5 OBSERVING THE UNOBSERVABLE

David Schmidtz[1]

5.1. Field Data and Controlled Experiments

The problem of producing public goods by collective action takes the form of a prisoner's dilemma when the marginal return per unit of contribution is less than one unit to the contributor but more than one unit to the group. In that case, the group is strictly better off if a given unit is contributed rather than withheld, but an individual is strictly better off withholding that unit. Because group members decide as individuals rather than as a group, the assumption that individuals act to maximize personal wealth implies that individuals in a prisoner's dilemma situation will contribute exactly nothing.

But theory and practice have a way of diverging. In 1980 and 1985, Canada's New Democratic Party attempted to raise $200,000 and $250,000, respectively, in campaign contributions. They succeeded both times, with $1,300 to spare in the second case (Bagnoli and McKee, 1991: 351). In 1979, the Association of Oregon Faculties wanted to raise $30,000 to hire a lobbyist to represent them at the state legislature. They sought $36, $60, or $84 contributions, depending on salary, from all faculty in the state. The drive was successful (Dawes, Orbell, Simmons, and Van de Kragt, 1986: 1172). These organizations solicited donations on the understanding that the money would be refunded if the target figure was not reached. Donors, in other words, had a money-back guarantee.

Why did these fundraising drives succeed? Was it the money-back guarantee? Unfortunately, we will never know. Of course, our problem is hardly unusual. Field experiments are almost always uncontrolled in the technical sense. That is, we want to hold all variables constant, except for a single target variable, so that we may learn how the outcome is affected specifically by changes in the target variable. But in the field, holding all other variables constant is next to impossible, which means there almost always will be a variety of explanations for the observed outcome. For this reason, assessing the importance of changes in the target variable (the fundraising mechanism in this case) is typically very difficult.

This is just one example of a traditional problem with the scientific approach to economics. When field data seem to disconfirm an applicable theory, it is often just as reasonable to reject the data as it is to reject the theory because the data are gathered under conditions that do not permit proper control over the set of possibly relevant variables. Nor is control the only problem. Some relevant variables, like the value people attach to the hiring of a state lobbyist, are not even *observable*, let alone controllable. In-principle testability has been a feature of most important theories in economics; that is to say, they have empirical content. Practical testability has been another matter entirely.

Laboratory conditions, however, offer the possibility of greater control, particularly over such ordinarily unobservable variables as preferences, knowledge endowments, and strategies of the agents involved.[2] To determine how behavior is affected by alternative institutions, we need controls for these key variables, which are not controllable or even observable in the field. In the laboratory, inherently private valuations can be stipulated by and hence known to the experimenter. That is, we can specify the resale value to each subject of a given unit at the same time as we strip the unit of all properties other than price and resale value. So valuations are private information in the sense that subjects do not know how much a unit is worth to other subjects. Nevertheless, the experimenter can, for example, see exactly how actual contribution levels compare with optimal contribution levels or how actual behavior compares with demand-revealing behavior. (In other words, do subjects offer as much as the unit is worth to them, or do they underbid?) We thereby learn about the strategies subjects employ. And because the experimental instructions are generally the only source of information subjects have about the nature of the experiment, knowledge endowments are well controlled too. Thus, the usual response to field experiments—that when a theory is not corroborated, something is wrong with the experiment—is not so easily defended as a response to data gathered in the laboratory.

5.2. The Isaac-Walker Design

A general (and hence powerful) model of the public goods problem has been tested in the laboratory by Isaac, Walker, and Thomas (1984) and by Isaac and Walker (1988). In these experiments, subjects (four or ten in a group) were given an initial endowment of tokens and asked to allot them between private and public exchanges. A token invested in the private exchange yielded one cent to the individual subject. A token invested in the public exchange yielded a sum to be divided equally among all subjects in the group. (The size of the sum was determined by the group's production function. Two such functions were used, both linear, yielding either 0.3 or 0.75 units per subject for each token invested in the public exchange.) A series of ten trials allowed subjects to learn that their single-period dominant strategy was to invest all of their income in the private exchange. Subjects were told in advance that the tenth period would be the final period so that their single-period dominant strategy would clearly be salient in the tenth period. (And in periods one to nine, for what it is worth, a set of zero contributions is the Nash equilibrium predicted by backward induction.)

The results of these experiments disconfirmed the prediction that investment in the public exchange in period 10 will be zero. Even in final periods preceded by what should have been ample opportunity to learn dominant strategies, contributions are positive and substantial, albeit suboptimal. Why do subjects contribute as much as they do? Why do they contribute anything at all? Since our set of assumptions implies a prediction that contributions will be precisely zero in period 10, the fact that this prediction is consistently wrong suggests that one of our assumptions is false. Because noncontribution is clearly a dominant strategy, a process of elimination locates the error in our behavioral assumption that subjects act so as to maximize personal income. The hypothesis that people will free ride whenever they have the opportunity is disconfirmed by experiments such as those conducted by Isaac and Walker.

So if subjects are not trying to maximize personal income, what are they doing? Perhaps some subjects simply do not behave strategically. That is, they may develop some expectation of how much of the public good will be produced and then simply contribute what they are willing to pay for that product. Some subjects may operate by rules of thumb (like "Make small contributions to charity") and thus may be oblivious to the incentives involved in particular situations.

Inability to understand the rules of the game could also lead some subjects to contribute, although one might expect that subjects who did not understand the situation would be more likely to sit back and contribute nothing. Moreover, the situation was by no means complicated. I think failure to comprehend does play a role in what we observe, but by the tenth period, only a very small one.

In any event, the nature of our concern about incomprehension in the laboratory is partly a matter of whether we mean to test descriptive theory or prescriptive policy. However much confusion there is in the laboratory, there is more confusion in more complicated environments. So the reality of incomprehension is something we have to be aware of when drawing conclusions about policy. (I will have more to say about the difference between tests of theory and tests of policy.)

- Perhaps the interests of others enter into subjects' preference functions. That is, subjects prefer to benefit other subjects. Or some subjects may not exactly have a preference for benefiting other subjects but may not consider it permissible to fail to do their share, whatever they perceive their share to be. Some subjects, if they are wealth maximizers at all, may place certain constraints on what they will do to maximize personal wealth. In other words, the externalities involved in the decision to contribute or not may enter into subjects' decisions as part of their set of constraints, along with income constraints and so on. To use David Gauthier's (1986: 160ff) terms, many of those who contribute even in final periods may be constrained rather than straightforward maximizers.

Contributing positive but suboptimal amounts in experimental situations is sometimes described as *weak* free riding. This label can mislead because the phenomenon so named may not be free riding at all. As I use the term, *free riding* is an attempt to enjoy the benefits of other players' contributions without responding in kind. Some individuals may contribute suboptimally not because they are free riders, weak or otherwise, but because they are averse to being, so to speak, taken for a ride. Let us call the aversion to being taken advantage of an *exploitation problem.* Subjects with an exploitation problem seek some assurance that the pattern of contributions within the group will be fair before they will contribute what they judge to be their share. Alternatively, a person may be willing to participate in a joint venture on mutually profitable terms. A person's contributions, however, may be limited to a level such that the sum of investments in the public exchange can be expected to yield a return sufficient to repay his own investment. Unless the person receives reasonable assurance that other people will contribute enough to ensure that his own contribution will not be wasted on a hopelessly underfunded cause, the person may decide to save his money. This limitation is an *assurance problem.*[3]

5.3. A Theory About Prisoner's Dilemmas

That leaves us with a puzzle about how to distinguish between defections according to what motivates them. The problem is important for policy because whether a money-back guarantee affects contribution levels depends entirely on

		How many do the other three players contribute?			
		None	One	Two	All Three
How many do you contribute?	No	0	+0.3	+0.6	+0.9
	Yes	−0.7	−0.4	−0.1	+0.2

Figure 5.1. Prisoner's dilemma (only "your" net gains and losses are shown)

whether the assurance problem is a real cause of defection. There is a theoretical issue here as well, because the standard theory implies that a money-back guarantee should not make any difference. It should not make a difference because, if the situation is a prisoner's dilemma, then adding a money-back guarantee cannot change the fact that noncontribution is a dominant strategy.

Figure 5.1 depicts the incentive structure in the Isaac-Walker experiments. There are four players with identical payoff functions, one of whom I refer to as *you.* If you contribute a unit, you lose that unit and each player including you gains 0.3 units. Not contributing is a dominant strategy, for no matter how many of the others contribute, you are always at least as well off not contributing. Indeed, you are strictly better off not contributing.

So goes the standard analysis of the prisoner's dilemma. The way I have defined free rider and assurance problems, however, sheds a new light on the dilemma. As defined, these two problems are complementary and essential components of the prisoner's dilemma. In effect, they constitute the two "halves" of the prisoner's dilemma. From your point of view, the left side of the matrix (the lower left three cells in the case at hand) represents your assurance problem. You may be willing to contribute so that (assuming the others also contribute) you get an extra twenty cents. But without assurance that the others will contribute, contributing may get you seventy cents less rather than twenty cents more. This gap, reflected in the difference between your payoffs in the upper left and lower left cells, constitutes an assurance problem. The right side (the upper right three cells in Figure 5.1) of the matrix, by contrast, represents your temptation to free ride. If the others contribute, then you get twenty cents by returning the favor. But there is no need for you to return the favor. In fact, by withholding, you get ninety cents rather than twenty. Thus the gap between the payoffs of the upper right and lower right cells constitutes a free-rider problem.

Consider the effect of the money-back guarantee. It solves the assurance problem by raising negative payoffs in the lower left cells to zero. On my analysis, that eliminates one of the two reasons for noncontribution that jointly

How many do you contribute?	How many do the other three players contribute? → None	One	Two	All Three
No	0	+0.3	+0.6	+0.9
Yes	0	0	0	+0.2

Figure 5.2. Your payoffs after adding money-back guarantee

constitute the prisoner's dilemma and thus should lead some people at least to contribute more than they otherwise would have. On the standard analysis, adding the money-back guarantee should not make any difference, since it does not change the fact that not contributing is the dominant strategy.

As we can see in Figure 5.2, defection is still a dominant strategy. When I was working on my master's thesis, though, I had a theory that people are not particularly responsive to the existence of dominant strategies. On my analysis, the prisoner's dilemma incentive structure is constituted by a conjunction of free-rider and assurance problems, and I believed that each problem has a distinct practical effect. I also thought that the money-back problem could solve the assurance problem, at least in some cases, thus leading in practice to a greater willingness to contribute notwithstanding the fact that noncontribution would remain a dominant strategy.

5.4. How to Observe Motives

In devising a test of a policy of offering money-back guarantees, the problem I initially faced was to determine the empirical status of my hypothesis that prisoner's dilemmas are conjunctions of free-rider and assurance problems. If a person does not contribute to a public goods project, how would we ever know what her reason was? And why would we care? Does a person's reason for withholding actually matter, or is this merely a distinction without a difference? The task was to design an experiment in which the two different kinds of withholding did not look the same.

I was familiar with the Isaac-Walker experiments, for at the time I was a research assistant for the Economic Science Laboratory and was enrolled in a course on experimental economics taught by Mark Isaac. My term paper for that course was a proposal for an experimental test of the assurance problem. The proposal was based on a modified Isaac-Walker design. My method was to

introduce a provision point into the Isaac-Walker design such that the good at stake became a *step good*—a good that cannot be provided at all until total contributions reach a certain critical level. Consider what happens if the provision point is set at 100 percent of the group's total endowment. In that case, free riding is impossible. If any subject withholds even a single token, then neither he nor anyone else can receive the group good. Because free riding is impossible, there is no way to be taken advantage of—hence no exploitation problem. Rational subjects have only two reasons to withhold: the assurance problem and failure to understand the incentive structure.

For purposes of comparison, suppose we combine the 100 percent provision point with a money-back guarantee (that is, a guarantee that contributions to the public exchange will be returned if the provision point is not met). In that case, *contribution* is a dominant strategy. So the presence of assurance problems in the former case (100 percent provision point) and their absence in the latter (100 percent provision point plus money-back guarantee) should account for any significant difference between contribution levels with and without the money-back guarantee. Thus, we can infer the empirical significance of the assurance problem and can distinguish this significance from that of free-rider problems and failures to understand.

Besides indicating the magnitude of the assurance problem, the modified Isaac-Walker design also tests, under laboratory conditions, the empirical effect of the money-back guarantee on voluntary mechanisms for funding public goods provision. Hence it constitutes a test of my theory about what leads people to defect in prisoner's dilemmas and, to a lesser extent, a test of the money-back guarantee as an economic policy. I will discuss this and one other modified Isaac-Walker design. First, however, I report results of two other laboratory research projects that also bear on the efficacy of the money-back guarantee and on the assurance problem's significance in public goods contexts. To my knowledge, these studies currently exhaust the experimental literature on the subject of money-back guarantees in the context of public goods production. (Section 5.7 explores the distinction between theory tests and policy tests and reports on a new Isaac-Walker design for testing voluntary contributions in large groups.)

5.5. Analyzing the Data

5.5.1. Bagnoli and McKee

Bagnoli and McKee (1991) sought to test the efficacy of a money-back guarantee in eliciting contributions in a public goods situation. Their groups consisted of either five subjects (seven groups in all) or ten subjects (two groups in all), with subjects having individual endowments ranging between seven and sixteen tokens. The decision space was continuous; subjects were allowed to contribute whatever portion of their endowment they desired. The production function was binary; the group good was exactly twenty-five tokens if the provision point was met and nothing otherwise. In the event of failure, contributions were returned to those who made them. The group good was divided among subjects according to a preset pattern. In some experiments, the division was equal. In two of the five-person experiments, subjects received from one to ten tokens as their share of the group good. This range was present within as well as across experiments. Individual endowments, the pattern of individual shares in the group good, and number of subjects were common knowledge. In all cases, the provision point was 12.5 tokens. (Subjects could contribute fractions of tokens.) The experiment consisted of fourteen periods.

The seven five-person groups met or exceeded the provision point in eighty-five of ninety-eight periods (fourteen periods in each of seven experiments). Total contributions were within a half token of the provision point in seventy-five of ninety-eight cases. Of the seven small groups, five earned 95 percent or more of the theoretical maximum for all fourteen periods taken together. (The theoretical maximum is the sum of individual incomes, plus the total value of the public good produced, minus the minimum cost of reaching the provision point.) Five of seven groups actually attained the theoretical maximum over the last five periods (thirty-three out of thirty-five periods in all) and a sixth was very close. The two ten-person groups met the provision point in seventeen of twenty-eight periods. They were at 95 percent efficiency over the last five periods, although it took longer for the pattern of contributions to converge on an equilibrium (Bagnoli and McKee, 1991: 359).

It is interesting that their design did not solve the free-rider problem at all; rather, it induced public goods provision in spite of an incentive structure with a built-in free-rider problem. The most pertinent feature of Bagnoli and McKee's institution seems to be its money-back guarantee as a solution to the assurance problem. Unfortunately, we cannot be certain of this. The extent to which solving the assurance problem was what led subjects to contribute in this experimental context remains unclear. For our purposes, it would have been ideal if the Bagnoli and McKee experiments had incorporated a control for the money-back variable. As it is, we do not know how much subjects would have contributed in this environment if they had not had a money-back guarantee. Hence, we have no way of knowing how much the money-back guarantee helped.

5.5.2. Dawes, Orbell, Simmons, and Van de Kragt

Dawes et al. (1986) explored how well contributions are promoted by two devices: the money-back guarantee and *enforced fairness*. The first device returns

contributions to donors if the group fails to meet its provision point. This solves the assurance problem. The second device requires contributions from those who did not contribute if and when the group succeeds in meeting its provision point. This solves the free-rider problem, albeit nonvoluntarily.

Dawes et al. hypothesized that contributions are better promoted by enforced fairness. Their rationale is this: the success of the money-back guarantee can be undermined by people's expectation that it will succeed. Why? Because as subjects become more confident that the money-back mechanism will succeed, free riding comes to seem less risky for it seems increasingly likely that their own contribution is not needed. Thus, the more the money-back mechanism succeeds, the more reason subjects have to withhold. In contrast, as subjects gain more confidence in enforced fairness, the assurance problem becomes less troublesome, for it seems increasingly likely that their own contribution will be sufficient (and the design rules out free riding). So the more the enforced fairness mechanism succeeds, the less reason subjects have to withhold.

Groups in these experiments consisted of seven subjects, each endowed with $5. Their task was to decide individually whether to contribute their endowment to the group project. The decision space was binary; subjects had two options: contribute $5 or contribute nothing. The production technology was also binary; it produced $10 for each subject if the provision point was met and nothing otherwise.

In the money-back dilemma, contributions were returned if the provision point was not met. The free-rider problem remained, but the assurance problem apparently was eliminated. In the enforced fairness dilemma, noncontributors had $5 taken from them if the provision point was met. (In effect, there are no noncontributors when enforced fairness succeeds; enforced fairness is forced unanimity.) The assurance problem remained because those who contributed voluntarily would lose their money if the provision point was not met. But the free-rider problem had been eliminated. Two provision points were examined. The first required $15; three of seven people had to contribute. The second required $25; five of seven had to contribute. These experiments consisted of a single period. The process was not iterated. As a control, full dilemmas, having neither money-back nor enforced fairness features, were also run with provision points.

As Table 5.1 shows, contributions in the Dawes et al. study were substantially higher in the enforced fairness dilemma than in the full dilemma. Dawes et al. (1986: 1183) concluded there was no statistically significant difference between full and money-back dilemmas: "There is no ambiguity whatever about the success of the money-back guarantee device for eliciting contributions compared with the success of the enforced contribution device: the enforced contribution is superior." They also conclude that "Fear of loss through contributing is not the critical motivation underlying defection" (1988: 1183).

Regarding their conclusion about the insignificance of the difference between

Table 5.1. Three dilemmas in Dawes et al. (percentage of subjects contributing)

	Provision Point = 3	*Provision Point = 5*
Full dilemma	51	64
Money-back dilemma	61	65
Enforced fairness dilemma	86	93

full and money-back dilemmas, we might note that the observed difference in the experiments with three-person provision points would have become statistically significant had the difference been replicated over a modest number of further trials. We might also note that if instead of looking at individual contributions, we look at the proportion of groups meeting the provision point, we get a slightly different picture. That is, eleven of twenty groups met their provision point in the full dilemma, whereas eleven of fourteen met their provision point in the money-back dilemma.

Regarding the difference found by Dawes et al. between money-back and enforced fairness dilemmas, their inference that fear of loss is not a critical motivation contradicts my hypothesis that the assurance problem is a major component of the public goods problem. But I believe their inference is invalid for the following reasons. First, enforced fairness does not function only by eliminating the chance to free ride. It also reduces the fear of loss by making it impossible for others to free ride. The design does not rule out the possibility that reducing fear in this way plays a critical role in the success of enforced fairness. Second, in the enforced fairness dilemma, unlike in the full or money-back dilemmas, there was no reason at all to fear that one's contribution would be wasted by virtue of being redundant. (Under the enforced fairness regime, when a subject's contribution is redundant, it would have been collected anyway, so it is not wasted in the way that redundant contributions are wasted in the money-back regime.) Thus, the enforced fairness device reduced fear in two ways, ways in which the alternatives did not, and either way might have helped the enforced fairness device to elicit contributions. Dawes et al. conclude that the superiority of the enforced fairness device shows that fear of loss is not a critical motivational factor, but this conclusion does not follow.

With respect to their claim that enforced fairness is the superior fundraising mechanism, an obvious issue is that enforced fairness is not a voluntary mechanism at all. We should not be too surprised that fundraising efforts are more likely to succeed when backed up by the threat of force. Suppose our faculty club wishes to raise money for the purpose of expanding the local museum. In one scenario, we agree to raise half a million dollars in order to get a matching grant from a large corporation, on the understanding that we will return faculty

contributions if we fail to reach the half-million-dollar target. In an alternative scenario, we say we will try to raise a half million voluntarily, and if we succeed we will seize another half million from faculty members who have not contributed yet. Even if enforced fairness is more effective, it raises rather urgent questions: How do we get that kind of power, and what gives us the right to use it in that way?

Moreover, aside from questions of legitimacy and problems in implementation, the conclusion that enforced fairness is superior does not follow from the data in the first place. In the Dawes et al. design, the efficient levels of contributions were $15 (three out of seven) and $25 (five out of seven) respectively for the full dilemma. Because redundant contributions were wasted, efficiency was reduced by any contributions made in excess of these figures. The same is true of the money-back dilemma. In the enforced fairness dilemma, however, efficiency was not reduced by excess contributions: $35 was an efficient level of investment because meeting the provision point meant that $35 was taken in any event. Obviously, this level of investment was not likely to be approached by the money-back mechanism; matching the contribution levels of enforced fairness (i.e., 86 percent and 93 percent) would have required people in the money-back experiments to throw their money away. (Subjects would not be certain they were throwing their money away, which mitigates my criticism, but they would be aware of the possibility.) When we consider that the efficient contribution levels in the money-back experiments were 43 percent and 71 percent, respectively, rather than 100 percent, the actual money-back contribution levels of 61 percent and 65 percent begin to look respectable, to say the least. Dawes et al. find "no ambiguity whatever" in the superiority of enforced fairness. However, although subtle, the ambiguity is real.

Dawes et al. hypothesized that the efficacy of the money-back guarantee would be undermined by the expectation of its success. This is true in their design. But perhaps real-world money-back guarantees are not subject to the same kind of undermining. In their design, increasing expectation of success also increased the expectation that one's contribution would be wasted, for there were two ways for it to be spent in vain. Both undercontribution and *over*-contribution entailed wasted investment in their design. Thus, the expectation of success not only creates a free-rider problem—as it might do in the real world—but creates its own assurance problem as well, which seems unrealistic.

5.5.3. Isaac, Schmidtz, and Walker

Mark Isaac, David Schmidtz, and James Walker (1989) used four-person groups, each subject initially being endowed with sixty-two tokens. Each subject had a continuous decision space (which means that each could contribute zero, sixty-two, or any whole number of tokens in between). The value of the public good was $G(x) = 1.2x$ if the provision point was met and zero otherwise. For example, if the sum of individual contributions is $2, and if $2 meets or exceeds the provision point, then the value of the public good produced is $2.40 (but if $2 falls short of the provision point, then the value of the public good produced is zero). This value is divided equally among subjects, so that, for example, if the sum of individual contributions is $2, then each of the four subjects receives one-fourth of the resulting G, which in this case means one-fourth of $2.40. Because the public good production function is linear (i.e., since the number 1.2 is a constant) and continuous for all values of x that meet the provision point, there is no possibility of contributions being wasted by virtue of being redundant (that is, because 1.2 is greater than 1 and constant for all values of x, there is no possibility of overcontributing). If the provision point is not met, contributions are returned in the money-back experiments and are not returned (are simply wasted) in the experimental controls, which lack the guarantee.

Three different provision points were used. (In a fit of enthusiasm for technical terms, we called them high, medium, and low.) The high provision point (HPP = 248) represents 100 percent of the group's endowment. Thus, the high provision point treatment eliminated the free-rider problem. The high provision point money-back experiments eliminated both free-rider and assurance prob- lems. The medium provision point (MPP = 216) represented the number of tokens sufficient to produce a return of at least sixty-five cents to each subject, thus ensuring that a given subject could contribute all sixty-two tokens and still make a profit if the provision point was met. The low provision point (LPP = 108) presented a provision level such that any subject making a contribution in excess of thirty-two tokens was not assured of earning a return worth more than thirty-two tokens even if the provision point was met. Thus, subjects contem- plating a contribution of more than half their income had assurance problems even given the money-back guarantee.

Subjects knew in advance that the experiment would run for ten periods. In early periods, there was a pronounced tendency for the four subjects to each make contributions of approximately one-fourth of the provision point. Isaac, Schmidtz, and Walker (1989: 223) call this a *focal point* contribution, a level of contribution that seems somehow obvious (see also Schelling, 1960). Table 5.2 pools data from eighteen experiments, six with each provision point. In keeping with the convention of Isaac, Walker, and Thomas (1984), the individual con- tributions are separated into five size categories ranging from zero to sixty-two.

Notice that contribution levels increased in proportion to increases in the provision point level. This is rather odd, because one would intuitively think assurance problems would worsen as the provision point rises. Table 5.2 suggests that

Table 5.2. First-period individual contributions in Isaac, Schmidtz, and Walker, no money-back guarantee (number of subjects contributing)

Size Range of Contributions	*Low Provision Point (focal pt. = 27)*	*Medium Provision Point (focal pt. = 54)*	*High Provision Point (focal pt. = 62)*
0	1	3	8
1–20	3	3	0
21–41	15	2	1
42–61	2	12	3
62	3	2	12
Total	24	24	24

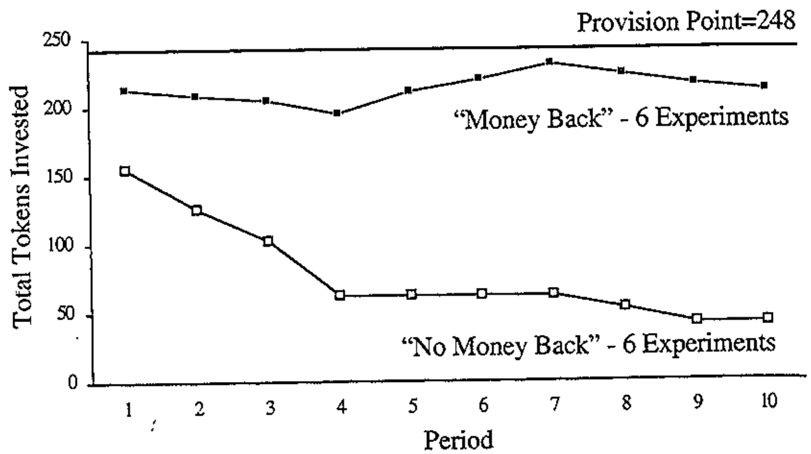
Table 5.3. Final-period individual contributions in Isaac, Schmidtz, and Walker, no money-back guarantee (number of subjects contributing)

Size Range of Contributions	*Low Provision Point*	*Medium Provision Point*	*High Provision Point*
0	19	20	19
1–20	1	0	1
21–41	3	0	0
42–61	0	3	0
62	1	1	4
Total	24	24	24

provision points may have served as focal points in early periods. As the experiment progressed, however, contributions usually collapsed completely. Table 5.3 shows the end-period result.

On the other hand, in three of the eighteen experiments (one for each provision point), contributions remained at higher levels than ever before observed in the general Isaac-Walker design. (The four contributions of sixty-two tokens in Table 5.3's data pool for the high provision point, for instance, all came from a single experiment.) Isaac, Schmidtz, and Walker's (1989: 228) conclusion is that without the money-back guarantee, the introduction of a provision point can dramatically increase contribution levels in a few cases, but generally does not succeed and probably makes matters worse. Thus, introducing the provision point itself made a difference, the extent of which depended on the contribution level at which the provision point was set.

Isaac, Schmidtz, and Walker found that Pareto-superior Nash equilibria tended to collapse to Pareto-inferior Nash equilibria. This was true with and without the possibility of "cheap riding" (contributing only the minimal amount necessary



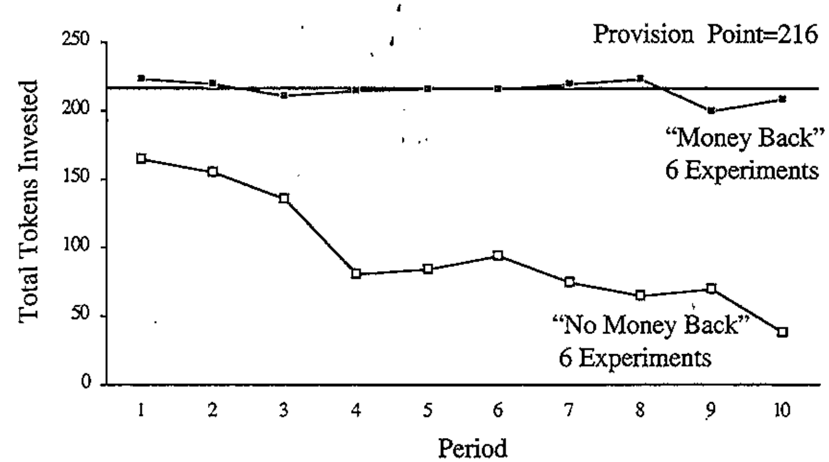
Source: Isaac, Schmidtz, and Walker (1989). Reprinted by permission of Kluwer Academic Publishers.

Figure 5.3 Mean number of tokens contributed in Isaac, Schmidtz, and Walker high provision point

to reach the provision point, given the expected pattern of contribution of the other members of the group). It does not appear to be dominance as such that causes the underprovision of public goods (Isaac, Schmidtz, and Walker, 1989: 229). Subjects' reasons for contributing are more complicated (and often considerably less rational) than that.

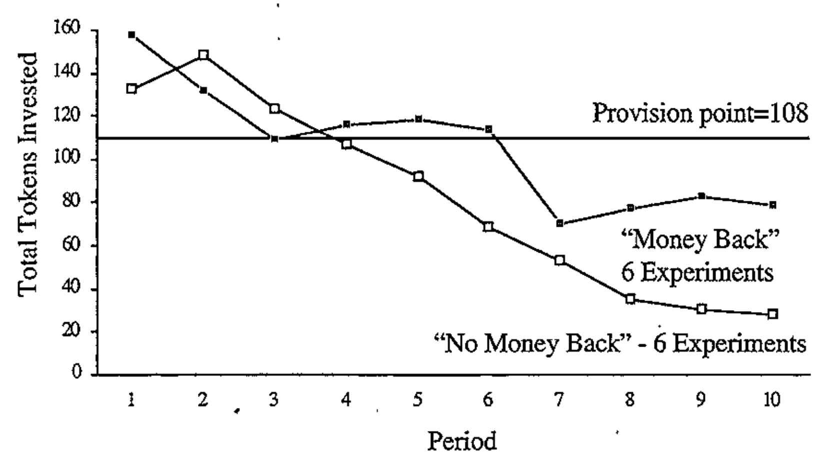
Introducing the money-back guarantee, however, made a substantial difference. Overall, success in meeting provision points went from 45 out of 180 to 93 out of 180 periods when a money-back guarantee was added. Figures 5.3, 5.4, and 5.5 compare average contribution levels for money-back versus no-money-back treatments on a per-period basis, for each of the three provision points.

These observations suggest that the assurance problem is observationally distinguishable from the free-rider problem and that it is a significant part of public goods problems. Further, the money-back guarantee that solves the assurance problem in theory also makes an empirical difference. Especially, note Figure 5.3. With or without the guarantee, the 248-token provision point renders exploitation and free riding impossible. The assurance problem is the sole reason for the difference between the money-back and no-money-back contribution levels. (Failures to understand presumably caused some of the failures to contribute in either case, but only the assurance problem can account for the *difference* between the two contribution levels.)



Source: Isaac, Schmidtz, and Walker (1989). Reprinted by permission of Kluwer Academic Publishers.

Figure 5.4 Mean number of tokens contributed in Isaac, Schmidtz, and Walker medium provision point



Source: Isaac, Schmidtz, and Walker (1989). Reprinted by permission of Kluwer Academic Publishers.

Figure 5.5 Mean number of tokens contributed in Isaac, Schmidtz, and Walker low provision point

Table 5.4. Guarantee versus no guarantee (number of periods in which groups met provision points)

	Money-Back	*No Money-Back*
Isaac, Schmidtz, and Walker	93/180	45/180
Dawes et al.	11/14	11/20
Bagnoli and McKee	85/98	no data

For practical purposes, imposing a provision point without the guarantee is a risky strategy. (Also, imagine people's reaction as you try to explain to them that you have decided not to produce any of the good but will not give their money back either.) Imposing a provision point together with a money-back guarantee appears to be a rather good idea, certainly better than simply asking for contributions. Table 5.4 shows the experimental success of the money-back guarantee in meeting provision points compared with that of the no money-back experimental controls.

In the money-back experiments with a high provision point, contributing sixty-two tokens was a dominant strategy. Contributing anything less was unquestionably a mistake, minimizing the subject's own income as well as that of the group. Nevertheless, it happened. Figure 5.3 shows the gap between the 248-token provision point and the line representing data from the money-back experiments. If all subjects had played their dominant strategies, there would have been no gap. In fact, however, in 56 out of 240 individual decisions made in the six ten-period high provision point experiments, subjects did not play their dominant strategy. Why? I have no truly satisfying answer, but we should note that nearly half of these failures to contribute occurred within a single experiment—one in which subjects reported responding to the group's initial failure by lowering their own bids either in frustration or in an attempt to prompt the others to action.

It also might have been important that those who irrationally made suboptimal contributions in the high provision point money-back experiments continued to earn an income of sixty-two tokens. (In contrast, those who contributed amounts between zero and sixty-two tokens in the no-guarantee, high provision point experiment lost their contributions, and the loss quickly taught them not to repeat their mistake.) Perhaps some subjects considered an income of sixty-two tokens per period perfectly satisfactory and for this reason felt no need to examine the possibility that *contributing* sixty-two tokens was actually a dominant strategy. This raises another question. It was assumed (at least by me) that a money-back guarantee would solve assurance problems without systematically affecting the amount of suboptimality attributable to simple failure to understand the game.

However, if a person could be hurt by his or her own misunderstanding in one game but could not be hurt in the same way in the other game, then perhaps my assumption is unwarranted. Perhaps there would be a difference in a person's tendency to learn from mistakes. If so, then this suggests that the gap between money-back and no-money-back experiments may understate the magnitude of the assurance problem. (Of course, because my main conclusion here is that the assurance problem is empirically significant, this possibility is not a threat.)

One thing the Isaac, Schmidtz, and Walker design does not do is allow subjects to observe each other's contribution within any given period. To the extent that they care about what other subjects are doing, they must respond to other subjects on the basis of cumulative information from past periods. In a given period, there is no way for them to learn about other subjects' intentions or to signal their own. In this respect, the Isaac, Schmidtz, and Walker design is a poor model of the telethon approach to fundraising, for example, because contributors to a telethon scheme have access to continuously updated information on what other contributors have so far pledged. Frequently, potential contributors also have information about *challenge pledges*—contributions promised pending the materialization of a certain level of contributions from others. (The fire department might challenge the police department, for example.) Of course, the main purpose of Isaac, Schmidtz, and Walker's experiments (from my point of view) was to verify the empirical significance of the assurance problem, and to show that the problem is observationally distinguishable from the free-rider problem. Our primary intent was to test theory, not policy. We were not out to test the efficiency of telethons. I will discuss the difference between tests of theory and tests of policy after taking note of recently completed experiments that do provide information about the efficiency of the telethon approach.

5.6. Contribution Patterns in Real Time

Robert Dorsey (1992) employed the basic design and incentive structure of Isaac, Schmidtz, and Walker. Dorsey's design, however, incorporates a major advance. It alters the decision-making environment in order to determine whether the decision-making process is sensitive to temporal factors and to the associated changes in the kind of information people have when making decisions. Dorsey's design gave subjects three minutes within each period to enter or update their contributions to the public exchange, in light of changing information about what the rest of the group had so far contributed in that period. This added time element reflects the kind of information and opportunity to update contributions that decision makers often have.

A second advantage of Dorsey's design over Isaac, Schmidtz, and Walker's

Table 5.5. The effect of real time (contributions as a median percentage of groups' total endowments; mean percentages shown in parentheses)

	Increase-Only Rule	*Increase-or-Decrease Rule*
No provision point	19.6 (23.12)	8.1 (11.53)
Provision point = 108	43.5 (34.78)	27.2 (26.99)

Source: Dorsey (1992: 277). Reprinted by permission of Kluwer Academic Publishers.

is that the extra information with which it provides subjects can alleviate assurance problems to some extent without resorting to provision points combined with money-back guarantees. This is so because a subject can easily see, given what other subjects have contributed, how much he or she can afford to contribute and still be assured of making a profit. Therefore, another way of solving the assurance problem (up to a certain contribution level) is to impose a rule that subjects can only add to previous pledges, not subtract from them. When subjects can add to or subtract from previous pledges, the assurance problem is present in full force.

By adding a provision point, Isaac, Schmidtz, and Walker also introduced a focal point, which meant that they needed to control for the possibility that the focal point rather than the money-back guarantee caused the rise in contribution levels. Because Dorsey's design dispensed with the need for a provision point, this control problem did not arise for him. Nevertheless, although Dorsey was able to design a fundraising mechanism that solved the assurance problem without introducing a focal point, he still wanted to test the effect of adding a provision point to what he called a "real time" environment—one with a time element that permits subjects to update contributions within periods. Unlike Isaac, Schmidtz, and Walker, who found that the provision point had a real but equivocal effect on contribution levels in the simultaneous game, Dorsey found that incorporating provision points in the real-time game had a substantial positive effect on contribution levels.

Dorsey contrasted the increase-only rule to a rule by which subjects could either increase or decrease commitments made previously in that period. In effect, the increase-or-decrease rule makes pledges nonbinding. Under such a rule, the temporal aspect of the real-time game does nothing to solve the assurance problem, even for last-second contributors. As we would expect, Dorsey found that in this environment contribution levels were much lower, sometimes collapsing entirely at the end of the three-minute period.

Table 5.5 reflects data gathered by Dorsey from sixteen experiments with the increase-only rule and eight experiments with the increase-or-decrease rule. In each treatment, half the experiments were run with no provision point and half

Table 5.6. Comparing Dorsey's increase-only and increase-or-decrease rules with Isaac, Schmidtz, and Walker (ISW) (contributions expressed as a median percentage of group's total endowments)

	Increase-Only	*Increase -or-Decrease*	*ISW*	*ISW (money-back)*
Provision point = 108	44.4	26.6	30.2	42.7

with a provision point of 108 tokens (equal to the low provision point of Isaac, Schmidtz, and Walker).

What we see is that adding a provision point in the real-time environment dramatically increases median contribution levels. (Even under the increase-or-decrease rule, the median contribution level was 27.2 percent with the provision point but only 8.1 percent without.) We also see that contribution levels are significantly higher under the increase-only rule than under the increase-or-decrease rule. Median contributions were roughly twice as high under the increase-only rule. Moreover, the real-time element itself makes a difference. Table 5.6 compares the data from Dorsey with data from Isaac, Schmidtz, and Walker. Isaac, Schmidtz, and Walker ran six experiments with 108 token provision points and another six experiments that also included a money-back guarantee against failure to meet the 108-token provision point. The median percentage of total tokens contributed in these experiments is presented in Table 5.6 for purposes of comparison.

As Dorsey notes, "Isaac, Schmidtz, and Walker, using identical parameters but without real time adjustments, reported that with a provision point of 108, only one experiment out of six achieved the provision point after the fourth period. With the real time environment, the provision point was reached at least once after the fourth period in eight out of 12 of the ten-period experiments" (Dorsey, 1992: 278).

Dorsey's real-time game is a different game. It has different rules, induces real people to play different strategies, and leads to different results. I once argued that the temporal factor (under an increase-only rule) could induce people to contribute more to a public goods project than they otherwise might. Against this, David Friedman (1987) argued that rational players in a real-time situation would wait until the last second. In the limit, everyone waits until the last second—thus, no one has any information about what anyone else is planning to do. "We all make our decisions in ignorance of what the rest are doing. We are back in the simultaneous game" (Friedman, 1987: 519). Dorsey's experiments, however, show that this is not how the temporal factor works. It may not be obvious that the real-time element has a bearing on the dominant strategy equilibrium, but nevertheless it does. When subjects have the opportunity within periods to lever up each other's contributions, the dominant strategy equilibrium is replaced by Nash equilibria. When players can respond to each other within a period, that means that a player's best strategy within the period—in terms of both what to contribute and *when*—depends on how other player's are prepared to respond during the period.

How, then, do we explain the fact that contribution levels rose when intraperiod decisions were made in real time? For one thing, the real-time element allows subjects to see whether total contributions are below the provision point at any particular moment during the period, while there is still time to do something about it. If we think of total contributions as being like a missile launched at the provision point, then the temporal factor allows players to use their contributions to guide the total toward the provision point. Without the temporal factor that allows subjects to adjust contributions, total contributions are like a ballistic missile. After the initial decision, no further control is possible. The analogy breaks down, however, when we use the increase-or-decrease rule, for with that rule, previous contributions can be withdrawn at a moment's notice. Contributions already announced are nonbinding and thus convey no real information about what adjustments will be required to meet the provision point as the period draws to a close.

I also suspect the increase-only format effectively converted the intraperiod game from a one-shot prisoner's dilemma to something that could be played as an iterated prisoner's dilemma. That is, a subject could contribute one penny, then step back until total contributions rose to a level sufficient to repay a one-penny contribution (say, four cents). Then the subject could contribute one more penny, wait until total contributions rose to a level adequate as a reciprocation for the subject's two-cent contribution, then take this as a cue to initiate another round by contributing one more, and so on. When Dorsey sequentially plotted intra-period contributions, the resulting curve in some cases showed just the series of regular small steps we would expect from such a strategy (Dorsey 1992: 279).

Whether this kind of strategy characteristically underlies the success of telethon fundraisers is far from clear, but it may be a part of it. Moreover, aside from the aspect of reciprocity permitted by the real-time environment, there is also the prospect of being able to make a decision with an assurance about the minimum level of total contributions one may expect, and this presumably is a big factor in both the Dorsey experiments and in the telethons that these experiments model.

Dorsey's results also have a significance that goes beyond their implications as a model of telethon fundraisers. Consider, for example, how it might be used to model negotiations over disarmament during the Cold War. Suppose the

situation was as follows. (It might have been; I am not qualified to say.) The Soviet Union (circa 1988) had an approach to disarmament that suggests an attempt to capitalize on the incentives associated with the real-time element. They apparently were taking unilateral steps to disarm and demilitarize. As the United States government liked to stress, it is indeed true that these steps were very small. But small steps were precisely what the situation called for. The Soviets had a very pressing assurance problem and could not be expected to take larger steps unilaterally. If the cuts were genuine, then the United States might be (or might have been) best served by cutting its own military strength by slightly more than the Soviets cut theirs, so as to put the onus on the Soviets to maintain the momentum.

One major obstacle to successful negotiations in such situations is the problem of monitoring compliance. It is relatively easy to ensure that a given missile has been rendered inoperative; the big problem is to ensure that replacements are not secretly being built. But notice the structure of this monitoring problem. The problem is that we need to prevent playing by an increase-or-decrease rule. Both sides need assurance that the only available moves will be subtractions from the arms stockpile. Covert additions must be ruled out. The game will not run smoothly until we are in a position to play by a decrease-only rule with incentive properties equivalent to those of Dorsey's increase-only rule for monetary contributions.

For public goods problems involving a need to secure monetary contributions, a logical next step in the line of research begun by Isaac and Walker would be to compare a real-time provision point environment to one that added a money-back guarantee in the event the provision point was not met. Although the provision point induced greater contributions in the Dorsey experiments, it also featured an assurance problem because to initiate the incremental process of building up to the provision point, subjects had to accept the risk of never reaching it, thereby wasting whatever they had contributed. Solving this assurance problem with a money-back guarantee resulted in an increase of nearly 50 percent in median contributions (from 30.2 to 42.7 percent of total endowment) in Isaac, Schmidtz, and Walker. Combining the real-time and money-back features might lead to still higher contribution levels.

5.7. Theory Tests and Policy Tests

Experimental methods can be used to test both economic theory and economic policy. But as Charles Plott has pointed out, testing descriptive theory and testing prescriptive policy are two different things.[4] We must be careful not to confuse them because experimental methods serve different functions in these two roles and are subject to different problems. The function of a test of an economic prescription is to shift the burden of proof onto those who believe that a certain policy will have different effects in the rest of the world than it does in the laboratory. The result of such a test is by no means decisive, but if the policy does not work in the laboratory, this should at least lead us to question our reasons for prescribing the policy.

In contrast, the intent of a test of descriptive theory is to confirm or disconfirm a theory insofar as it yields testable predictions when applied to the experimental design. This is not to say that single tests are typically decisive. In fact, they seldom are. The point, rather, is that a descriptive theory can be confirmed or disconfirmed in a way that a prescriptive policy cannot.

For example, we might have a descriptive theory that people are exclusively self-interested. The data discussed in preceding sections, however, weigh heavily against the theory that people are exclusively self-interested. We may not want to say the data decisively refute the theory, but the theory is certainly misleading in a variety of laboratory situations in which people are called on to engage in collective action. At the same time, we might prescribe a policy of treating people as if they were exclusively self-interested when devising mechanisms for providing public services in large urban communities. Disconfirming the descriptive theory does not show that the prescriptive policy is false, or even that it is bad policy. Implementing the policy might, after all, yield just the results it was intended to yield in the particular kind of situation for which it was prescribed.

The pitfall to avoid in policy tests is careless extrapolation from the laboratory to situations that are different in kind. Laboratory simulations of large-scale policy problems are inevitably unrealistic. The possibility that some of the disanalogies will be relevant is inescapable. This is a problem inherent in any field of empirical research. But policy prescriptions are not intended to apply to laboratory situations. The laboratory policy test must be understood as essentially a *simulation* of the situation for which the policy was really proposed.

In contrast, the danger of inappropriate extrapolation does not arise in the same way for tests of theory. Disconfirmation in the laboratory is disconfirmation, period. We do not need to know if the experimental design is realistic: we just need to know if the theory applies to the design. One may suspect that the theory yields true predictions in other situations, but in any event, following up this suspicion requires one to modify the theory (preferably not in an ad hoc way) so that it is no longer systematically disconfirmed by what has been observed in the laboratory. Laboratory situations are real situations involving real people and real money. Theory tests in the laboratory are not "dry runs" in the way policy tests are. They are not mere simulations. Laboratory situations are themselves situations for which theories can have testable implications. Whether or not they

simulate the real world, they are capable of disconfirming a theory, provided the theory yields testable implications when applied to them.

The possibility remains, however, that the theory will not apply at all in this sense. This is the real problem with tests of economic theory. If, for example, subjects receive unintelligible instructions, or if the monetary incentives are insignificant, then most economic theories will say nothing about what behavior we should expect to observe. The problem will not be that the situation is unrealistic. On the contrary, confused subjects and insignificant monetary rewards are features of many real-world situations. The problem will simply be that the theory has no testable implications under such circumstances.

One particular cause for concern when evaluating these experiments as a test of policy is that the experiments use small numbers. The conventional wisdom is that cooperation is easier to achieve in groups with small numbers. Would the results we got in groups of four and ten also be found if we had used groups of 40 or 100? Recent tests by Isaac, Walker, and Williams (unpublished) suggest that they would.

Isaac, Walker, and Williams ran tests with groups of 40 and 100. A second interesting feature of some of these experiments is that the ten-period sequences were spaced out over several days. Subjects left the lab between periods. (The periods themselves were not in real time. Subjects received information about contributions by the rest of the group only after the period was over.) Thus, subjects had to make an effort to enter a contribution. Following the first period, they were informed that if they did not show up at the laboratory for a given period, a default contribution of zero tokens would be invested for them. This feature adds another touch of realism that should interest those wanting to derive policy implications, for surely real-world defections sometimes are due not to greed or fear but rather to laziness or perceived inconvenience. Some people fail to contribute money not because they do not want to contribute money but rather because they do not want to spend the time and energy that it takes to contribute money. In many real-world situations, free riding saves time.

Under these conditions, Isaac, Walker, and Williams (unpublished: 10) discovered that with a marginal per capita return of 0.30—the same rate of return used in Isaac, Schmidtz, and Walker's groups of four, "groups of size 40 and 100 allocate more tokens to the group account on average than do groups of size 4 and 10." The larger groups contributed 40 to 50 percent of their tokens in the early rounds and 35 to 40 percent of their tokens in round ten. Groups of four, in contrast, contributed 20 to 30 percent of their tokens in early rounds and 15 percent or less in round 10 (Isaac, Walker, and Williams, unpublished: fig. 6).[5] So these data are encouraging in terms of its policy implications for voluntary mechanisms in general, although the effect of money-back guarantees in particular on contribution levels in large groups remains to be seen.

5.8. Conclusion

Data from literally thousands of experiments show that many subjects contribute substantial amounts even when defecting is a dominant strategy. On the other hand, Isaac, Schmidtz, and Walker have also seen that subjects occasionally fail to contribute even when contributing is a dominant strategy. We could talk all day about what might have gone wrong with the experimental design, but what was really wrong was the subjects, and the subjects were typical.

Solving the assurance problem makes a substantial difference, but even a design that made optimal contributions a dominant strategy failed to elicit optimal contributions consistently. Still, the laboratory data yielded one interesting result that is clear: dominance as such is not what leads people to pick one strategy over another. The decision procedures people actually employ are much harder to characterize than that. Subjects contribute to public goods projects even when noncontribution is a dominant strategy. Their contributions, however, tend to be suboptimal because of free-rider, exploitation, and assurance problems. Also, some subjects probably fail to make optimal contributions simply because they fail to comprehend the request.

As a policy tool, money-back guarantees generally should, when feasible, be incorporated into attempts to fund public goods production by voluntary contributions. Of course, in many cases, it is fairly obvious that a certain fundraising effort will meet its target in any event. In these cases, there is no point in complicating the issue with a superfluous guarantee. In other cases, it is not at all obvious that the target will be met. For these cases, the money-back guarantee should be used, but the provision point must be chosen with great care. It must be set at a level of funding that could reasonably be considered both necessary and sufficient for the good in question to be successfully produced. In addition to the obvious point that a necessary and sufficient level of funding is the amount of funding we want in the first place, a provision point set at such a level will be taken more seriously by prospective contributors.

The disadvantage of setting a provision point at all is that if it is not met, the fundraisers end up with nothing. But if the provision point is set so that a funding level below it would be inadequate anyway, the disadvantage may not matter. When there is no natural provision point and the adequacy of funding is strictly a matter of degree, then the task for fundraisers is simply to maximize revenue. The evidence is that the combination of a well-chosen provision point and a money-back guarantee increases expected revenue. It does not guarantee increased revenues, however. Contributors, not fundraisers, get the guarantee.[6]

Notes

1. Associate Professor, Philosophy Department, Bowling Green State University, Bowling Green, OH 43403. This article expands and updates material originally appearing in Chapter 6 of David Schmidtz (1991), *The Limits of Government: An Essay on the Public Goods Argument*. Adapted by permission of Westview Press, Boulder, Co.

2. I borrow this point from Vernon Smith's address to the Public Choice Society annual meeting in 1988.

3. Note that if we were somehow to solve the free-rider problem that constitutes one essential component of the prisoner's dilemma on my analysis, we would be left with the assurance problem as the only reason for noncontribution. This game's payoff rankings would be that of an assurance game. That is, if free riding on other people's contributions was ruled out, we would have a situation in which people would prefer to contribute if and only if others were contributing as well. See Sen (1967).

4. This section is heavily indebted to Plott (1982). See also Roth (1986).

5. I do not know why average contributions were higher in the larger groups, but it occurs to me that, for a given marginal rate of return per contribution, constrained maximizers would tend to contribute more as group size rose. Constrained maximizers (as defined by David Gauthier) are players who contribute when reasonably certain that the resulting payoff will be higher than the baseline payoff of mutual defection. Note, then, that the group contribution needed to ensure a return in excess of the mutual defection payoff is not determined by group size. It is determined by the marginal rate of return. So if the marginal return for a dollar contributed is thirty cents, then a total contribution of $2.34 from the rest of the group lifts me above the baseline, regardless of how large the rest of the group is. If there are thirty-nine rather than three other people in the group, then I can be much more confident that there will be three other constrained maximizers (or contributors of any kind) among them. Thus, the presence of constrained maximizers is possibly why contributions rose with group size.

6. I want to thank Elizabeth Willott for helpful comments on an earlier draft of this essay.

References

Bagnoli, Mark, and Michael McKee. (1991). "Voluntary Contribution Games: Efficient Provision of Public Goods." *Economic Inquiry* 29, 351.

Dawes, Robyn M., John M. Orbell, Randy T. Simmons, and Alphons J. Van de Kragt. (1986). "Organizing Groups for Collective Action." *American Political Science Review* 80, 1171–1185.

Dorsey, Robert E. (1992). "The Voluntary Contributions Mechanism with Real Time Revisions." *Public Choice* 73, 261–282.

Friedman, David. (1987). "Comment: Problems in the Provision of Public Goods." *Harvard Journal of Law and Public Policy* 10, 505–520.

Gauthier, David. (1986). *Morals by Agreement*. New York: Oxford University Press.

Isaac, R. Mark, David Schmidtz, and James M. Walker. (1989). "The Assurance Problem in a Laboratory Market." *Public Choice* 62, 217–236.

Isaac, R. Mark, and James M. Walker. (1988). "Group Size Effects in Public Goods Provision: The Voluntary Contributions Mechanism." *Quarterly Journal of Economics* 103, 179–199.

Isaac, R. Mark, James M. Walker, and Susan Thomas. (1984). "Divergent Evidence on Free Riding: An Experimental Examination of Some Possible Explanations." *Public Choice* 113–149.

Isaac, R. Mark, James M. Walker, and Arlington W. Williams. (Unpublished). "Group Size and the Voluntary Provision of Public Goods: Experimental Evidence Using Large Groups."

Plott, Charles R. (1982). "Industrial Organization Theory and Experimental Economics." *Journal of Economic Literature* 20, 1485–1527.

Roth, Alvin E. (1986). "Laboratory Experimentation in Economics." *Economics and Philosophy* 2, 245–273.

Schelling, Thomas. (1960). *The Strategy of Conflict*. Cambridge: Harvard University Press.

Schmidtz, David. (1991). *The Limits of Government*. Boulder: Westview Press.

Sen, Amartya. (1967). "Isolation, Assurance, and the Social Rate of Discount." *Quarterly Journal of Economics* 81, 112–124.

Jeff