CREMA

Center for Research in Economics, Management and the Arts

Rewarding Honest Taxpayers? Evidence on the Impact of Rewards from Field Experiments

Lars P. Feld

Bruno S. Frey

Benno Torgler

Working Paper No. 2006 - 16

REWARDING HONEST TAXPAYERS? EVIDENCE ON THE IMPACT OF REWARDS FROM FIELD EXPERIMENTS

by

Lars P. Feld, Bruno S. Frey and Benno Torgler*

Abstract: This paper analyzes the impact of rewards on tax compliance as an additional instrument to take into account. While social psychologists and neuroscientists have emphasized the importance of rewards, the tax compliance literature has strongly disregarded the possibilities of rewards. The use of field experiments presents an alternative "carrot" strategy for tax policy. Design mechanisms to conduct a field experiments focusing on the impact of rewards on tax compliance are discussed.

JEL classification: *H260, H410, C900, D630*

Keywords: *positive rewards, tax compliance, tax evasion, field experiments*

* Lars P. Feld, Philipps-University of Marburg, Public Finance Group, Am Plan 2, D-35037 Marburg (Lahn), Germany, Feld@wiwi.uni-marburg.de; Bruno S. Frey, Institute for Empirical Research in Economics, University of Zurich. Bluemlisalpstrasse 10, CH-8006 Zürich, Switzerland, e-mail: bsfrey@iew.unizh.ch; Benno Torgler, The Whitney and Betty MacMillan Center for International and Area Studies at Yale, Leitner Program in International & Comparative Political Economy, New Haven, CT 06520 (USA), e-mail: benno.torgler@yale.edu. All three authors are also associated with CREMA – Center for Research in Economics, Management and the Arts, Switzerland. The authors have benefited from the comments of participants of the Conference on "Managing and maintaining compliance", Leiden, April 9-11, 2006 (especially Henk Elffers, Erich Kirchler and Willy Rovers).

I. INTRODUCTION

Why do people pay taxes? This question has attracted increased attention in the tax compliance literature over the last few years. Allingham and Sandmo (1972) presented a formal model wishowing that the extent of tax evasion is negatively correlated with the probability of detection and the degree of punishment. However, this seminal model has been criticized by many authors (see, e.g., Graetz and Wilde 1985, Alm, McClelland and Schulze 1992, Frey and Feld 2002). A main point connected to the empirical and experimental findings, is that these deterrence models predict far too little compliance and far too much tax evasion (for an overview see Alm 1999 and Torgler 2002). In many countries the level of deterrence is too low to explain the high degree of tax compliance. Moreover, there is a big gap between the amount of risk aversion that is required to guarantee such a compliance and the effectively reported degree of risk aversion. For the United States, the estimated Arrow-Pratt measure of risk aversion is between one and two, but only a value of 30 would explain the observed compliance rate (see Graetz and Wilde 1985, Alm, McClelland and Schulze 1992). Similarly, in Switzerland the relative risk aversion varies between 1 and 2, but a value of 30.75 would be necessary to reach the observed level of tax compliance of 76.52 percent (see Frey and Feld 2002).

Elffers (2000) shows that it is a long way before a person becomes a tax evader. He defines three steps in the staircase to tax evasion: (i) taxpayers have to have the will not to comply, (ii) not everyone with the inclination to evade taxes is able to translate the intention into action, and (iii) individuals inclined to evade taxes check for the opportunity to do so. In the third staircase standard economic theory comes into play and individuals evaluate the expected value of evasion. Similarly, other researchers argue that many individuals do not even think of tax evasion. Pyle (1991) criticizes the assumption that individuals are amoral utility maximisers: “Casual observation suggests that not all individuals think quite like that.

Indeed, it seems that whilst the odds are heavily in favor of evaders getting away with it, the vast majority of taxpayers behave honestly" (p. 173). Frey (1999) uses the expression "ipsative possibility set" (p. 196) and shows that there are taxpayers who do not even search for ways to cheat at taxes. Long and Swingen (1991, p. 130) argue that "some individuals are simply predisposed NOT to evade". Experiments indicate that there are individuals who always comply, that is, a certain compliance exists even without (low) penalties and audits (Feld and Tyran 2002).

In general, Elffers (2000) suggests to reduce the significance of coercive instruments to resolve the social dilemma of tax payments. His conclusion ('policy advice') is to try to prevent people from reaching the final step of the staircase. Thus, the instrument of deterrence is not the only instrument to make individuals comply. The theoretical models of individual choice using the economics of crime approach are too simple. There are numerous factors that affect the reporting decision of individuals. The Internal Revenue Service (1978) listed 64 potential factors that could affect tax compliance. Governments and tax administrations have an incentive to search for tax policy strategies that generate additional revenues, especially in times with large and persistent deficits. There is a persistent theme in the tax compliance literature in the last few years to move away from deterring noncompliance toward positive encouragement for compliance and therefore emphasizing "the "carrot" for compliance rather than the "stick" for noncompliance… This insight is especially important because, from the tax collection standpoint, it is extraordinarily expensive to arrange an enforcement regime so that, from a strict cost-benefit calculus, noncompliance does not appear attractive to many citizens" (Slemrod 1992, p. 7).

This paper focuses on rewards, which may influence individuals' compliance behavior as a "carrot". Instead of raising the relative cost of not paying taxes, the instrument of rewards raises the benefits of paying taxes. Currently, there is limited amount of empirical and experimental evidence that investigates in detail the impact of positive rewards on tax

compliance. Section II introduces the concept of positive rewards and provides an overview of the current literature. Field experiments may provide a good tool to generate evidence about the impact of positive rewards. Section III discusses the pros and cons of field experiments, summarizing also the results obtained in the limited amount of previously conducted field experiments in the area of tax compliance. Section IV provides possible design mechanism for such a field experiment focusing on rewards and Section V finishes with some concluding remarks.

II. THE IMPORTANCE OF REWARDS

It is a relatively novel approach to investigate the impact of rewards on tax compliance. There exists some anecdotal evidence about the implementation of rewards to enhance tax compliance, especially in Asian countries. For example, Japan offers to have your picture taken with the Emperor if you were found to be honest. The Philippines put your name into a lottery if you were found to be compliant with the VAT. South Korea considers the allowance to airport VIP rooms, certificates or awards, and discusses the possibilities of free parking in public parking facilities[1].

Rewards could be more effective than punishments for eliminating undesired behavior or for motivating desired behavior because it is perceived as supporting (see, e.g., Nuttin and Greenwald 1968). Indeed, the role of rewards in shaping human and also animal behavior has long been a subject among social psychologists (see, e.g., Thorndike 1911, 1932, Postman 1947, Skinner 1953, Nuttin and Greenwald 1968). Early exchange theorists excluded punishment from the scope of social exchange relations (see, e.g., Blau 1964, Homans 1974).

[1] We are thankful to Jim Alm and Hyung-Wook Kang for providing us these anecdotes.

Punishment seemed to be less effective than reinforcement (Estes 1944, Skinner 1938, Thorndike 1932).

Molm (1988) criticizes that these forms of power have been studied largely by separate scientific disciplines such that little is known about how they interact with one another and what their strengths and weaknesses are under equivalent conditions. Molm (1994) reports that a series of experiments comparing reward-based power with punishment-based power in nonnegotiated exchange relations in which all actors have the capacity to reward and punish their exchange partners, the effects of punishment power are consistently weak: “The distribution of exchange is almost entirely a function of reward power; actors with greater power to punish do not receive increased benefits from their exchange partners” (p. 75). Sims (1980, p. 136) summarized the literature on punishment in organizations focusing on cross-sectional and longitudinal psychometric research studies undertaken in both laboratory and field settings stating that some preliminary conclusions indicate that in most of the studies rewarding behavior tends to have a much stronger effect on subordinate performance. Several areas of psychology and organizational behavior suggest an asymmetrical effect of rewards and punishment; they are therefore not equally efficient to influence workers’ behavior such as, for example, reducing loafing (see, George 1995). Such an asymmetrical effect of rewards and punishments are supported from neuroscience. Studies suggest that rewards and punishment are processed in different systems in the brain and therefore have differential effects on behavior (Gray 1981, Larsen and Katelaar 1991).

To the authors’ knowledge, there is only one detailed theoretical study in economics (Falkinger and Walther 1991) analyzing the possibility of pecuniary rewards as an economic incentive for taxpayers to be honest. In their model a taxpayer who is investigated has to pay a penalty for the evaded tax and receives a reward for the paid tax. The authors show that on the one hand a mixed penalty-reward system improves the taxpayer’s position and on the other hand does not lower the tax revenues of the government. Thus, introducing rewards coupled

with an increase of the penalty constitutes a welfare improvement. This study shows that the analysis of rewards might be an important topic in the tax compliance literature. A rational choice approach would allow to consider the impact of both, rewards and sanctions. However, investigations on illegal activities solely emphasize the deterrence by sanctions. For example, it can be argued that sanctions can be problematic and damaging even in dealing with terrorism (Frey 2004). It is highly relevant to consider possible effects of rewards on tax compliance behavior and thus move beyond standard theories of tax evasion.

In psychology and behavioral economics *crowding out* and *crowding in* effects have received considerable attention (Frey 1997, Le Grand 2003, Bénabou and Tirole 2003, Fehr and Rockenbach 2003, Falk and Kosfeld 2006). The theory suggests that outside interventions perceived to be controlling such as deterrence tend to crowd-out intrinsic motivation. On the other hand actions perceived to be supporting tend to crowd-in intrinsic motivation. A punishment for not behaving as a "good" taxpayer is felt to be controlling in particular if the charges raised do not fully apply (Feld and Frey 2002). Taxpayers who are mistakenly accused to cheat on their taxes may perceive the intervention by the tax office to be controlling. Therefore their tax morale decreases or even erases. Similarly, increasing monitoring and penalties for noncompliance, individuals notice that extrinsic motivation has increased, which on the other hand crowds out intrinsic motivation to comply with taxes. Thus, the net effect of a stricter tax policy is unclear. If the intrinsic motivation is not recognized, taxpayers get the feeling that they can as well be opportunistic. This puts the relevance of policy instruments in supporting or damaging the intrinsic motivation to the fore. Intrinsic motivation depends on the application of policy instruments. Tax morale is not expected to be crowded out if the honest taxpayers perceive the stricter policy to be directed against dishonest taxpayers. Regulations which prevent free riding by others and establish fairness and equity help preserve tax morale. In contrast, receiving certain types of rewards for being a good taxpayer may be perceived as supporting and tends to bolster and raise tax

morale. This motivational effect thus works in the same direction as the relative price effect, and strengthens the attractiveness of giving rewards to "good" taxpayers. In the case of the normally applied punishment for failing to pay the taxes due the relative price effect and the motivational crowding-out effect work in opposite directions. This may explain why the empirical evidence on the effect of punishment on tax evasion is inconclusive, and the respective econometrically estimated parameters often are not statistically significant, or are even of the wrong sign. If the crowding-out effect is stronger than the relative price effect of punishment, tax evasion is raised rather than lowered.

Currently, two experiments have investigated the impact of rewards on tax compliance. Both allow to a certain extent to analyze the impact of positive rewards in relation to other tax policy strategies. In both cases, support has been found that rewards are a very strong policy instrument to enhance tax compliance. Alm, Jackson and McKee (1992) use experiments to analyze the effects of positive inducements upon tax compliance behavior. They designed: i) a lottery treatment where subjects who were checked and found to be fully compliant for the current and the previous four rounds could enter in a lottery in which the chances of winning were 1 in 25 and with a lottery prize that was roughly equal to the average earnings of a subject for the entire session (50 tokens), ii) a fixed reward session where fully compliants received a reward of 2 token which was equal to the expected value of the lottery, iii) an audit reduction session, where an individual found to be compliant has the future probability of audit reduced from 0.04 to 0.027 on the first occurrence of an audit and from 0.027 to 0.013 on the second occurrence. In addition they introduced a public good session, where the public good is created by sum of the taxes collected from the group in a given period, multiplying this sum by 2 to reflect the consumers' surplus generated by the public good and then dividing the amount equally among the five people in the group. The results indicate that positive inducements have a significant and positive impact on compliance.

However, although i) and ii) have the same expected value, the lottery session had the largest effect on compliance.

Torgler (2003a) conducted an experiment in Costa Rica with taxpayers, holding traditional factors, such as the probability of detection and the fine rate, constant and thus analyzing to which extent other factors as fiscal exchange, moral suasion, and positive rewards systematically influence tax compliance. The findings indicate that these factors ceteris paribus increase the compliance rate. In the positive reward session, a subject audited and found to be fully honest received a monetary reward. Such a reward can also be seen as a compensation for the burden of investigation which the taxpayer has to pass if he or she is audited (see Falkinger and Walther 1991). Interestingly, the highest tax compliance rate was found in the positive reward session followed by the moral suasion session and the fiscal exchange treatment. It seems that the norm of reciprocity in the degree of tax compliance is followed by taxpayers where the government creates positive rewards or a fiscal exchange. The more the governments provide public services corresponding to taxpayers' preferences in exchange for an adequate tax price, and the more they honor honesty, the more taxpayers are willing to comply. These results support the previous findings of Alm, Jackson and McKee (1992) that positive incentives seem to be a good instrument to enhance tax compliance.

In sum, laboratory experiments enable a good research design to continue the investigations of rewards on tax compliance. Alternatively, the research design of field experiments is, as we will discuss now, an excellent tool to investigate the impact of rewards on compliance.

III. THE USE OF FIELD EXPERIMENTS TO INVESTIGATE ALTERNATIVE TAX POLICY STRATEGIES

Whereas much work in the tax compliance literature has concentrated on standard factors as audit, penalty, and tax rate, it is useful to evaluate alternative policy instruments. Field experiments might be a useful approach to investigate the relevance of such instruments.

Using controlled field experiments has many advantages. Compared to laboratory experiments, real tax authorities instead of experimenters are involved, which evokes real processes in the usual environment outside a laboratory setting. It helps to better test the effects of different instruments on taxpayers in the real situation of filling out the tax form and paying the taxes. This helps to formulate practical advice on tax policy, based on a scientific test. Certainly, compared to lab experiments, field experiments allow social and economic interactions and are thus less controlled, but causality can be better determined than in non-experimental studies (see Burtless 1995 and Harrison and List 2004 about the advantages and problems of field experiments).

There is no observable effect of an artificial experimental environment, as subjects were completely unaware of having taken part in the field experiment. The experiments are thus conducted in the usual environment where social and economic interactions occur (see Burtless 1995). This has the advantage that the subject pool is more representative than in laboratory experiments. The results have a strong policy implication and might be relevant for policymakers. However, it is surprising that there is hardly any field experiments in the tax compliance literature. Higher transaction costs in organizing a cooperation between the tax administration and the researchers, compared to laboratory experiments, and the sensitivity of the tax filing data according to privacy protection laws might be reasons why field experiments are less frequently used. Field experiments consume a great deal of real resources. First cooperation between the tax authorities must be established. It is difficult to

develop and implement a treatment, as it has to be approved by the tax administration and other government authorities. Thus, it may be supposed that sensitive or unorthodox treatments cannot be developed. Secondly, compared to laboratory experiments such experiments are costly in terms of time. The experiment has to be prepared before individuals receive their tax forms. It takes almost a year until all tax forms are returned to the tax administration and are thus ready to be evaluated. Moreover, field experiments are limited in duration. While experiments can analyze inter-temporal aspects, field experiments are normally conducted only once. For some questions it might be interesting to analyze to which extent a policy instrument works over time. A short-duration intervention might have an immediate effect, but long-term effects are unknown. Furthermore, questions as what would happen if a policy instrument as, e.g., moral suasion were used regularly, remain unanswered.

To our knowledge, only few studies have used this instrument. Slemrod, Blumenthal and Christian (2001) use a controlled field experiment in Minnesota to analyze taxpayer response to an increased probability of audit. 1724 randomly selected taxpayers were informed by letter that the return they were about to file (state and federal) would be closely examined. They used 2 years' income return data from the same taxpayers which enabled them to compare changes in reported income, deductions and tax liability between those taxpayers who received the treatments and similar groups of taxpayers who were not subject to any treatment. They found that the treatment effect varies depending on income. In the treatment group, low and middle income taxpayers increased their reported income between 1993 and 1994 relative to the control group. The effect was much stronger for those with a higher opportunity to evade. In 1994, the reported income of high income taxpayers fell sharply in relation to the control group. The perception that tax evasion will not be detected and punished automatically, according to the authors could be a reason for these results and thus they propose that "heightened audit threat should be carried out simultaneously with a rethinking of how the audits themselves are carried out" (p. 482). As the authors state, the

analysis had a comparably small sample size of high-income taxpayers, which reduces the inference to be drawn. Follow-up experiments should start the field experiment at the beginning of the tax year to analyze avoidance behavior as well.

Similarly, Blumenthal, Christian and Slemrod (2001) worked also together with the Minnesota Department of Revenue to analyze the impact of moral suasion on voluntary income tax compliance with a field experiment. They used the difference-in-difference approach with data for the tax years 1993 and 1994. Compliance behavior was measured by the income reported or the tax paid and was compared with the reference group (no communication). They found that the average compliance rate of those in the treatment group was 220$ higher compared to the control group (0.08 percent of average income). However, the coefficient was not statistically significant. Hence, this study did not find a significant effect of moral appeals. In a second step, Blumenthal et al. (2001) conducted a multiple regression in which they used the treatments as dummy variables to check other variables. The results indicate that people with greater opportunities to evade or avoid taxes (e.g., the self-employed) are less susceptible to normative appeals.

With a similar approach Torgler (2004) analyzes the effects of moral suasion focusing on two different compliance variables: the timely paying and filling out of the tax form. In cooperation with a local tax administration in Switzerland a controlled field experiment with taxpayers was undertaken. Contrary to the previous controlled experiment done by Blumenthal et al. (2001), which found little or no evidence of a positive effect of normative appeals on tax compliance, Torgler (2004) chose to cooperate with a *local* tax administration, because moral suasion efforts might be more effective at the lower government level. The results show that the moral suasion treatment group has a higher compliance rate than the reference group. The findings also indicate an increasing effect over time in the treatment groups. In general, the strongest treatment effect was observed for the variable that measured taxpayers' payment morality. However, the difference-in-differences approach and the

multivariate regressions indicate that the treatment effect was not statistically significant. Thus, results are in line with the Blumenthal et al. (2001) findings indicating that moral suasion has hardly any effect on taxpayers' compliance behavior.

IV. DESIGNING A FIELD EXPERIMENT TO INVESTIGATE THE IMPACT OF REWARDS ON TAX COMPLIANCE

This section discusses different aspects of a possible field experimental design on rewards. Compared to previous experiments, field experiments offer the great chance to observe the taxpayers' behavior in a non-artificial environment using a representative sample of taxpayers and working with relatively large samples. As we will see, there are specific aspects that are challenging when designing a field experiment on the impact of rewards

Topic

A field experiment requires the cooperation between the tax administration and the researchers. Using a field experiment in the tax compliance area is faced with many restrictions. First of all, the sensitivity of the tax filing data reduces the incentives of the tax administration to cooperate in such a project. Contrary to a lab experiment, a field experiment has to be realistic. It is for example highly problematic to develop treatment designs that do not correspond to official (tax) law. This reduces the possibility set of conducting experiments. Thus, traditional parameters such as the tax rate are hardly an instrument to investigate in a field experiment. However, alternative tax policy strategies such as positive incentives might be more attractive for a field experiment to investigate as they are less affected by the restrictions the tax administration encounters.

Location

Field experiments can in principle be undertaken in any country. However, some circumstances may provide specific advantages. For example, Switzerland is particularly suited to conduct a controlled field experiment at the local level because of the highly decentralized Swiss tax system. Basically, the Swiss cantons have the power to tax personal and corporate income. Direct taxation at the federal level is only derived from the basic power of the cantons. The local jurisdictions levy surcharges on cantonal income and wealth taxes. More importantly in our context, the Swiss cantons and local jurisdictions collect taxes in a very decentralized fashion. Tax evasion and tax fraud are regulated by cantonal tax codes and only recently and lightly harmonized by federal law. While local jurisdictions cope with the 'normal cases' and are thus in contact with the majority of taxpayers, cantonal tax administrations cover the 'special cases' and collect taxes from only a minority. A controlled experiment is thus feasible and credible in Switzerland. Moreover, it can be done with quite low costs and may reduce the transaction costs of the interaction between the tax administration and the researchers. A decentralized tax system furthermore allows to conduct field experiments in different environments and to check whether the environment also affects the tax policy instrument in a well-controlled "real" environment.

Lottery, Monetary Payment and Non-Monetary Rewards

There are various possibilities to apply rewards for paying taxes. They may range from direct monetary payments, participating in a lottery offering a sizeable sum of money, to getting different kind of gifts. A field experiment can be conducted in order to test whether rewards are indeed an effective and low cost instrument to reduce tax evasion. It can be expected that the reward structure systematically affects tax compliance.

As mentioned, Alm, Jackson and McKee (1992) investigated four different forms of positive inducements in their laboratory experiment. The lottery had the highest average compliance rate among all sessions (0.513) followed by the fixes reward session (0.448) and the audit reduction session (0.369). In all cases compliance was statistically significant and higher than the baseline case of 0.332. Interestingly, the lottery mechanism lead to a higher compliance than the fixed reward session, even though their expected returns are identical. Two aspects are essential for rewarding taxpayers via random allocation. It induces the *chance* of getting rewarded and allows for relatively *high rewards*. Both factors may produce supportive effects. Uncertainty and unpredictable rewards generate the attention, which is enforced by the larger size of rewards. A large prize with a low probability of success is more attractive than a smaller, more certain prize. Such an effect was recently supported by evidence from neuroscience showing that a reward schedule in which subjects knew the outcome in advance produced only modest dopamine transmissions (that are responsible for behavioral responses) while an unpredicted monetary reward produced significant dopamine transmissions (see Zald et al. 2004).

We may also observe differences between direct monetary payments and rewards in non-monetary form. Direct monetary payments can be proportional to the size of the tax payment (i.e. a percentage rebate), or in the other extreme may take the same size for all "good" taxpayers. The relative price effect is larger in the first case, but this beneficial effect may easily be overcompensated by a crowding-out effect. A reward received in the same monetary dimension as the tax payments is likely to be discounted by the taxpayers as a "claim", and then does not positively influence tax compliance. In contrast, a reward on purpose distinguished from the taxes due tend to be perceived as a sign of acknowledgement. If this is indeed the case, it is even better to provide a reward in *non-monetary* form. The idea of a gift may emphasize the exchange relationship between taxpayers and the state and therefore enhance reciprocity that positively affects social exchanges (e.g., Falk and

Fischbacher 2006, Fehr and Gächter 2000). It is a sign of appreciation that may work more strongly than just a reduction in taxes. Gifts can have different dimensions, for example, better and less costly access to public services (in the case of private taxpayers for instance the receipt of a voucher for public transport), free entry to cultural activities in the neighboring area, more favorable access to government services entering public museums and similar institutions), free entrance in recreation areas, food coupons for local festivities etc. The *way* rewards are handed out to "good" taxpayers is essential for its effects on taxpayer behavior and therefore different treatments should be included in a field experiment.

Income Effect

In general, a given reward for correctly fulfilling their duties changes the relative prices in favor of paying taxes and against evading them. However, it requires that the income effect induced by a higher wealth position does not work in the opposite direction. In general, the effect of income on tax compliance is difficult to assess as it depends, for example, on risk preferences and the progression of the income tax schedules. However, the reward is very small in relation to the tax liability so that a possible income effects tends to be small.

Detection of a "Good" Taxpayer

A key aspect in such a field experiment is how to detect a "good taxpayer". The experiment and also the use of a rewarding system strongly depend on tax administration's assessment of a good taxpayer. The reputation of the tax administration may decrease, if (notorious) tax evaders are rewarded by mistake. A lack of adequate assessment reduces therefore the strength of a rewarding system. Besides that, there are different levels of compliance such as declaring all the income, not overstating any deductions, sending back the tax form on time with all the required documents, paying the taxes due etc. Thus, it may be interesting to use different compliance variables when investigating the impact of rewards on compliance.

However, this leads to higher costs of using rewarding schemes as a tax policy instrument and may also lead taxpayers to behave strategically. The optimal scenario would be that the tax administration takes a detailed look at all taxpayers involved in the field experiment. However, such a procedure cannot be expected. Endogenous audit selection rules can also bias field experiments. Tax agencies seldom select tax returns randomly for audit but use instead information from the returns to determine audit[2]. So in this case the probability of an audit is endogenous, depending on the behavior of taxpayers and tax agencies. Thus, it is important to use the same audit probability for all the different experimental groups. Possible biases can be controlled through representative samples of taxpayers in all experimental groups.

Size of the Reward

According to standard economic theory rewards are expected to change the relative prices such that paying taxes becomes a more attractive alternative compared to evading taxes. This does, however, not necessarily mean that the effect is so large that it can be *identified empirically*. This holds if the reward given is of *small* size. The tax administration faces the trade-off between the costs and benefits of giving rewards compared to the costs and benefits of other incentives, in particular with the cost involved with punishment. To be cost effective, rewards must raise *net* tax revenues, i.e. gross revenues after deducting the cost of rewards.

Strategic Behavior

Rewards can induce strategic behavior by the taxpayers. For example, if rewards are provided due to behavioral *changes* (e.g., depending on the reduction of evasive behavior), it could be

[2] Internal Revenue Service (IRS), e.g., uses the Discriminant Index Function (DIF, formula) based on items reported on current tax returns in its selection of returns (Alm, Cronshaw and McKee 1993). Other countries follow similar practice (Roth, Scholz and Witte 1989).

rational to increase tax evasion in a first step in order to reduce it in a second step to generate higher benefits from the rewards. Thus, it is relevant to make the rewards dependent on whether a taxpayer is fully honest or not. This would reduce the incentives of behaving strategically. However, it should be noted that rewards therefore have the tendency to increase compliance primarily by altering the frequency of extreme behaviors, shifting possibly individuals from very low to very high compliance rate (Alm, Jackson and McKee 1992). On the other hand, there are taxpayers that do not even search for ways to cheat on taxes. Their behavior does not respond to changes in the tax policy parameters and therefore to the relative price effect, and is therefore not subject to marginal but rather an absolute evaluation (see Frey 1997 and Long and Swinger 1991). Thus, relative price changes, as a reason of higher punishment or a higher reward, are only considered by cheating taxpayers with a relatively low tax morale.

Control Group

If the design of the field experiment includes a control group, the taxpayers in the control group are excluded from the possibility to obtain rewards. This may induce some problems, especially when the field experiment is conducted at the local level. It can be expected that taxpayers discuss this issue among each other. Individuals in the control group may get emotional as they are not treated equally. Thus, compliance may decrease in the control group which leads to biases when comparing the treatment groups with the reference group (stronger rewards effects). On the other hand, a field experiment without a control group would reduce the strength of conducting a difference-in-difference approach as such a design would only allow to observe changes over time between the different reward groups. As an alternative, another community can be used as a reference group. However, besides the difficulties to convince another community to cooperate in such a project, a fully comparable community is difficult to find.

Design of the Letter

It may be useful to send the letter to taxpayers suggesting the chances to be rewarded in a separate envelope and not together with the tax documents to increase the probability that taxpayers (who use professional assistance) read the letter. For example one could choose a colored sheet (e.g. pink) so that individuals become better aware of it (see Torgler 2004). Furthermore, the chief tax administrator in person could sign the letter to reduce anonymity. The style (easy to read and to understand) and an adequate letter length (not too long) should be chosen to make it easier to capture the taxpayers' attention. Otherwise individuals are likely not to read the letter to the end. To reduce biases, individuals should not be informed that they are selected randomly for a tax compliance study.

Time Effects

Alm, Jackson and McKee (1992) stress that rewards must be both immediate and salient to have a quantitatively significant effect. The reward sessions indicate that there is a tendency that compliance decreases over time. Thus, long-term effects should be taken into account which suggests the relevance of observing the panel of taxpayers over a certain time period. There is the danger that taxpayers get used to the chance of obtaining rewards. A one period field experiment may catch a certain "surprise effect" that disappears over time. Furthermore, a random audit selection system induces additional problems. If only a limited amount of good taxpayers are evaluated and rewarded, it is possible that previously rewarded taxpayers are not rewarded in the future. What sort of reaction can be expected from these taxpayers? Additionally, what happens if the reward system is not established after the controlled field experiment? Tax administrations could fear possible negative effects and oppose such a field experiment in advance. It is also interesting to check whether some sub-groups of taxpayers react differently across time.

Questionnaire

The field experiment can be accompanied by an ex post experiment questionnaire. For many years many experimentalists (especially economists) disregarded the possibility to combine surveys with experiments. Certainly, the questionnaires should be designed so that questions with a strong relation to the previous experiment should be avoided, because of possible biases. Nothing speaks against the use of surveys, which allow to generate several control variables. It also allows to compare tax compliance results from the field experiment with information about subjects' attitudes toward paying taxes. To resolve this puzzle of surprisingly high tax compliance, many researchers argue that tax morale can help us to explain the high degree of tax compliance. However, there is only a limited amount of studies that have investigated in detail the correlation between tax morale and tax evasion (see, e.g., Weck 1983 and Torgler 2003b). Thus, data generated with field data also allows to provide more evidence to which extent tax morale affects tax compliance

Persons versus Firms

Different subject groups may react differently to a reward system. Similar behavioral responses could be expected also for firms, as in firms, also individuals decide about the level of tax compliance (Fehr and List 2004). However, firms are subject to important additional constraints due to the competitive environment they are acting in. This produces incentives among the decision makers to quickly discount a monetary reward into total tax liability. In such a case, only the relative price of rewards would work. Nevertheless, also non-monetary rewards may have a high attraction to firms. It may be useful to generate a reward to the firm per se rather than to specific leaders. Especially in complex firm structure, it is difficult to find the adequate reward system that considers the value of the individuals in a firm. Providing some with a relative advantage to others may lead to different kinds of emotions.

One useful form would be that the tax office issues a *certificate* indicating that the taxes, to the best of their knowledge, have been correctly declared, that the firm has been cooperative, and that the taxes due have been paid in time. Such a certificate demonstrates that the firm acts as a "good" taxpayer. Firms' reputation and image increase. Shareholders may respond in a positive way, raising share prices; the firm may get more favorable conditions on the capital market; and the customers' trust in the firm's products may increase. Thus, a field experiment that is able to differentiate between individual taxpayers and firms would allow to generate additional insights.

IV. CONCLUSIONS

This paper analyzes the impact of rewards on tax compliance as an additional instrument to punishments. While social psychologists and neuroscientists investigated the impact of rewards in detail, the topic is novel in the area of tax compliance. We suggest that field experiments are highly relevant to investigate a variety of strategies that governments and tax administrations can pursue to increase tax compliance. Rewards could be an effective tool to increase compliance. Two previous laboratory experiments show that compliance increases significantly when individuals found to be compliant are rewarded for their honesty. We propose design mechanism to investigate the impact of rewards on compliance with field experiments. We also show that some institutional factors enhance the advantages of using such a field experiment. For example, Switzerland seems to provide for a useful case study due to the highly decentralized Swiss tax administration and tax law. In sum, we believe that future tax compliance studies should pay more attention to the impact of rewards also taking into account the largely neglected tool of field experiments.

REFERENCES

Allingham, M. G. and A. Sandmo (1972). Income Tax Evasion: A Theoretical Analysis, *Journal of Public Economics.* 1: 323-338.

Alm, J. (1999). Tax Compliance and Administration, in: W. Bartley Hildreth and James A. Richardson (eds.), *Handbook on Taxation*. New York: Marcel Dekker: 741-768.

Alm, J. , G. H. McClelland and W. D. Schulze (1992). Why Do People Pay Taxes?, *Journal of Public Economics.* 48: 21-48.

Alm, J., B. Jackson, and M. McKee (1992). Deterrence and Beyond: Toward a Kinder, Gentler IRS, in: J. Slemrod (ed.), *Why People Pay Taxes. Ann Arbor*: University of Michigan Press: 311-329.

Alm, J., M. B. Cronshaw and M. McKee (1993). Tax Compliance with Endogenous Audit Selection Rules, *Kyklos.* 1: 27-45.

Bénabou, R. and J. Tirole (2003). Intrinsic and Extrinsic Motivation, *Review of Economic Studies*. 70: 489-520.

Blau, P. (1964). *Exchange and Power in Social Life*. New York: Wiley.

Blumenthal, M., C. Christian and J. Slemrod (2001). Do Normative Appeals Affect Tax Compliance? Evidence from a Controlled Experiment in Minnesota, *National Tax Journal*. 54: 125-138.

Burtless, G. (1995). The Case for Randomized Field Trials in Economic and Policy Research, *Journal of Economic Perspective*. 9: 63-84.

Elffers, H. (2000). But Taxpayers Do Cooperate!, in: M. Van Vught, M. Snyder, T. R. Tyler and Anders Biel (eds.), *Cooperation in Modern Society. Promoting the Welfare of Communities, States and Organizations*. London: Routledge: 184-194.

Estes, W. K. (1944). An Experimental Study of Punishment, Psychological Monographs. No. 263.

Falk, A. and M. Kosfeld (2006). Distrust – The Hidden Cost of Control, forthcoming in: *American Economic Review*.

Falk, A. and U. Fischbacher (2006). A Theory of Reciprocity, *Games and Economic Behavior*. 54: 293-315.

Falkinger, J. and H. Walther (1991). Rewards versus Penalties: On a New Policy Against Tax Evasion, *Public Finance Quarterly.* 19: 67-79.

Fehr, E. and S. Gaechter (2000) Fairness and Retaliation: The Economics of Retaliation, *Journal of Economic Perspectives*. 14: 159-181.

Fehr, E. and J. A. List (2004). The Hidden Costs and Returns of Incentives – Trust and Trustworthiness among CEOs, *Journal of the European Economic Association*. 2: 743-771.

Fehr, E. and B. Rockenbach (2003). Detrimental Effects of Sanctions on Human Behavior, *Nature*. 422: 137-140.

Feld, L. P. and B. S. Frey (2002). Trust Breeds Trust: How Taxpayers are Treated, *Economics of Governance.* 3: 87-99.

Feld, L. P. and J.-R. Tyran (2002). Tax Evasion and Voting: An Experimental Analysis, *Kyklos.* 55: 197-222.

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

Frey, B. S. (1999). *Economics as a Science of Human Behaviour,* Boston/Dordrecht/London: Kluwer.

Frey, B. S. (2004). *Dealing with Terrorism – Stick or Carrot?.* Cheltenham, UK: Edward Elgar.

Frey, B. S. and L. P. Feld (2002). Deterrence and Morale in Taxation: An Empirical Analysis, CESifo Working Paper No. 760, August 2002.

George, J. M. (1995). Asymmetrical Effects of Rewards and Punishments: The Case of Social Loafing, *Journal of Occupational and Organizational Psychology*. 68: 327-338.

Graetz, M. J. and L. L. Wilde (1985). The Economics of Tax Compliance: Facts and Fantasy, *National Tax Journal.* 38: 355-363.

Gray, J. A. (1981). A Critique of Eysenck's Theory of Personality, in H. J. Eysenck (ed.). A *Model for Personality*. New York: Springer: 246-276.

Harrison, G. W. and J. A. List (2004). Field Experiments, *Journal of Economic Literature*. 42: 1009-1055.

Homans, G. C. (1974). *Social Behavior: Its Elementary Forms*. New York: Hartcourt Brace Jovanovich.

Internal Revenue Service (1978). A Dictionary of Compliance Factors. Washington: United States Department of the Treasury.

Larsen, R. J. and T. Katelaar (1991). Personality and Susceptibility to Positive and Negative Emotional States, *Journal of Personality and Social Psychology*. 61: 132-140.

Le Grand, J. (2003). *Motivation, Agency, and Public Policy*. *Of Knights and Knaves, Pawns and Queens*. Oxford: Oxford University Press.

Long, S. and J. Swingen, J. (1991). The Conduct of Tax-Evasion Experiments: Validation, Analytical Methods, and Experimental Realism, in: P. Webley, H. Robben, H. Elffers and D. Hessing, *Tax Evasion: An Experimental Approach*. Cambridge University Press, Cambridge: 128-138.

Molm, L. D. (1988). The Structure and Use of Power: A Comparison of Reward and Punsihment Power, *Social Psychology Quarterly*. 51: 108-122.

Molm, L. D. (1994). Is Punishment Effective? Coercive Strategies in Social Exchanges, *Social Psychology Quarterly*. 57: 77-94.

Nutin, J. and A. G. Greenwald (1968). *Reward and Punishment in Human Learning*. New York: Academic Press.

Pyle, D. J. (1991). The Economics of Taxpayer Compliance, *Journal of Economic Surveys.* 5: 163-198.

Roth, J. A., J. T. Scholz and A. D. Witte (eds.) (1989). *Taxpayer Compliance,* Vol. 1 and Vol. 2. *Philadelphia*: University of Pennsylvania Press.

Sims, H. P., Jr. (1980). Further Thoughts on Punishment in Organizations, *The Academy of Management Revie*w. 5: 133-138.

Skinner, B. F. (1938). *The Behavior of Organism*. New York: Academic Press.

Skinner, B. F. (1953). *Science and Human Behavior*. New York: Macmillan.

Slemrod, J. (ed.), *Why People Pay Taxes. Tax Compliance and Enforcement,* Ann Arbor: University of Michigan Press: 193-218.

Slemrod, J., M. Blumenthal and C. Christian (2001). Taxpayer Response to an Increased Probability of Audit: Evidence from a Controlled Experiment in Minnesota, *Journal of Public Economics.* 79: 455-483.

Thorndike, E. L. (1911). *Animal Intelligence*. New York: Macmillan

Thorndike, E. L. (1932). *The Fundamentals of Learning*. New York: Teachers College, Columbia University.

Torgler, B. (2002). Speaking to Theorists and Searching for Facts: Tax Morale and Tax Compliance in Experiments, *Journal of Economic Surveys*. 16: 657-684.

Torgler, B. (2003a). Beyond Punishment: A Tax Compliance Experiment with Taxpayers in Costa Rica, *Revista de Análisis Económico*. 18: 27-56.

Torgler, B. (2003b). Tax Morale and Tax Evasion: Evidence from the United States, WWZ-Discussion Paper 03/01, WWZ, Basel.

Torgler, B. (2004). Moral Suasion: An Alternative Tax Policy Strategy? Evidence from a Controlled Field Experiment in Switzerland, *Economics of Governance*. 5: 235-253

Weck, H. (1983). *Schattenwirtschaft: Eine Möglichkeit zur Einschränkung der öffentlichen Verwaltung? Eine ökonomische Analyse*. Finanzwissenschaftliche Schriften 22, Peter Lang, Bern.

Zald, D.H., I. Boileau, W. El-Dearedy, R. Gunn, F. McGlone, G.S. Dichter and A. Dagher (2004). Dopamine Transmissions in the Human Striatum during Monetary Reward Tasks, *Journal of Neuroscience*. 24: 4105-4112.