

The following resources related to this article are available online at www.sciencemag.org (this information is current as of April 1, 2009):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/cgi/content/full/320/5883/1605>

Supporting Online Material can be found at:

<http://www.sciencemag.org/cgi/content/full/320/5883/1605/DC1>

This article **cites 26 articles**, 2 of which can be accessed for free:

<http://www.sciencemag.org/cgi/content/full/320/5883/1605#otherarticles>

This article appears in the following **subject collections**:

Economics

<http://www.sciencemag.org/cgi/collection/economics>

Information about obtaining **reprints** of this article or about obtaining **permission to reproduce this article** in whole or in part can be found at:

<http://www.sciencemag.org/about/permissions.dtl>

Policies Designed for Self-Interested Citizens May Undermine “The Moral Sentiments”: Evidence from Economic Experiments

Samuel Bowles^{1,2}

High-performance organizations and economies work on the basis not only of material interests but also of Adam Smith’s “moral sentiments.” Well-designed laws and public policies can harness self-interest for the common good. However, incentives that appeal to self-interest may fail when they undermine the moral values that lead people to act altruistically or in other public-spirited ways. Behavioral experiments reviewed here suggest that economic incentives may be counterproductive when they signal that selfishness is an appropriate response; constitute a learning environment through which over time people come to adopt more self-interested motivations; compromise the individual’s sense of self-determination and thereby degrade intrinsic motivations; or convey a message of distrust, disrespect, and unfair intent. Many of these unintended effects of incentives occur because people act not only to acquire economic goods and services but also to constitute themselves as dignified, autonomous, and moral individuals. Good organizational and institutional design can channel the material interests for the achievement of social goals while also enhancing the contribution of the moral sentiments to the same ends.

David Hume (1711–1776), the Scottish philosopher and economist, cautioned legislators that constitutions and public policies should be designed for “knaves” motivated only by their “private interest” (1). Over the past century, economics, embracing Hume’s axiom, has devised ingenious ways that taxes, subsidies, tournaments, auctions, and other incentives can be structured to induce self-regarding individuals to act in the common interest when market competition alone would fail to accomplish this (2, 3). This past October, three of its leading practitioners—Leonid Hurwicz, Eric Maskin, and Roger Myerson—were awarded the Nobel Memorial Prize in Economic Sciences for their work in what is now called mechanism design. Dismissed by some as an arcane branch of applied mathematics, the field is, to the contrary, of immense practical importance for the public good: the invisible hand needs a helping hand.

But what if citizens are not knaves, or at least not all of them, all of the time? In this case, policies designed to harness self-interest to public ends may be counterproductive (4, 5). As Bruno Frey warned, a constitution for knaves may produce knaves (6). A real-life experiment (7) provides an example.

In Haifa, at six day care centers, a fine was imposed on parents who were late picking up their children at the end of the day. Parents responded

to the fine by doubling the fraction of time they arrived late. When after 12 weeks the fine was revoked, their enhanced tardiness persisted unabated. While other interpretations are possible, the counterproductive imposition of the fines illustrate a kind of negative synergy between economic incentives and moral behavior. The fine seems to have undermined the parents’ sense of ethical obligation to avoid inconveniencing the teachers and led them to think of lateness as just another commodity they could purchase.

The example points to a shortcoming in the conventional economic approach to policy design: It overlooks the possibility that economic incentives may diminish ethical or other reasons for complying with social norms and contributing to the common good. Where this is the case, the kinds of incentives stressed by economists may have counterproductive effects. (By “incentives” without adjective, I mean those appealing to self-regarding preferences.)

The critical assumption in the conventional approach is not that other-regarding motives are absent but that policies that appeal to economic self-interest do not affect the salience of ethical, altruistic, and other social preferences. According to this view, the effects of material interests and “moral sentiments” on behavior are additive rather than interactive. This is called the assumption of separability; a mathematical formulation is provided in (8).

Incentives and Market Failures

When individuals do not take into account the effects of their actions on others (called external

effects or spillovers), the result of private decentralized decision-making will be inefficient in the sense that by implementing some other feasible outcome, at least one individual could be made better off without anyone being made worse off. These inefficient outcomes are termed market failures (environmental degradation or traffic congestion, for example). They would be avoided if people were held liable for the costs that their actions inflict on others (and were recompensed for the benefits conferred on others). What economists call complete contracts do just this: They eliminate the spillovers, internalizing the external effects by assigning claims and liabilities so that each actor “owns” all of the benefits and costs resulting from his or her actions, including those conferred or imposed on others.

Thus, if contracts were complete, the invisible hand would work: Self-interested individuals would implement outcomes that are efficient in the above sense (9). This is the economic alchemy by which entirely self-regarding individuals are induced to act as if they cared about the effects of their actions on others. Prices do the work of morals, recruiting shabby motives to elevated ends. A consequence, according to the philosopher David Gauthier, is that if contracts are complete, “morality has no application to market interaction under the conditions of perfect competition” (10).

Contracts are rarely complete, however, in part because information about the amount and quality of the good or service provided is either asymmetric or nonverifiable, that is, it is not known to both parties, or even if known it cannot be used in the courts to enforce a contract. As a result, market failures are not confined to the well-known cases of environmental spillovers but occur in the workaday exchanges essential to the functioning of a capitalist economy: labor markets and credit markets. Contractual incompleteness occurs in these cases because of the impossibility of writing an enforceable contract that specifies that the employee will work hard and well and the fact that credit contracts cannot be enforced if the borrower is broke (11). Contracts are also incomplete (or nonexistent) in team production processes and the voluntary provision of public goods such as neighborhood amenities or adherence to social norms.

The labor and credit market examples share a common structure: A principal (the employer or the lender) wishes to induce the agent (the employee or the borrower) to act in a way beneficial to the principal, but the conflict of interest between the two cannot be resolved by specifying the terms of a complete and enforceable contract. The de facto terms of the exchange are determined by the strategic interaction among the parties, not by the courts. The same problem arises when a farmer pays a share of his crop to the landowner. The problem common to these cases is that the agent does not own the results of his or her actions: The lender takes the loss if the

¹Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA. ²Dipartimento di Economia Politica, Università degli Studi di Siena, Siena, Italy.

E-mail: samuel.bowles@gmail.com

borrower cannot repay because of the agent's choice of an overly risky project; the employer enjoys most of the benefits of the employee's hard work.

The task of the mechanism designer is to find a way to assign to each actor the entire benefits and costs (to themselves and to others) of his actions, thereby providing a surrogate for complete contracts. For example, assigning ownership of the land to the sharecropper (who would then own the entire crop) would accomplish this. Replacing sharecropping by a fixed rent that does not depend on how much is produced would do the same.

This emphasis on mechanisms to the exclusion of morals is new. Before the advent of economics in the 18th century, it was more common to appeal to civic virtues: fellow feeling toward one's neighbors, the work ethic, and the moral obligation to repay. These motives are hardly adequate to avoid market failures (especially among strangers in the global marketplace), but morals and other-regarding motives are essential to the performance of firms, communities, and other institutions. Examples include the payment of taxes [in the United States, far in excess of the amounts that would maximize one's expected income (12)] and the positive influence of good will toward management on employee effort (13). Behavioral experiments that model the voluntary provision of public goods and relationships between principals and agents show that substantial fractions of most populations adhere to moral rules, willingly give to others, and punish those who offend standards of appropriate behavior, even at a cost to themselves and with no expectation of material reward (14–16).

Thus, societies address market failures through some combination of incentive-based design and other-regarding motives (9). Likewise,

organizations such as the Ford plant depicted in Diego Rivera's mural (Fig. 1) motivate their members in part by the carrots and sticks of incentives but also by appealing to other-regarding motives such as the desire to do a good job and a sense of reciprocal obligations among members of a firm. However, recent advances in experimental economics provide convincing evidence that the joint effect of these two kinds of motivation is not simply the sum of their effects considered separately.

The 41 experiments on which this review is based [listed with technical details in (17)] show that the separability assumption commonly fails. In a few cases, explicit incentives and ethical motives are complements, the former enhancing the salience of the latter. In most cases, though, separability fails in the opposite way: Incentives undermine ethical motives. As is standard in behavioral economics, most of the experiments were played anonymously for real (and often substantial) money stakes.

Four reasons have been suggested for the failure of the separability assumption: incentives may frame a decision problem and thereby suggest self-interest as the appropriate behavior, or affect the long-term development of preferences, or compromise the individual's sense of autonomy, or convey information affecting behavior. These processes—termed framing, endogenous preferences, overdetermination, and the information content of incentives, respectively—often work jointly and sometimes with opposite effect. Experiments illustrating these four explanations follow.

Framing

Incentives are part of how a decision situation is represented and may signal appropriate behavior (18), as seems to have been the case for the Haifa parents, for example. Framing is also at

work when simply using market terminology (“exchange”) to describe an experiment reduces fair-minded behavior (19) or in which market-like competition “offers justifications for actions that in isolation would be unjustifiable” (20).

The frame-shifting effects of incentives may occur in cases of government-imposed incentives, too. Here is an example. Experimental subjects whose livelihoods depend on easily depleted forest resources in rural Colombia were asked by Cardenas and his colleagues to individually and anonymously choose how much to withdraw from a mutually beneficial common pool analogous to “the forest” (21). Payoffs were such that the level of withdrawal that maximized the gains of the group as a whole was substantially less than the level that maximized the gains of the individual acting singly. The experiment thus captured a common market failure in which self-interested actions by each would overexploit a common pool resource (the forest) and reduce the well-being of all.

Groups of subjects played eight rounds of this game without communication, withdrawing on average amounts that were about midway between the individually self-interested and the group-beneficial levels (Fig. 2). Their substantial deviation from the individually selfish level is a measure of the subjects' other-regarding or ethical values. The experimenters then changed the rules. In subsequent play, for some groups face-to-face communication was allowed (but there was no way to make binding promises). Groups in this “communication” treatment improved their performance, extracting less from the “forest,” thereby deviating more from self-interest, and gaining higher benefits.

The other treatment precluded communication but simulated “government regulation.” Withdrawals were not to exceed the announced



Fig. 1. Diego Rivera's mural of factory workers at Ford's River Rouge assembly plant (detail). Modern economies require cooperation toward common ends among countless individuals, often occurring as the result of both self-interested and

ethical motives. Recent behavioral experiments show that organizational strategies may backfire if they rely solely on explicit economic incentives and seek to limit the options of group members.

PHOTO: COURTESY OF THE DETROIT INSTITUTE OF ARTS

group-optimum level, and subjects would be monitored and fined for overexploitation. The regulation reduced the level of withdrawal that would be chosen by an entirely selfish individual, but the expected fines were such that some over-exploitation of the common pool remained the payoff maximizer's optimal choice. In this "regulation" treatment, subjects initially responded by restricting their withdrawals to close to the group optimum, but after two periods their behavior increasingly conformed to self-interest, and for the last three rounds their choices were almost entirely self-interested (Fig. 2), sacrificing only one-fifth as much individual payoff to protect the "forest" as subjects in the communication treatment. The fine, although insufficient to enforce the social optimum, apparently all but extinguished the subjects' ethical predispositions that in the earlier rounds had induced them to withdraw much less than would maximize their own payoffs.

Endogenous Preferences

Incentives may also induce long-term change in motivations. Preferences are endogenous if one's experiences result in durable changes in motivations and hence a change in behavior in given situations. A number of experiments have documented durable learning effects (22, 23). In these experiments, as in the case of the fines for tardiness at the Haifa day care centers, incentives induced more self-interested behavior, even after they were withdrawn. In the public goods experiment designed by Falkinger *et al.* (23), subjects who had experienced an incentive system that was very effective in increasing contributions to the public good later played the same game without the incentives: They contributed 26 percent less than subjects who had not been exposed to the incentives.

However, experiments of just a few hours duration are unlikely to uncover the causal mechanisms at work. This is because adopting new preferences is often a slow process more akin to acquiring an accent than to choosing an action in a game. The developmental processes involved typically include population-level effects such as conformism, schooling, religious instruction and other forms of socialization that are not readily captured in experiments. However, historical, anthropological, social, psychological, and other data (5) show that economic structures affect parental child-rearing values, personality traits rewarded by higher grades in school, and other developmental influences. Thus, economies structured by differing incentives are likely to produce people with differing preferences (24).

Incentives change preferences because they affect key aspects of how we acquire our motivations. These effects include the fact that incentives influence both the range of alternative preferences to which one is exposed and the economic rewards and social status of those with preferences different from one's own (11). For example, if the relevant incentives allow the

selfish to exploit the civic-minded, then the latter are less likely to be copied. Other effects are less obvious: A competitive market with complete contracts leaves little scope for acting on ethical, reciprocal, or generous preferences, even among those so inclined. Moreover, if such markets were, as Gauthier says, morality-free zones, then nothing would be lost if people regarded markets as off limits for morality, other than the possibility

of the groups, consistent with the idea that developmental processes that affect adult behavior are linked to economic structure. We found that in anonymous experimental settings, individuals from the more market-integrated societies were also more fair-minded in that they made more generous offers to their experimental partners and more often chose to receive nothing rather than accept an unfair offer. A plausible explanation is that this kind

of fair-mindedness is essential to the exchange process and that in market-oriented societies individuals engaging in mutually beneficial exchanges with strangers represent models of successful behavior who are then copied by others.

Self-Determination

Where people derive pleasure from an action per se in the absence of other rewards, the introduction of explicit incentives may "overjustify" the activity and reduce the individual's sense of autonomy. The underlying psychological mechanism appears to be a fundamental desire for "feelings of competence and self-determination that are associated with intrinsically motivated behavior" (25). There is a substantial empirical literature on the psychology of intrinsic motivations (26, 27), as well as nonexperimental studies in economics [surveyed in (28)]. Recent experiments by economists are consistent with this view.

For example, Falk and Kosfeld (29) explored the idea that "control aversion" may be a reason that incentives degrade performance. Experimental agents in a role similar to an employee chose a level of production that was costly to them and beneficial to the principal (the employer). The agent's choice effectively determined the distribution of gains between the two, with the agent's maximum payoff occurring if he produced nothing.

Before the agent's decision, the principal could elect to leave the choice of the level of production completely to the agent's discretion or impose a lower bound on the agent's production (three bounds—low, medium, and high—were varied by the experimenter across treatments; the principal's choice was whether or not to impose it). The principal could infer that a self-regarding agent would perform at the lower bound and thus



Fig. 2. Effects of social preferences with communication or fines. Shown are two measures of the salience of social preferences in an experiment replicating the problem of cooperation to protect an environmental resource (21). The top panel shows the average deviation (in months of exploitation of the forest) from the level that would have maximized the individual's material payoff, given what others in the group did. For example, in round 1 of stage I, both groups exploited the forest a little more than 3 months less than would have maximized their individual payoffs. Implementing the social optimum would have required deviating about 6-1/2 months from the self-interested response. The second measure is the income foregone by the individual by withdrawing less from the forest than would have maximized his income. The first eight rounds were the same treatment for both groups (no communication, no regulation), and there were no significant group differences in behavior by either measure. The groups diverged sharply in Stage II. In the communication treatment, subjects deviated from the self-interest optimum by an average of 4 or more months (top right), more than they had in the absence of communication, and at considerable cost to themselves (bottom right). In the regulation treatment, subjects initially conformed closely to the social optimum, but in successive rounds increasingly acted in an own-payoff-maximizing manner. [Used by permission.]

that the moral dispensations claimed for the marketplace will be generalized to other arenas of life.

However, markets may have quite the opposite developmental impact. Behavior by subjects in experiments among hunter-gatherers, herders, farmers, and others in 15 small-scale societies in Africa, Asia, and Latin America (24) closely reflected the highly diverse economic livelihoods

that imposition of the bound would maximize the principal's payoffs.

In the experiment, however, agents chose a lower level of production when the principal imposed the bound. Apparently anticipating this response, fewer than a third of the principals opted for its imposition in the medium- or low-bound treatments. The minority of "untrusting" principals earned on average half of the profits of those who did not seek to control the agents' choice in the low-bound treatment, and a third less in the medium-bound condition. In postplay interviews, most agents agreed with the statement that the imposition of the lower bound was a signal of distrust.

Control aversion and the desire for self-determination are not the only effects of the principal's seeking to bound the agent. The imposition of the minimum in this experiment gave the agents remarkably accurate information about the principals' beliefs concerning the agents: Those who imposed the bound had substantially lower expectations of the agents. Their consequent attempt to control the agents' choices induced over half of the agents (in all three treatments) to contribute minimally, thereby affirming the principals' pessimism. This illustrates our fourth reason that the separability assumption may fail.

Incentives Convey Information

Principals select incentives based on their own objectives and their beliefs about how well the agent will perform his task under each possible incentive. Thus, the incentives selected necessarily reveal information about the principal's preferences, the nature of the task, and his beliefs concerning the agent (30, 31). The incentives selected may indicate that the principal is seeking to profit at the expense of the agent, or that the principal believes the agent to be otherwise not committed to performing well, or that the job is onerous, or, as we have seen, that he does not trust the agent.

This predicament for the principal is nicely illustrated in an experiment by Fehr and Rockenbach (32). German students in the role of "investor" chose a costly action benefitting the other player, the "trustee," who, knowing the investor's choice, could in turn provide a personally costly "back-transfer," returning a benefit to the investor. When the investor transferred money to the trustee, he also specified a desired level of the back-transfer. The experimenters implemented an incentive condition in which the investor had the option of declaring that he would impose a fine if the trustee's back-transfer were less than the desired amount. The investor could also decline the use of the fine, the choice of using or declining the fine option being taken before the trustee's decision. There was also a "trust" condition in which no such incentives were available to the investor.

The use of the fine reduced return transfers, whereas renouncing the fine when it was available increased return transfers (Fig. 3). Only one-third of the investors renounced the fine; their payoffs were 50 per cent greater than the in-

vestors who threatened use of the fines. The authors' interpretation is that trusting elicited a positive reciprocal response that was extinguished by the threat of the fine. This was especially the case when it appeared that the intent of the fine was to impose an unfair outcome. Where the investor announced desired returns that would have shared the benefits equally, the use of the fines reduced back-transfers by only 8 percent. Where the announced desired back-transfer would have allowed the investor to capture most of the benefits had the trustee complied, however, the reduction in back-transfers was 38 percent.

The fact that incentives reveal that the principal is untrusting or self-aggrandizing helps explain a common pattern of experimental results. Incentives imposed by peers who do not stand to benefit personally do not compromise social preferences and are often synergistic with them. An example comes from a public goods experiment in which fellow group members have the opportunity to reduce their own payoffs in

would otherwise contribute to the performance of the agent often forgo the use of explicit incentives and sanctions even in cases where the latter are feasible (35). Why then do we ever observe counterproductive incentives in practice?

An experiment by Fehr and Gächter [reported in (36)] with Swiss students suggests an answer: Even if incentives reduce the total gains associated with a project, their use may give the principal a sufficiently larger slice of the smaller pie to motivate the principal to use them. The experiment, similar to the Fehr and Rockenbach experiment above, was constructed so that had subjects responded optimally on the basis of self-regarding preferences, the total surplus (sum of payoffs of employer and employee) would have been more than twice as great under the incentive treatment as under the trust treatment. Negative synergy between the incentive and social preferences was so strong, however, that the total surplus was much higher in the trust treatment than when incentives were introduced. This was

true even in those cases where principals offered exactly the kind of contract that a mechanism designer would recommend. Under these "optimal" contracts, profits were more than double those in the trust treatment, whereas the payoffs to employees were less than half. The incentive treatment allowed employers to save enough in wage costs to offset the reductions in work effort.

Thus, one of the reasons agents respond negatively to incentives—that they benefit the principal at the agent's expense—also explains why incentives may nonetheless

be used by profit-maximizing principals, despite the fact that they shrink the pie. If a mutually acceptable division of the pie could be decided in advance (and enforced ex post) this problem would not arise, but such ex ante agreements are typically not feasible in real economies.

Why Moral Sentiments and Material Interests Are Not Separable

According to the conventional economic approach, individuals process raw materials and make exchanges so as to get something; the things acquired do not include one's moral standing or sense of having acted well. Behavior is not only acquisitive, however; it is also constitutive. People act also so as to be or to become a good person or one who is esteemed by others. When one's person itself is the raw material and its transformation or affirmation is the objective, the presence of explicit economic incentives may have unintended effects. The

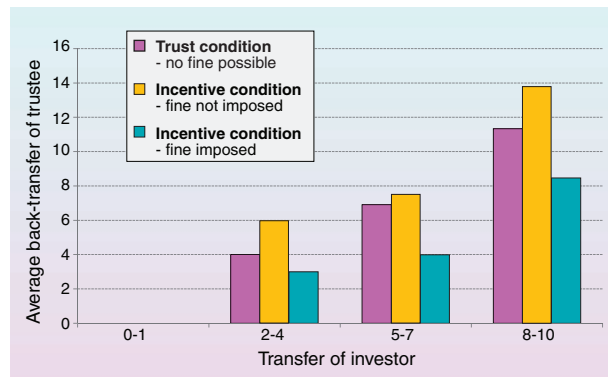


Fig. 3. Average trustee's back-transfer by level of investor's transfer (32). Larger investors' transfers are reciprocated by larger trustees' back-transfers, but the average back-transfer is least when the fine is imposed and greatest when the fine was available to the investor but was renounced. [Used by permission.]

order to punish (reduce the payoffs of) others in their group once each member's contributions are revealed and in which group membership is shuffled so that a punisher cannot benefit from the target's response in subsequent periods. In contrast to the case where punishments are imposed by a principal who will benefit at the agent's expense if the agent responds positively, here there is a strong positive response by low contributors (15). The most plausible explanation of the effectiveness of peer punishment is that when punished, those who have contributed less than others feel shame, which they redress by contributing more subsequently (33). If instead those punished experience anger (if they have contributed more than others, for example), contributions may fall, and costly retaliatory punishment escalations result (34).

Why Are Counterproductive Incentives Common?

We have seen that principals who understand that incentives may undermine social preferences that

individual who would give to charity so as to affirm his status as a generous person may find that even a small material reward overjustifies his good deed and hence degrades its value as a signal (37).

Constitutive action appears to be at work also in the cases where incentives and morals work synergistically as complements. The rule of law and other institutional designs limit the more extreme forms of antisocial behavior and facilitate mutually beneficial interactions beyond the family. This may enhance the salience of social preferences by assuring people that those who conform to moral norms will not be exploited by their self-interested fellow citizens. This phenomenon may have been at work among the Hokkaido University subjects who cooperated more in a public goods experiment when assured that others who did not cooperate would be punished (38), despite the fact that this had no effect on the subjects' own material incentives. They apparently wanted to be cooperative but wished even more to avoid being the sucker who is exploited by defectors. Similar synergies occur in natural settings: Social norms support observance of traffic regulations, but these may unravel in the absence of state-imposed sanctions on flagrant violations.

The proximate reasons for nonseparability may reflect the fact that appeals to self-interest and to morality activate different cognitive and neurological processes, appeals to the former competing with the latter by a process of cognitive overload.

Discussion

Although standard in economics, reliance solely on self-interest in the design of policies has never won universal assent. Until recently, however, dissenting views, like Titmuss' (39) celebrated claim that paying for blood donations degrades the willingness to contribute, were thought to lack either empirical support or a coherent account of why separability might fail. But an experiment reported in (40) suggests that Titmuss may have been right, at least for women. Other experiments surveyed here provide additional evidence that material interests and moral sentiments are not separable in the sense required by the conventional economic approach to policy-making. Economists, psychologists, and others, in part stimulated by these new empirical data, are well on their way to constructing an economic psychology of the interplay of self-regarding and other-regarding motivation that may eventually enlighten mechanism design and public policy.

Is there a message for policy-makers? There is nothing about mechanism design (or economics as a whole) that would preclude more realistic psychological assumptions. Where explicit incentives so strongly compromise social preferences that the effect is the opposite of that intended, the message is clear enough: Other

means of affecting citizens' actions must be sought. However, where the effect of incentives is merely blunted rather than reversed, surprisingly, either greater or lesser use of economic incentives may be optimal (8). If the mechanism designer seeks to implement a target level of compliance (for example, that a given fraction of the population should be inoculated), greater use of the incentive is called for unless alternative means to the same end are available. However, lesser use of the incentive may be optimal in other cases, as the mechanism designer must weigh the value of additional contributions to the public good against the administrative costs, taking account of both the reduced effectiveness of the incentive and the fact that it reduces the utility that citizens experience from the act of contributing.

Policy-makers could learn valuable lessons, too, from experimental evidence that some mechanisms induce even the civic-minded to act as if they were selfish. Examples include anonymous competitive markets and the public goods game without opportunities for peers to monitor and punish free riders. Other mechanisms, such as the public goods game with peer punishment or opportunities for communication among participants, may induce the self-interested to act as if they were civic-minded. Thus, even for a population with a given distribution of self-regarding and other-regarding motives, small differences in institutional design may lead to very different outcomes at the societal level.

This suggests an extension of Hume's maxim about citizens as knaves and a challenge to the next generation of mechanism designers: Good policies and constitutions are those that support socially valued ends not only by harnessing selfish preferences to public ends but also by evoking, cultivating, and empowering public-spirited motives. The modest tax on plastic grocery bags enacted in Ireland in 2002 that resulted in a 94 per cent decline in their use appears to have had just this effect (41): Carrying a plastic bag joined wearing a fur coat in the gallery of antisocial anachronisms.

References and Notes

1. D. Hume, *David Hume, The Philosophical Works*, T. H. Green, T. H. Grose, Eds. (Longmans, Green, and Co, London, 1898), pp. 117–118.
2. J. J. Laffont, E. Maskin, in *Advances in Economic Theory*, W. Hildenbrand, Ed. (Cambridge Univ. Press, Cambridge, 1982) pp. 31–92.
3. L. Hurwicz, in *Economic Institutions in a Dynamic Society: Search for a New Frontier*, T. Shiraishi, S. Tsuru, Eds. (Macmillan, London, 1989).
4. A. O. Hirschman, *Econ. Philos.* **1**, 7 (1985).
5. S. Bowles, *J. Econ. Lit.* **36**, 75 (1998).
6. B. S. Frey, *Econ. J.* **107**, 1043 (1997).
7. U. Gneezy, A. Rustichini, *J. Legal Stud.* **29**, 1 (2000).
8. S. Bowles, S.-H. Hwang, *J. Public Econ.* **92**, 8 (2008).
9. K. J. Arrow, in *Frontiers of Quantitative Economics*, M. D. Intriligator, Ed. (North Holland, Amsterdam, 1971), pp. 3–23.

10. D. Gauthier, *Morals by Agreement* (Clarendon Press, Oxford, 1986), p. 93.
11. S. Bowles, *Microeconomics: Behavior, Institutions, and Evolution* (Princeton University Press, Princeton, NJ, 2004).
12. J. Andreoni, B. Erand, J. Feinstein, *J. Econ. Lit.* **36**, 818 (1998).
13. T. F. Bewley, *Why Wages Don't Fall During a Recession* (Harvard Univ. Press, Cambridge, MA, 1999).
14. E. Fehr, U. Fischbacher, *Nature* **425**, 785 (2003).
15. E. Fehr, S. Gaechter, *J. Econ. Perspect.* **14**, 159 (2000).
16. H. Gintis, S. Bowles, R. Boyd, E. Fehr, Eds., *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life* (MIT Press, Cambridge, MA, 2005).
17. Additional information about the studies reviewed here is available as supporting material on Science Online.
18. A. Tversky, D. Kahneman, *Science* **211**, 453 (1981).
19. E. Hoffman, K. McCabe, K. Shachat, V. L. Smith, *Games Econ. Behav.* **7**, 346 (1994).
20. A. Schotter, A. Weiss, I. Zapater, *J. Econ. Behav. Organ.* **31**, 37 (1996).
21. J. C. Cardenas, J. K. Stranlund, C. E. Willis, *World Dev.* **28**, 1719 (2000).
22. B. Irlenbusch, D. Sliwka, *IZA Discussion paper No. 1758*; <http://ideas.repec.org/p/iza/izadps/dp1758.html> (Institute for the Study of Labor, Bonn, 2005).
23. J. Falkinger, E. Fehr, S. Gaechter, R. Winter-Ebmer, *Am. Econ. Rev.* **90**, 247 (2000).
24. J. Henrich et al., *Behav. Brain Sci.* **28**, 795 (2005).
25. E. L. Deci, *Intrinsic Motivation* (Plenum, New York, 1975).
26. J. Cameron, K. Banko, W. D. Pierce, *Behav. Anal.* **24**, 1 (2001).
27. E. L. Deci, R. Koestner, R. M. Ryan, *Psychol. Bull.* **125**, 627 (1999).
28. B. Frey, R. Jegen, *J. Econ. Surv.* **15**, 589 (2001).
29. A. Falk, M. Kosfeld, *Am. Econ. Rev.* **96**, 1611 (2006).
30. D. Sliwka, *Am. Econ. Rev.* **97**, 999 (2007).
31. R. Benabou, J. Tirole, *Rev. Econ. Stud.* **70**, 489 (2003).
32. E. Fehr, B. Rockenbach, *Nature* **422**, 137 (2003).
33. A. Hopfensitz, E. Reuben, *Tinbergen Institute Working Paper 05-0751*; www.tinbergen.nl/discussionpapers/05075.pdf (2006).
34. B. Herrmann, C. Thöni, S. Gächter, *Science* **319**, 1362 (2008).
35. M. Brown, A. Falk, E. Fehr, *Econometrica* **72**, 747 (2004).
36. E. Fehr, A. Falk, *Eur. Econ. Rev.* **46**, 687 (2002).
37. R. Benabou, J. Tirole, *Am. Econ. Rev.* **96**, 1652 (2006).
38. M. Shinada, T. Yamagishi, *Evol. Hum. Behav.* **28**, 330 (2007).
39. R. M. Titmuss, *The Gift Relationship: From Human Blood to Social Policy* (Pantheon, New York, 1971).
40. C. Mellstrom, M. Johannesson, *J. Eur. Econ. Assoc.* **6**, 4 (2008); available at <http://gupea.ub.gu.se/dspace/handle/2077/2741>.
41. E. Rosenthal, "Motivated by a Tax, Irish Spurn Plastic Bags," *New York Times*, 2 February 2008; www.nytimes.com/2008/02/02/world/europe/02bags.html.
42. The author declares no competing interests and would like to thank M. Alexander, J. C. Cardenas, J. Carpenter, C. Efferson, E. Fehr, M. Hauser, S.-H. Hwang, S. Gaechter, H. Gintis, G. Loewenstein, K. McCabe, S. Polania, T. Taylor, and E. Wood for their contributions to this research and the Behavioral Sciences Program of the Santa Fe Institute, the Russell Sage Foundation, the University of Siena, the European Science Foundation, and the U.S. National Science Foundation for support of this work.

Supporting Online Material

www.sciencemag.org/cgi/content/full/320/5883/1605/DC1
Table S1
References
10.1126/science.1152110