CBESS Discussion Paper 09-05

Instrument Choice and Motivation
Evidence from a Climate Change Experiment

by Timo Goeschl* and Grischa Perino**

* University of Heidelberg
** CBESS, University of East Anglia

9 June 2009

Abstract
Are prices or quantities the best regulatory instrument to align private actions with public interests in the presence of externalities? We add another dimension to this ongoing debate by experimentally analyzing the interaction between instrument choice and intrinsic motivation of regulated agents. The response of subjects facing a trade-o_ between real CO2 emissions and private monetary payoffs to both a price and a quantity instrument are tested. We find evidence that taxes crowd out intrinsic motivation while emission standards are neutral. Crowding is short term persistent and not well explained by established cognitive theories of motivational crowding.

JEL classification codes
H23, H41, Q58, C91

Keywords
Instrument Choice, Motivation Crowding, Externalities

CBESS
University of East Anglia
Norwich NR4 7TJ
United Kingdom
www.uea.ac.uk/ssf/cbess

1 Introduction

The choice of the appropriate regulatory instruments to align private actions with public interests in the presence of externalities is one of the enduring themes of debate among public economists. In this debate, it has become customary to classify candidate instruments into one of two groups: One consists of so-called 'quantity' or 'command-and-control' instruments that specify actions in terms of outcomes. Empirically, quantity instruments dominate regulatory policy in the form of technical prescriptions and limits on activity levels. The other group comprises so-called 'price' or 'economic' instruments that specify incentives. Its most frequent empirical incarnation are tax instruments.

The debate on the relative merits of quantity and price instruments has a long and distinguished history. On basic efficiency grounds, the theoretical case for price instruments has been made forcefully over the years (Baumol and Oates 1988). At the same time, the theoretical models used for deriving the efficiency of instruments have yielded more equivocal results on the relative merits of quantity and price instruments when extended to include abatement cost uncertainty (Zhao 2003), impacts on technological change (Requate and Unold 2003, Krysiak 2008), and enforcement considerations (Montero 2002). Political economy considerations constitute another source of subtlety when considering instrument choice (Dijkstra 1999, Keohane et al. 1998, Hahn 1990). Finally, comprehensive assessments of the practical dimensions of regulatory implementation also deliver a less clear-cut verdict on prices versus quantities in regulatory policy (Harrington et al. 2004, Goulder and Parry 2008). The notion of a single 'best' regulatory instrument to align private actions with public interests would therefore appear to remain elusive.[1]

The present paper revisits the question of instrument choice by focusing on the fundamental drivers of how individuals respond to different forms of regulation. Specifically, this paper is interested in the possible interactions between instrument choice and the pro-public motivation of the regulated. The pro-public motivation in question is the willingness of economic agents to align their private actions with public interests voluntarily. This focus traces

[1] See Kerr and Newell (2003), Austin and Dinan (2005), Fischer and Newell (2008). Goulder and Parry (2008) provide an up-to-date literature review on instrument choice in environmental economics.

its origins to an empirical literature gathering evidence on the way in which explicit incentives interact with behavior that generates either positive or negative externalities, starting with Titmuss (1970). More specifically, there is a bold conjecture by Frey (1997) that in the presence of preexisting private contributions above the non-cooperative level, introducing price and quantity instruments to increase contributions will have opposite effects on motivation: Explicit price incentives will lead to a crowding-out of motivation while quantity instruments behave neutrally or even crowd in motivation. Further, these crowding effects are conjectured to persist after the explicit incentives are removed. Frey (1992) applies this general conjecture to the specific case of environmental externalities that is also studied here.

What makes the question of an interaction between instrument choice and motivation newly salient are two developments: One is a recent literature of new experimental findings and theoretical foundations on which this present paper builds. This recent literature examines how explicit incentives interact with the intrinsic motives and preferences that determine people's conduct. Price incentives have been found to give rise to problematic effects (Bénabou and Tirole 2006, Fehr and Falk 2002, Frey and Oberholzer-Gee 1997). But also other types of explicit incentives can suffer from similar deleterious interactions with intrinsic motivation: Fines for missing performance targets decreased the average agent's effort in experimental contract settings (Fehr et al. 2007). Likewise, introducing fines for late pick-up of children from daycare centers led to an increase in late pick-up in a field experiment (Gneezy and Rustichini 2000a). Small changes to the contract environment appear to matter: Merely restricting the agent's choice set in a principal-agent setting can trigger more selfish behavior (Falk and Kosfeld 2006). Ellingsen and Johannesson (2008) present additional experimental evidence that control systems and pecuniary incentives both lead to what has become known as 'motivation crowding' (MC). Alongside the evidence in favor of 'motivation crowding-out', there is also evidence for *positive* interactions between explicit incentives and intrinsic motivation (Frey and Jegen 2001, Charness and Gneezy 2009). Against this background of new results, the question of instrument choice and motivation can therefore be asked with new methods and concepts at hand.

The second salience of the question of motivation and regulatory instruments arises from the fact that policy - among the EU and its members in particular - has set itself ambitious

targets in areas such as climate change policy that will require additional environmental regulations to be imposed not only on businesses, but also on individual consumers (EU 2008, CCC 2008). For these economic agents intrinsic motivation is clearly relevant. Moreover, individuals' attitudes and preferences do also influence decisions of firms too a larger degree than often presumed by economists (NCF 2006, Brekke and Nyborg 2008).

In this paper we develop an experimental setting in which the presence and magnitude of MC effects associated with the two archetypical instruments of regulating environmental externalities, taxes and standards, can be compared directly. In an experiment where subjects face a decision between private money payoffs and contributions to mitigate climate change we are able to study how the introduction of a tax and an emission standard (enforced by a fine) affect subjects willingness to contribute to the environmental public good. To our knowledge, this is the first paper to provide a direct comparison of the behavioral impact of these common regulatory instruments in a controlled setting.

In the chosen setting, subjects take a sequence of three anonymous and independent decisions that involve a simple trade-off between providing a global public good in the form of greenhouse gas emissions reductions and collecting personal payoffs in the form of cash pay-outs. Subjects are randomly selected into one of four groups. All groups go through an initial round of choice without any explicit regulatory interventions in place. Two of the four groups make their second-round decision with explicit price incentives or quantity restrictions imposed, while the other two groups act as controls. In the third round, all four groups choose again without explicit interventions. We compare the choices between rounds and test whether and how experiencing an episode of explicit incentives in the second round changes subjects' behavior. This design choice makes it possible to focus not only on the instantaneous effects of regulation, but also on the intertemporal dimension of the interaction between explicit incentives and motivation that is the subject of considerable debate (Frey 1992, Bénabou and Tirole 2003, 2006).

The experimental evidence that this paper reports supports the notion of crowding out of motivation by price instruments but not by quantity restrictions. We also find that the crowding out of the price instrument spills over from the regulated round into the round when incentives have been removed. Together, these findings can be taken as providing support

for the 'Frey (1992) Conjecture'. The findings are derived using methods such as differences-in-differences that provide an unambiguous link between instruments and crowding effects. This approach demonstrates - as a by-product - the presence of deviations from conventional theory other than motivation crowding: Subjects do not fully adjust to changes in relative prices regardless of whether they are triggered by a regulatory intervention or not. The presence of such deviations from conventional theory underlines the need to disentangle motivation crowding from other, contaminating effects. We provide an example of the danger of erroneously attributing observed variations to motivation crowding.

The final contribution of the paper is to take steps to explain the crowding effects observed. A critical review of the existing cognitive theories (Bénabou and Tirole 2003, 2006, Sliwka 2007) that explain MC reveals that due to the experimental design, their mechanisms have limited traction in explaining the present evidence. As an additional check, we test for the longer-term persistence predicted by cognitive theories by performing a follow-up session one week after the initial session. We find that the effects observed in the initial session are no longer detectable after this comparatively short period. Other mechanisms may be necessary to explain the experimental evidence, among them affective mechanisms studied in the psychological literature on social motivation (Forgas and Laham 2005).

The following section discusses the experimental design. Section 3 sets out the behavioral predictions before turning to the key results in section 4. Section 5 evaluates the results in light of theoretical considerations and related experimental findings. Section 6 concludes.

2 The Experiment

2.1 The General Set-Up

The basic experimental design is simple. It consists of a sequence of three anonymous and independent choice experiments (rounds). Anonymity of choice here means anonymity with respect to other subjects and the experimenter. Among subjects, there was no interaction either in terms of communication or information at any stage of the experiment. This total anonymity was announced in advance. With respect to the experimenter, subjects

were randomly assigned IDs when entering the room and were seated according to the first letter of their ID. Independence of choice here means that there was no relationship among subjects in actions or payoffs or between a subject's characteristics, her choice in previous rounds and her set tasks in subsequent rounds. Subjects knew that there might be more than one round, but not how many. All material used in the experiment was in hard-copy, pre-printed and not individualized. Anonymity and independence together mean that subjects could not reasonably believe that the experiment was anything else but a sequence of choice experiments. Payoffs and contributions are accumulated over the three rounds.

The purpose of the choice experiment is to reveal subjects' willingness to pay for an environmental public good in each of the three rounds. It does so by observing individuals' choice among alternatives that differ in the relative price of public goods contributions vis-à-vis private gains. The four treatment groups that will be introduced below take the identical choice experiment in every round except for round 2. In round 2, two treatment groups will face archetypical regulatory interventions, with one group subject to taxes and the other subject to restrictions backed by a sanction. The other two groups offer appropriate controls. The heuristic strategy is then to *test for the presence of significant differences between subjects that experience an episode of explicit incentives in round 2 and those that do not, using changes in willingness to pay derived from the choice experiments as a measure.* Testing is done at two points: (i) when the explicit interventions are applied in round 2, and (ii) in the subsequent round 3 when regulatory interventions have been removed.

We first focus on the design of the choice experiment, before discussing the treatments and appropriate controls.

2.2 The basic choice experiment

The basic choice experiment is the baseline experiment and it is used for all four treatments in round 1, all four treatments in round 3, and one treatment in round 2 (*BASELINE*). In round 2, specific variations of the baseline are used in three treatments explained below.

The public good used in the experiment has to ensure that the trade-off between the public good and private gain is real while minimizing the (strategic) interaction between participants. We achieve this by using abatement of greenhouse gas emissions that contribute

to global climate change as the public good. The choice of a global public good such as greenhouse gas abatement eliminates any strategic interaction between subjects due to the enormous number of people involved in its provision (about six billion). Contributions towards the public good purchase and retire carbon dioxide (CO_2) permits from phase II of the European Union Emission Trading System (EU ETS) and thus restrict the EU's regulated greenhouse gas emissions by the same amount. Subjects therefore decide on real CO_2 reductions.

In each round, subjects choose a production level that determines both their private payoff and their contribution to the public good. Feasible production levels, corresponding payoffs and contributions for the baseline case are given in Table 1. The permit price at the time of the experiment was EUR 24.50 per ton of carbon dioxide emissions. Production level 5 generates the maximum of the sum of payoff and contribution, implying that some amount of production may be socially desirable. During the introduction to the experiment subjects were informed about the EU ETS and the process of buying and irrevocably withdrawing permits from the aggregate quota. The feasibility of abating EU CO_2 emissions through this mechanism was demonstrated during the introduction by logging into the experimenter's carbon trading account and purchasing and retiring one permit. The acquisition and retirement of EU ETS permits is not possible for private individuals without significant upfront cost.[2]

The mechanics of the experiment involve two sheets for each round. One instruction sheet that explains the choice experiment and one answer sheet along the lines of Table 1 on which subjects enter their choice of production level into a matrix showing for each production level the associated private payoff, contribution, and amount of CO_2 abated.

By design, the relative price between private payoffs and contributions changes monotonously along the activity level. As in other choice experiments, subjects' choice thus reveals their respective willingness to pay: The additional EUR 3 in cash payout from a one-point increase in activity between $[0; 9]$ requires a strictly non-decreasing reduction in contributions and

[2]For individuals to carry out similar activities requires the completion of bureaucratic processes in order to establish a personal trading account at the German Emissions Trading Authority plus a registration fee of EUR 200. In total, at most 31 private individuals had a personal trading account in Germany in 2008.

1	2	3	4
Production Level	Your Payoff	Climate Contribution	CO_2 Abated
0	0.00 EUR	30.00 EUR	1.22 t
1	3.00 EUR	29.50 EUR	1.20 t
2	6.00 EUR	28.50 EUR	1.16 t
3	9.00 EUR	27.00 EUR	1.10 t
4	12.00 EUR	25.00 EUR	1.02 t
5	15.00 EUR	22.50 EUR	0.92 t
6	18.00 EUR	19.00 EUR	0.78 t
7	21.00 EUR	14.00 EUR	0.57 t
8	24.00 EUR	8.00 EUR	0.33 t
9	27.00 EUR	2.00 EUR	0.08 t
10	27.00 EUR	0.00 EUR	0.00 t

Table 1: Payoffs and climate contributions

hence amounts of CO_2 abatement. On this basis, one can convert the discrete price schedule into upper and lower bounds for a subject's marginal willingness to pay (see columns 2 and 3 in Table 2). For instance, a subject choosing production level 1 in round 1 is willing to give up EUR 3 for at least $20kg$ and at most $41kg$ reduction in CO_2 emissions. Converting this into upper and lower bounds, this implies a marginal willingness to pay for one ton of CO_2 abatement in a EUR range of $[73.50; 147]$. It is therefore possible to observe - within a range - individual's willingness to pay and therefore the trade-off between real emissions reductions and personal gain.

This completes the description of the basic set-up.

2.3 Treatments

The four treatments are three variants of the basic choice experiment plus a baseline with no variation. The variations occur only in round 2. Two treatments, *TAX* and *COMMAND & CONTROL*, involve regulatory interventions through explicit incentive schemes. One

Production Level	Round 1 (min)	Round 1 (max)	Round 2 (min)	Round 2 (max)
0	147.00 EUR/t		122.50 EUR/t	
1	73.50 EUR/t	147.00 EUR/t	67.38 EUR/t	122.50 EUR/t
2	49.00 EUR/t	73.50 EUR/t	36.75 EUR/t	67.38 EUR/t
3	36.75 EUR/t	49.00 EUR/t	12.25 EUR/t	36.75 EUR/t
4	29.40 EUR/t	36.75 EUR/t	9.80 EUR/t	12.25 EUR/t
5	21.00 EUR/t	29.40 EUR/t	3.43 EUR/t	9.80 EUR/t
6	14.70 EUR/t	21.00 EUR/t	-2.45 EUR/t	3.43 EUR/t
7	12.25 EUR/t	14.70 EUR/t	-4.08 EUR/t	-2.45 EUR/t
8	12.25 EUR/t	12.25 EUR/t	-4.08 EUR/t	-4.08 EUR/t
9	0.00 EUR/t	12.25 EUR/t	-18.38 EUR/t	-4.08 EUR/t
10		0.00 EUR/t		-18.38 EUR/t

Table 2: Implied MWTP for choices in round 1 (all treatments) and round 2 (*TAX* and *EX*)

treatment *EXOGENOUS PAYOFF CHANGE* (*EX*) is the control for the treatment *TAX*. The fourth treatment, *BASELINE*, serves as control for *COMMAND & CONTROL* (*C&C*).

The explicit interventions in *TAX* and *C&C* change the basic choice experiment subjects face in round 2. Subjects are instructed that in order to protect the environment individual emissions are metered and regulated. In the *TAX* treatment, a corresponding tax level is set for each level of production(see column 2 in Table 3). It is made clear that tax revenues are neither redistributed nor spent on abating carbon dioxide. The instruction sheet presents the amount of tax payable at each production level (see Table 3) and the resulting net payoff. The answer sheet contains only the net payoff and is therefore otherwise identical to the one in round 1. In the *C&C* treatment, a permitted threshold of 6 is stipulated in the instructions and for each production level above this threshold, a penalty is payable that matches the tax due at this level (see Table 3). It is explained that fine revenues are neither redistributed nor spent on abating carbon dioxide. The instruction sheet presents the amount of fine payable at each production level (see column 3 Table 3) and the resulting net payoff. The answer sheet contains only the net payoff and is therefore otherwise identical to the one in round 1. Both regulatory instruments are equivalent in the sense that both

1	2	3
Production Level	Tax	Penalty (*C&C* treatment)
0	0.00 EUR	0.00 EUR
1	0.50 EUR	0.00 EUR
2	0.75 EUR	0.00 EUR
3	1.50 EUR	0.00 EUR
4	3.50 EUR	0.00 EUR
5	5.50 EUR	0.00 EUR
6	8.00 EUR	0.00 EUR
7	11.50 EUR	11.50 EUR
8	15.50 EUR	15.50 EUR
9	19.50 EUR	19.50 EUR
10	21.00 EUR	21.00 EUR

Table 3: Taxes and penalties

shift the payoff maximizing production level from 9 and 10 to 6.

The control treatments for the explicit interventions introduced in *TAX* and *COMMAND & CONTROL* are *EX* and *BASELINE*, respectively. Subjects in the *BASELINE* treatment face the same relative prices between the private and the public good as subjects in *COMMAND & CONTROL* for all production levels in the range of $[0; 6]$. For the range $[7; 10]$, appropriate adjustments need to be introduced in order to make choices comparable. This is because this range undergoes a shift in relative prices that results from the introduction of penalties for exceeding the stipulated threshold of 6. Exceeding that level is economically sensible only for those with a negative marginal willingness to pay, i.e. subjects that expect to benefit from climate change.

The *EX* treatment mimics the change in relative prices resulting from the tax in the *TAX* treatment. Compared to the baseline, contributions to climate change abatement in the *TAX* treatment become relatively less expensive. This can be seen in the reductions in the lower and upper bound of the MWTP for each activity level reported by Table 2 for round 1 (columns 2 and 3) and round 2 (columns 4 and 5). Behavior in round 2 of the *TAX*

treatment can therefore not be directly compared with behavior in round 1 or in *BASELINE*. The appropriate control for *TAX* is therefore a treatment in which subjects face the exact same shift in payoffs and, hence, relative prices. This is achieved in the *EX* treatment, where this shift is motivated on the instruction sheet not by reference to a tax, but appears merely as an experimental payoff variation. The answer sheets of *TAX* and its control, *EX*, are identical. The design thus situates subjects in both treatments in a choice situation that is indistinguishable from an economic point of view.[3] What distinguishes the treatments is the presence of an explicit intervention for environmentally more benign behavior in one of them. This allows us to disentangle the effects of the change in the relative price of carbon abatement from any effect due to the interventional character of this change - here the imposition of explicit abatement incentives in the form of a tax in the *TAX* treatment. This is indispensable for identifying the effect of interventions on intrinsic motivation as opposed to other imperfections in response to changes in relative prices that might occur for reasons unrelated to motivational factors.

In round 3, all subjects face the same decision problem as in round 1. This continuation in case of the *BASELINE* treatment and the reversion to round 1 in the *EX* treatment was implemented without further explanation. In the *TAX* and *C&C* treatments, subjects were informed that emissions would neither be measured nor sanctioned in this round and that the payoff matrix was therefore exactly the same as in round 1. They also knew - as before - that their action would not affect future payoffs.

A total of 172 students participated spread over two sessions each lasting about one hour. Both sessions included all four treatments. For session 1 we recruited 43 subjects by an open call to students of different departments (e.g. economics, political science, physics, humanities etc.). The 129 subjects of session 2 were recruited from a single lecture. We found no evidence that behavior of subjects in the two sessions differed significantly. 10 participants had to be excluded from the analysis since their participation in more than one session could not be ruled out with certainty. Choices of eight further subjects were not included since they failed to fill out the answer sheet or had misread instructions. Hence,

[3]This includes the use of tax revenues or, respectively, price differentials: Both are retained in the experimental fund.

the results presented in what follows are based on 154 independent observations. After the third round some participants (5 in session 1 and 10 in session 2) were selected at random and both their private payoffs and contribution to abate CO_2 materialized. This payment procedure was also announced in advance. Individual payoffs ranged from EUR 0.- to EUR 81.- and 30 permits for a total of 30 tons of CO_2 were bought and retired. The experiment was conducted between May and July, 2008.

3 Behavioral Predictions

In order to unambiguously identify motivation crowding effects of explicit incentive schemes on behavior in rounds 2 and 3, we require behavioral predictions that can account for two types of changes between rounds: The first are those that can be explained on the basis of conventional theory, the second those that deviate from conventional theory, but are not treatment-specific. Only deviations from the predictions of conventional theory that vary systematically with treatments can provide evidence of motivation crowding. While results for round 3 can be compared directly under this caveat, results for round 2 are properly interpreted using differences-in-differences.

We start with changes between rounds that can be explained on the basis of conventional theory. These are based on the interaction between three elements.

1. Subjects' willingness to pay for contributions to climate change abatement: Subjects reveal their marginal willingness to pay by way of their choice in round 1 where they give up cash pay-outs in exchange for contributions to emissions reduction (see columns 2 and 3 in Table 2).

2. Changes in the relative price schedule: With the exception of the *BASELINE* treatment, all treatments in round 2 alter the relative price of contribution or payoffs over at least part of the range $[0; 10]$ relative to round 1.

3. Stability of preferences: Conventional theory maintains that subjects' willingness to exchange payoffs for contributions at the margin should not change simply because

relative prices have changed. The circumstances of the price change should be immaterial for their marginal willingness to pay as long as the economic nature of the choices remains the same. Preferences should therefore be *invariant with respect to instruments*, even if those instruments include explicit abatement incentives.

Taken together, these three elements provide the basis for inferring a choice in round 2 that is consistent with a choice in round 1 on the grounds of conventional theory. How do we arrive at this inference? The inference of round 2 behavior relies on the observed approximate MWTP of subjects in round 1 and mapping those into round 2. Since round 1 and round 2 intervals can overlap (see Table 2), the mapping is not always unique. In cases where mappings are not unique, the working hypothesis is that individuals are operating at the lower bound of the MWTP for climate change abatement. The justification for the working hypothesis is that subjects in the *BASELINE* treatment tend to reduce their MWTP from round 1 to 2. We provide robustness checks on this assumption below. Figure 1 visually summarizes how the choice in round 2 (on the vertical axis) is inferred on the basis of the choice in round 1 (on the horizontal axis) for each treatment. Four curves are shown: The 45-degree line is the inferred choice for the *BASELINE* treatment. Relative prices remain constant from round 1 to round 2 and so does the inferred choice. The inferred choice in the *COMMAND & CONTROL* treatment is represented by a 45-degree line for first round choices in the interval $[0; 6]$ that subsequently kinks and runs horizontally at a second round choice of 6 for first round choices within $[7; 10]$. This reflects the fact that for those subjects with a production level greater than 6 in round 1, the change in relative prices introduced by the production restriction backed by penalties makes a choice of 6 the most preferred option given their MWTP.[4] The stepwise functions, finally, represents the inferred choice in round 2 in the *TAX* and *EX* treatment. Both treatments are equivalent under economic criteria and give rise to the same inferred change in choice. Since the relative cost of cash payouts is greater in round 2, the inferred activity levels are correspondingly reduced. The black dashed

[4]When inferring choices we censor MWTP to abate CO_2 at zero, i.e. do not consider preferences that induce people to emit CO_2 for its own sake. However, a choice of 10 in round 1 reveals a MWTP of at most zero and hence round 2 choices above 6 would be consistent with that. None of the subjects in the *C&C*, *TAX* or *EX* treatments that chose 10 in round 1 actually did so.

(gray dotted) line indicates the inferred choice based on the minimum (maximum) MWTP revealed in round 1. For four out of ten production levels the predictions differ depending on which bound is used. In what follows we use the mapping based on the minimum MWTP (the black dashed line) but provide robustness checks for other mappings.

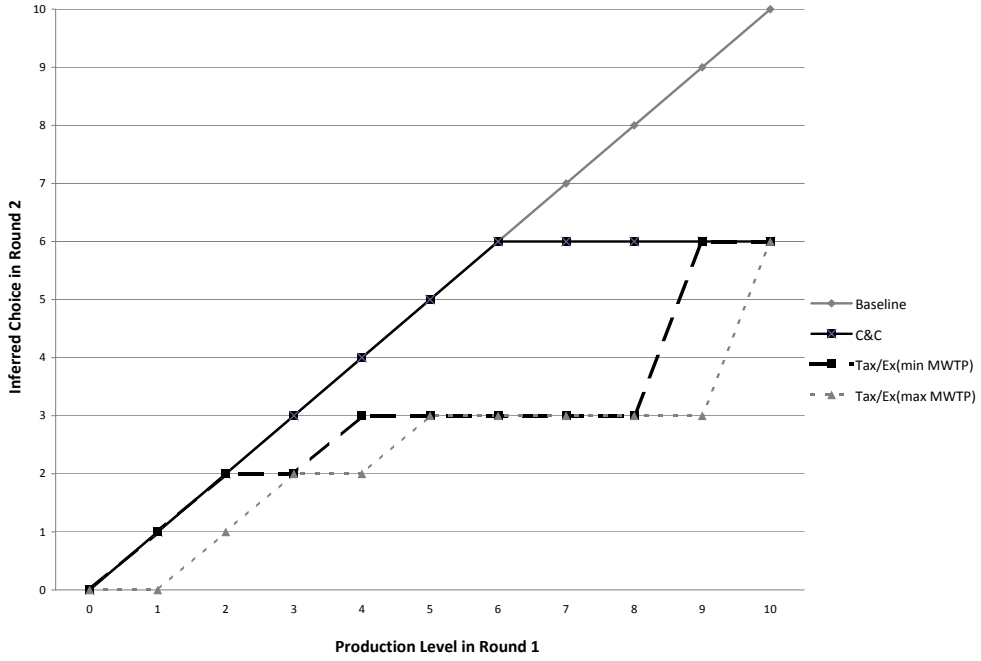

Figure 1: Mapping from production levels in round 1 to inferred choices in round 2.

To complete the behavioral predictions, we finally consider those changes between rounds that deviate from conventional theory, but are not specific to treatments. Such factors that are common to all treatments (e.g. boredom or a general failure to adjust to changes in relative prices) need to be discarded. The proper approach is therefore one of differences-in-differences: We compare treatment-specific deviations from inferred behavior between treatments that have the same change in relative prices. Alternative approaches such as comparing absolute production levels or changes in these levels are potentially contaminated by fixed effects and are therefore less powerful. In fact, their results can be misleading as we show below. Instead, the differences-in-differences will allow us to differentiate between imperfections in adjustments to changes in relative prices and motivation crowding triggered

by explicit regulatory interventions. On this basis, we formulate our first three hypotheses regarding deviations of observed choices from predictions in round 2.

Hypothesis 1 (Round 2: *TAX* vs. *EX*) *The deviation from predicted production levels and actual choices in round 2 is the same for an explicit price intervention (TAX) and for an equivalent price shift without an explicit interventional character (EX).*

The *TAX* treatment and its control, *EX*, are subject to an identical price shift between rounds 1 and 2. This price shift renders private payoffs more expensive in terms of contributions to climate change abatement. For both treatments, we would therefore predict a reduction in production levels (see Figure 1). Given the identical variation in relative prices, the respective deviations can be compared directly, without further adjustments. In the light of the predicted invariance ceteris paribus of choices with respect to the source of a price shift, hypothesis 1 predicts that subjects in both treatments will adjust in the same way to the same change in relative prices. In other words, we should find no difference in the deviations between the two treatments if motivation crowding is irrelevant.

Hypothesis 2 (Round 2: *C&C* vs. *BASELINE*) *The deviation from predicted production levels and actual choices in round 2 is the same for an explicit quantity restriction backed by a pecuniary sanction (COMMAND & CONTROL) and for an unrestricted production (BASELINE).*

In round 2 subjects in the *BASELINE* and *COMMAND & CONTROL* treatments face the same relative prices for production levels below 7. For levels of 7 and above relative prices in the *C&C* treatment change due to the penalty imposed in round 2. In the *C&C* treatment predicted production levels in round 2 are the same as in round 1 except that all choices above 7 are reduced to 6. In the *BASELINE* treatment there is no change in relative prices at any production level. In order to ensure comparability between treatments we therefore apply the Falk and Kosfeld (2006) adjustment procedure to the *BASELINE* treatment. This procedure adjusts the observations in the *BASELINE* treatment such that all production levels above 6 are treated as 6 in both predictions and actual choices in round 2. Since relative prices below 7 do not change from round 1 to round 2 in either treatment,

there is no a priori reason to expect that the changes from round to round (i.e. the deviations from predictions) should differ between the treatments once a Falk-Kosfeld adjustment is implemented.

As a cautionary tale, we also include - as a third hypothesis - a direct comparison of the instruments of a tax and a standard. Such a comparison would be of obvious policy interest, pitting the two regulatory approaches into a direct face-off. On the basis of conventional theory, the hypothesis will be one invariance with respect to instruments.

Hypothesis 3 (Round 2: *TAX* vs. *C&C*) *The deviation from predicted production levels and actual choices in round 2 is the same for an production restriction backed by a pecuniary sanction (COMMAND & CONTROL) and for an explicit price incentive (TAX).*

Both treatments covered by hypothesis 3 change the relative price of contributions from round 1 to round 2 by introducing explicit abatement incentives. The resultant change in relative price is identical over the production interval $[7; 10]$ since the penalty and the tax rate involve identical payoff deductions by design. In the production interval $[0; 6]$, *COMMAND & CONTROL* retains the relative prices of round 1 while subjects in the *TAX* treatment face lower relative prices for contributions. It would be tempting to conclude that if individuals have stable preferences, deviations from predicted choices should be indistinguishable between the two treatments and if statistically significant deviations were present, this would be evidence of motivation crowding. This would be predicated, however, on an assumption that subjects perfectly adjust to changes in relative prices. Since the changes in relative price schedules differ significantly between treatments, differences in deviations from predictions might also be driven by other factors not captured by conventional theory that affect adjustments to changes in relative prices. We test hypothesis 3 in the following section and show a meaningful interpretation is difficult.

Note that hypotheses 1 through 3 do not make a prediction on levels in round 1. Levels will depend on individual's preferences for climate change abatement given their interpretation of the public goods situation. Even subjects with preferences for the public good can follow the 'large economy' logic (Andreoni 1988) of the climate change game and arrive at a zero willingness to pay on account of the literally billions of players involved in the public

good provision. Alternatively, subjects can conclude that the experiment offers an offset opportunity for their consumption of private goods that have an adverse effect on the public good (Vicary 2000, Kotchen 2009). In this case, even with a very large number of players, individuals will rationally choose production levels below 9 (Kotchen 2009). Changes in the setting from round 1 to round 2 are restricted to simple changes in relative prices. The fundamental characteristics of the public goods situation remain therefore unaffected.

Hypotheses 1 through 3 will fail if there is a sufficiently strong interaction between the regulated individual's willingness to pay and the explicit abatement incentives provided by the regulatory instrument. The presence of such an interaction is at the heart of experimental and theoretical contributions to understanding MC. A prerequisite is that the experimental public goods situation is interpreted as an opportunity for prosocial or moral behavior such that individuals can act upon their pro-public motivation. If present, this has two implications. One is that subjects will rationally choose production levels below 9 in round 1 in line with the strength of their pro-social motivation. The second is that since this motivation relates to subtle inter-personal relationships, altering these relationships by introducing explicit incentives can give rise to an unintended reduction in cooperative behavior. Both in vertical (Bénabou and Tirole 2003, Sliwka 2007) and horizontal relationships (Bénabou and Tirole 2006), the introduction of externally provided incentives interacts with preexisting information imperfections to generate reductions in cooperative effort by individuals. Jointly, these theoretical perspectives on MC give rise to cognitive explanations for motivation crowding and could explain why the above hypotheses fail. We discuss in section 5 how the experimental set-up relates to the mechanisms required for a cognitive interaction between treatments and the relevant motivational preferences.

Turning to predictions on relative changes from round 1 to round 3, we arrive at the final hypothesis.

Hypothesis 4 (Round 3 behavior) *The change in production levels from round 1 to round 3 is identical for all treatments.*

Recall that round 3 presents the original price schedule of the baseline and consists of the same answer sheet as round 1. Therefore, observed production levels in round 1 can

serve as predictions of round 3 choices for all treatments and no adjustment is required for comparing the changes in production levels across treatments. By comparing the changes across all treatments, any common factors that induce a change in revealed MWTP (e.g. boredom) are taken into account.

Note that hypothesis 4 can hold even if hypotheses 1 through 3 fail: If preference shifts occur in round 2 despite the behavioral predictions of the relevant hypotheses, these shifts may only be effective within the round. The relationship between the two hypotheses is therefore a question of short-term persistence of incentive-induced changes in preferences.

The presence of such persistence - even in the long run - has been postulated by theoretical models of MC (Bénabou and Tirole 2003, 2006). The reason is that the cognitive processes modeled there are based on Bayesian learning. Even if regulation has been removed, subjects still take any information previously revealed into account. Cognitive motivation crowding (CMC) is therefore a candidate explanation if hypothesis 4 fails. Evidence from field experiments provides grounds for predicting such a failure: Short-term spillovers from incentive episodes into later non-incentive episodes have been observed in other examples of public goods contribution (Meier 2007).

We provide a more detailed discussion of the contribution of CMC to explaining the experimental evidence in section 5 and now turn to the experimental evidence used to test these hypotheses.

4 Results

We start with some basic information on the behavior of the four groups in the three rounds before testing the key behavioral hypotheses set up in the preceding section.

Figure 2 reports on average production levels of each treatment group in each round, with 95% confidence intervals added. Round 1 offers information on subjects' willingness to pay for abatement before explicit incentive schemes are introduced. While mean production levels are different, these first-round differences are not statistically significant as the confidence intervals indicate: p-values of pairwise Mann-Whitney tests range from 0.2281 to 0.9038 and a Kruskal-Wallis test of equality of all four means gives a p-value of 0.5181. Figure 3 shows

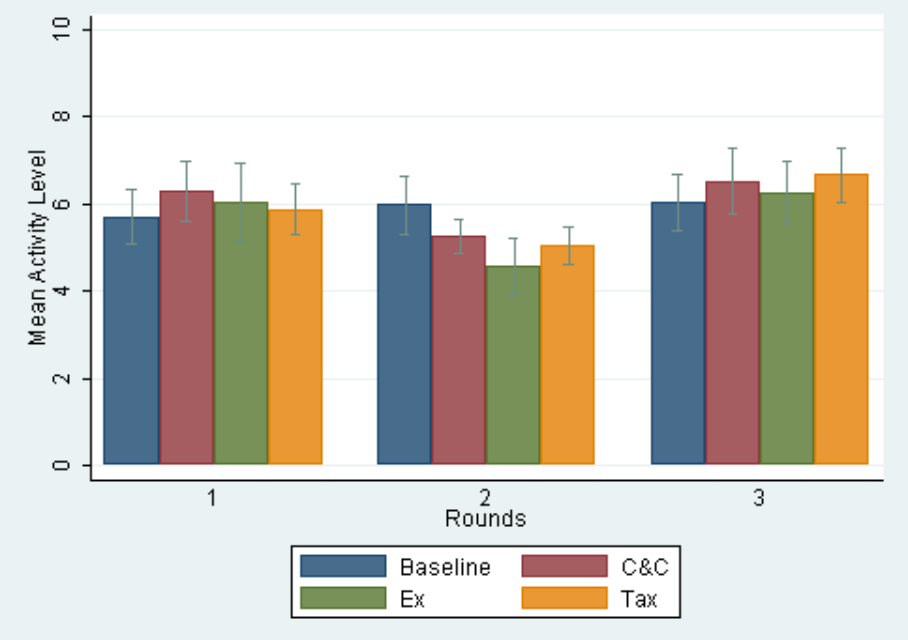

Figure 2: Production levels

that choices in round 1 are dispersed over the entire interval of feasible choices with modals between 5 and 7. Hence, average production levels are substantially below the pure private payoff maximum of 9: The median production level is 6 in all groups (see Table 4). While not *per se* conclusive evidence for the presence of intrinsic pro-social motivation (rather than selfish offset motives) among subjects, this provides a pool of subjects for whom crowding out could conceivably be observed when explicit interventions are introduced in round 2.

4.1 Instantaneous impacts of regulatory interventions

The effects of explicit regulatory interventions to induce a more environmentally benign behavior are laid out in hypotheses 1 - 3 for the case with no interaction between explicit incentive schemes and intrinsic motivation. We now proceed to test these hypotheses.

For presentational purposes, statistical tests are combined with scatter plots that visualize subjects' choices in different rounds. The size of the dots represents frequency of observation, lines represent the inferred choices for each treatment per Figure 1. Figure 4, for instance, visual compares for each of the four treatments inferred and observed choices between rounds

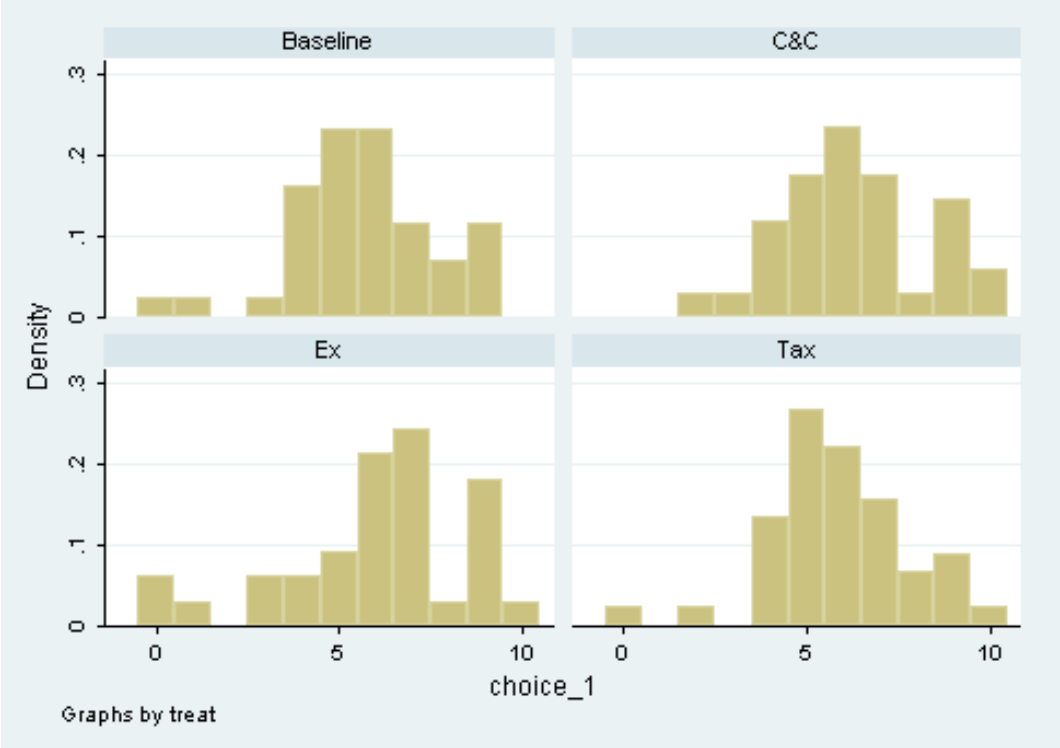

Figure 3: Distribution of production levels in round 1 by treatment

1 and 2. This reveals the following basic pattern: (a) Preferences are approximately - but not perfectly - stable. In the *BASELINE* treatment 29 out of 42 subjects stick to their first round choice, 4 increase and 9 decrease production levels in round 2. On average production levels increase by 0.26. The p-value of a Wilcoxon signed-rank test of equality of choices in both rounds is 0.1494. (b) Predictions for the *C&C* treatment are even more accurate since 29 out of 34 subjects behave in line with the inferred choices based on revealed MWTP and changes in relative prices. (c) While only 13 out of 45 (12 out of 33) subjects in the *Tax* (*Ex*) treatment exactly match the inferred choice in round 2, Figure 4 clearly shows that subjects initially in the middle range choose on average lower production levels in round 2 than those in the other treatments. This is in line with the shift in relative prices (see Table 2).

These findings support our starting points of stable preferences and relevance of changes in relative prices. Next we search for systematic deviations from this rules stipulated by conventional theory that are treatment specific, i.e. to whether changes in relative prices are

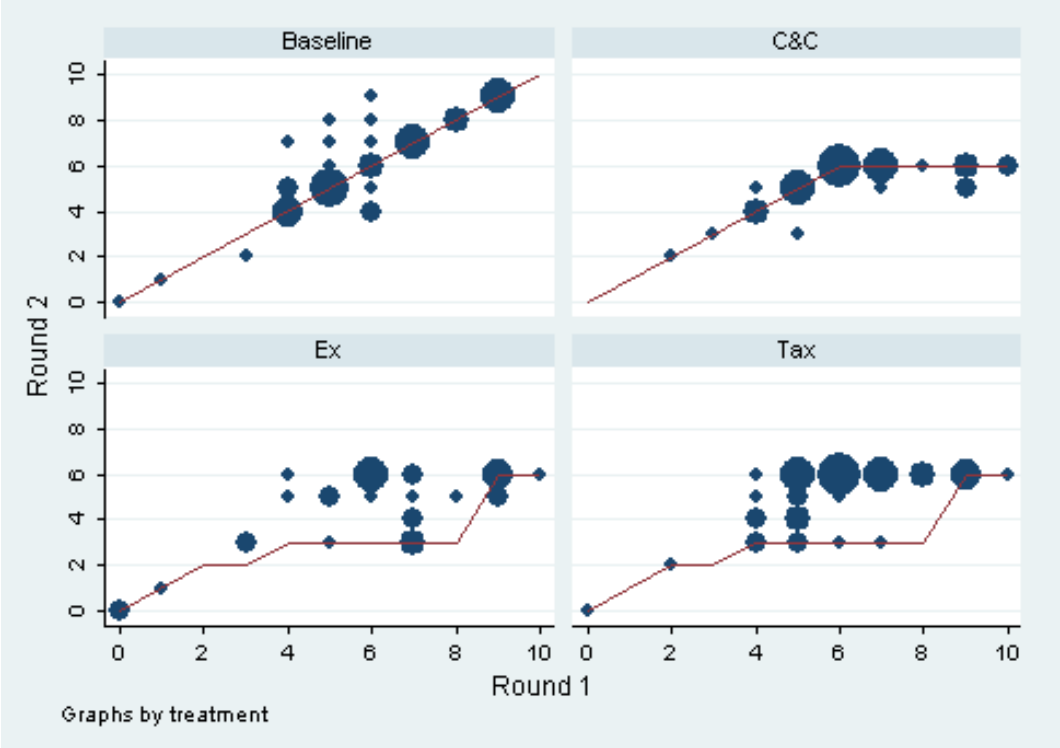

Figure 4: Scatter plot of round 1 and 2 choices by treatment. Lines indicate inferred choices.

explicitly framed as regulatory interventions, and can, hence, be attributed to motivation crowding.

Testing hypothesis 1 checks for such differences by comparing whether subjects in the *TAX* and *EX* treatments respond differently to the same change in relative prices experienced in round 2. Comparing the deviations from inferences based on the lower bound on the MWTP revealed in round 1 across treatments, a Mann-Whitney test gives a p-value of 0.0529. Subjects in the *TAX* treatment on average choose higher production levels than their *EX* counterparts. While the average deviation from predicted production levels was -1.82 in the *TAX* treatment it amounted to only -1.24 in *EX*. Using the mean of the MWTP intervals in Table 2 instead of the lower bound the p-value increases to 0.0740. Since we observe an increase in production levels, i.e. a reduction in MWTP in the *BASELINE* treatment, using the lower bound is more appropriate.

Result 1 *In round 2 subjects in the TAX treatment reduce their MWTP for CO_2 abatement by more than subjects in the EX treatment. This is evidence for crowding out of intrinsic*

motivation by an explicit price intervention compared to an equivalent but non-interventive change in relative prices.

We now proceed to test whether there is a similar effect for the quantity restriction backed by a fine in the *C&C* treatment. Hypothesis 2 conjectures that in the absence of motivation crowding there should be no difference between the deviation in the *C&C* and in an adjusted *BASELINE* treatment. In the *C&C* treatment inferred choices for round 2 are equal to production levels in round 1 capped from above at 6. In the *BASELINE* treatment the Falk and Kosfeld (2006) adjustment does the same for round 1 production levels that serve as predictions of round 2 levels and for the actual choices in round 2. Comparing the difference between inferred and actual choices across treatments a Mann-Whitney test gives a p-value of 0.1983. Hypothesis 2 can therefore not be rejected at standard levels of significance.

Result 2 *There is insufficient evidence for the presence of crowding effects induced by an explicit quantity restriction backed by a fine.*

This finding accords with Falk and Kosfeld (2006). They observe crowding out of intrinsic motivation of agents to co-operate with a principal if the principal herself restricts their choice set. However, the effect vanishes if the restriction originates from the experimenter. This is a first indication that subjects do not regard the present experiment as a principal-agent situation with the experimenter.

Next we provide the cautionary tale of directly comparing the two regulatory interventions by testing hypothesis 3. A Mann-Whitney test for differences in the deviations between inferred and actual choices gives a p-value of 0.0000. Hence, we can reject hypothesis 3 on all conventional levels of significance. Individuals subject to price regulation exhibit a significant drop in their MWTP which is not the case for those regulated by a quantity restriction. In fact, actual production levels in round 2 are on average 0.12 below predictions in the *C&C* treatment and 1.82 above predictions in the *TAX* treatment.

Result 3 *In contrast to quantity restrictions, price regulation is associated with a significant drop in subjects' MWTP for abatement.*

However, as explained above this is not conclusive evidence of motivation crowding since it can not distinguish between a general distortion in reactions to relative price changes and

		Treatment			
		BASELINE	*C&C*	*EX*	*TAX*
Round 1	Average	5.72	6.29	6.06	5.89
	Median	6	6	6	6
Round 3	Average	6.02	6.53	6.27	6.67
	Median	6	6	6	7
# Subjects		42	34	33	45

Table 4: Mean and median production levels in rounds 1 & 3

MC. Results 1 and 2, however, strongly indicate that there are several deviations from classic theory operating at the same time. Attributing the entire difference summarized in Result 3 to motivation crowding would in this case drastically overestimate the relevance of MC. The distortion in adjustments to relative price changes is confirmed by comparing the deviation in the *C&C* treatment to that in the *Ex* treatment. Testing for equality of deviations also yields a p-value of 0.0000 although there is no explicit intervention in the *Ex* treatment and hence no basis for motivation crowding. This shows that it is crucial to separate MC from other behavioral effects in order to truly provide evidence for its presence and quantify its impact.

4.2 Short-term persistence of crowding effects

We now turn to the question of how an episode of explicit regulatory incentives influences subsequent behavior once the incentives have been removed. Results 1 to 3 establish the presence of MC vis-à-vis explicit price incentives in round 2. If the MC detected in round 2 are manifestations of a (Bayesian-) learning process, that suggests intertemporal persistence of the change in preferences of those subjects that have experienced an emissions tax. These behavioral changes may therefore spill over into the subsequent period, even though the incentives are no longer present. Comparing production levels in rounds 1 and 3 provides evidence on short term persistence of behavioral changes triggered by explicit interventions applied in round 2. In round 3, subjects face again an identical decision to round 1. The only difference lies in their regulation experience, i.e. their treatment in round 2. As Table

4 and Figures 2 and 5 show, average production levels in the regulated treatments readjust overall - as expected - in the direction of levels chosen in round 1. Hypothesis 4 predicts

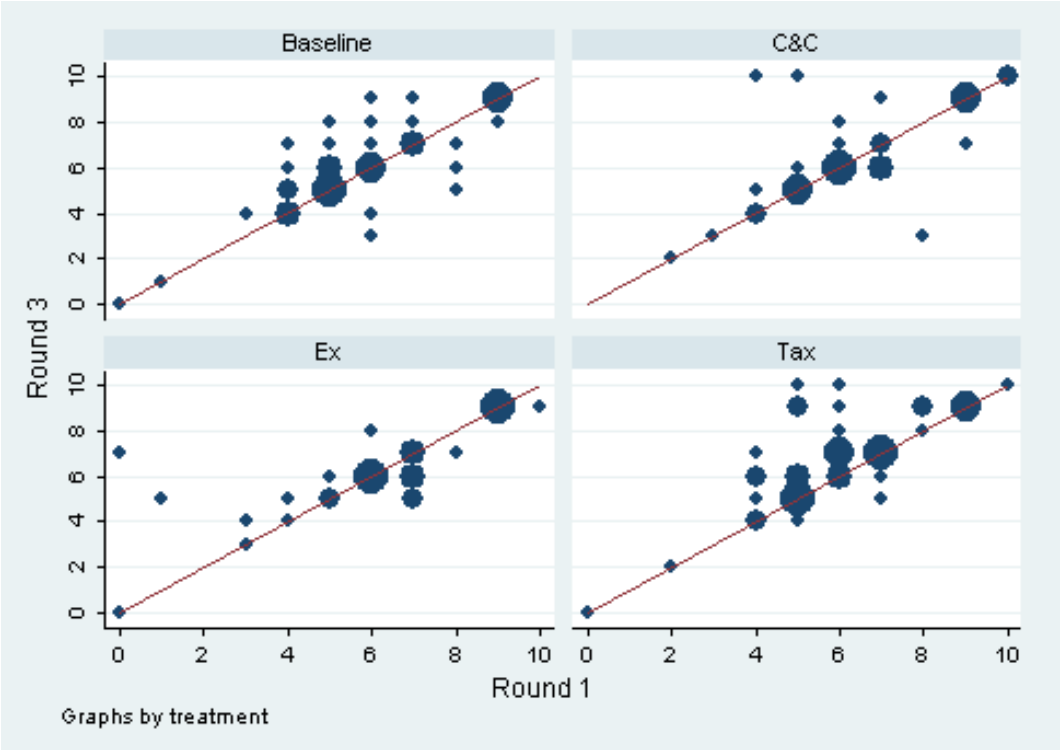

Figure 5: Scatter plot of round 1 and 3 choices by treatment. Lines indicate inferred choices.

that the choice in round 3 is not affected by what happened in round 2, i.e. that any changes in revealed MWTP for abatement are common across treatments. Using a Kruskal-Wallis rank test to test for the equality of differences between round 1 and 3 choices across all treatments yields a p-value of 0.0599. Next we conduct pairwise comparisons using Mann-Whitney tests. Production levels increase significantly more in the *TAX* than in the *EX* treatment ($p = 0.0129$) from rounds 1 to 3. This can also be seen in the lower two panels of Figure 5. Crowding out induced by explicit price interventions is persistent in the short run. There is no significant difference in the difference between round 1 and 3 choices for the *BASELINE* and *C&C* treatments ($p = 0.3427$) confirming the absence of motivation crowding established in Result 2. Moreover, the difference between the *TAX* and the *C&C* treatments remain significant ($p = 0.0453$). Hypothesis 4 is rejected since we do find significant treatment effects in how choices in rounds 1 and 3 relate to each other.

Result 4 *The crowding out of intrinsic motivation induced by explicit price regulation is short term persistent although the intervention is no longer present.*

5 Discussion

5.1 Evaluation of the experimental evidence

Summarizing the results of the experiment, we find that explicit price incentives trigger crowding-out of contributions to a global public good, even after taking into account imperfect adjustments to changes in relative prices in a neutrally framed context. By contrast, an episode of explicit quantity controls enforced in a way that - in economic terms - is broadly equivalent to the price interventions does not lead to crowding-out. Testing for short-run persistence, the crowding effect of the price incentive survives into a subsequent round in which the tax is no longer operational.

It is important to put these results into the broader context of the recent literature on motivation crowding. Empirically the presence of crowding effects has been established for a wide range of situations and instruments (Gneezy and Rustichini 2000a,b, Falk and Kosfeld 2006, Meier 2007). However, both field and laboratory experiments focus on small scale principal-agent relationships (with the exception of Meier (2007)) more suitable e.g. for workplace settings than for public policy. Moreover, none of these studies compares quantity and price based interventions. The theoretical literature explaining crowding phenomena relies on preexisting informational imperfections. The introduction of incentive instruments provides new information in vertical relationships (Bénabou and Tirole 2003, Sliwka 2007) or changes the signal value of contributions to the public good in horizontal relationships (Bénabou and Tirole 2006). This literature models the reduction in motivation as a rational response to a change in the informational context triggered by the introduction of an incentive scheme. The interaction between incentives and motivation is therefore purely cognitive.

The theoretical literature provides both concepts and terminology that appear to fit our experimental findings. We proceed by testing whether the appearance survives scrutiny.

5.2 Can Cognitive Theories Explain the Evidence?

Cognitive motivation crowding (CMC) theories demonstrate how the interaction of explicit incentives and informational imperfections can give rise to MC. Their explanatory power with respect to the experimental evidence presented here is determined by two aspects. One is the relationship between the mechanisms required for CMC and the experimental design. The other is the relationship between the predictions of CMC and the experimental evidence. We address these two aspects in the following.

CMC requires horizontal or vertical interaction channels for the interaction between information and incentives to play out. Through a number of experimental design features, however, these channels are shut down in the present experimental setting. One feature is the common knowledge at the outset of the experiment on (a) the randomness of group assignment, (b) the independence of actions between rounds, and (c) on the independence of the type and sequence of incentives provided (if any) from individual or group behavior. On the basis of this common knowledge, participants have no reason for expecting that the incentives are in some way conditioned on their individual or aggregate behavior (Bénabou and Tirole 2003). Also, the randomness of assignment and independence of choices across rounds rules out the conditional behavior studied in Sliwka (2007) as an explanation as subjects cannot learn something about the environment from observing the explicit incentives provided. The second feature is the anonymity of decisions among participants. This anonymity shuts down reputation effects along the lines of Bénabou and Tirole (2006). The anonymity combined with the independence of choices also rules out horizontal interactions between subjects that would be a precondition for peer pressure effects (Kandel and Lazear 1992, Fischer and Huddart 2008). In short, there are no obvious cognitive reasons to our knowledge for expecting MC to be present in this setting. This contrasts with the settings in Andreoni (1993), Cardenas et al. (2000), Gneezy and Rustichini (2000a) and Gneezy and Rustichini (2000b).

One exception to the above are any effects subjects' self-image might have on their behavior (Bénabou and Tirole 2006). In general, self-image effects would seem almost impossible to avoid in experiments of this type. However, if self-image is the driving force behind any observed crowding effect, we should be able to distinguish it from other drivers if the

predicted direction and persistence of effects differ. Note that the experimental set-up also should not give rise to a vertical relationship between subjects and experimenter as it is common knowledge in the experiment that the experimenter is completely precommitted, demonstrates that he has no outcome-specific interest, and that the experimenter cannot condition on observed individual or group behavior in any way. However, subjects decisions are not perfectly anonymous to the experimenter who could conceivably identify the winning individuals during the pay-off procedure. This might interfere with subjects' choices if they have beliefs about the experimenter's expectations.[5] This potential interference is of relevance only as far as the effects of regulatory interventions are concerned, whereas its potential impact on the absolute level of contributions is immaterial to the subsequent analysis. If subjects believe that the explicit incentive schemes in round 2 signal experimenter's expectations or a social norm, this would result in a shift of contributions towards six in both the *TAX* and *C&C* treatments for rounds 2 - 3. However, neither Figure 4 nor 5 indicate such a trend. Production levels do not move towards six, but quite the opposite. Round 3 in the *TAX* treatment is the only incidence across all treatments in rounds 1 & 3 where production levels are significantly different from (i.e. above) six ($p = 0.0262$). Moreover, the experimenter's expectations or social norm conjecture fails to explain the difference between the command-and-control and the tax intervention.

As a control, subjects were given an exit questionnaire that tested for attitudes vis-à-vis emissions trading and other relevant factors that might interact with subject's observed behavior. For example, subjects may object to the idea that polluters can continue polluting as long as they purchase permits rather than being obliged to abate. The questionnaire asked subjects about their views on the idea of trading pollution rights. There is no detectable relationship between these views and behavior in the laboratory. Alternatively, subjects may find that the aggregate emissions quota is set too high or low, potentially biasing their decisions. The only finding in this context is the predictable results that subjects with a stronger environmental orientation did provide more contributions to climate change abatement than those with a low level. While contributing to our understanding of the level

[5]There is mixed evidence on the relevance of experimenter blindness (Hoffman et al. 1994, Bolton et al. 1998).

of contributions, though, environmental orientation did not interact with changes in behavior in any detectable way. The experimental set-up was also careful to introduce explicit prices regarding the environmental good for all treatments at the same time, namely at the start of the experiment. This set-up ensures that the results pick up only intervention-specific effects rather than a possible opposition of subjects to the general idea of putting a price on environmental goods. These attitudes can therefore not explain observed effects and to not correlate with behavior.

With the obvious channels of most cognitive mechanisms of MC shut down and evidence inconsistent with communication of experimenter's expectation or social norms via interventions only self-image concerns remain of the established drivers of MC. To check whether the Bénabou and Tirole (2006) type of self-image concern can explain the crowding out of price instruments in this setting we perform a 'long-run' persistence test.

5.3 Testing for long-term persistence

To test for the long-term persistence of the observed crowding effects participants of session 2 were invited to participate in a follow-up session exactly one week later. This was feasible without a previous announcement since they were recruited from within a weekly lecture course. They again faced the same payoff matrix as in period 1. All treatments had the same instructions. 70 students participated in the follow-up with 62 being included in the analysis.[6]

Hypothesis 5 (Round 4 behavior) *The change in production levels from round 1 to round 4 is identical for all treatments.*

Using standard theory to predict round 4 behavior the same arguments apply as for hypothesis 4. The difference is that round 4 tests for evidence of long-term persistence. Rejecting hypothesis 5 implies that the preference changes induced by round 2 persist for a non-trivial amount of time and have therefore altered something more fundamental in subjects' valuation of climate change abatement contributions in either direction. Self-image concerns based

[6]The remaining 8 subjects were already excluded from the analysis of rounds 1-3 for reasons specified in section 2.

		Treatment			
		BASELINE	*C&C*	*EX*	*TAX*
Round 1	Average	5.80	6.67	6.29	6.00
	Median	7	7	7	6
Round 4	Average	6.27	7.27	6.93	7.00
	Median	7	8	7.5	9
# Subjects		15	15	14	17

Table 5: Mean and median production levels in rounds 1 & 4

on cognitive mechanisms would predict a long-run persistence of any difference between instruments observed in the short-run. This is a result of the Bayesian updating process that takes into account all information previously revealed.

With little theory to provide guidance on the time scales of behavioral changes, round 4 is spaced one week from the session containing rounds 1 through 3. The smaller sample (62 subjects) of returned subjects is statistically not distinguishable from the larger population that participated in rounds 1 to 3. Inspection of the average and median production levels in Table 5 seems at first to suggest significant effects. The median production level has increased in all treatments, most notably in the *TAX* treatment. Comparisons of average production levels in period one and four confirm this observation of increased production levels only for the the *TAX* ($p = 0.0668$) treatment and for the *BASELINE* ($p = 0.0260$). However, variances within treatment groups are considerable. Comparing effects across treatments fails to pick up longer-term treatment effects between the initial and final round of the experiment one week later. The p-values of pairwise Mann-Whitney tests are 0.6964 for *TAX* vs. *EX*, 0.5715 for *BASELINE* vs. *C&C* and 0.7700 for *TAX* vs. *C&C*.

Result 5 *There are no significant differences in the adjustments from round 1 to round 4 between treatments.*

Hypothesis 5 can therefore not be rejected. When testing for longer-run persistence of these effects by repeating the last round one week later, crowding effects are no longer present. This makes it unlikely that subjects' concern about their self-image (Bénabou and Tirole

2006) drive the crowding effects observed in this experiment. This renders the last of the established drivers of MC implausible. However, some caution is in order due to the smaller number of observations in round 4.

5.4 Other candidate explanations

The limited traction of established cognitive drivers of MC raises the question of what alternative mechanism is likely to be activated by price incentives rather than quantity incentives, but would fail to lead to a longer term change in behavior. One plausible candidate for an alternative mechanism that combines an instrument-specific response with a short-, but not long-term persistence is affect-based behavior, leading to affective motivation crowding (AMC) as opposed to CMC.

Affect-based behavior requires a 'valence' basis, i.e. an association of positive or negative emotions with an event (Forgas and Laham 2005). A plausible valence basis for the effects observed in the experiment derives from the fact that its choice situation can be readily regarded as a moral decision in which dichotomous categories of 'good' and 'bad' apply. If this moral dichotomy is at the root of emotional attitudes vis-á-vis regulation, different forms of regulation can differ by how well they map into those categories. If they fail to map well along subjects' moral views, this can generate feelings of being treated unfairly and therefore lead to a behavioral reversal (De Dreu and Steinel 2006, Pillutla and Murnighan 1996).

A quantity control instrument of the type applied in round 2 creates two distinct regimes and thus maps well into a dichotomous view of 'good' and 'bad' behavior. It creates one regime for those within the threshold and another for those outside. In rounds 1 and 3 we asked subjects what they regard as the socially optimal production level.[7] The mean across all subjects was 4.98 in round 1 and 5.19 in round 3, the median in both rounds was 5. Since the threshold level of 6 is close to what subjects regard as the 'social optimum', it is likely to divide the set of subjects close to where subjects themselves would conjecture that a threshold between good and bad would lie.

The price instrument, on the other hand, is only marginally selective rather than divid-

[7]The question was on the same answering sheet subjects used to choose their production level (see the instructions in the appendix).

ing behavior into dichotomous moral categories. If subjects regard themselves as incurring sacrifices in order to 'do the right thing', the introduction of a tax in round 2 can appear to condemn every choice in round 1, however prosocial, with the exception of a level of zero. The implicit message of a tax instrument that no behavior is good enough not be penalized will then plausibly conflict with subjects' attitudes regarding their behavior because it fails to acknowledge the moral categories that may motivate individuals. Requiring a decision on production levels while such sentiments of being treated unfairly do linger in subjects' minds will plausibly lead to an reduction in the willingness to contribute.

To test for this we compare the mean distance between a subject's choice and what he or she regards as socially optimal. Note that choice and self-stated 'social optimum' are significantly different for all treatments ($p = 0.0000$ in round 1 and $p < 0.001$ in round 3 for all treatments except the *BASELINE* where $p = 0.0796$). The only treatment where this distance significantly increases ($p = 0.0157$) from round 1 to 3 is the *TAX* treatment. The p-values for the other treatments are between 0.7039 and 0.9667. Note that there is no significant change in what subjects in the *TAX* treatment consider to be socially optimal ($p = 0.7232$). Hence, exposure to an explicit price incentive scheme does not crowd out moral standards themselves but (temporarily) the motivation to adhere to them in a public goods context. The stability of moral standards fits well with the short term persistence of crowding effects observed. Once the affective response to the tax has ebbed away, subjects revert to their previous relation between moral sentiment and actual behavior. This accords with the predictions of the theoretical framework introduced by Loewenstein and O'Donoghue (2007) to represent the interaction between deliberate processes including moral standards and affective responses to stimuli of different intensity and valence.

Both the observation of price-induced MC at the same time as the absence of control-induced MC in this public goods choice situation and the line of reasoning on the valence basis of this difference are very close to the original discussions by Frey (1992, 1997) that first brought the notion of motivation crowding into the economic literature. 'Frey motivation crowding' (FMC) consists of deleterious price-based incentives for environmental public goods alongside rather benign effects of command and control on intrinsic motivation. These effects are also observed in the experiment presented here. At the same time, FMC is ex-

plicitly related back to a psychological literature that emphasizes not only cognitive, but in particular also affective drivers of human behavior (DeCharmes 1968). AMC is therefore a plausible alternative as well as complementary explanation for FMC alongside CMC.

6 Implications

The focus of our present experiment is - in our mind - novel for three reasons: The first is that it directly compares two archetypical policy instruments and finds that - contrary to the conventional prediction - instrument choice and motivation appear to interact in a way predicted by Frey (1997). The second is that these effects are generated by an experimental design that provides only limited traction for the cognitive mechanisms developed in the theoretical literature. Affective mechanisms offer a plausible alternative explanation of the evidence. The third is that quantity and price based instruments can differ due to deviations from conventional theory other than motivation crowding. We find evidence that, in contrast to quantity restriction, price regulation is affected by imperfect adjustments to relative prices over and above any changes in marginal willingness to contribute caused by reduced intrinsic motivation.

What is the significance of these findings? We believe that there are two areas of significance. One area is the development of a comprehensive theory of MC. Given the extensive empirical literature, the need to develop a precise understanding of its nature is imperative. Cognitive theories of MC are obviously important steps towards developing such an understanding. At the same time, MC is not sufficiently well understood to be attributed to cognitive mechanisms alone: MC of the type described by Frey is found in settings in which CMC would not predict its presence. Our experimental results show that FMC may be attributable to either CMC, AMC, or both. The second area is policy. Given the incompleteness of current MC theories, it may be too early to formulate policy recommendations to tackle MC: Behavior under regulation may systematically differ from its predictions. Alternative remedies might be passed over at significant cost.

The link between instrument choice and motivation established in this paper adds an important new dimension to the problem of 'prices versus quantities'. Further research

should shed more light on the nature and underlying drivers of this type of interaction.

References

Andreoni, J.: 1988, Privately provided public goods in a large economy: The limits of altruism, *Journal of Public Economics* **35**, 57–73.

Andreoni, J.: 1993, An experimental test of the public-goods crowding-out hypothesis, *American Economic Review* **83**(5), 1317–1327.

Austin, D. and Dinan, T.: 2005, Clearing the air: The costs and consequences of higher cafe standards and increased gasoline taxes, *Journal of Environmental Economics and Management* **50**, 562–582.

Baumol, W. J. and Oates, W. E.: 1988, *The theory of environmental policy*, Cambridge University Press, Cambridge, UK.

Bénabou, R. and Tirole, J.: 2003, Intrinsic and extrinsic motivation, *Review of Economic Studies* **70**, 489–520.

Bénabou, R. and Tirole, J.: 2006, Incentives and prosocial behavior, *American Economic Review* **96**(5), 1652–1678.

Bolton, G. E., Katok, E. and Zwick, R.: 1998, Dictator game giving: Rules of fairness versus acts of kindness, *International Journal of Game Theory* **27**, 269–299.

Brekke, K. A. and Nyborg, K.: 2008, Attracting responsible employees: Green production as labor market screening, *Resource and Energy Economics* **30**, 509–526.

Cardenas, J. C., Stranlund, J. and Willis, C.: 2000, Local environmental control and institutional crowding-out, *World Development* **28**(10), 1719–1733.

CCC: 2008, Building a low-carbon economy - the UK's contribution to tackling climate change, *Report*, Committee on Climate Change, London.

Charness, G. and Gneezy, U.: 2009, Incentives to exercise, *Econometrica* **forthcoming**.

De Dreu, C. and Steinel, W.: 2006, Social decision making in fuzzy situations: Motivated decision making and strategic choice, *in* D. De Cremer, M. Zeelenberg and J. K. Murnighan (eds), *Social Psychology and Economics*, Routledge, London, pp. 55–78.

DeCharmes, R.: 1968, *Personal Causation: The Internal Affective Determinants of Behavior*, Academic Press, New York.

Dijkstra, B. R.: 1999, *The Political Economy of Environmental Policy: A Public Choice Approach to Market Instruments*, Edward Elgar, Cheltenham, UK.

Ellingsen, T. and Johannesson, M.: 2008, Pride and prejudice: The human side of incentive theory, *American Economic Review* **98**(3), 990–1008.

EU: 2008, On the effort of member states to reduce their greenhouse gas emissions to meet the communitys greenhouse gas emission reduction commitments up to 2020. COM(2008) 17 final, *Proposal for a decision of the european parliament and of the council*, Commission of the European Communities, Brussels.

Falk, A. and Kosfeld, M.: 2006, The hidden costs of control, *American Economic Review* **96**(5), 1611–1630.

Fehr, E. and Falk, A.: 2002, Psychological foundations of incentives, *European Economic Review* **46**, 687–724.

Fehr, E., Klein, A. and Schmidt, K. M.: 2007, Fairness and contract design, *Econometrica* **75**(1), 121–154.

Fischer, C. and Newell, R. G.: 2008, Environmental and technology policies for climate mitigation, *Journal of Environmental Economics and Management* **55**(2), 142–162.

Fischer, P. and Huddart, S.: 2008, Optimal contracting with endogenous social norms, *American Economic Review* **98**(4), 1459–1475.

Forgas, J. P. and Laham, S. M.: 2005, The interaction between affect and motivation in social judgments and behavior, *in* J. P. Forgas, K. D. Williams and S. M. Laham (eds),

Social Motivations. Conscious and Unconscious Processes, Cambridge University Press, Cambridge, pp. 168–193.

Frey, B. S.: 1992, Pricing and regulating affect environmental ethics, *Environmental and Resource Economics* **2**, 399–414.

Frey, B. S.: 1997, *Not Just for the Money: An Economic Theory of Personal Motivation*, Edward Elgar Publishing, Cheltenham, UK.

Frey, B. S. and Jegen, R.: 2001, Motivation crowding theory, *Journal of Economic Surveys* **15**(5), 589–611.

Frey, B. S. and Oberholzer-Gee, F.: 1997, The cost of price incentives: An empirical analysis of motivational crowding-out, *American Economic Review* **87**(4), 746–755.

Gneezy, U. and Rustichini, A.: 2000a, A fine is a price, *Journal of Legal Studies* **29**(1), 1–17.

Gneezy, U. and Rustichini, A.: 2000b, Pay enough or don't pay at all, *Quarterly Journal of Economics* **115**(2), 791–810.

Goulder, L. H. and Parry, I. W.: 2008, Instrument choice in environmental policy, *Review of Environmental Economics and Policy* **2**(2), 152–174.

Hahn, R. W.: 1990, The political economy of environmental regulation: Towards a unifying framework, *Public Choice* **65**(1), 21–47.

Harrington, W., Morgenstern, R. D. and Sterner (eds.), T.: 2004, *Choosing Environmental Policy. Comparing Instruments and Outcomes In the United States and Europe*, Resources for the Future, Washington.

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

Kandel, E. and Lazear, E. P.: 1992, Peer pressure and partnership, *Journal of Political Economy* **100**(4), 801–817.

Keohane, N., Stavins, R. N. and Revesz, R. L.: 1998, The positive political economy of instrument choice in environmental policy, *in* P. Portney and R. Schwab (eds), *Environmental Economics and Public Policy: Essays in Honor of Wallace E. Oates*, Edward Elgar, Cheltenham, UK.

Kerr, S. and Newell, R. G.: 2003, Policy-induced technology adoption: Evidence from the U.S. lead phasedown, *Journal of Industrial Economics* **51**(3), 317–343.

Kotchen, M. J.: 2009, Voluntary provision of public goods for bads: A theory of environmental offsets, *Economic Journal* **119**, 883–899.

Krysiak, F. C.: 2008, Prices vs. quantities: The effects on technology choice, *Journal of Public Economics* **92**(5-6), 1275–1287.

Loewenstein, G. and O'Donoghue, T.: 2007, The heat of the moment: Modeling interactions between affect and deliberation, *Working paper*, Cornell University, Ithaca, NY.

Meier, S.: 2007, Do subsidies increase charitable giving in the long run? matching donations in a field experiment, *Journal of the European Economic Association* **5**(6), 1203–1222.

Montero, J.-P.: 2002, Prices versus quantities with incomplete enforcement, *Journal of Public Economics* **85**, 435–454.

NCF: 2006, Phase I EU ETS trading decisions, *Report*, New Carbon Finance, London.

Pillutla, M. M. and Murnighan, J. K.: 1996, Unfairness, anger, and spite: Emotional rejections of ultimatum offers, *Organizational Behavior and Human Decision Processes* **68**(3), 208–224.

Requate, T. and Unold, W.: 2003, Environmental policy incentives to adopt advanced abatement technology: Will the true ranking please stand up?, *European Economic Review* **47**, 125–146.

Sliwka, D.: 2007, Trust as a signal of a social norm and the hidden costs of incentive schemes, *American Economic Review* **97**(3), 999–1012.

Titmuss, R.: 1970, *The Gift Relationship: From Human Blood to Social Policy*, Allen and Unwin, London.

Vicary, S.: 2000, Donations to a public good in a large economy, *European Economic Review* **44**, 609–618.

Zhao, J.: 2003, Irreversible abatement investment under cost uncertainties: tradable emission permits and emission charges, *Journal of Public Economics* **87**(12), 2765–2789.