

Global environmental problems, voluntary action and government intervention*

Andries Richter^{1*} and Daan van Soest²

¹CEES, Oslo University

²VU University Amsterdam and Tilburg University

*contact author: a.p.richter@bio.uio.no

February 28, 2011

Draft version prepared for E. Brousseau, T. Dedeurwaerdere, P.A. Jouvét and M. Willinger (eds), 2012. *Global Environmental Commons: Analytical and Political Challenges in Building Governance Mechanisms*, Oxford University Press, pp. 223-248.

Abstract:

The global community faces several very pressing environmental challenges such as climate change, depletion of the high-sea fisheries, and unprecedented rates of biodiversity loss. Governments are in the process of designing environmental policies to address these problems unilaterally, but also collectively (in the form of international agreements). Meanwhile, private citizens and firms are observed to voluntarily take protective action. Whereas standard game theory would predict that formal government intervention can only provide an extra stimulus for protective action, there are many examples of external interventions decreasing agents' propensity to undertake socially desired activities. This chapter provides an overview of the literature on the circumstances under which formal interventions can crowd out voluntary contributions to the common good. Furthermore, it is discussed how the effectiveness of government intervention may be improved by preserving the agents' intrinsic motivation to contribute to the common good.

*The authors gratefully acknowledge financial support by the Netherlands Organization for Scientific Research (NWO) as part of the program on Evolution and Behavior. They also wish to thank Marc Willinger and an anonymous referee for their insightful and inspiring comments on an earlier version of this chapter.

1. Introduction

Climate change, the depletion of (high sea) fisheries, and biodiversity loss rank high on the list of environmental problems the global community is confronted with. Over the past decade these problems have received ample attention in the media (at least in developed countries), and as a result firms and consumers are very much aware of their existence and also – albeit to a varying degree – of the necessity to take mitigating action.

Even if individual agents are well aware of these problems, it is not necessarily the case that they do take protective action. Environmental quality depends on the aggregate behavior of all agents on this planet, while each individual agent's actions have a negligible impact on any of the above global environmental problems. Indeed, all these problems are classic examples of the so-called “tragedy of the commons” as the benefits of protective actions are enjoyed by everyone, while the costs of taking them are private.

This analysis suggests that without government intervention, the prospects for mitigating or preventing climate change, fisheries depletion and biodiversity loss are bleak. However, casual observation of the behavior of people around us reveals that many people take at least some protective action. When buying a new electric appliance people do take into account the energy consumption of the various competing brands: energy efficiency and amounts of CO₂ emitted per kilometer driven are among the criteria on the basis of which people choose a new car, and many households voluntarily separate their waste flows for recycling purposes. Also within the business community the concept of corporate social responsibility receives increasing attention (Khanna, 2001).

These examples of private mitigation activities suggest that people are motivated to take protective environmental action even if extrinsic incentives – such as environmental taxes or quota – are absent. This raises two questions. The first is whether or not the amount of action taken is equal to the socially optimal level. And if not, the second question is how government intervention should be designed to bring protective action to its socially optimal level.

The first question is easy to answer – for most environmental problems we cannot rely on the voluntary actions of consumers and firms alone. If we could, these

environmental problems would have been solved a long time ago. But the second question is much more relevant, because of the following reason. In the standard game-theoretic framework – that is, assuming that all agents are exclusively self-interested – government intervention is always welfare enhancing as own-profit maximizing agents do not voluntarily contribute to the common good (if there are net costs involved in doing so). However, because we observe that at least some people voluntarily undertake protective action, the possibility arises that government intervention may be counter-effective, because it may result in crowding out of the agent's intrinsic motivation to voluntarily contribute. And this is what this chapter aims to analyze; under what circumstances does government intervention strengthen (“crowd in”) or weaken (“crowd out”) the regulated agents' intrinsic motivation to act pro-socially, and under what circumstances is the change in behavior permanent?

While the issue of stimulating crowding in and preventing crowding out of intrinsic motivation is important in all policy domains they are especially important when facing global environmental problems like climate change, depletion of large-scale fisheries and biodiversity loss. These are all examples of environmental problems with many stakeholders (virtual everyone on the planet), and both the benefits and costs of reducing the problem are substantial. Typically benefits and costs are distributed unevenly across countries and citizens, which makes striking international agreements very difficult (see Finus 2008). In general, governments are unwilling to impose large costs on their citizens (Nordhaus and Yang 1996), and hence regulations are unlikely to be very strict in spite of the fact that most citizens perceive these to be serious problems. Due to this problem-awareness, citizens are willing to undertake voluntary actions to solve large-scale problems, such as sustain the global climate (Milinski et al., 2006, 2008). The combination of a high willingness to solve a problem voluntarily and rather mild government regulations are the circumstances under which crowding out is most likely to occur, and hence analyzing how crowding out can be prevented is especially important in these global commons problems.

The set-up of this chapter is as follows. In section 2 we provide an overview of the economic and psychological literature on the interaction between formal government intervention and the regulated agents' intrinsic motivation to contribute, resulting either

in crowding in or out. In section 3 we turn to examples of crowding out in (real-world) environmental problems. Having established that crowding out is observed to occur in the real world, we analyze how this is related to the institutional setting, and we do so in section 4. In section 5 we identify the factors that may lead to crowding out. Section 6 shows how these factors affect formal and informal institutions. The corresponding policy recommendations are provided in section 7, and finally section 8 concludes.

2. Intrinsic motivation and crowding out

One is said to be intrinsically motivated to perform an activity when one receives no apparent reward except the activity itself (Deci, 1971). Employees in firms may not be just motivated to work hard to capture incentive payments; they also take pride in their work. In a similar vein, people are willing to abstain from certain activities that are harmful to their environment, even if the private returns to these activities are negative (think of voluntarily reducing the amount of waste produced by one's household, which requires effort to search for environmentally friendly products that are, in many cases, more expensive than other products).¹ The intrinsic motivation to contribute to the provision of public goods (like reducing one's emissions of greenhouse gases, or purchasing sustainably produced tropical hardwoods to prevent species extinction) may arise because of moral considerations (Ayala, 2010) or because of a sense of fulfillment of having contributed to something constructive (cf. Ariely et al. 2008 for experimental evidence on this).²

The reason why the introduction of extrinsic incentives does not always result in increased provision of public goods is because the extrinsic reward may negatively affect an individual's intrinsic motivation 'to do the good'. This is best illustrated using Figure 1. In this figure, an agent's contribution to a public good is measured along the horizontal axis, while the vertical axis reflects the incentives provided by the government to induce

¹ Clearly, the extent to which people are willing to voluntarily contribute to a specific public good varies with its characteristics such as the amount of private benefits the good provides, how costly it is to contribute to its provision, the number of stakeholders involved, whether the benefits and costs of the good are distributed uniformly over the stakeholders or highly unevenly, etc. For an overview, see Barrett (2007) and Ostrom (2003).

² The utility derived from the process of contributing to something constructive is closely related to the idea that people may derive 'warm glow' from contributing to public goods, expressing a positive willingness to pay for projects in valuation studies. That these preferences should be taken seriously, is forcibly argued in the chapter by Nunes and Onofri in Part 3 of this volume.

the agent to act cooperatively (a subsidy per unit of contribution to the public good, or a tax per unit of pollution generated). If the agent is intrinsically motivated to contribute to the public good, her contributions will be non-zero even if there are no incentives provided by the government to do so. Suppose that this contribution level is equal to point A in Figure 1. Now the standard line of reasoning in economics would be: “we do not know why the agent provides a non-zero contribution to the public good, but if the government provides incentives (in the form of a subsidy or a tax), the stimulus for the agent to contribute is larger and hence contributions will go up” (see also Ostrom, 2000). Assuming that the agent’s contribution schedule is stable, the policy maker would expect that contributions would be increased to point B if the incentive (on the vertical axis) is set equal to, for example, s^* .

<Insert figure 1 about here>

However, the assumption of a stable contribution schedule is often not met in practice. External intervention may crowd out the intrinsic motivation to contribute, resulting in a leftward shift of the individual’s contribution schedule from S to, for example, S' . If so, setting the incentive equal to s^* would result in contributions ending up in point C rather than in point B. The monetary incentive itself increases the marginal cost of shirking and/or increases the marginal benefit of good provision, but this does not only result in a movement along the original contribution schedule (from A to B), but also in a leftward shift to point C. Hence, crowding out occurs because external incentives are not separable from intrinsic motivation (Bowles, 2008).

This simple figure suggests two things. First, government intervention is effective if the incentive is sufficiently strong, but it may be countereffective if the external incentive is too small (see also Frey 1997, Ch. 11, and Nyborg and Rege, 2003). In our example introducing s^* results in contributions falling from A to C, but if $s \geq s^{**}$, contributions end up in D, or better. This implies that the issue of preventing crowding out of intrinsic motivation is of special importance in case of global environmental problems – as argued in the introduction. Absent any international agreements effectively addressing the issues of climate change, species extinction and fisheries depletion in the

oceans, governments may still decide to unilaterally implement policies – but they will necessarily be less stringent or encompassing than if proper international agreements were in place. For example, emissions of CO₂ may be taxed at about 15 euros per ton (the current price at which tons of CO₂ are traded at the European Trading System), while a non-negligible share of the currently available global damage estimates 80 euros or higher (Tol, 2008). If the tax rate imposed is too low to counteract a possible leftward shift of the “emission reduction supply schedule”, the government’s unilateral intervention may even be countereffective. And the same holds for possible interventions regarding trade in tropical hardwoods (to reduce the rate of species extinction) or regarding the consumption of high-sea fish species.

The second conclusion is that it is probably cheaper to use policy instruments that preserve the agent’s intrinsic motivation to contribute; when preserving the agent’s intrinsic motivation, the same level of contributions can be achieved at much lower costs. This is because (i) the loss of intrinsic motivation constitutes a direct decrease in utility, and (ii) the enforcement costs are likely to be an increasing function of the level of stimulus provided.³ And clearly government intervention would be even more effective if it were able to crowd *in* the agent’s intrinsic motivation to contribute, inducing her supply curve *S* to shift to the right. Hence, governments should be interested in finding out which policy instruments are likely to result in crowding out, and which are able to increase contributions by strengthening the agent’s intrinsic motivation to contribute.

According to psychologists, crowding out is likely to occur if external intervention is perceived to reduce people’s self-determination and/or their self-esteem because they feel that their involvement and competence is not appreciated (Frey and Jegen, 2001, p.8). But the extent to which crowding out occurs in actual practice, is hard to establish because of the many confounding factors. Kreps (1997) gives the example of the difficulty of identifying the existence of intrinsic motivation in the workplace. Employees may work long hours because they take pride in their work, but also because

³ Note that this is even the case if the stimulus is provided in the form of taxes. For the same level of monitoring, higher taxes make tax evasion more profitable and hence enforcement needs to be increased. This is a welfare cost while the tax revenues themselves are a transfer from the agent to the government, and hence do not constitute a welfare gain.

they respond to fuzzy extrinsic motivators including fear of discharge, scrutiny by fellow employees or even the desire for their coworkers' esteem.

Whereas strong empirical support of the crowding out theory is hard to find, anecdotal evidence abounds. Children are willing to mow the lawn, but after they have been paid once to do it, they are only willing to do it again if they receive monetary compensation (Deci, 1971). The pharmaceutical company Merck decided to invest in developing an unprofitable drug against river blindness in order to increase its workforce's intrinsic motivation for research (Murdock, 2002). Imposing a fine on parents arriving late to collect their children at day care increased the number of late-coming parents (Gneezy and Rustichini, 2000a). Small honoraria for seminar speakers may increase the probability of declining the invitation (Gneezy and Rustichini, 2000b). Unit pricing of unsorted waste does not necessarily result in a substantial increase in the amount of material offered for recycling (Ackerman, 1997; Berglund, 2006). Survey results suggest that people may actually decrease their contribution to voluntary work if a fee for non-participation is introduced (Brekke et al., 2003). Opposition against the implementation of "not in my backyard" (NIMBY) projects may increase when monetary compensation is offered to the local community. An example is the case with a mid-level radioactive nuclear waste repository in Switzerland (Frey and Oberholzer-Gee, 1997; Frey et al., 1996).

3. Crowding out in environmental problems

As Kreps (1997) argues, "abundant smoke signifies a fire, and the assertion is too strongly rooted in folk wisdom to be entirely hot air." While monetary incentives do not crowd out intrinsic motivation per se (Cameron and Pierce, 1994), it is by now well-established it can happen for a wide range of conditions (Bowles, 2008; Deci et al., 1999, 2001; Frey and Jegen, 2001; Frey and Stutzer, 2006; Gintis et al., 2005; Ostrom, 2005; Volland, 2008).

However, this list of instances of crowding out raises two questions. The first question concerns the circumstances under which crowding out is observed to take place. The examples presented are of a parent who wants his lawn mowed by his child, a boss of a firm who wants a project to be implemented, an owner of a day care centre who wants

to induce the parents to pick up their kids on time, etc. These are all examples of so-called principal-agent problems where the principal has certain objectives for which she is dependent on the efforts of someone else (the agent) to have them achieved. In these situations the interests of the principal and the agent do not completely coincide. So how does this relate to environmental problems where there is not just one principal to be “served”, but a larger group – or even society? Indeed, failing to take protective action in environmental problems results in damages accruing to a large and diverse group of agents (citizens, firms, or consumers) as is the case with, for example, overharvesting fisheries, failing to invest in water capture in arid regions, destroying biodiversity by converting forests into arable land, and the use of fossil fuels contributing to global warming. These environmental problems have in common that all agents involved would be better off if they collectively undertook protective action, but given that all others do so, it is in each agent’s private interest not to do it. In that sense, they are all examples of social dilemmas where the objective functions of individuals and society are not perfectly aligned – see below. The second question regarding the examples presented refers to the temporal nature of the change in motivation: if crowding out occurs, is this change permanent and (largely) irreversible, or is it just transitory?

Before addressing the issue of whether crowding out may occur in (global) environmental problems too, we investigate whether social interaction in multiplayer social dilemma situations is properly described by the neoclassical assumption of selfish actors. Here, social dilemmas are situations in which actions that maximize the sum of payoffs of all stakeholders do not coincide with the actions that maximize the payoffs of an individual agent, and vice versa. Hence, in a social dilemma the social and private objectives are not perfectly aligned. Consider the fishery. After every fishing trip, the agent pockets the sales revenues of his harvests, but he bears only part of the costs. When catching a fish one does not only remove that fish from the pond or sea, but also all offspring generated by that fish. If all fishermen decided to catch fewer fish, more offspring would be produced and all fishermen would be better off. However, given that all other fishermen restrained their fishing effort, each individual fisherman would profit

from increasing his fishing effort. And similar considerations apply to other environmental problems such as global warming and biodiversity loss.⁴

Whether or not agents are willing to voluntarily contribute to a public good is often difficult to observe in practice. Data on individual harvests or catch, for example, is hard to obtain, and also it is difficult to establish what baseline level to compare actual harvests with. One more easily observable act of voluntary cooperation is agents' propensity to discipline their peers. For example, Brazilian fishermen in the Bahia region destroy the nets of fellow fishermen who do not respect the catch quotas (Cordell and McKean, 1992). Sanctioning one's peers is an example of a second-order public good provision because there are often – hidden – costs associated with imposing punishments (for example because one exposes oneself to retaliation) while – if effective – the benefits of the punished individual restraining his fishing effort level accrues to all individuals having access to the fishery. When agents impose punishments on their peers to improve public good provision, they offer a benefit to the rest of the group while they themselves incur a personal cost. Hence, disciplining one's peers is a second-order public good.

Because it is hard to establish to what extent observed pro-social behavior is the result of intrinsic motivation or of other considerations (including the threat of social punishment or the presence of hidden rewards), economists and psychologists alike have studied this behavior by means of controlled experiments. In social laboratory experiments subjects are confronted with a (computerized) game that captures the essence of the decision problem the researcher wants to analyze (such as the fishery problem described above). Subjects are induced to think carefully about how to behave as decisions have actual financial consequences. The propensity of humans to act cooperatively has, among others, been studied in Public Goods (PG) games and in Common Pool Resource (CPR) games. In both games the costs of contributing to the common good is larger than the private benefits but smaller than the resulting increase in

⁴ While the term 'crowding out' typically refers to government interventions reducing *individual* agents' motivation to contribute to public goods directly, it can also refer to the situation where the self-regulatory capacity of a *community* is negatively affected. Indeed, even if not all members of a community contribute to the (local) public good, some may decide to form a coalition to enhance its provision; see van Soest et al. (2010). In that setting, the policy issue is how to design government interventions that strengthen rather than weaken the self-regulatory capacity of the community. For an analysis of coalition formation regarding climate change and biodiversity conservation, see the chapter by Bréchet and Eyckmans in Part 2 of this volume.

aggregate payoffs (i.e., the sum of payoffs of all group members). In the PG game the decision to be made is how much to contribute to a public good, where the associated benefits accrue to all members of the group. In the CPR game subjects need to decide whether they are willing to put in less than the privately optimal amount of harvesting effort in order to reduce the negative consequences of their catch on the payoffs of their peers. It is easy to see that homo economicus would act non-cooperatively, deciding not to contribute to the public good in the PG game and choosing the privately optimal extraction effort level in the CPR game. This prediction also holds in case of multiple rounds of interaction as long as the number of periods the game is played is finite (and subjects are informed about this). In such a finitely repeated setting, there is no reason to “invest” in trying to maintain cooperation in the last round (because there are no future decisions to be affected), and hence also not in the round before that (because there will be zero cooperation in the last round anyway). But that means that there is no reason to invest in maintaining cooperation in the second-but-last round either, and hence, on the basis of backward induction, the conclusion is that it does not pay to invest in even the first round. The typical pattern that emerges when PG games and CPR games are played over multiple rounds is that cooperation declines quite steeply over time (see for example Fehr and Gächter, 2000, Gächter et al., 2008, and Ostrom et al., 1992). This pattern emerges if university students are used as subjects, but it is also observed with a wide variety of other subject pools too. This suggests that humans are willing to act cooperatively, but that defection by others results in cooperation declining over time.

Interestingly, both Fehr and Gächter (2000) and Ostrom et al. (1992) also implemented treatments in which subjects can punish their peers for acting non-cooperatively. In these treatments, each round consists of two stages. The first is the “social dilemma stage” as described above (i.e., modeled in the form of either a PG game or a CPR game), and the second is the sanctioning stage. In the latter stage subjects can decrease the payoffs of any of their peers at positive costs to themselves. Game theory predicts that sanctioning would never occur because of a backward induction argument very similar to the one above. In the last period there is no reason to sanction because the punisher incurs costs while it is impossible to affect the punished individual’s behavior because the game ends. Hence there is no incentive to act cooperatively in the first stage

(the social dilemma stage) of the last round either. That means that there is no reason to punish one's peers in the second stage (the punishment stage) of the one-but-last round, which implies that there is also no incentive to act cooperatively in the first stage (the social dilemma stage) of that one-but-last round. Continuing reasoning backwards, the standard assumption of players being purely selfish results in the prediction that there will be zero contributions to the common good and also no sanctioning in any of the rounds the game lasts. The experimental evidence gathered by Ostrom et al. (1992) and Fehr and Gächter (2000), however, refute these predictions. Punishments do take place, and the threat of being sanctioned raises the level of cooperation almost to the level that maximizes group payoff. And similar results are found when self-regulation is by means of rewards (as opposed to punishments), although the results tend to be a little less strong (Vyrastekova and van Soest, 2008).

Given that humans (fishermen in the real world, student subjects in controlled economic experiments) act more cooperatively than predicted by standard economic theory, the question is to what extent formal government intervention can be counterproductive by crowding out the regulated agents' propensity to cooperate voluntarily. In the environmental economics literature several cases have been documented of crowding out occurring in environmental problems. One of the earliest examples is by Anderson and Lee (1986, p.690) who observe that "the suggestion that policies be implemented assuming that people will not comply with them has the potential for eroding social capital which depends on respect for the law". This claim has been corroborated by, among others, Hatcher et al. (2000) and Sutinen et al. (1999).

A second example of formal rules crowding out informal norms is provided by Cárdenas et al. (2000). In this study experiments were run with people in rural Colombia who are confronted with a common pool problem in their daily life. The game used by Cárdenas et al. was a CPR game in which subjects were asked to decide how much timber to extract from a forest. The scenario presented was that harvesting had an adverse effect on water quality (as is actually the case in the study region), posing a cost on everyone in the group. The game was first played without any regulations while at a later stage an extraction norm was introduced that was enforced by a mild probabilistic fine. Cárdenas et al. (2000) find that subjects reduce their extraction level after the regulation

is introduced, but start extracting more aggressively after realizing that consequences are rather mild. Strikingly, in the last rounds, extraction levels were higher with the regulation than without. As a result, payoffs are significantly lower when individuals are confronted with a formal rule than in its absence; the weak official rule interacted with the internal norms of the subjects and crowded out their intrinsic motivation to cooperate. Therefore, it can be seen as a warning towards indiscriminately introducing regulatory intervention without a proper understanding of how it might undermine norms already operating in the field.

These examples suggest that crowding out can occur in multiperson environmental problems too, but the studies presented above do not answer the question whether crowding out is likely to be permanent or just transitory. Some evidence on the issue of the irreversibility of crowding out is available from Bouma et al. (2008). This study addresses the issue of farmers' propensity to invest in the construction (as well as the maintenance) of soil and water conservation structures. The study sites were five villages in three different watersheds in (predominantly arid) rural India. Investments in soil and water conservation structures (as well as in their maintenance) provide private and public benefits, and one can hypothesize that the same set of factors determines whether households invest in the structures themselves, or whether they put effort into maintaining them. In practice there is a crucial difference between the two activities, though. The investments themselves are highly subsidized (either by government organizations or by NGOs), whereas the maintenance activities are not. As expected, the households' propensity to invest depends on the extent to which they are dependent on income from agriculture, on the size of their land holdings, etc. Interestingly, the authors found that the decision to invest also depends on the community's amount of social capital. To measure social capital the authors use several proxies including social homogeneity (as given by the relative size of the largest caste in the community), but social capital was also quantified by having villagers participate in a simple game, a so-called Trust game. This game is played by two players and is set up as follows. The player that moves first, the investor, has to decide how to allocate an amount of money between him/herself and the second mover, the trustee. The investor can keep all money, give the total amount to the trustee, or anything in between. The amount of money sent –

if any – is tripled by the experimenter, and given to the trustee. The trustee then has to decide how to divide the amount received between him/herself and the first mover, the investor. The trustee can keep all money, give the total amount to the trustee, or anything in between.

Behavior in the Trust game is expected to reflect social capital because the standard game theoretic prediction is that the investor will not send any money. The reason is that the trustee will not give back any money if he/she is purely selfish, and hence the investor will pocket the investment fund him/herself. The Trust game therefore provides two measures of social capital – trust/altruism as measured by the amount sent by the investor, and altruism/reciprocity as measured by the share returned by the trustee; see also Cox (2004).

Bouma et al. (2008) use straightforward regression analysis to analyze the villagers' propensity to invest in soil and water structures themselves, as well as in their maintenance. Explanatory variables include the various measures of social capital (social homogeneity, amount sent, and share returned) as well as a large vector of subject-specific and village-specific control variables. In the results of the regression explaining the villagers' propensity to invest in the structures themselves, only variables that reflect private stakes are significant (per capita land holdings, household size) and none of the "social capital" variables. In contrast, the "social capital" ones are the most significant variables explaining the propensity to undertake maintenance activities. This is salient because the former investment activities are subsidized whereas the latter maintenance activities are not. Interestingly, the probability of a household contributing decreases substantially if there are maintenance funds that support material costs but that do not compensate individual effort.

These results reflect two things. First, it seems that the formal intervention by the government organizations and NGOs crowded out the households' propensity to voluntarily contribute to a public good – given that social capital indeed matters, as evidenced by its role in the maintenance activities. Second, it is noteworthy that this crowding out in the investment phase did not spill over to maintenance activities, suggesting that crowding out may be highly context-specific, but also does not result in

permanent crowding out (as the maintenance activities obviously took place after the investments in the structures had been made).

4. Crowding out and the design of environmental policies

Having established that indeed formal intervention may be counterproductive in terms of resource conservation objectives, the question arises whether government policies can be designed such that formal and informal institutions are mutually reinforcing. It seems that three institutional characteristics are especially important: (i) the extent to which the external intervention is perceived to be legitimate and adequate (or proportional), (ii) the extent to which participation is voluntary, and (iii) to what extent the institution is perceived to be supportive (rather than restrictive).

When an institution is perceived to be legitimate and fair, participants are much more inclined to obey the rules (Frey, 1997, Ch.6). A striking example supporting this finding comes from Danish fisheries, where “fishers feel they are taken hostage by an illegitimate management system, and thus feel it is morally correct not to comply” (Raakjær Nielsen and Mathiesen, 2003). Somanathan (1991) describes how state intervention in Central Himalaya “directly weakened villagers’ incentives to allow regeneration and conserve forests”. As a result, a well-functioning informal system based on social arrangements was crowded out. Similar phenomena can be observed in many African societies (Vatn, 2007). This raises the question how legitimacy can be achieved. Sometimes it seems to be enough to convince the individuals about the usefulness of the rule and that obeying it is in everyone’s interest (Rodriguez-Sickert et al., 2008). Reeson and Tisdell (2008), however, found that moral suasion does indeed promote cooperation, but only in the very short run.

One way to achieve legitimacy is involving stakeholders in the process of designing formal institutions (Dankel, 2009; Hatcher et al., 2000; Jentoft et al., 1998). This does not only include the set-up of the monitoring and enforcement mechanisms, but also on the way in which the benefits and costs of the institution’s interventions are allocated between all stakeholders (see also the chapter by Groom, Gatti, Goeschl and Swanson in Part 2 of this book). Such a participatory approach may build trust between users themselves, but also between users and central authorities. This may crowd in

stewardship motives, and increase compliance. This form of co-management has the additional advantage that stakeholders possess important knowledge which may help crafting better institutions (Jentoft et al., 1998). An active dialogue between stakeholders and decision makers can also help identifying and overcoming potential conflicts of objectives and stakeholders (Dankel, 2009). Many economists are somewhat skeptical about involving stakeholders too closely in the process of designing institutions, as it gives them the possibility to seek rents (Bergland et al., 2002; Johnson and Libecap, 1982). This is indeed problematic, especially when stakeholders differ in the amount of resources they have to lobby for their interests. In many cases, the voices that shout loudest are most heard (Hatchard, 2005). One could overcome this by making stakeholders more responsible and accountable (Mikalsen and Jentoft, 2008). This is especially necessary when the local users value various aspects of the environment differently than society does. For example, local users may be interested in having a well-functioning ecosystem (which ensures income in the future), while they do not necessarily care about biodiversity as such.

Regarding this, several economic experiments have been conducted to test whether user participation does indeed increase the effectiveness of the institution under consideration – i.e., does it mitigate the extent to which agents' intrinsic motivation are being crowded out? In a laboratory setting, this can be tested by allowing regulated subjects to vote on the details of the enforcement institution's design, and subsequently compare their behavior in the social dilemma activity to behavior of agents participating in a treatment in which regulation is exogenously imposed. Voting serves a dual purpose. First, the voting outcome (for example based on a majority voting rule) affects the design of the institution, and hence its direct effectiveness. But voting outcomes also provide information about the intentions and preferences of the community's majority to effectively protect the resource and to maximize group payoff (as opposed to trying to non-cooperatively maximize one's individual payoff). Therefore, in voting experiments we can observe whether a group of people is able to find consensus on designing effective institutions, but also whether voting itself affects the compliance of the institution that has been agreed upon.

Sutter and Weck–Hannemann (2004) provide an example where a failure to obtain majority agreement for the socially optimal action is detrimental to social welfare, as behavior in the social dilemma activity becomes significantly less cooperative. In their experimental study, subjects have the possibility to vote on a minimum contribution level to a public good, upon which they make their decisions about how much to contribute. When the group fails to achieve a majority vote in favor of the rule, contributions are significantly lower than in the treatment without. This makes intuitive sense because even though there are no binding rules in either treatment, a failure to reach consensus reveals information about the lack of cooperativeness of the co–players.

Obviously, the consequences of not achieving a majority vote are even more detrimental if the voting outcome results in the abolishment of formal institutions, as is uncovered by Tyran and Feld (2006). In this study, subjects can vote on the level of a (deterministic) sanction in a public goods environment. As is the case in Sutter and Weck–Hannemann (2004), subjects tend to contribute significantly less (more) when the majority vote was against (in favor of) the presence of an enforcement institution empowered to impose fines on those who contribute less than a certain level.

Having established that introducing voting with respect to details of the enforcement institution’s design can either improve or reduce welfare and conservation depending on the voting outcome (directly via the change in the institution, but also indirectly by crowding in voters’ intrinsic motivation to contribute to the common good), the question arises what factors determine voting behavior. Vyrastekova and van Soest try to answer this question in two related papers (Vyrastekova and van Soest (2003), and van Soest and Vyrastekova (2008)). In these two papers, subjects are allowed to vote on whether the enforcement institution should be provided with sufficient incentives to actively sanction excessive extraction, or not. More specifically, one subject was assigned to take the role as policy enforcer. The other subjects voted on whether or not the subject representing the enforcement institution is allowed to keep the fine revenues. If a majority votes against this, any collected fines are removed from the game. In this case, the enforcer is not expected to actively impose fines when observing violations of the formal rule because there are fixed costs associated with punishing. In this setting the weakly dominant strategy is to vote in favor of the enforcer receiving the fine revenues.

In Vyrastekova and van Soest (2003) two treatments were compared. In the first treatment, the policy enforcer always receives the revenues of her sanctioning activity (i.e., the fines imposed on those resource users who extract more than is prescribed by a rule). In the other treatment, the enforcer is only allowed to keep the fines if the majority votes in favor of this, as described above. Vyrastekova and van Soest find support for the hypothesis that voting actually improves efficiency of resource use as compared to the treatment in which incentives are assigned exogenously. Casting their vote serves as a means for resource users to communicate their stance with respect to the need for reduced aggregate extraction. Conditional on a majority having voted in favor of implementing an appropriate incentive structure, extraction behavior was significantly more cooperative in the voting treatment (i.e., much closer to the level that maximizes group welfare) than in the treatment where the enforcer is always allowed to pocket the fine revenues. Hence, allowing subjects to vote on the regulatory regime crowds in their propensity to act cooperatively.

In a companion paper, van Soest and Vyrastekova (2008) analyze to what extent actual voting outcomes depend on the characteristics of the enforcement institution. The specific characteristic they focus on is the probability that when engaging in enforcement, the institution is indeed able to successfully impose fines. Keeping the expected fine constant, they compared the impact of a 50% chance of conviction (and a specific fine level) on voting behavior to that in case of a 90% chance of conviction (and a lower fine level). In both cases, the weakly dominant strategy is to always vote in favor of the enforcer receiving the fine revenues, because of the arguments given above. Van Soest and Vyrastekova actually find marked differences between the 50% and 90% probability treatments. Whereas in the latter treatment resource users almost always vote in favor of the enforcer receiving the fines, a favorable majority voting outcome is achieved in less than 40% of the cases in the former treatment.

These results are striking as they imply that trying to save on enforcement costs by reducing the probability of conviction (with a concomitant increase in the fine level such that the expected fine is kept constant) is hazardous if the enforcement institution's effectiveness is at least to some extent dependent on the support of the regulated individuals. If the intervention is insufficiently effective, intrinsic motivation to

contribute to the public good is reduced and the regulated agents decide to vote against the government regulation. These findings suggest that individuals will not support an institution that is perceived to be unfair. A similar study has been undertaken by Kosfeld et al. (2009), where individuals could choose to become member of a sanctioning institution. The authors show formally that a likely equilibrium outcome will be that such an institution is formed, resulting in increased efficiency. These findings have been corroborated in an experimental setting. This study showed that institution formation can be an effective tool for solving a social dilemma, but fairness issues can be serious obstacles, confirming the results obtained by van Soest and Vyrastekova (2008).

An interesting case arises when individuals can communicate with their peers regarding what they perceive to be appropriate behavior. An experimental regularity is that communication alone is often sufficient to promote cooperation, even if any agreements made are non-binding. In many instances the social pressure arising from “cheap talk” is more effective in correcting behavior than a fine that could serve as a price. Even more surprising is the fact that voluntary participation can foster cooperation even without social pressure. Di Falco and van Rensburg (2008) analyze the effect of governmental subsidies on livestock farmers in Ireland. Farmers receive livestock premiums based on the number of cattle, but they can choose to sign up for a rural environmental protection scheme (REPS) as well. The authors analyze the effect of both payments on cooperation, but also on conservation effort. While the livestock premiums have no effect on cooperation and a negative effect on conservation, the payments from REPS lead to more cooperation and higher conservation effort. This is remarkable as encouraging cooperation is not an explicit aim of REPS. An open question remains whether the voluntary nature of the program makes users more cooperative, or just attracts users that have more cooperative attitudes. While it is well established that there are important feedbacks between institutions, preferences, and economic outcomes, further research is needed to identify the causal relationships between those elements.

5. The theoretical foundations of crowding out

In the previous section we have documented that the occurrence of crowding out is determined by (i) the legitimacy of the institution and the level of involvement of the

individuals, (ii) the voluntary nature of it, (iii) and the enforcement structure. While these are all properties of an institution, the mechanisms behind crowding out must be identified at the individual level. Microeconomic models that assume agents to be exclusively motivated by material interests are undoubtedly very useful, but they are not necessarily capable of describing behavior of the average person, who is concerned about her identity, embedded in social structures, and equipped with a moral compass. Even worse, “policies designed for self–interested citizens may undermine the moral sentiments”, as Samuel Bowles (2008) has pointed out. Therefore, formal models of moral motivation help us to understand the interactions between extrinsic and intrinsic motivations, while taking into account the corresponding feedbacks between the individual and the institution.

The fact that voluntary contributions to public goods are so omnipresent suggests that individuals derive some benefit from it. This raises the question whether people care about the public good itself or whether they enjoy the act of giving. The first is sometimes referred to as “pure” or output–oriented altruism, while the second is referred to as “impure” or action–oriented altruism (Francois and Vlassopoulos, 2008). This form of altruism is “impure” because it is not the result that makes people happy, but the act of giving itself (see also the discussion of “warm glow” in the chapter by Nunes and Onofri in Part 3 of this volume). Whether this makes the deed less altruistic is part of a lively ongoing debate⁵, and may explain why some people find the term “altruistic” misleading and prefer to use terms like pro–social or other–regarding behavior.

Behavior is always the result of preferences and beliefs, embedded in certain institutions (Bowles, 2003). One way to account for pro–social behavior is to assume that agents have “social preferences”, such as inequity aversion or care about the payoff of other people in general (Fehr and Fischbacher, 2002). Some authors have criticized that explaining social behavior with social preferences is a tautology (see for example Baland and Platteau 1996, Ch 6). This is certainly a valid concern for any model with a limited strategy space and does of course also apply to models of moral constraints – if the

⁵ This is nicely illustrated in the American sitcom “Friends”, in the following conversation between Joey and Phoebe. Joey: Look, there's no unselfish good deeds, sorry. Phoebe: Yes there are! There are totally good deeds that are selfless. Joey: Well, may I ask for one example? Phoebe: Yeah, it's... Y'know there's...no you may not! (Friends, Season 5, Episode 4).

researcher imposes them, it is no surprise that model outcomes reflect “moral behavior”. One could overcome this problem by developing very flexible models that allow for a whole array of strategies, such as pro- or anti-social preferences. Nyborg and Rege (2003) analyze how different models of moral motivation, based on altruism, social norms, fairness considerations and conditional cooperation can explain crowding out. In the literature several mechanisms have been suggested that give rise to crowding out (cf. Bowles, 2008). First, a loss of self-determination triggers some loss of motivation. We hypothesize that this is linked to the fact that humans may undertake voluntary action to signal their pro-social stance and trustworthiness. Second, incentives convey information which changes the beliefs, and hence, the choices of an agent. And finally, incentives change the context frame of a decision or trigger a complete preference change. Let us discuss each of them in more detail.

5.1. Crowding out and costly signals

Costly signaling theory suggests that behavior which seemingly fails the cost-benefit test occurs because such behavior conveys reliable information from the sender to the receiver. Contributing to the ‘common good’ can improve one’s reputation in the community (or one’s social status), which may yield future benefits. In this view acting pro-socially is an investment in one’s reputation or self-image, which can be profitable because it gives cooperative individuals the possibility to identify each other in social interactions, thus avoiding being exploited by non-cooperators. Trust plays a crucial role in economic exchange (Fehr, 2009), and one’s contributions to public goods can be interpreted as a signal that the person is likely to be trustworthy in bilateral exchange situations too. Dynamic models have been developed that show that investing in one’s reputation or self-image can enhance one’s long-term payoff and hence contributing to the common good can be rational after all (Brandt and Sigmund, 2005; Gintis et al., 2001; Nowak and Sigmund, 1998).

Costly signals have been attributed to crowding out before, though often implicitly as impaired expression possibility (Frey, 1997), but also explicitly (Posner, 2000a, 2000b; Smith and Bird, 2005). When this signal gets blurred one may as well stop investing in it. This can be illustrated with a simple example. Consider a population that

consists of three types of people: pure altruists, “strategic altruists” and selfish individuals. In society, being an altruist is perceived to be a good thing, and it leads to a good reputation or high social status (at least within the group of the altruistically-minded individuals). The pure altruists care only about the result, i.e. the public good, but not about the social consequences, while selfish individuals are purely financially motivated. Strategic altruists contribute to the public good when it leads to higher social status and reputation. When no material incentives are attached to the provision of a public good, the selfish would free ride, while the pure and strategic altruists would contribute. When a material incentive is introduced, the selfish increase their contributions. This implies that altruists can no longer be distinguished from selfish individuals. This could be one reason for the strategic altruists, who are concerned about their reputation, to stop signaling their good intentions. The same would occur if, more realistically, individual preferences are determined by material interest, altruistic motivations and reputational or self-image concerns. Bénabou and Tirole (2006) have shown formally how such a model can explain several aspects of crowding out. These theoretical predictions have been confirmed not only in laboratory settings, but also in a field experiment: a financial incentive increases the willingness to contribute to a good cause in private (i.e. when nobody is watching), while it actually decreases contributions in public when it is an observable signal (Ariely et al., 2009).

In reality, the benefits from a good self-image are not constant, but more realistically depend on the composition of the population. Janssen and Mendys-Kamphorst (2004) modeled a situation where individuals are either altruists or egoists and choose whether to contribute to a public good. Social rewards depend positively on the number of contributing cooperators and negatively on the number of selfish people. A financial incentive for contributing induces more selfish people to contribute and hence lowers the social reward for altruists. As a result, altruists may cease contributing and aggregate provision may decrease.

5.2. Crowding out and beliefs

The information content of an incentive is related to beliefs, which have often been attributed to the crowding out phenomenon. This is especially the case in traditional

principal–agent settings, where contracts are usually incomplete, and an extrinsic reward or control could change the perceived nature of the task. One reason for this to occur is that the reward reveals that the task requires much more effort or is much less fun than previously thought (Bénabou and Tirole, 2003). Therefore, beliefs play a big role in explaining crowding out in bilateral interactions, and it is possible to extend this line of reasoning to the delivery of public goods. In most societies there is much debate on the scope of civic duties and therefore people may update regularly what they are expected to do as committed citizens. If one receives payment for donating blood one may infer that it is something one is not expected to do by default. Scientists, politicians or celebrities may accept to give a talk without any compensation, but after having received a honorarium a couple of times, they may think twice whether or not to give a free talk. Brekke et al. (2003) have developed a model in which utility depends on leisure, the consumption of a private good, the consumption of a public good, and a self–image as a socially responsible person (given by how actions deviate from some socially desired effort level). A sufficient extrinsic incentive will make individuals feel that they are no longer morally obliged to contribute, and hence contributions may go down if the unit value of leisure is higher than the unit value of private consumption.

In certain cases, beliefs are linked to the pro–social signals described in section 5.1. If people are conditionally cooperative, not trusting reveals information about the expected share of selfish individuals in the population, making conditional cooperators not cooperate (Sliwka, 2007). In a similar vein, Ellingsen and Johannesson (2008) show how crowding out can easily occur in bilateral interactions, where players are either altruistic or selfish. Utility depends on material payoffs and on the warm glow from giving, but also on how actions are perceived by the others. Actions depend on one’s own preferences, on one’s beliefs about the preferences of the persons one is matched with, and on their observed prior actions. A key mechanism is that one obtains a higher utility by being nice to a good person (i.e. an altruist), and hence signaling to be an altruist may pay back.

5.3. Crowding out and context–dependent preferences

In many situations behavior is observed to be context–dependent and takes place in a “decision frame” that “is controlled partly by the formulation of the problem and partly by the norms, habits, and personal characteristics of the decision maker” (Tversky and Kahneman, 1981). Evidence from social experiments and the field suggests that many social preferences are conditional (Cox et al., 2008; Fischbacher et al., 2001). For this reason many people play fair only as long as the opponent reciprocates. More generally, preferences depend on the situation the agent faces and also on the process that has led to the situation (Bowles, 2003). Situation–dependent preferences are not unique for social preferences, as choices are always the result of given preferences in a certain environment. Process–regarding preferences, however, are special in the sense that they do not depend only on the outcome, but also on the chain of events that led to this outcome (Ben-Ner and Putterman, 1998). People may be reluctant to help someone who took some foolish decisions that brought him into trouble while they do help someone who was just extremely unlucky or unfortunate. Thus, one may conclude that preferences are higher–dimensional.

While many studies draw attention to the multi–dimensional nature of preferences verbally, they are hardly used in formal analyses as the results may be rather complex. Assuming preferences to be higher–dimensional implies that the corresponding equilibria are higher–dimensional as well. When an individual stops cooperating, we may be inclined to detect a preference change. This may be the correct inference in some cases but not in all. An alternative explanation is that preferences are stable but the environmental context has changed, thus resulting in the agent changing his/her behavior. That means that people may decrease their contributions to the public good in response to the introduction of formal government intervention if they interpret the regulation as reflecting a lack of trust (Fehr and Fischbacher, 2005).

Laboratory experiments are useful for unraveling many of these situations. One example is the fact that players make different choices when they face a human opponent or a computer. Another frequent observation is that cooperative individuals cease cooperating after having been exploited by defectors. What looks like a true preference

change may just be a change in behavior given that a subject finds herself in a less cooperative environment than expected or believed.

This implies that it is very difficult to identify a true preference change. This is illustrated in Figure 2, which plots an individual's contribution to the public good on the horizontal axis against the sum of contributions of his/her fellow group members on the vertical axis. In this stylized example we assume that the decision maker's contribution to the public goods is conditional on the number of other people contributing – she is assumed to be conditionally cooperative.

<Insert figure 2 about here>

Suppose that an experimenter observes a high level of cooperation by both the individual and by the rest of the group (say point H) in the early stage of a game; see the left-hand side panel of Figure 2. During the game, the sum of contributions by the rest of the group (on the vertical axis) decreases and the individual agent reduces his/her cooperation too, resulting in a move to point L. The experimenter thus observes the shift from point H to point L in the left-hand panel, but she cannot determine whether this change in behavior is only a response to a change in the environment, or whether the individual is so frustrated that his willingness to contribute in general has changed; see the right-hand side panel in Figure 2. The first situation is represented by a shift along the contribution schedule $S1$, where one could speak of a reversible change. If the number of cooperators returns to the initial high level, individual contribution will also be high again. An irreversible change would occur if the contribution schedule shifts to the left, from $S2$ to $S2'$. This implies that a return to the initial level of cooperation in the group will not be sufficient to restore the initial level of contribution, because the new level will be at H' . Note that observing a drop in cooperation followed by lower individual contribution, is not sufficient to recognize which situation applies, because one does not know the slope of the contribution schedule. To be able to do so, one would need sufficiently detailed data for single individuals, and that is typically not available. In real-world behavior it is even more difficult to distinguish temporary change from a long-lasting “once bitten, twice shy” effect.

Another way of capturing the context dependency of decisions is to assume that individuals hold multiple preferences that are weighted according to the situation. A good example is the model by Nyborg (2000), where individuals hold a preference as consumers, but also as good citizens. Individuals apply different preferences in different contexts. Therefore, choices depend highly on the situation. It is a priori unclear how consumers weigh these different preferences. In any case, an external intervention may lead to a shifting of weights. This is especially relevant when intervention influences the perception of a social dilemma. If an individual faces several co-players that are “in the same boat” she may be more inclined to infer some form of Kantian moral imperative: “if everyone cooperates, we will all be better off.” If, however, a government regulator is imposing some law, the agent probably views the game very differently. The stage is now completely set, so her actions will most likely neither affect her future payoff, nor the rules set by the regulator. Therefore she would probably try to get the best out of the situation, given the presence of the regulator.

6. Repercussions on institutional constraints

Let us now try to map these microeconomic fundamentals to the experimental results in order to understand the implications for institutional design – focusing on how these microfoundations interact with the institution’s legitimacy, voluntariness, and enforcement structure.

Legitimacy is strongly linked to context-dependency, but also to the endogeneity of preferences. When the institution lacks legitimacy individuals may infer that being cooperative does not pay off, or even lose all confidence in cooperative behavior altogether. Note that it matters whether the institution is perceived to be legitimate, not whether it is legitimate by objective standards. Raakjær Nielsen and Mathiesen (2003) present a case from Danish fisheries, where rule-compliance went down after the fishermen felt decoupled from the decision-making process. Fairness issues should be taken seriously as they may not only influence the distribution of rents, but also the success of the institution, and therefore, its efficiency. In South Africa the government intended to reduce illegal fish landings by establishing *de jure* rights for the local

fishermen. Some fishermen had the feeling that the process was not very fair and expressed their discontent by “protest fishing” (Hauck, 2008).

The voluntary nature of an institution is important for three reasons. First, it reveals important information about the intentions of other individuals. Second, it gives individuals the possibility to signal their social attitude or build up a good image or reputation. These can only work when individuals have actually the choice to do so, otherwise it would be impossible to distinguish true signals or images from forced ones. Because of this loss of information, individuals may as well cease investing in a social image at all. The third reason why a voluntary institution may perform better than a compulsory one, has been pointed out by Hauert et al. (2007). When individuals have the possibility to withdraw from some joint activity, it is much harder for others to take advantage of them. Therefore, self-determination provides an escape route to being exploited by defectors. An example where a new law crowded out existing norms of reciprocity is presented by Borges and Irlenbusch (2007). When the German government introduced a law that made it possible to return any product just bought (for example via the internet), the number of products returned upon purchase increased sharply. This happened in spite of the fact that most sellers offered the same refund possibility even before the introduction of the law.

Concerning the importance of the enforcement structure, it is difficult to pinpoint the exact micro-foundations as many mechanisms are at work, making the interaction between incentives and enforcement a complex one. Chhatre and Agrawal (2008) analyzed 152 forests that were common property in 9 countries. They found, as expected, that the more valuable forests were depleted faster when local enforcement is absent. Interestingly, in the presence of enforcement the opposite holds, as regeneration is higher when forests are more valuable.

It seems to be important to distinguish enforcement mechanisms that are centralized from decentralized ones. Most experiments focus on the role of punishments and rewards in decentralized peer-to-peer enforcement. While most studies show that peer-to-peer enforcement is very effective in inducing cooperation, its impact on welfare is more ambiguous (Egas and Riedl, 2008). Dreber (2008) found that costly punishment raises cooperation, but the cost may be so high that the community as a whole may be worse off

and therefore “winners don’t punish”. This finding has been challenged by Gächter et al. (2008) who found that punishment pays off, but only in the long run. Nikiforakis (2008) finds that adding a second stage of peer-enforcement may reduce all efficiency gains. This happens because the possibility of counter-punishments induces most players not to punish in the first place, because they fear retaliation. The same breakdown can be observed when rewards are used as enforcement mechanism, albeit for different reasons. A second stage of rewarding gives defectors the chance to build a profitable rewarding network (Stoop et al., 2008).

Hence, both the punishment and reward mechanisms have their drawbacks, as the desire to punish non-cooperators can be very resource consuming, while rewarding or reputation does not always help sanctioning the bad guys. Rockenbach and Milinski (2006) have looked at this issue in detail, and concluded that it is the combination of punishment and rewarding based on reputation performs best in an experimental setting. Ohtsuki et al. (2009) have analyzed a model where two players of the population randomly meet and can either cooperate, defect, or punish. Individuals have a good reputation, or a bad one. There is, however, the chance that someone mistakenly identifies a partner as good when he is bad or the other way round. They find that when the probability to correctly identify someone is high, defecting with bad guys is the best strategy, while one should cooperate with the good ones. When the probability of correctly identifying someone is low, always defecting is superior. When the probability is in a very narrow parameters space between these two cases, punishing defectors and cooperating with cooperators is the winning strategy. This may indicate that increasing information (either through monitoring or gossip) makes punishment redundant, provided that one can actually refuse to interact with bad people. If one cannot, as typically the case in a common pool problem, the answer is less straightforward. In such a case, an appropriate weapon against defectors is needed, be it either punishment or ostracism. This theoretical finding is confirmed by Rustagi et al. (2010) who combine laboratory experiments and field data to reveal the key factors that determine success in forest management in Ethiopia. The authors identify the interplay of norm-compliance, costly monitoring and a sufficiently large share of cooperatively minded individuals to explain

whether forests are successfully preserved or not – hinting at a complex interaction between social preferences, institutions, and the renewable resource at stake.

7. Recommendations for policy design

The previous sections have shown that crowding out actually occurs in the real world, and they also give hints regarding the way in which the phenomenon can be prevented from occurring. The most important one are that the external intervention must be such that (i) it does not reduce people's sense of self-determination, (ii) it is perceived as legitimate and fair, and (iii) it supportive social norms in place. In many instances the most effective way to meet these three criteria is to allow for direct stakeholder participation. That means that various fairly simple recommendations to prevent crowding out from occurring.

First, communication and monitoring are important mechanisms for self-regulation, and often gossip suffices as an enforcement instrument. If this is not the case and more drastic punishments are needed, the welfare effects are more ambiguous. Therefore, lack of communication or monitoring possibilities can form an obstacle to the evolution of effective social norms. In such a case, formal institutions may be more efficient. When a central authority steps in, lack of involvement, but also fairness considerations can be an important reason why individuals actively try to undermine the institution. Including individuals in the rule designing process is often a solution, but may sometimes have drawbacks. This is especially the case when some individuals perceive the institution to be unfair; either no consensus is reached or individuals who opposed the rule feel not committed to follow it. Sometimes the social norms in place are maladaptive, because the environment or the technology has changed, while the social norms have not (Posner, 1996). In such a case, government intervention may be necessary, but should be done very carefully. Any attempt to manage or regulate social norms may backfire, because the authority is perceived to be part of the “game” (Posner, 1998). In the same vein, any attempt made by a central authority to strengthen existing social norms or “invest in social capital” may be well-intended, but may have unpredictable consequences.

Financial incentives are problematic because individuals take into account that this will affect their self-image. Incentives in the form of public goods could be a solution, because individuals can then signal that they have a pro-social attitude. Indeed, provision of compensation in the form of publicly provided goods are more effective in increasing support for “not in my backyard” projects, such as a noisy road, than money (Mansfield et al., 2002).

Once the government decides to impose an external incentive, it is important to identify the nature of the product/task that is targeted. Subsidizing a good that has signaling character, like a hybrid car, may be counter-effective as pointed out by Ariely et al. (2009). If a small symbolic tax tries to underline the fact that a certain behavior is unacceptable, this may work especially well with goods that have signaling character. For example, the use of plastic bags in Ireland dropped by 94% after the introduction of a small tax on their purchase because using them became highly stigmatized (Bowles, 2008).

In general, the regulated agents’ perception of the institution is crucial. In the famous child care study by Gneezy and Rustichini (2000a), a small fine was perceived as a price that parents were more than happy to pay. Therefore, the conventional economic wisdom of “getting the incentives right” only works as intended if governments send an unambiguous message. In that sense, tradable emission rights, for instance, may be especially susceptible to crowding out, because the owner holds the “right” to do something, taking away the negative connotation from polluting (Frey and Stutzer, 2006). This seems to be relevant as well in the recent discussion on the effectiveness of individual transferable quotas (ITQ) in fisheries (Costello et al., 2008). When social issues are important, as is typically the case in small-scale fisheries, ITQs may undermine local stewardship (Ban et al., 2009). In the worst case, ITQ regimes crowd out ecosystem responsibility of the ITQ holders (who hold the “right” to fish), while leading to “protest fishing” among the non-holders.

8. Conclusions

In this chapter we have presented evidence that we can neither rely on external regulation, nor on voluntary actions alone to solve many environmental problems –

especially so when global environmental problems are at stake. Absent rigorous and effective international cooperation governments are likely to undertake unilateral environmental policies that are insufficiently stringent from a global point of view. Introduction of a lenient policy may be countereffective because the increase in “public good provision” by purely extrinsically motivated citizens or agents may not compensate the decrease in voluntary provisions by the citizens or agents who are intrinsically motivated to care for the global commons.

The first-best solution would be that governments take their responsibilities and start cooperating at the international level to address the global commons problems we are currently confronted with. Increasing public goods provision to the socially optimal level requires a substantial increase in efforts of all agents in the economy. A substantial increase in stringency of national policies will most likely not prove to be countereffective. First, intrinsically motivated agents would welcome a strong worldwide shift towards environmental policy that seriously attempts to save the global commons. Second, extrinsically motivated agents are confronted with such a substantial increase in policy stringency that it is in their best interest to obey the law and take the necessary actions. Absent effective international cooperation, however, the probability of crowding out intrinsic motivations is non-negligible. This may especially happen if the environmental costs are not shared fairly among citizens or countries, or regulations are so weak that governments fail to convince citizens that they are willing to craft effective regulations. In this case, intrinsically-motivated citizens may lose their motivation for voluntary action, while the incentives for extrinsically motivated agents are not strong enough.

While there is not one standard recipe for solving diverse social dilemmas, the overview provided by this paper identifies several regularities that determine success or failure of policy design. These regularities can help synthesizing important core principles for designing effective governance structures. While it is unlikely that global problems will be solved without active government actions, any governmental actions will be much more effective if it these are linked to lower institutional scales.⁶ Therefore,

⁶ see the chapter by Ostrom in part II of this volume.

it is imperative that governments are aware that any implemented law will affect the complex motivational structure of its citizens.

Experimental and theoretical work has shown that decentralized arrangements, based on voluntary action, communication, peer control, and reputation can be very effective, but also highly fragile. As a warning, it must be emphasized that imposing external interventions may not strengthen these arrangements, but replace them, leading to crowding out and a loss of social capital. Even when a certain regulation is established to formalize a certain right that *de facto* already exists, unexpected outcomes may materialize. In any case, governments need to be aware that institutional changes imposed may lead to unintended consequences that will be very difficult to reverse.

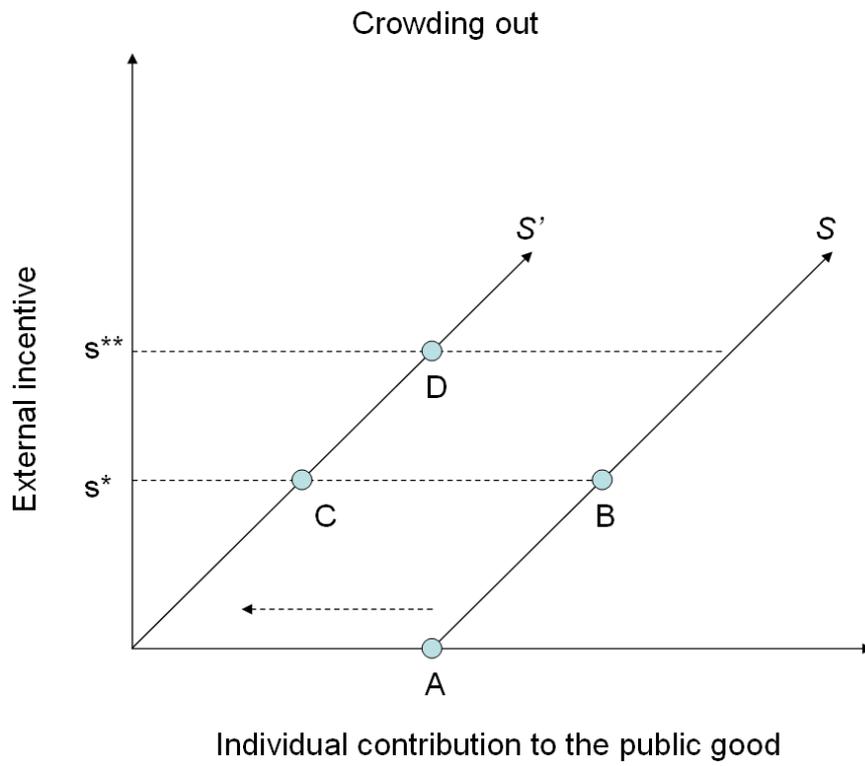


Figure 1. Crowding out occurs if the supply curve of an individual shifts to the left and the individual contributes less to the public good for a given external incentive.

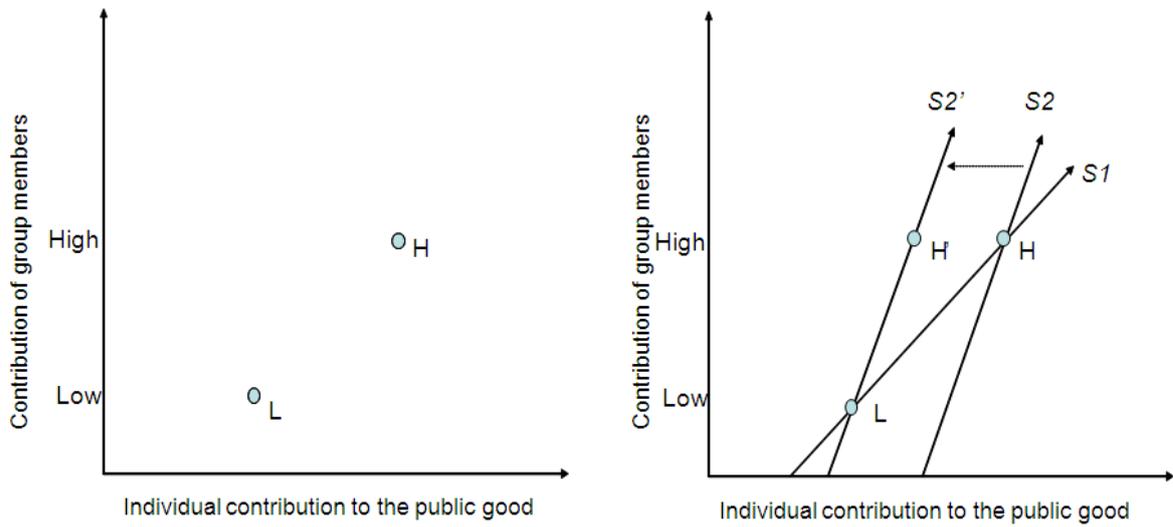


Figure 2 Reversible and irreversible changes in behavior. The left panel shows the observed contribution to a public good. The right panel illustrates that this information is not sufficient to distinguish a shift in the supply curve from a contribution change along a stable supply curve.

References

- Ackerman, F., 1997, *Why Do We Recycle: Markets, Values, and Public Policy*. Island Press, Washington D.C.
- Anderson, L.G. and Lee, D.R., 1986. Optimal Governing Instrument, Operation Level, and Enforcement in Natural Resource Regulation: The Case of the Fishery. *American Journal of Agricultural Economics*, 68(3), 678-690.
- Ariely, D., Bracha, A. and Meier, S., 2009. Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially. *American Economic Review*, 99(1), 544-555.
- Ariely, D., Kamenica, E., Prelec, D. 2008. Man's search for meaning: The case of Legos. *Journal of Economic Behavior & Organization*, 67(3-4), 671-677.
- Ayala, F. J., 2010. The difference of being human: Morality. *Proceedings of the National Academy of Sciences*, 107, 9015-9022.
- Baland, J.M. and Platteau, J.P., 1996, *Halting degradation of natural resources*. Clarendon Press for FAO, Oxford.
- Ban, N.C., Caldwell, I.R., Green, T.L., Morgan, S.K., O'Donnell, K., Selgrath, C., J., Lynham, J., Costello, C., Gaines, S.D., Grafton, R.Q. and Prince, J., 2009. Diverse Fisheries Require Diverse Solutions. *Science*, 323(5912), 338-339.
- Barrett, S., 2007. *Why Cooperate? The Incentive to Supply Global Public Goods*. Oxford: Oxford University Press.
- Ben-Ner, A. and Putterman, L.G., 1998, *Economics, Values, and Organization*. Cambridge University Press, Cambridge.
- Bénabou, R. and Tirole, J., 2003. Intrinsic and Extrinsic Motivation. *The Review of Economic Studies*, 70(3), 489-520.
- Bénabou, R. and Tirole, J., 2006. Incentives and Prosocial Behavior. *The American Economic Review*, 96(5), 1652-1678.
- Bergland, H., Clark, D.J. and Pedersen, P.A., 2002. Rent-seeking and quota regulation of a renewable resource. *Resource and Energy Economics*, 24(3), 263-279.

- Berglund, C., 2006. The assessment of households' recycling costs: The role of personal motives. *Ecological Economics*, 56(4), 560-569.
- Borges, G. and Irlenbusch, B., 2007. Fairness Crowded Out by Law: An Experimental Study on Withdrawal Rights. *Journal of Institutional and Theoretical Economics*, 163, 84-101.
- Bouma, J., Bulte, E. and van Soest, D., 2008. Trust and cooperation: Social capital and community resource management. *Journal of Environmental Economics and Management*, 56(2), 155-166.
- Bowles, S., 2003, *Microeconomics: Behavior, Institutions and Evolution*. Princeton University press, Princeton.
- Bowles, S., 2008. Policies Designed for Self-Interested Citizens May Undermine "The Moral Sentiments": Evidence from Economic Experiments. *Science*, 320(5883), 1605-1609.
- Brandt, H. and Sigmund, K., 2005. Indirect reciprocity, image scoring, and moral hazard. *Proceedings of the National Academy of Sciences of the United States of America*, 102(7), 2666-2670.
- Brekke, K.A., Kverndokk, S. and Nyborg, K., 2003. An economic model of moral motivation. *Journal of Public Economics*, 87(9-10), 1967-1983.
- Cameron, J. and Pierce, W.D., 1994. Reinforcement, Reward, and Intrinsic Motivation: A Meta-Analysis. *Review of Educational Research*, 64(3), 363.
- Cárdenas, J.C., Stranlund, J. and Willis, C., 2000. Local Environmental Control and Institutional Crowding-Out. *World Development*, 28(10), 1719-1733.
- Chhatre, A. and Agrawal, A., 2008. Forest commons and local enforcement. *Proceedings of the National Academy of Sciences*, 105(36), 13286-13291.
- Cordell, J.C. and McKean, M.A., 1992, *Sea Tenure in Bahia, Brazil*. ICS Press, San Francisco, USA.
- Costello, C., Gaines, S.D. and Lynham, J., 2008. Can Catch Shares Prevent Fisheries Collapse? *Science*, 321(5896), 1678-1681.
- Cox, J., Sadiraj, K. and Sadiraj, V., 2008. Implications of trust, fear, and reciprocity for modeling economic behavior. *Experimental Economics*, 11(1), 1-24.
- Cox, J.C., 2004. How to identify trust and reciprocity. *Games and Economic Behavior*, 46(2), 260-281.

- Dankel, D.J., 2009, Building Blocks of Sustainability in Marine Fisheries Management: Stakeholders, objectives, and strategies, PhD thesis. University of Bergen, Norway.
- Deci, E.L., 1971. Effects of Externally Mediated Rewards on Intrinsic Motivation. *Journal of Personality and Social Psychology*, 18(1), 105-115.
- Deci, E.L., Koestner, R. and Ryan, R.M., 1999. A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation. *Psychological Bulletin*, 125, 627-668.
- Deci, E.L., Koestner, R. and Ryan, R.M., 2001. Extrinsic Rewards and Intrinsic Motivation in Education: Reconsidered Once Again. *Review of Educational Research*, 71(1), 1.
- Di Falco, S. and van Rensburg, T.M., 2008. Making the Commons Work: Conservation and Cooperation in Ireland. *Land Economics*, 84(4), 620-634.
- Dreber, A., Rand, D.G., Fudenberg, D. and Nowak, M.A., 2008. Winners don't punish. *Nature*, 452(7185), 348-351.
- Egas, M. and Riedl, A., 2008. The economics of altruistic punishment and the maintenance of cooperation. *Proceedings of the Royal Society B: Biological Sciences*, 275(1637), 871-878.
- Ellingsen, T. and Johannesson, M., 2008. Pride and Prejudice: The Human Side of Incentive Theory. *The American Economic Review*, 98, 990-1008.
- Fehr, E., 2009. On the Economic and Biology of Trust. Institute for Empirical Research in Economic, University of Zurich, Working Paper Series, No. 399.
- Fehr, E. and Fischbacher, U., 2002. Why social preferences matter - The impact of non-selfish motives on competition, cooperation and incentives. *Economic Journal*, 112(478), C1-C33.
- Fehr, E. and Fischbacher, U., 2005, The economics of strong reciprocity. In: H. Gintis, S. Bowles, R. Boyd and E. Fehr (Eds.), *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life*. MIT Press, Cambridge, Massachusetts. pp. 151–191.
- Fehr, E. and Gächter, S., 2000. Cooperation and punishment in public goods experiments. *American Economic Review*, 90(4), 980-994.
- Fischbacher, U., Gächter, S. and Fehr, E., 2001. Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, 71(3), 397-404.

- Finus, M., 2008. Game Theoretic Research on the Design of International Environmental Agreements: Insights, Critical Remarks, and Future Challenges. *International Review of Environmental and Resource Economics*, 2(1), 29-67.
- Francois, P. and Vlassopoulos, M., 2008. Pro-social Motivation and the Delivery of Social Services. *CESifo Economic Studies*, ifn002.
- Frey, B.S., 1997, *Not Just for the Money: An Economic Theory of Personal Motivation*. Edward Elgar Publishing, Cheltenham, UK and Brookfield, USA.
- 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 Motivation Crowding- Out. *The American Economic Review*, 87(4), 746-755.
- Frey, B.S., Oberholzer-Gee, F. and Eichenberger, R., 1996. The Old Lady Visits Your Backyard: A Tale of Morals and Markets. *The Journal of Political Economy*, 104(6), 1297-1313.
- Frey, B.S. and Stutzer, A., 2006, *Environmental Morale and Motivation*, Institute for Empirical Research in Economics, University of Zurich, Working Paper Series ISSN 1424-0459. Working Paper No. 288.
- Gächter, S., Renner, E. and Sefton, M., 2008. The Long-Run Benefits of Punishment. *Science*, 322(5907), 1510.
- Gintis, H., Bowles, S., Boyd, R. and Fehr, E. (Eds.), 2005. Moral sentiments and material interests: Origins, evidence, and consequences. *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life*. MIT Press, Cambridge, Massachusetts.
- Gintis, H., Smith, E.A. and Bowles, S., 2001. Costly signaling and cooperation. *Journal of Theoretical Biology*, 213(1), 103-119.
- Gneezy, U. and Rustichini, A., 2000a. A Fine Is a Price. *The Journal of Legal Studies*, 29(1), 1-17.
- Gneezy, U. and Rustichini, A., 2000b. Pay Enough or Don't Pay at All. *The Quarterly Journal of Economics*, 115(3), 791-810.
- Hatchard, J., 2005, *Engaging Stakeholder Preferences Through Deliberative Democracy in North Sea Fisheries Governance*. In: Tim S. Gray (Ed.), *Participation in Fisheries Governance*. Springer, Dordrecht, the Netherlands, pp. 45-64.

- Hatcher, A., Jaffry, H., Thébaud, O.B. and Bennett, E., 2000. Normative and Social Influences Affecting Compliance with Fisheries Regulations. *Land Economics*, 76(3), 448–461.
- Hauck, M., 2008. Rethinking small-scale fisheries compliance. *Marine Policy*, 32(4), 635-642.
- Hauert, C., Traulsen, A., Brandt, H., Nowak, M.A. and Sigmund, K., 2007. Via Freedom to Coercion: The Emergence of Costly Punishment. *Science*, 316(5833), 1905.
- Janssen, M.C.W. and Mendys-Kamphorst, E., 2004. The price of a price: on the crowding out and in of social norms. *Journal of Economic Behavior & Organization*, 55(3), 377-395.
- Jentoft, S., McCay, B.J. and Wilson, D.C., 1998. Social theory and fisheries co-management. *Marine Policy*, 22(4-5), 423-436.
- Johnson, R. and Libecap, G., 1982. Contracting problems and regulation: the case of the fishery. *The American Economic Review*, 1005-1022.
- Khanna, M., 2001. Non mandatory approaches to environmental protection. *Journal of Economic Surveys*, 15(3), 291-324.
- Kosfeld, M., Okada, A. and Riedl, A., 2009. Institution Formation in Public Goods Games. *American Economic Review*, 99(4), 1335-1355.
- Kreps, D.M., 1997. Intrinsic Motivation and Extrinsic Incentives. *The American Economic Review*, 87(2), 359-364.
- Mansfield, C., Van Houtven, G.L. and Huber, J., 2002. Compensating for Public Harms: Why Public Goods Are Preferred to Money. *Land Economics*, 78(3), 368.
- Mikalsen, K.H. and Jentoft, S., 2008. Participatory practices in fisheries across Europe: Making stakeholders more responsible. *Marine Policy*, 32(2), 169-177.
- Milinski, M., Semmann, D., Krambeck, H.-J. and Marotzke, J., 2006. Stabilizing the Earth's climate is not a losing game: Supporting evidence from public goods experiments. *Proceedings of the National Academy of Sciences of the United States of America*, 103(11), 3994-3998.
- Milinski, M., Sommerfeld, R. D., Krambeck, H.-J. r., Reed, F. A. and Marotzke, J., 2008. The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proceedings of the National Academy of Sciences*, 105(7), 2291-2294.

- Murdock, K., 2002. Intrinsic Motivation and Optimal Incentive Contracts. *The RAND Journal of Economics*, 33(4), 650-671.
- Nikiforakis, N., 2008. Punishment and Counter-Punishment in Public Good Games: Can We Really Govern Ourselves? *Journal of Public Economics*, 92(1-2), 91-112.
- Nordhaus, W.D. and Yang, Z., 1996. A regional dynamic general equilibrium model of alternative climate change strategies. *American Economic Review*, 86(4), 741-765.
- Nowak, M.A. and Sigmund, K., 1998. Evolution of indirect reciprocity by image scoring. *Nature*, 393(6685), 573-577.
- Nyborg, K., 2000. Homo Economicus and Homo Politicus: interpretation and aggregation of environmental values. *Journal of Economic Behavior & Organization*, 42(3), 305-322.
- Nyborg, K. and Rege, M., 2003. Does Public Policy Crowd Out Private Contributions to Public Goods. *Public Choice*, 115(3), 397-418.
- Ohtsuki, H., Iwasa, Y. and Nowak, M.A., 2009. Indirect reciprocity provides only a narrow margin of efficiency for costly punishment. *Nature*, 457(7225), 79-82.
- Ostrom, E., 2000. Crowding out Citizenship. *Scandinavian Political Studies*, 23, 3-16.
- Ostrom, E., 2003. How Types of Goods and Property Rights Jointly Affect Collective Action. *Journal of Theoretical Politics* 15, 239-270.
- Ostrom, E., 2005, Policies that crowd out reciprocity and collective action. In: H. Gintis, S. Bowles, R. Boyd and E. Fehr (Eds.), *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life*. MIT Press, Cambridge, Massachusetts. pp. 253-275.
- Ostrom, E., Walker, J. and Gardner, R., 1992. Covenants with and without a Sword - Self-Governance Is Possible. *American Political Science Review*, 86(2), 404-417.
- Posner, E.A., 1996. Law, Economics, and Inefficient Norms. *University of Pennsylvania Law Review*, 144(5s), 1697-1744.
- Posner, E.A., 1998. Symbols, Signals, and Social Norms in Politics and the Law. *The Journal of Legal Studies*, 27(s2), 765-797.
- Posner, E.A., 2000a, *Law and social norms*. Harvard University Press.
- Posner, E.A., 2000b. Law and social norms: The case of tax compliance. *Virginia Law Review*, 86, 1781-1819.

- Raakjær Nielsen, J. and Mathiesen, C., 2003. Important factors influencing rule compliance in fisheries lessons from Denmark. *Marine Policy*, 27(5), 409-416.
- Reeson, A.F. and Tisdell, J.G., 2008. Institutions, motivations and public goods: An experimental test of motivational crowding. *Journal of Economic Behavior & Organization*, 68(1), 273-281.
- Rockenbach, B. and Milinski, M., 2006. The efficient interaction of indirect reciprocity and costly punishment. *Nature*, 444(7120), 718-723.
- Rodriguez-Sickert, C., Guzmán, R.A. and Cárdenas, J.C., 2008. Institutions influence preferences: Evidence from a common pool resource experiment. *Journal of Economic Behavior & Organization*, 67(1), 215-227.
- Rustagi, D., Engel, S., Kosfeld, M., 2010. Conditional Cooperation and Costly Monitoring Explain Success in Forest Commons Management. *Science* 330, 961-965.
- Sliwka, D., 2007. Trust as a signal of a social norm and the hidden costs of incentive schemes. *The American economic review*, 97(3), 999-1012.
- Smith, E.A. and Bird, R.B., 2005, Costly signaling and cooperative behavior. In: H. Gintis, S. Bowles, R. Boyd and E. Fehr (Eds.), *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life*. MIT Press, Cambridge, Massachusetts, pp. 115-148.
- Somanathan, E., 1991. Deforestation, Property Rights and Incentives in Central Himalaya. *Economic and Political Weekly*, 26.
- Stoop, J., van Soest, D. and Vyrastekova, J., 2008, Two Carrots and No Stick: The Effectiveness of Second-Order Rewarding in Social Dilemma Situations, Tilburg university, mimeo.
- Sutinen, J.G. and Kuperan, K., 1999. A socio-economic theory of regulatory compliance. *International Journal of Social Economics*, 26(1/2/3), 174-193.
- Sutter, M. and Weck-Hannemann, H., 2004. An Experimental Test of the Public Goods Crowding Out Hypothesis when Taxation Is Endogenous. *Finanz Archiv*, 60, 94-110.
- Tol, R.S.J., 2008. The Social Cost of Carbon: Trends, Outliers and Catastrophes. *Economics: The Open-Access, Open-Assessment E-Journal*, Vol. 2, 2008-25.
- Tversky, A. and Kahneman, D., 1981. The framing of decisions and the psychology of choice. *Science*, 211(4481), 453-458.

- Tyran, J.-R. and Feld, L.P., 2006. Achieving Compliance when Legal Sanctions are Non-deterrent. *The Scandinavian Journal of Economics*, 108(1), 135-156.
- van Soest, D.P., de Zeeuw, A. and Bouma, J., 2010. Voluntary cooperation and local public goods provision: Coalition formation and the role of external labor markets. VU University Amsterdam, mimeo.
- van Soest, D. and Vyrastekova, J., 2008, Higher fines, lower conviction probabilities, and the support for government regulation, Tilburg University, mimeo.
- Vatn, A., 2007. Resource regimes and cooperation. *Land Use Policy*, 24(4), 624-632.
- Vollan, B., 2008. Socio-ecological explanations for crowding-out effects from economic field experiments in southern Africa. *Ecological Economics*, 67(4), 560-573.
- Vyrastekova, J. and van Soest, D., 2003. Centralized Common-Pool Management and Local Community Participation. *Land Economics*, 79(4), 500-514.
- Vyrastekova, J. and van Soest, D., 2008. On the (in)effectiveness of rewards in sustaining cooperation. *Experimental Economics*, 11(1), 53-65.