designed to test for the relative importance of distributional and reciprocal motivations. The conclusions are generally that both distributional and reciprocal motivations play a role, although the reported findings differ somewhat in the details.

The main difference between these experiments and the present one is in the generation of initial entitlements. Here, the entitlements are generated endogenously through a common interactive game rather than the more common practice of exogenously imposed distributions. This design is motivated by an interest in eliciting preferences during a regulatory change where agents have a common history. We characterize play during the interactive game in terms of how cooperative players were. Below we discuss how the cooperativeness of the player during stage one depends on the player's social preferences. Here we will argue that the cooperativeness of an individual during stage one can correlate with choices in stage two in two opposing directions, one being due to distributional preferences and the other to reciprocal preferences. These opposite effects make it difficult to sort out to what extent choices are motivated by distributional versus reciprocal concerns, although we show that interaction effects between the cooperativeness of the individual and of the group as a whole can make such identification possible.

The reported findings of reciprocity in the social preference literature lead us to expect that such preferences, in addition to purely distributional ones, are present in our experiment as well. A choice of the trust fund solution implicitly punishes those who otherwise would get a high income and reward those who otherwise would get a low income. We expect that reciprocal preferences for the trust allocation will be increasing in the cooperativeness of the individual since the choice serves as a means to reward others who cooperated or to punish those who did not. This is the first of the two opposing effects.

The second of the two opposite effects may arise due to preferences in favor of reduced inequality held by individuals who have a relatively high initial income position. Even though the individual's second stage own income is unaffected by his choice over the grandfathering or trust solution, Charness and Rabin (2002) demonstrate that people hold dis-interested social preferences where they act to maximize the minimum payoff among players. There are at least two reasons why the expressions of such preferences may be strengthened by the income of the individual. First, there may simply be income effects. Second, since the individual actually has to pay something for the trust fund solution, he may see this payment as a reduction in the

inequality between himself and the lower income earners in the group. In a negative externality game like ours, the final income position in the interactive game (which determines the initial income position in stage two) depends on how cooperative the individual was; a less cooperative individual will have higher relative earnings while a more cooperative individual will have lower relative earnings since cooperation is individually costly. Based on this reasoning, we therefore predict that demand for the trust fund may be decreasing in the cooperativeness of the individual.

Since the cooperativeness of individuals can thus have opposite influences on the demand for the trust solution when arising from reciprocal as opposed to distributional preferences, all we can observe is the net effect of the two types of preferences. Nevertheless, we may be able to identify reciprocal and distributional preferences separately by recognizing that they are expressed not only through the cooperativeness of the individual, but also through the relationship between the cooperativeness of the individual and of the group. For example, expressions of reciprocal preferences may be strengthened by being in a group that did not succeed in cooperating. In such groups cooperative individuals may have a stronger desire to punish the non-cooperators and reward the cooperators than in groups where everyone generally cooperated. In this case we would predict a negative interaction effect between the cooperativeness of the individual and the group. Similarly, the relative income position (in monetary terms) of a non-cooperative person depends also on the cooperativeness of the others. For any given percentage income distribution, a more cooperative group will generate higher monetary earnings for everyone, increasing the monetary distance between high and low earners. Based on this increasing monetary distance, we predict a positive interaction effect between the cooperativeness of the individual and the group as long as the expressions of distributional preferences are indeed correlated with own income. With these interaction effects we would expect to have a better chance at identifying reciprocity preferences in the least cooperative groups and redistribution preferences in the most cooperative groups.

So far we have focused the discussion on the expression of social preferences through the stage two choice. Nevertheless, social preferences also play a role in the interactive play during the first stage of the experimental game. For example, preferences for egalitarian distributions can be expressed during stage one by choosing activity levels that match the choices of others since coordination results in a more equal income distribution. This may be very difficult to do, however, since successful coordination also depends on choices made by others. Expressing dis-interested distributional preferences during stage one may be even more difficult since the choice of activity level only impacts the relationship between one's own and all others' income, leaving the distribution among others the same. Additionally, the activity choice affects the efficiency level in the group, so subjects may have to trade off efficiency and distributional goals during stage one. The choice in stage two, on the other hand, provides subjects with an opportunity to express dis-interested distributional preferences, unconfounded by efficiency concerns.

There are also problems with inferring reciprocal preferences from stage one. Subjects have ample opportunity in stage one both to signal the intent to cooperate in the hope that others will reciprocate and to reciprocate cooperative actions taken by others. Both of these types of reciprocal actions may be motivated by a simple desire to generate higher earnings over time by strengthening group cooperation. Further, this strategy suffers from strategic uncertainty since the benefit in terms of future earnings depends on the actions of others. The expression of these reciprocal preferences may therefore be affected by an individual's risk attitude. Very risk averse individuals are less likely to take cooperative actions even when they hold reciprocity preferences. Other reciprocal preferences may be present that suffer less from strategic uncertainty, such as a desire to punish and reward other members even when there is no expectation of increased future

Table 3a
Characteristics of subject pool

	Mean	Median	Standard deviation	Minimum	Maximum
Propensity	0.44	0	0.50	0	1
Percent WTP[a]	44.4	37.0	34.45	1	100
$WTP[a]	2.11	1.49	1.84	0.07	8.7
Rank (percentage)	16.7	17.20	4.40	1.2	24.2
GroupEff	27.5	28.2	7.88	11.4	38.1
GroupVar	2.8	1.5	4.22	0.3	19.2
Rank–GrEff	4.6	4.5	1.8	0.4	8.8
Rank–GrVar	46.6	22.4	82.7	3.1	465

Stage one and two characterization ($N = 168$).

[a] These statistics are calculated only based on those who actually selected the trust and therefore bid a non-zero amount. All regression analyses, however, include all bids.

earnings. The reciprocity preferences that can be expressed in stage two are of the second kind and should therefore not depend on risk attitudes. In contrast, these preferences may not be fully expressed in the externality game since actions can be misperceived as non-cooperating acts and lead to further defections by others, ultimately unraveling cooperation entirely in the group. To address these confounds our stage two design removes both strategic uncertainty of other players' actions and self-serving motivations. What is left to express is a residual demand to punish or reward other players for their actions taken during stage one.

In summary, expressing a willingness to pay for the trust fund in stage two can signal either disinterested distributional preferences, perhaps strengthened by one's own relatively high income position at the end of stage one, or a residual reciprocal preference to punish or reward others based on a perception of how cooperative they were in stage one.

4. Results

We recruited 96 student subjects at the University of Central Florida and 72 at Macon State College from various undergraduate classes. We therefore collected data from a total of 28 groups and 168 subjects. Each session consisted of either 3 or 4 groups, and subjects stayed in the same group throughout the entire experiment.

Table 3a provides summary statistics for the main variables.[10] Our first observation is that we succeeded in creating variation across the cooperativeness of subjects during stage one. The stage one earnings rank of a person in any given group, Rank, measures how un-cooperative the individual was in stage one and is calculated as the percentage of the group income that the individual received, so it is normalized by group income, or efficiency.[11] Rank varies from 1 to 24 percent in our sample with a mean of 17 percent. Perfect equality would have resulted in uniform ranks of 17 percent.

[10] Complete statistical analyses can be found in the digital archive at ExLab: http://exlab.bus.ucf.edu.

[11] Using an ordinal measure of rank, such as the rank order of the individual in the group in terms of income, would not capture how well the individual compares to participants in other groups. A cardinal measure, such as the one we use, allows an interpersonal comparison across groups. As an alternative to this normalized rank, we also ranked participants according to the absolute value of their average activity level during stage one. Our conclusions are not affected by the use of this alternative rank measure.

Table 3b
Characteristics of subject pool

	Macon State College ($N=72$)	UCF ($N=96$)
Male (%)	42	53
White (%)	71	64
African–American (%)	15	9
Married (%)	42	2
Work (%)	78	51
Hours worked per week (h)	27	21
Hourly pay ($)	15	9
Highest expected education. Bachelor's degree (%)	61	75
GPA	2.9	3.0
Age	28	19

Demographic questionnaire responses ($N=168$).

We use two variables to capture the extent to which groups are cooperative, GroupEff and GroupVar. GroupEff measures the extent to which the group was able to achieve the goal of cooperation, viz. efficiency. It is measured as the percentage of the maximum possible group earnings ($135) that was achieved in stage one. GroupVar measures the dispersion, or inequality, of stage one earnings within a group (in dollars), showing how well group members coordinated in their action choices. Groups with only a few cooperating individuals are not considered to be equally cooperative as groups where most or all members cooperate, even if they achieve the same efficiency level. Coordination is an important property of cooperation.

Controlling for GroupVar, we take GroupEff as a good measure of cooperation history. Nevertheless, the same is not true of GroupVar. The coefficient on GroupVar reflects not only group history, but also the extent to which preferences are unconditionally egalitarian since the unequal distribution of the default choice in stage two is identical to the final income distribution in stage one. Thus, GroupVar characterizes not only the degree of coordination in stage one, but also the extent to which the grandfathering allocation results in an unequal income distribution. GroupVar varies between 30 cents and $19.00, with a mean of $2.80. GroupEff varies between 11 and 38 percent with a mean of 28 percent. Efficiency is therefore quite low, resulting in stage one earnings that average only $6.20 per participant. The lowest income was 50 cents and the highest one was $12.00, which can be compared to the maximum possible earnings of $22.50 in the social optimum.

Fig. 1 displays the stage one activity choices for all six-group members by period, separately for each of the 28 groups. The common pattern of increasing defection over time can be seen in the fitted lines, but there is also a fair amount of variation across groups, as evidenced by both the scatter plots and the lines. We conclude that our design successfully created a diverse history of cooperation across the different individuals and the different groups and that all groups experienced a serious social dilemma. The demographics summary in Table 3b shows that we have a broadly representative sample.

In stage two, we observe a propensity to choose the trust scheme over the grandfathering scheme, with 44 percent of our subjects expressing a WTP for the trust, and an average WTP amongst these subjects of $2 (or, coincidentally, 44 percent of stage one earnings). In the experiment the subjects bid in percentages of stage one earnings, but we remain agnostic as to whether it is the percentage bid or the implied dollar amounts that they are focusing on. We report findings using both measures. Fig. 2 shows a distribution of all the bids evaluated as dollar amounts,

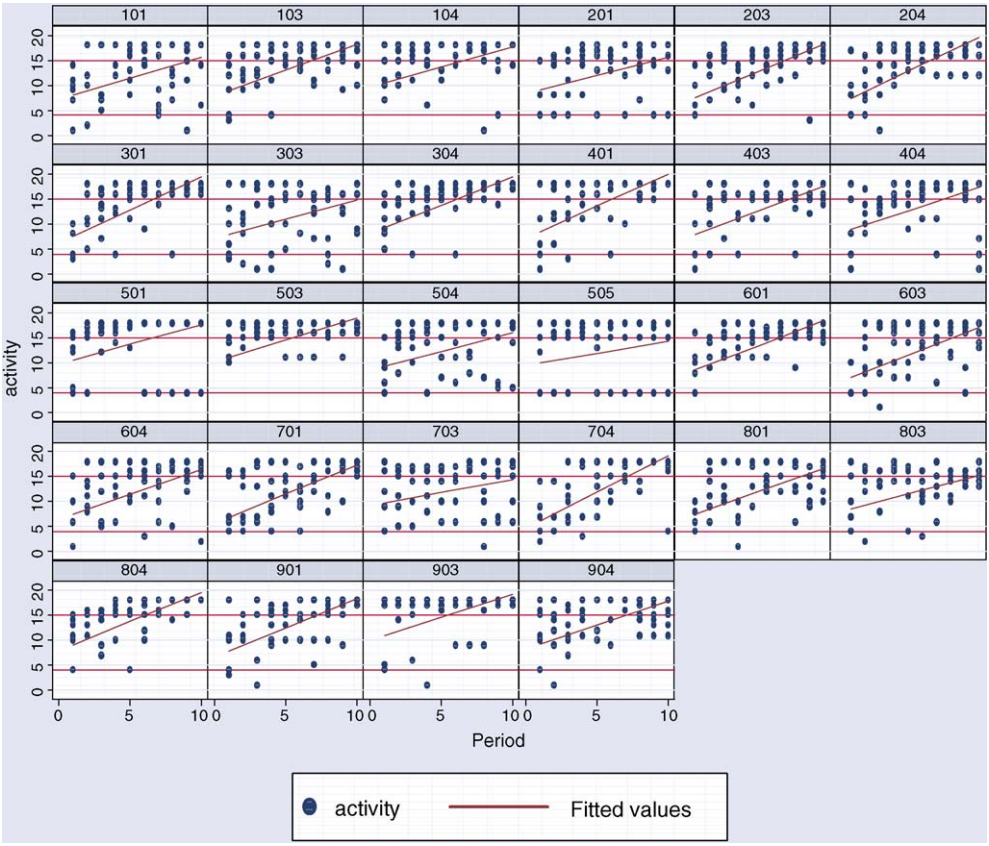

Fig. 1. Activity levels by period separately for each group. We show the Nash equilibrium as a horizontal line at activity level 15 and the social optimum at four.

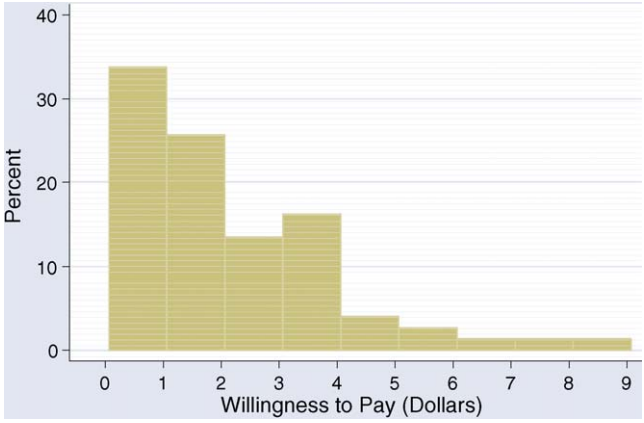

Fig. 2. Distribution of dollar WTP. This graph does not include the $0 bids.

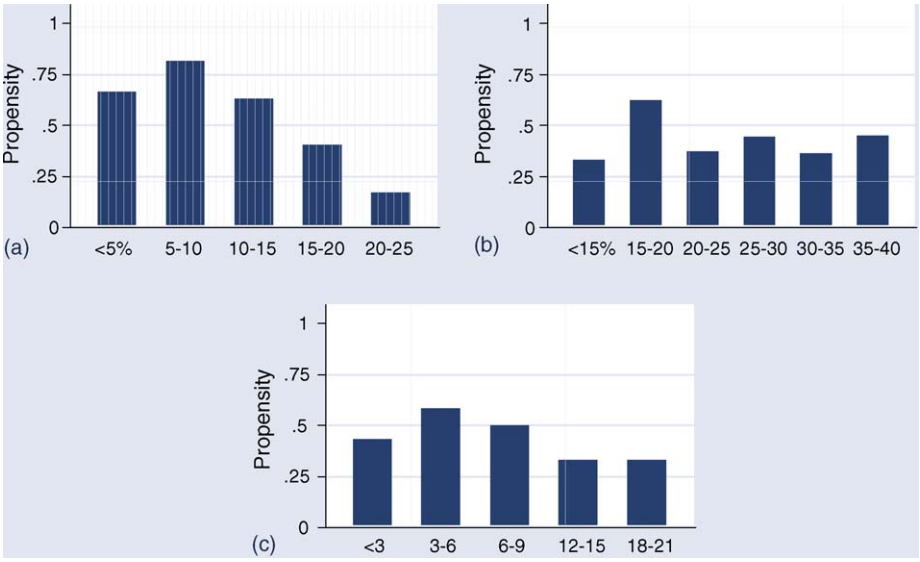

Fig. 3. Propensity to select the trust. (a) By individual's Rank; (b) by Group Efficiency; (c) by Group Income Variance.

not including those who bid \$0. Most bids are \$4 or less, which compares to average stage one earnings of \$6.20. We conclude that agents have social preferences that include considerations for income distribution and reciprocity other than for self-serving reasons.

Figs. 3 and 4 relate the expressed preferences to the cooperativeness of the individual and the group, based on our raw observations. We see some preference for the trust fund amongst all individuals, although it gets weaker amongst the least cooperative ones, as measured by their

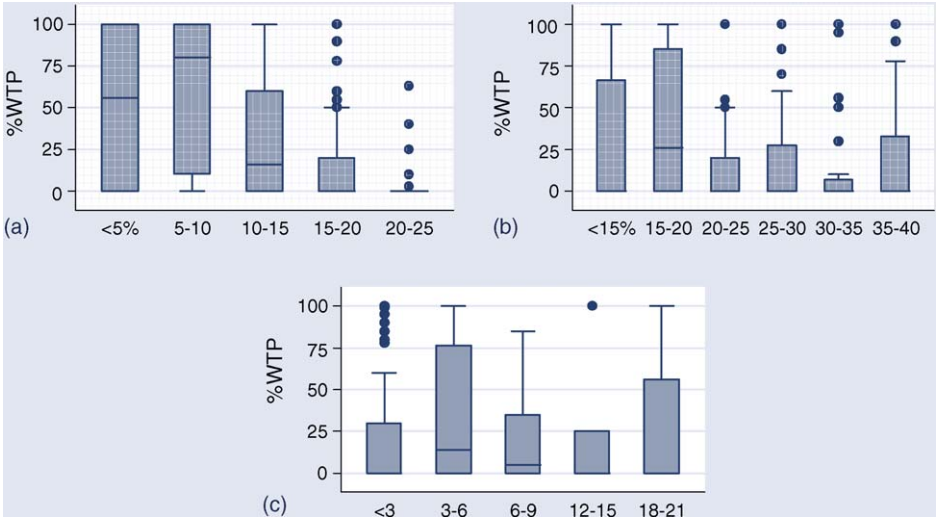

Fig. 4. Distribution of percent WTP for subjects selecting the trust. This graph does not include the \$0 bids. Intervals are constructed based on actual observations (e.g. there were no observations between 9 and 12 for Group Income Variance). (a) By individual's Rank; (b) by Group Efficiency; (c) by Group Income Variance.

high Rank. Similarly, groups characterized by different degrees of cooperation and by different degrees of inequality in the final distribution all include individuals who have a preference for the trust fund. From Fig. 4 it is also clear that there is a great deal of variability across participants in the intensity of their expressed preferences, as measured by their WTP. We therefore turn to our conditional analysis where we control for the hypothesized interaction effects between the cooperativeness of the individual and the group, as well as a number of demographic and attitudinal characteristics of the person.

We report regression models for propensity to select the trust fund, as well as for each of percent WTP and $WTP. These models allow for the censoring that occurs due to subjects not being allowed to enter negative bids or bids exceeding 100 percent of their stage one income. We use a Probit model with Huber–White robust standard errors for propensity. We use a censored regression model with individual-specific upper censoring bounds corresponding to each subject's stage one earnings for $WTP with a lower bound of $0 and a Tobit model censored at 0 and 100 percent for percent WTP. Since many of the explanatory variables were generated endogenously during stage one, we also test for the presence of endogeneity bias in our coefficient estimates, but find none significant.[12]

Table 4 summarizes our regression findings. Since the coefficient on Endow (the random endowment) is insignificant, there is no strong evidence in favor of income effects confounding the tests of our hypotheses.

Consider first the coefficients on Rank and GroupEff. Across all specifications, both variables are negative and significant. The interaction effect, Rank–GrEff, however, is positive and significant. We therefore conclude that the interest in the trust fund decreases as individuals become less cooperative, particularly when they are in groups that were relatively less cooperative, as measured by GroupEff. As the efficiency of the group increases, the cooperativeness of the individual for the expression of the preferences matter less and less. We therefore find support for our prediction that the degree of cooperativeness of the individual should increase the reciprocal preferences, at least as long as the groups are not too cooperative. As the cooperativeness of the group increases, measured by GroupEff, this effect weakens.[13] Recall, however, that GroupEff only measures one aspect of group cooperation: the efficiency of the outcome. The other aspect, coordination of actions (measured by the variance in the group income distribution, GroupVar), is significant in the percent WTP regression. We discuss GroupVar further below.

The results implied by the interaction term Rank–GrEff are illustrated in Fig. 5(a)–(c), which show predicted choices as a function of Rank. For these predictions we set GroupEff to either its maximum or minimum value, hence capturing the most versus the least efficient groups in our experiment. The label LE refers to low efficiency and the label HE to high efficiency groups, with a prefix C for cooperative and N for non-cooperative individuals. In the predictions we set all other explanatory variables equal to their mean values. Fig. 5(a) contrasts how predicted propen-

[12] Detailed regression results are available in ExLab. We performed a version of the Hausmann specification tests for endogeneity. This test is described in Gujarati (1995, pp. 670–671). These tests indicate that we do not have a strong endogeneity problem. Nevertheless, we also tested our hypotheses using instrumental variable and simultaneous equation techniques. These regressions confirm the findings we report here.

[13] We also ran regressions that include GroupEff and Rank calculated separately for the first and last five periods of stage one. We find the same qualitative effects for these variables as we do for the aggregate ones shown in Table 4. This verifies that our conclusions are not confounded by how the cooperativeness of individuals and groups change through the game.

Table 4
Regression results

Variable	Propensity	Percent WTP	$WTP
Rank	−0.18 (0.001)***	−22.68 (0.000)***	−0.92 (0.001)***
GroupEff	−0.10 (0.008)***	−12.06 (0.000)***	−0.45 (0.016)**
GroupVar	+0.09 (0.104)	+6.77 (0.051)*	+0.26 (0.232)
Rank–GrEff	+0.5 (0.011)**	+63.33 (0.000)***	+2.55 (0.019)
Rank–GrVar	−0.005 (0.107)	−0.32 (0.136)	−0.01 (0.361)
Endow	−0.02 (0.101)	−0.66 (0.647)	−0.03 (0.746)
Session 2	+0.10 (0.570)	−7.85 (0.620)	−0.33 (0.731)
Session 3	−0.24 (0.076)*	−40.47 (0.013)	−2.51 (0.010)
Session 4	+0.31 (0.091)*	+11.62 (0.489)	+0.56 (0.577)
Session 5	−0.28 (0.030)**	−48.56 (0.035)	−2.77 (0.045)
Inst	−0.08 (0.496)	+16.63 (0.190)	+0.86 (0.258)
Age	+0.002 (0.853)	+0.23 (0.831)	+0.03 (0.663)
Male	−0.02 (0.812)	+10.50 (0.308)	+0.79 (0.203)
White	+0.13 (0.283)	+21.09 (0.099)	+0.63 (0.406)
African–American	−0.17 (0.309)	−16.66 (0.394)	−1.63 (0.177)
Hours work per week	−0.002 (0.632)	+0.06 (0.856)	+0.006 (0.777)
Hourly pay	+0.004 (0.106)	+0.54 (0.050)	+0.04 (0.031)
Highest education	−0.171 (0.005)***	−9.14 (0.161)	−0.59 (0.136)
Married	−0.448 (0.015)**	−22.21 (0.277)	−1.42 (0.249)
Gpa	−0.03 (0.365)	−4.68 (0.144)	−0.27 (0.162)
Extrovert	+0.10 (0.301)	−2.06 (0.805)	−0.04 (0.942)
Agreeable	−0.03 (0.808)	−1.78 (0.861)	+0.003 (0.996)
Conscientious	+0.04 (0.559)	+0.02 (0.998)	−0.09 (0.850)
Openness	−0.02 (0.827)	+0.86 (0.928)	−0.09 (0.871)
Moral absolutist	−0.07 (0.007)***	−4.80 (0.095)	−0.27 (0.113)
Mother morality	−0.15 (0.002)**	−13.92 (0.003)	−0.82 (0.003)
Father morality	+0.04 (0.138)	+6.07 (0.055)	+0.37 (0.055)
Science	+0.07 (0.062)*	+3.29 (0.333)	+0.19 (0.343)
Moral theist	+0.05 (0.071)*	+3.07 (0.239)	+0.18 (0.261)
Non-religious	+0.02 (0.517)	+0.47 (0.907)	+0.06 (0.795)

Hausmann specification tests do not detect significant endogeneity due to the stage one choice variables used as explanatory variables in these stage two regressions.
*** p-values in parentheses at 1% level.
** p-values in parentheses at 5% level.
* p-values in parentheses at 10% level.

sity differs between groups with the highest and lowest cooperation as individual cooperation decreases (Rank increases). The solid line predicts the propensity to choose the trust for groups with efficiency levels as low as that of our least efficient group in the experiment, which captured only 11 percent of the socially optimum group earnings. The dashed line predicts the propensity to choose the trust for groups with efficiency levels as high as that of our most efficient group, which captured 38 percent of the socially optimum earnings. As one moves from the left to right along the horizontal axis Rank increases, indicating increasingly non-cooperative behavior in stage one.

In the least cooperative groups the most cooperative individuals (CLE) have a higher predicted propensity to choose the trust than the least cooperative individuals (NLE) (Fig. 5(a)). Conversely, in the most cooperative groups it is the least cooperative individuals (NHE) who are predicted to have the higher propensity, although the difference is small. The other two graphs in

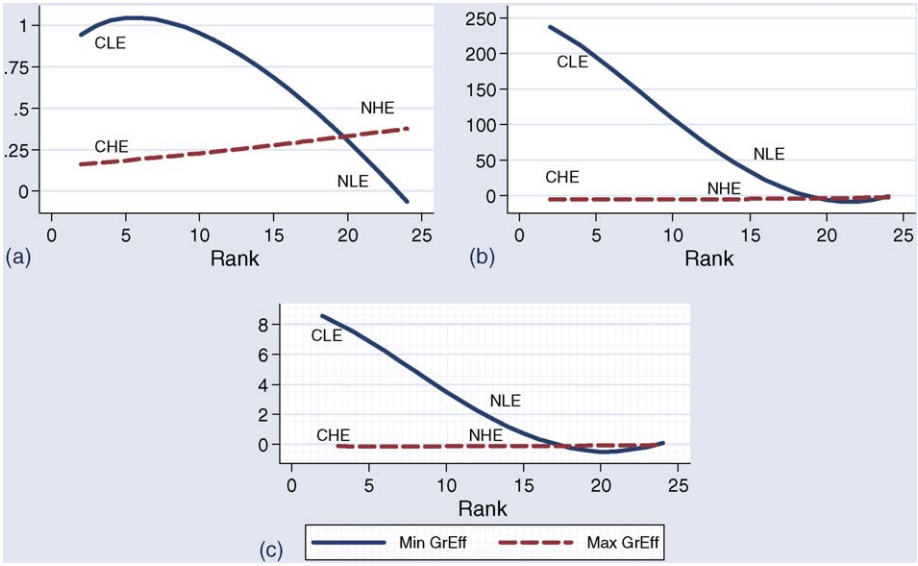

Fig. 5. Earnings Rank, Group Efficiency and predicted preferences for the trust. (a) Propensity; (b) expected percent WTP; (c) dollar expected WTP. CLE, cooperative individual in low efficiency group; CHE, cooperative individual in high efficiency group; NLE, non-cooperative individual in low efficiency group; NHE, non-cooperative individual in high efficiency group. The expected bids are calculated as the bid times the propensity to select the trust fund.

Fig. 5 show the same qualitative pattern for percent WTP and $WTP, weighted by the predicted propensity.

Regression results for our last two variables, GroupVar and Rank–GrVar, are mixed. GroupVar is positive and significant in specification 2 (percent WTP) and weakly significant and positive in specification 1 (propensity). It thus appears that increased inequality, as measured by the variance, has a tendency to strengthen preferences for the trust, although perhaps not very strongly. This may be due to the fact that we do not see much variation in GroupVar across most groups.[14] Finally, we do not find a significant interaction effect between a person's cooperativeness and the income inequality of the group (Rank–GrVar). Recall that the coefficient on GroupVar may reflect either purely distributional preferences or reciprocity preferences since it also measures the degree of coordination in the group.

In summary, we find that when facing a regulatory change, such as the change from stage one to two in our experiment, a sizeable portion of our participants express preferences in favor of the policy solution with the egalitarian income distribution. The preferences expressed intensify with the degree of cooperativeness of the individual during the unregulated part of the interactions, particularly in groups that were not successful at cooperating. Nevertheless, we observe some preferences in favor of the trust fund even amongst individuals who were not very cooperative in the past, and these intensify as the groups are more successful at cooperating. We therefore find some evidence of both distributional and reciprocal preferences in the expression of preferences over social policies.

[14] The distribution of GroupVar is quite skewed. Twenty-six of the 28 groups have a GroupVar of less than seven, with 21 of these 26 being less than two. The remaining two have values of 14 and 19, respectively.

5. Conclusions

We investigate the social preferences of individuals who face a regulatory change after a common interactive history of negative externalities. We model an egalitarian solution to mirror a trust fund approach to distributing emission permits and an unequal distribution to mirror a grandfathering approach. Our design reflects recent policy debates over regulatory change in social dilemma situations, where distributional issues are often a central issue in lobbying around such policies. In particular, we focus on recent policy proposals that explicitly include provisions for redistributing scarcity rents that result from the initial allocation of tradable pollution licenses. To reflect this context, we design a controlled laboratory experiment in which participants first play an interactive negative externality game and then make income distribution decisions in a separate non-interactive policy making task.

We find that a sizeable portion of our participants express preferences in favor of the egalitarian policy solution. We conclude that agents have social preferences that include considerations for income distribution and reciprocity other than for self-serving reasons. The preferences expressed intensify with the cooperativeness of the individual as expressed during the unregulated stage. This is particularly the case in groups that were not successful at cooperating and is consistent with reciprocity concerns. Increased inequality of the grandfathering solution also results in an increase (although perhaps less significantly) in demand for the trust.[15] Finally, we observe some preferences in favor of the trust fund even amongst individuals who were quite un-cooperative in the past, and these intensify in the groups that were more successful at cooperating. This is consistent with our hypothesis on distributional preferences.

Our findings broadly support inferences about social preferences drawn elsewhere in the experimental literature. Our experiment differs significantly from others that elicit distributional preferences because the initial income positions from which redistribution choices are made are endogenously determined through a common history of interactions rather than exogenously imposed by the experimenter. Further, we elicit social preferences in the absence of influences from self-serving interests, efficiency concerns, and strategic uncertainty. These influences cannot be controlled in purely interactive settings or in less controlled settings in the field. Our design is motivated by our interest in evaluating the possibility that regulatory changes that include distributional concerns, other than those based on competing self-serving interests, have some support in existing social preference structures. We find that this is the case in our laboratory environment.

These findings stand in marked contrast to those reported in Rutström and Williams (2000) who also ask whether past actions have an effect on distributional choices. Nevertheless, they employed a context where subjects did not have a common and interdependent history. Their subjects went through a stage one task individually and then made distributional choices in a stage two. Stage two efficiency, as captured by the aggregate group earnings, depended on the combination of independent actions taken in stage one by all group members. Despite the link to history that was provided through the efficiency effect, they found choices to be consistent with a pure self-regard. These contrasting results indicate that distributional and reciprocal preferences may depend on perceptions of a common past where the nature of players as cooperative or not is revealed in a transparent way.

Our findings provide a step toward a better understanding of the preferences underlying the economic and political process of policy design and open the way for future research to address

[15] This could be due to inequity aversion, or reciprocity, since more unequal groups are also less coordinated.

other issues. Such issues include the trade off between self-interest and distributional preferences for the political actors, the differences in the characteristics of interested parties such as economic size, and issues of political influence. The experimental design introduced here, where the initial income positions are endogenously generated, provides a promising framework for continued research on the preferences underlying the political process of regulatory change.

Acknowledgements

Johnson thanks the University of Denver, Faculty Research Fund and the Political Economy Research Institute at the University of Massachusetts, Development, Peacebuilding, and Environment Program of the University of Massachusetts. Rutström thanks the U.S. National Science Foundation for research support under grants NSF/IIS 9817518, NSF/MRI 9871019 and NSF/POWRE 9973669. All supporting documentation can be found at ExLab: http://exlab.bus.ucf.edu.

References

Ahn, T.K., Ostrom, E., Schmidt, D., Shupp, R., Walker, J., 2001. Cooperation in PD games: fear, greed, and history of play. Public Choice 106, 137–155.

Andreoni, J., Miller, J., 2002. Giving according to GARP: an experimental test of the consistency of preferences for altruism. Econometrica 70, 737–753.

Becker, G.M., DeGroot, M.H., Marschak, J., 1964. Measuring utility by a single-response sequential method. Behavioral Science 9, 226–232.

Bolton, G., Ockenfels, A., 2000. ERC: a theory of equity, reciprocity and competition. American Economic Review 90, 166–193.

Bovenberg, A., Goulder, L., 2001. Neutralizing the adverse industry impacts of CO_2 abatement policies: what does it cost? In: Carraro, C., Metcalf, G.E. (Eds.), Behavioral and Distributional Effects of Environmental Policy. University of Chicago Press, Chicago, pp. 45–90.

Cason, T.N., Plott, C.R., 1996. EPA's new emissions trading mechanism: a laboratory evaluation. Journal of Environmental Economics and Management 30, 133–160.

Charness, G., Rabin, M., 2002. Understanding social preferences with simple tests. Quarterly Journal of Economics 117, 817–869.

Clark, J., 2002. House money effects in public goods experiments. Experimental Economics 5, 223–231.

Cox, J., 2004. How to identify trust and reciprocity. Games and Economic Behavior 46, 260–281.

Cox, J., Friedman, D., Gjerstad, S., 2004. A tractable model of reciprocity and fairness. Unpublished Manuscript, Department of Economics, University of Arizona.

Dufwenberg, M., Kirchsteiger, G., 2004. A theory of sequential reciprocity. Games and Economic Behavior 47, 268–298.

Ellerman, D.A., 2004. The US SO2 Cap-and-Trade Programme. In: Tradable Permits: Policy Evaluation, Design and Reform. Industry, Services, and Trade 9. Organization for Economic Cooperation and Development (OECD), Paris, pp. 77–106.

Engelmann, D., Strobel, M., 2004. Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. American Economic Review 94, 857–869.

Falk, A., Fischbacher, U., 2006. A theory of reciprocity. Games and Economic Behavior 54, 293–315.

Fehr, E., Schmidt, K.M., 1999. A theory of fairness, competition, and cooperation. Quarterly Journal of Economics 114, 817–868.

Fischbacher, U., 1999. Z-Tree: Toolbox for Ready Made Economic Experiments. Working Paper 21, University of Zurich.

Gujarati, D.N., 1995. Basic Econometrics, 3rd ed. McGraw-Hill, New York.

Harrison Jr., D., 2004. Ex-post evaluation of the reclaim emissions trading programmes for the Los Angeles Air Basin. In: Tradable Permits: Policy Evaluation, Design and Reform. Industry, Services, and Trade 9. OECD, Paris, pp. 49–76.

Hoffman, E., Spitzer, M.L., 1985. Entitlements, rights, and fairness: an experimental examination of subjects' concepts of distributive justice. Journal of Legal Studies 14, 259–297.

Hoffman, E., McCabe, K., Shachat, K., Smith, V., 1994. Preferences, property rights, and anonymity in bargaining games. Games and Economic Behavior 7, 346–380.

Intergovernmental Panel on Climate Change (IPCC), 2001. Climate Change 2001: Mitigation. Cambridge University Press, New York.

Isaac, R.M., Holt, C. (Eds.), 1999. Emission Permit Experiments. Research in Experimental Economics, vol. 7. JAI Press, Stamford, CT.

John, O., Srivastava, S., 1999. The big five-trait taxonomy: history, measurement, and theoretical perspectives. In: Pervin, L.A., John, O.P. (Eds.), Handbook of Personality: Theory and Research, 2nd ed. Guilford Press, New York, pp. 102–238.

Kaufmann, B., Geen, G., Sen, S., 1999. Fish Futures: Individual Transferable Quotas in Fisheries. Fisheries Economics, Research and Management Pty. Ltd., Kiama, NSW, Australia.

Najam, A., Huq, S., Sokona, Y., 2003. Climate negotiations beyond Kyoto: developing countries concerns and interests. Climate Policy 3, 221–231.

Organization for Economic Cooperation and Development (OECD), 2004. Tradable Permits: Policy Evaluation. Design and Reform. OECD, Paris.

Ostrom, E., Gardner, R., Walker, J. (Eds.), 1994. Rules, Games, and Common-Pool Resources. University of Michigan Press, Ann Arbor, MI.

Rabin, M., 1993. Incorporating fairness into game theory and economics. American Economic Review 83, 1281–1302.

Rutström, E., 1998. Home-grown values and incentive compatible auction design. International Journal of Game Theory 27, 427–441.

Rutström, E., Williams, M., 2000. Entitlements and fairness: an experimental study of distributive preferences. Journal of Economic Behavior and Organization 43, 75–89.

Schlager, E., 1994. Fishers' institutional responses to common-pool resource dilemmas. In: Ostrom, E., Gardner, R., Walker, J. (Eds.), Rules, Games, and Common-Pool Resources. University of Michigan Press, Ann Arbor, MI, pp. 247–265.

Tang, S.Y., 1994. Institutions and performance in irrigation systems. In: Ostrom, E., Gardner, R., Walker, J. (Eds.), Rules, Games, and Common-Pool Resources. University of Michigan Press, Ann Arbor, MI, pp. 225–245.

Tietenberg, T., Johnstone, N., 2004. Ex post evaluation of tradable permits: methodological issues and literature review. In: Tradable Permits: Policy Evaluation, Design and Reform. Industry, Services, and Trade 9. OECD, Paris, pp. 10–48.

US Congressional Budget Office, 2000. Who Gains and Who Pays Under Carbon-Allowance Trading: The Distributional Effects of Alternative Policy Designs. Unpublished Manuscript, http://www.cbo.gov/ftpdocs/21xx/doc2104/carbon.pdf.