Tax Morale and Optimal Taxation

Werner Güth and Rupert Sausgruber*

February, 2004

Abstract

We study experimentally how taxpayers choose between two tax regimes to fund a public good. The first-best tax regime imposes a general, distortion-free income tax. However, this tax cannot be enforced. The second-best alternative supplements the income tax by a specific commodity tax. This tax cannot be evaded but distorts optimal consumption choices, instead. The result is that a large majority of subjects prefer the general income tax regime. The bulk of votes is consistent with actual payoffs. We isolate tax morale as cause for payoffs above theoretical predictions.

Keywords: Optimal taxation, tax evasion, voting, experiments

JEL classification: H26, H21, C91.

*Güth: Max Planck Institute for Research into Economic Systems, Strategic Interaction Group, Kahlaische Strasse 10, D-07745 Jena, Germany. E-mail: gueth@mpiew-jena.mpg.de. Sausgruber (corresponding author): University of Innsbruck, Department of Public Economics, Universitätsstr. 15, A-6020 Innsbruck. E-mail: rupert.sausgruber@uibk.ac.at. We thank Wolfram Richter, Gerald Pruckner, and Steven Beckman for helpful comments. We are grateful for financial support by the Austrian National Bank, Jubiläumsfonds, under Project No. 9134.

1 Introduction

The theory of optimal taxation is about the distortions introduced by taxes. The interest of this theory is normative: based on the canonical assumptions of rational and self-interested taxpayers it gives advice on how to design a tax system such that the efficiency cost of taxation is reduced in the best possible way. The theory of optimal taxation has significantly contributed to clarify the objectives of tax policy and to understand the economic properties of tax instruments. Nevertheless, we have to admit that, relative to the theory's significance in the academic field of public finance, its impact on practical tax policy is meagre (e.g., Slemrod 1990).

What limits the relevance of optimal tax formulas for tax policy? Part of the answer may be that the theory has put too little emphasis on the motives why people pay taxes. Generally, the theory of optimal taxation presumes that taxes can be fully enforced. A large body of literature on tax compliance illustrates that this assumption is not warranted empirically (for a survey see, Andreoni, Erard, and Feinstein 1998). This literature broadly falls into two groups. The first adheres to the standard assumptions in economics and regards the decision of paying taxes as one that depends on the deterrent effects of legal sanctions (Allingham and Sandmo 1972, Yitzhaki 1987). According to this view, people face a decision under risk when they trade off the cost of paying taxes against the benefits of remaining undetected. For a short time now the theory of optimal taxation pays attention to this literature. For instance, the models by Cremer and Gahvari (1993), Boadway, Marchand, and Pestieau (1994) and Richter and Boadway (2003) acknowledge that there exists an efficiency cost of tax evasion, which ought to be traded off against that of tax distortion.

The second group of the literature on tax compliance has developed from the behavioral branch of economics. Most taxpayers face small expected penalties from tax evasion. Nevertheless, most people abide by the tax law (e.g., Slemrod and Yitzhaki 2002). There is now abundant empirical evidence that people in many instances do not act in their pure self-interest and, leaning against other social sciences, the behavioral literature on tax compliance incorporates social motives to explain a sense of morale or social duty to pay taxes (for references see,

Andreoni, Erard, and Feinstein 1998: 851 ff.). Recently, behavioral economics has gained enormous momentum and is on the verge of becoming mainstream in some fields within economics (for surveys see, e.g., Rabin 2002, Fehr and Falk 2002, Camerer 2003). The theory of optimal taxation accounts for social preferences by the concept of social welfare functions. But this concept is based on ad hoc reasoning rather than on real behavior. So far the theory of optimal taxation has not dealt with the question how new assumptions about individual preferences would modify traditional optimal tax rules.

In this paper, we illustrate that the normative policy implications derived from the existing framework of optimal taxation theory may be grossly misleading unless it accounts for basic insights from recent research in behavioral economics. To achieve this aim, instead of studying normative standards of what tax authorities might do, in an experiment we observe the behavior of taxpayers who are motivated by real incentives. Our study is particularly centered around a behavioral pattern in public goods experiments, which is highly robust and well documented in the literature (see, Ledyard 1995). In these experiments subjects typically contribute 40 to 60% of their endowment to the public good although selfish individuals would contribute nothing. In our setup a public good is funded by individual taxes and we will refer to tax payments above theoretical predictions as "tax morale".[1] We illustrate the welfare implications of tax morale in our design and contrast them against the normative predictions of optimal taxation theory.

We apply experimental methods since, for obvious reasons, data on tax compliance is difficult to come by in the field. In contrast, in an experiment we know people's incomes and we can observe tax payments. In the field the economic effects of taxation typically interact with numerous circumstances that are difficult to control. As a consequence, attempts to evaluate tax instruments with observational data are prone to substantial difficulties (for discussions see, e.g., Creedy 2000, Saez 2001, Fullerton and Metcalf 2002). A virtue of experimental methods is that we can compare different tax regimes under true *ceteris paribus* conditions (see, e.g., Quirmbach, Swenson, and Vines 1996). For instance, in an

[1]This is in line with previous contributions, e.g., by Alm, McClelland, and Schulze (1992).

experiment it is particularly straightforward to measure the relative efficiency of alternative tax regimes.

Another advantage of experimental methods is that it enables us to shed some light on the role of tax morale for the political feasibility of tax reform and the endogenous evolution of the tax system.[2] For this reason, we allow taxpayers to vote between two tax regimes to fund the public good. This choice is characterized by a tradeoff between tax distortion and tax evasion. The theory provides a clear-cut prediction of the size of this tradeoff. However, the actual weight of the tradeoff is endogenously determined by tax morale, i.e., the amount of taxes paid in excess of theoretical predictions. Our interest is to see whether tax morale can induce taxpayers/voters to implement a tax regime that is not feasible in the standard framework of optimal taxation theory.

The basic design is as follows: Participants in the experiment gain utility from consuming a set of private goods and a public good. In a referendum they choose between two tax schemes to fund the public good. The *G-scheme* imposes a *[G]eneral income tax* on endowments that subjects declare for taxation. The tax is efficient in the sense that it does not distort consumption choices. But the tax cannot be enforced and opportunistic individuals will not declare any taxes. Under the rules of the *S-scheme* the income tax applies together with a *[S]pecific commodity tax*. The commodity tax bears an efficiency cost from distorting the private consumption choice. Nevertheless, assuming rational behavior, the *S-scheme* increases overall efficiency by raising receipts for funding the public good without enforcement.

In the experiment we implement conditions such that rational self-interested agents would prefer the distortionary *S-scheme*. This prediction is compared against an alternative prediction based on abundant evidence regarding the voluntary contribution of public goods. In line with the insights from this literature, we hypothesize that to fund the public good people exhibit tax morale, i.e., they are inclined to pay income taxes in spite of an individual incentive to free-ride.

[2]For discussions on how non-selfish preferences as well as heuristics and biases in judgement may translate into policy outcomes see, e.g., Quattrone and Tversky (1988), McCaffery (2000), Tyran and Sausgruber (2002), Slemrod and Krishna (2003).

If tax morale is high, the *G-scheme* will be overall more efficient. On the other hand, if tax morale is humble, the *S-scheme* will dominate the *G-scheme*.

As we have expected, the results show that tax morale is essential in trading off the efficiency cost of the two tax systems: against the predictions of standard theory the income tax (*G-scheme*) as the first-best solution, which cannot be formally enforced, results in the same overall efficiency as the specific commodity tax (*S-scheme*), which sets a limit to free-riding. Moreover, in voting subjects reveal a strong preference for the *G-scheme*, i.e., a majority of subjects rejects the theoretically favored alternative. We interpret these results as strong indication for the necessity to incorporate behavioral aspects of taxation into an encompassing framework of optimal taxation.

The paper is structured as follows: In section 2 we explain the experimental design and discuss our hypotheses. Section 3 presents the results and section 4 concludes.

2 Experimental Design

Each participant in the experiment is a taxpayer. Taxpayers are organized into groups of size N. Every subject receives an endowment E_i drawn from a known distribution with support $[\underline{E}, \overline{E}]$. Subjects can declare any share of their endowment for taxation. Call d_i the endowment subject i declares for taxation at rate $\tau \in [0, 1)$. Disposable income thus is $e_i = E_i - \tau d_i$. The taxpayer can use disposable income to purchase individual quantities of two commodities x_i and y_i. Goods are priced at p_x and p_y. We think of x as a specific commodity that can be taxed whereas y is some composite numeraire good. Rational taxpayers maximize payoff subject to the budget constraint $e_i = x_i p_x(1 + t) + y_i p_y$, where $t \geq 0$ is the rate of the specific tax that marks up to the price of commodity x. The subject's payoff is determined according to:

$$\pi_i = x_i y_i + G. \tag{1}$$

Here, $x_i y_i$ is the utility from privately consuming (x_i, y_i) and $G = \alpha T$ the utility of public good provision financed by tax receipts T raised from N subjects.

The parameter α, with $0 < \alpha < 1$, is the public good's marginal per capita return. Endowment points not collected as taxes and not spent on commodity purchases result in zero payoff.

Experimental subjects vote between two tax schemes to fund the public good: the *G-scheme* imposes a non-enforceable income tax and places emphasis on tax morale as social mechanism to sustain efficiency. More specifically, we implement this by purely appealing to tax morale, i.e., by not monitoring tax compliance at all.[3] In the *S-scheme* commodity x is taxed. Since it is in the self-interest to purchase this commodity, the *S-scheme* provides an effective mechanism to prevent free-riding on taxes raised from other members of the society. We now explain the two tax schemes in full detail.

i) G-scheme: general income tax ($\overline{\tau} > 0$, $t = 0$): Assume that all tax receipts are raised efficiently via a general tax $\overline{\tau}$ on endowments. For any disposable income e_i, and monetary amounts e_i^x and $(e_i - e_i^x)$ spent on commodities x and y, respectively, payoff under a general tax is given by:

$$\pi_i^G = \frac{e_i^x}{p_x}\left(\frac{e_i - e_i^x}{p_y}\right) + \alpha \sum_{j=1}^{N} \overline{\tau} d_j \tag{2}$$

Regardless of tax declarations d_i, a subject maximizes payoff when half of disposable income is spent on commodity x, i.e. $e_i^{x*} = e_i/2$. Optimal consumption choices are $x_i^* = \frac{e_i}{2p_x}$ and $y_i^* = \frac{e_i}{2p_y}$. In the optimum, a subject's payoff is:

$$\pi_i^{G*} = \frac{e_i^2}{4p_x p_y} + \alpha \sum_{j=1}^{N} \overline{\tau} d_j. \tag{3}$$

From $\frac{\partial \pi_i^{G*}}{\partial d_i} = \overline{\tau}\left[\alpha - \frac{e_i}{2p_x p_y}\right]$, it follows that a rational taxpayer reduces tax declarations as long as $\alpha < \frac{e_i}{2p_x p_y}$. If, for instance, $\alpha < \frac{E(1-\overline{\tau})}{2p_x p_y}$, the optimal declaration is $d_i^* = 0$ always. Declaring taxes would, however, be efficiency enhancing

[3]Existing experimental studies on tax compliance implement penalties for underreporting along with certain probabilities of detection. These features have been criticized to lead astray because designing tax compliance as a decision under risk impedes separating risk preferences from people's predispositions to comply with a norm. See, Alm, McClelland, and Schulze (1992).

if the total value generated from funding the public good by an additional unit of tax exceeds the reduction of benefits from private consumption, i.e., $\alpha N > \frac{e_i}{2p_x p_y}$. If, for instance, $\alpha N > \frac{\overline{E}}{2p_x p_y}$ efficiency considerations always suggest truthful declarations, i.e., $d_i^+ = E_i$ for all $E_i \in [\underline{E}, \overline{E}]$. For the experiment we choose parameters such that a rational taxpayer declares zero income although positive declarations would raise efficiency, e.g., in the sense of $\alpha < \frac{\underline{E}(1-\overline{\tau})}{2p_x p_y} < \frac{\overline{E}}{2p_x p_y} < \alpha N$.

ii) S-scheme: specific commodity tax ($\underline{\tau} > 0, t > 0$): Assume that in addition to an income tax with rate $\underline{\tau} < \overline{\tau}$ there is now a specific tax $t > 0$ on the taxpayer's spending on commodity x. Payoff is:

$$\pi_i^S = \frac{e_i^x}{p_x(1+t)} \left(\frac{e_i - e_i^x}{p_y} \right) + \alpha \sum_{j=1}^{N} \left(\underline{\tau} d_j + \frac{t}{1+t}\, e_j^x \right). \tag{4}$$

In the optimum it holds: $e_i^{x*} = e_i/2 + z$, where $z = \frac{\alpha t p_x p_y}{2}$. Payoff in the optimum with a specific tax is:

$$\pi_i^{S*} = \frac{e_i^2 - 4z^2}{4p_x(1+t)p_y} + \alpha \sum_{j=1}^{N} \left[\underline{\tau} d_j + \frac{t}{1+t} \left(\frac{e_j}{2} + z \right) \right]. \tag{5}$$

Our parameters assure that $\alpha < \frac{e_i - t p_x p_x}{2p_x(1+t)p_y} < \alpha N$ for all possible choices of d_i and endowments $E_i \in [\underline{E}, \overline{E}]$. Again, under these conditions rational taxpayers declare zero income ($d_i^* = 0$) although everyone would be better off under truthful tax telling ($d_i^+ = E_i$).

2.1 Details of the Design

There are 4 phases. Each phase consists of 15 periods. In phase 0 in every period subjects choose the quantities of the goods x and y they wish to buy. This phase serves to make subjects familiar with the consumption choice task, so yet no taxes apply. In phase 1 the members of each group at the ballot first choose between the two alternatives, general tax on declared income at rate $\overline{\tau}$ (*G-scheme*) or specific tax at rate t plus a general tax at rate $\underline{\tau}$ (*S-scheme*). The regime supported by simple majority is then implemented for the subsequent 15 periods. Phase 1 ends after 15 periods. Phases 2 and 3 are one to one repetitions

of phase 1. Thus, there is a total of 3 votings followed by 15 periods of buying commodities each.

The experiment is parameterized in the following way: Endowments are drawn from a uniform distribution $E_i \sim U(110, 190)$ points. Points are converted into cash at the exchange rate of points 100 = Cents 25 (Euro 0.25). Points accumulated during the experiment are paid out to participants immediately after the experiment. For the income tax rate we choose $\overline{\tau} = 0.5$ in the *G-scheme* and $\underline{\tau} = 0.05$ in the *S-scheme*. Prices are set at $p_x = 1.1$ and $p_y = 50$. The rate for the specific tax is $t = 0.27$ such that the after tax price of commodity x is $p_x(1 + t) \approx 1.4$. The size of a group is five, i.e., $N = 5$. The marginal per capita return of the public good is set at $\alpha = 0.5$. By this parameter choice, the restrictions discussed in the previous section are satisfied.

2.2 Predictions

In our design tax declarations are voluntary and the experiment is parameterized such that a rational taxpayer abstains from paying taxes on income. Thus, if people behave rationally, the *S-scheme* pareto-dominates the *G-scheme*. In other words, with zero declarations payoffs for everyone are smaller under the general tax [see eq. (3)] than under the specific tax [see eq. (5)].

To illustrate, *Table 1* shows the payoff of an individual under the assumption that everyone in the group has the same mean endowment, $\hat{E} = 150$, and consumption choices are optimal. If nobody declares any taxes, the *S-scheme* results in 22 % higher payoff than the *G-scheme* (124 vs. 102). The reason is that the efficiency cost arising from commodity taxation is small relative to the efficiency cost that prevails if the public good is not provided. Consequently, rational voters would strictly prefer the *S-scheme*.

Against this prediction, we hypothesize that people declare substantial fractions of their endowments for taxation. There is now abundant empirical evidence that people stick to socially desirable behavior even if it is individually costly to do so. This evidence has brought about recent theories of social preferences, which prove to be relevant in numerous instances (for a recent survey, see, Camerer

Table 1

Own payoff under optimal consumption choices depending on own (i) and others' ($-i$) tax declarations in percent of $\hat{E} = 150$. (*G-Scheme*/*S-Scheme*)

		own (i)		
		0.00	0.265	1.00
	0.00	(102/124)	(87/123)	(63/120)
others' ($-i$)	0.265	(142/128)	(127/127)	(103/123)
	1.00	(252/138)	(237/136)	(213/133)

2003, Ch. 2). Since in our design voluntary tax declarations induce efficiency gains, social preferences will give rise to tax morale (for a similar argument, see, Alm, McClelland, and Schulze 1992).

Finally, if taxpayers anticipate tax morale, taxpayers may even vote for the general tax scheme. In *Table 1*, if everybody declares more than 26.5 % of endowment, the *G-scheme* dominates the *S-scheme*. If everybody declares taxes honestly, payoff is even 60 % higher in the general than in the specific tax scheme (213 vs. 133). Moreover, voting for the *G-scheme* may serve people as a signal to establish tax morale. Indeed, previous experimental studies have found that voting in favor of socially desirable rules is capable to generate norm compliant behavior even if the norm ex post cannot be formally enforced (see, Alm, McClelland, and Schulze 1999; Tyran and Feld 2001).

3 Results

Experiments were run in May 2003 at the University of Innsbruck. In total we had 75 subjects participating in four sessions. Subjects were undergraduate students from various majors. The experiment was programmed and conducted using the software z-Tree (Fischbacher 1999). Including Euro 4 for show up on time, the average subject earned Euro 20.7 within approximately 2 hours.

In section 3.1 we evaluate the efficiency of the *G-scheme* relative to the *S-scheme* based on actual behavior. In section 3.2 we discuss voting behavior.

3.1 Efficiency

In our setup we measure efficiency simply in terms of payoffs. *Figure 1* depicts payoffs averaged across subjects and differentiated by tax schemes. In phases 1 and 2, subjects on average earn slightly more in the *G-scheme* than in the *S-scheme*; but as the figure indicates, the differences are small and insignificant (Phase 1: 121 versus 117 experimental points, p= 0.397, Mann-Whitney test, one-sided, based on statistically independent group observations; Phase 2: 120 versus 116 points, p= 0.443). Only in phase 3 payoffs are smaller in the *G-scheme* than in the *S-scheme* (115 versus 124, p= 0.056). To test whether overall there is a difference in efficiency between the two schemes we calculate a regression based t-test: we regress individual payoffs on variables for the period, tax declarations relative to endowment, and dummies for phases. To account for statistical dependence within groups we calculate robust standard errors adjusted for clustering on groups. According to this test, there is no difference in payoffs between the tax schemes ($p = 0.568$). We summarize:

Result 1 *Payoffs do not differ between tax schemes, i.e., the G-scheme is equally efficient as the S-scheme.*

In *Figure 1*, the grey dotted line at 102 and the black dotted line at 124 indicate the equilibrium predictions for the *G-scheme* and *S-scheme*, respectively (compare, *Table* 1). Apparently, result 1 would not prevail under standard predictions. We now turn to the question why payoffs deviate from theoretical predictions.

As we have discussed in section 2.2, tax morale provides a reason for payoffs to exceed theoretical predictions. If taxpayers voluntarily declare endowment points for taxation, this has large effects when the tax rate is high ($\overline{\tau}$), as in the *G-scheme*. In contrast, it has only modest effects when the tax rate is low ($\underline{\tau}$), as in the *S-scheme*. In the *G-scheme*, subjects on average have declared 21% of their

endowment for taxation. In the *S-scheme* the respective number is 32%. Due to the difference in the tax rate on declared endowment, nevertheless, average taxes from declared endowment are much higher in the *G-scheme* (15.3 points) than in the *S-scheme* (2.3 points). Taking as a basis for statistical testing 5 groups that have experienced both tax schemes, a Wilcoxon signed-rank test reveals that this difference is significant ($p = 0.043$, two-sided). This means that the amount of taxes that taxpayers pay voluntarily is larger in the *G-scheme* than in the *S-scheme*. In other words:

Result 2 *Tax morale is higher in the G-scheme than in the S-scheme.*

Tax morale induces efficiency gains from the provision of the public good. In addition, result 2 suggests a stronger demand effect for tax morale in the *G-scheme* than in the *S-scheme*.[4] This explains why payoffs are above equilibrium in the *G-scheme*. On the other hand, we still need to explain why payoffs stay behind equilibrium in the *S-scheme*. The straightforward answer to this question is that taxpayers do not buy optimal consumption bundles. We define *consumption efficiency* as the actual payoff from consumption over the maximum possible payoff, i.e., $Eff^G = x_i y_i / \frac{e_i^2}{4p_x p_y}$ in the *G-scheme* and $Eff^S = x_i y_i / \frac{e_i^2 - 4z^2}{4p_x(1+t)p_y}$ in the *S-scheme*. Our data reveals that there is some inefficiency in consumption in both tax regimes: In phase 1 efficiency is 90.6% in the *G-scheme* and 89.5% in the *S-scheme*. The difference between schemes is insignificant (p=0.602, Mann-Whitney test). In phase 2 efficiency is 93.2% in both schemes; in phase 3 the respective numbers are 93.3% and 95.4% ($p = 0.312$).

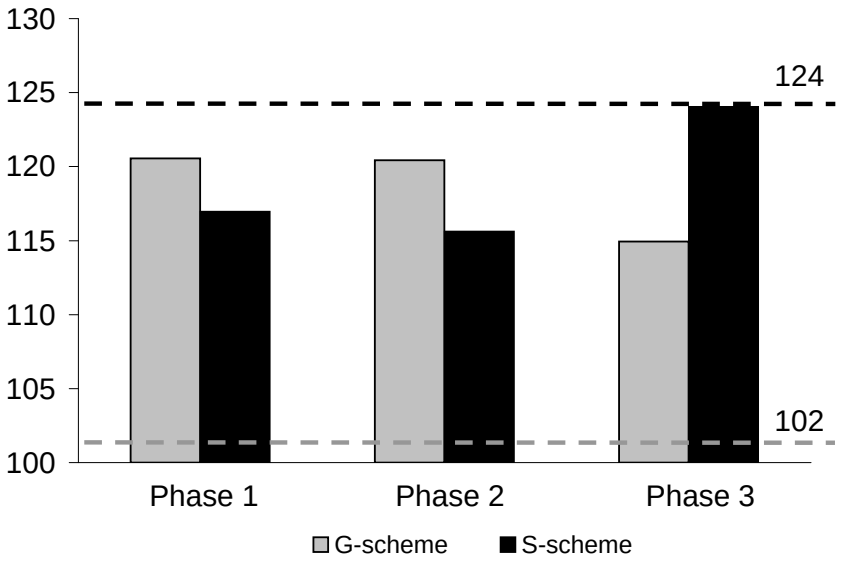
Non-optimal consumption choices cut down on payoffs regardless of the tax regime. However, tax morale offsets these inefficiencies in the *G-scheme*, whereas this is not the case in the *S-scheme*. This explains why payoffs stay behind theoretical predictions in the *S-scheme*, but are above predictions in the *G-scheme*. We state this as our next result:

[4] Another way to see this is that although switching between the *S-scheme* and the *G-scheme* increases the income tax rate by factor ten (from $\underline{\tau} = 0.05$ to $\overline{\tau} = 0.5$), actual tax declarations relative to endowment differ only by 11 percentage points on average (32% in the *S-scheme* and 21% in the *G-scheme*).

Result 3 *Tax morale is the cause for result 1. In our design, tax morale is sufficient to suspend with the theoretical need to supplement an efficient income tax by a specific commodity tax.*

Figure 1

Average payoffs in points per period



A final observation deserves attention: In the last phase average payoffs have been smaller in the *G-scheme* than in the *S-scheme* (see, *Figure 1*). The cause for this observation is that the commodity tax of the *S-scheme* provides constant funding for the public good, whereas tax morale decreases during the course of the experiment. In the *S-scheme*, average per period individual commodity-tax payments are 15.2, 15.3, and 16.7 in phases 1, 2, and 3, respectively. In contrast, in the *G-scheme* average tax declarations start at 25.1 % in phase 1, and go down to 20.6% and 16.9% in phases 2 and 3, respectively. As a result, individual income-tax payments drop from 18.5 points in phase 1 to 15.2 and 12.3 points in phases 2 and 3, respectively. Because of this dynamic the *S-scheme* eventually raises higher total tax receipts to fund the public good and, as a consequence, payoff-dominates the *G-scheme* in phase 3. We conclude:

Result 4 *Tax morale erodes over time, what eventually suggests to supplement the non-enforceable income tax by a specific commodity tax in our design.*

3.2 Voting Behavior

With respect to voting we report the following result:

Result 5 *Subjects overwhelmingly vote in favor of the general income tax regime (G-scheme). There is no empirical tendency for choices to converge towards the specific commodity tax regime (S-scheme).*

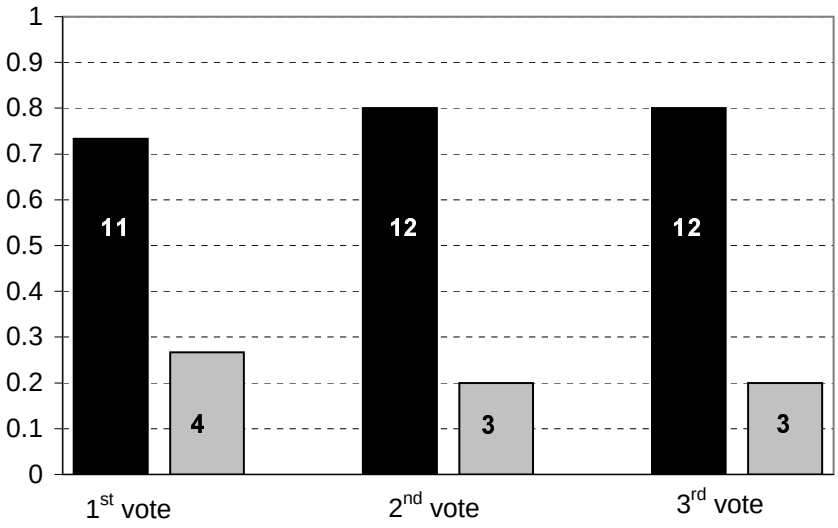
Figure 2 provides support for this result. The figure shows the frequency of groups that have implemented the *G-scheme* and the *S-scheme*, respectively. The vertical axis depicts relative frequencies; absolute frequencies are written on the bars. In the *first vote*, 11 out of 15 groups (73 percent) approved of the *G-scheme* (black bar); only 4 groups chose the *S-scheme* (shaded bar). According to a binomial test, approval of the *G-scheme* is significantly higher (p=0.059). Taking individual behavior, there are 49 out of 75 subjects voting in favor of the *G-scheme*. Again, this result reveals a bias in subjects' decisions in favor of the *G-scheme* (p=0.003).[5] These findings reject the hypothesis that taxpayers support the commodity-income tax mix of the *S-scheme*.

Figure 2 furthermore shows that this pattern persists along the sequence of the experiment: in both, *the second* and *third* vote, 80 percent (13/15) of groups vote for the *G-scheme*. Thus, we do not observe choices to converge towards the *S-scheme*. At the level of individuals there is no statistically detectable learning behavior, neither. Between the *first* and the *second* vote, 11 subjects who have previously voted in favor of the income tax scheme switched their the decision towards the commodity tax scheme. There were, however, 11 other subjects who switched choices into the reverse direction. Between the *second* and the *third* vote, 9 subjects who have previously voted for the *G-scheme* switched to the *S-scheme*, against 9 subjects who switched into the reverse. These numbers illustrate that choices do not converge to any particular tax design.

[5]We provide this test for illustrative reasons only. Note, however, that individual observations are statistically dependent within groups.

Figure 2

Frequency of groups approving off the *G-scheme* [black] and the *S-scheme* [shaded]. N=15.



In the previous section we have reported that payoffs hardly differ between schemes. Therefore, it is not clear a priori why so many subjects favor the *G-scheme*. To further explore this issue we provide a closer analysis of individual payoffs. In particular, we ask whether it has been individually rational to support the income tax scheme. To answer this question we take actual income tax declarations in the *G-scheme* and compare payoffs against hypothetical payoffs as they would result from commodity taxation in the *S-scheme*. Since for this comparison we do not observe consumption choices under the *S-scheme*, we assume that subjects buy the optimal commodity bundles in both schemes; i.e., payoffs in the *G-scheme* are recalculated based on optimal consumption choices and actual tax-declarations according to eq. (3). These payoffs are then compared against payoffs according to eq. (5) assuming the same individual tax declarations.

There are 14 out of 15 groups who voted at least once in favor of the *G-scheme*. Based on the above calculation, after implementing the scheme for the first time, 9 of these groups by a majority of their subjects do better in the *G-scheme* than they hypothetically would have done in the *S-scheme*. According to a binomial

test this number is different from random (p=0.090). Taking individual data, this corresponds to 40 out of 70 subjects, whose payoffs under the rules of the *G-scheme* exceed hypothetical ones of the *S-scheme*. Again, this number differs from random (p=0.094).

Does this pattern persist? To answer this question we look at the behavior of subjects who voted in favor of the income tax also a second and third time. 12 out of 14 groups that have accepted the *G-scheme* once implement it also for a second time. Only 5 out of these 12 groups (p=0.212) by majority of their members manage to sustain profits higher than the hypothetical ones under the rules of the *S-scheme*. Finally, 9 out of 12 groups that have accepted it twice also vote in favor of the *G-scheme* a third time. In phase 3, there remain only 3 out of these 9 groups (p=0.746) that by majority of their members would still have hypothetically done better in this scheme. We conclude:

Result 6 *In the first vote, votes are consistent with actual payoffs, i.e., for a majority of subjects voting in favor of the G-scheme can be ex post rationalized by payoff considerations. In the subsequent two votes many subjects continue voting in favor of the G-scheme even if tax declarations are low and their earning would have been higher under the alternative S-scheme.*

Why does the *S-scheme* receive so little support? One potential reason is that subjects hold ex-ante beliefs of high tax morale on behalf of their fellow taxpayers. Such a belief would rationalize a vote for the *G-scheme*. In the light of result 6 it appears odd, however, that many subjects continue to vote in favor of the *G-scheme* in phase 3 even after they have experienced that tax morale went down.

An alternative explanation comes from cognitive psychology. Recent contributions by Slemrod and Krishna (2003), and McCaffery (2000) emphasize the relevance of cognitive theory to taxation. In our setup, subjects earn payoffs from two sources: from private consumption and from consuming a public good. In comparing the two tax regimes, subjects know for sure that the payoff from private consumption will be smaller in the *S-scheme* than in the *G-scheme*.[6] In

[6]Subjects are explicitly told that the tax induces an increase in the price of commodity x. See appendix for instructions.

contrast, subjects may find it difficult to predict how a switch between the tax regimes would change their earnings from the public good. Such a difference in the saliency of expected changes in payoffs may contribute to explain why subjects' choices are biased in favor of the *G-scheme*.

Our experiment was not aimed to rigorously test for the validity of such arguments. Whatever explanation is valid, result 5 is incompatible with standard theory. Our study, therefore, exemplifies a case where predictions and guidelines derived from standard theory of optimal taxation would fail in an attempt to render tax reform feasible or to explain the endogenous evolution of tax design.

4 Conclusion

Slemrod (1990) in a general assessment argues that the theory of optimal taxation is incomplete as a guide to action unless it accounts for tax evasion. Recent contributions that have met this concern challenge long standing insights of optimal taxation theory regarding the optimal tax mix. Particularly, distortionary commodity taxes may become part of the optimal tax structure when tax evasion is possible. In this experiment we have designed a case for which the theory makes a strong claim to supplement an efficient income tax by a specific commodity tax. This is the case when the income tax as the first-best solution can be evaded whereas the commodity tax cannot.

The main result is that tax morale plays an essential role in trading off tax evasion against tax distortion. In our experiment, tax morale materializes as disposition to pay taxes to fund a public good against an individual incentive to free-ride. We find that tax morale overall is sufficient to suspend with the need for additional commodity taxation. This observation indicates a need to incorporate behavioral motives of tax compliance into the debate on optimal tax systems.

A second result is that tax morale erodes over time. Our design in many instances is similar to a voluntary contribution experiments, which reveal a typical decay in cooperation in repeated interaction (see, Ledyard 1995). Recent experimental research has shown that cooperation is highly sensitive to factors

like punishment opportunities, social identification, or communication (for a survey, see, e.g., Camerer 2003). In context of tax evasion even weak means of economic deterrence have shown to result in sustained tax morale in multi-round experiments (Alm, McClelland, and Schulze 1992, Tyran and Feld 2001). On the other hand, countries nowadays frequently bemoan a decline in tax morale (see, e.g., Schneider and Enste 2003). Therefore, we may wish to learn more on how manage a norm of tax compliance in the field (for some discussion, see, Besley, Preston, and Ridge 1997, Kahan 1997, Fehr and Falk 2002).

Our research has highlighted the role of tax morale within the quest for an optimal tax design. Institutional design assuming agents who react optimally to the substitution of a commodity tax by an income tax (or vice versa) may recommend the wrong or at least an inadequate measure when agents systematically deviate from rational choice. In the case at hand opportunistic optimality would mean to avoid income tax payment and thereby to question the provision of public goods. But this is not what we observe. Thus the behavioral approach to institutional or mechanism design, which puts emphasis on actual decision behavior, may come to different conclusions and recommendations. This stresses the necessity to supplement optimal tax design by one paying tribute to how people actually react to public good provision and their tax financing. Our study is a step into this direction. We hope that our study will motivate further research on behavioral aspects of taxation in an attempt to develop an encompassing framework of optimal taxation.

References

Alm, J., G.H. McClelland, and W.D. Schulze (1992): Why do People Pay Taxes? *Journal of Public Economics* 48: 21-38.

Alm, J., G.H. McClelland, and W.D. Schulze (1999): Changing the Social Norm of Tax Compliance by Voting. *Kyklos* 52: 141-71.

Allingham, M.G., and A. Sandmo (1972): Income Tax Evasion: A Theoretical Analysis. *Journal of Public Economics* 1: 323-38.

Andreoni, J., B. Erard, and J. Feinstein (1998): Tax Compliance. *Journal of Economic Literature* 36: 818-60.

Besley, T., I. Preston, and M. Ridge (1997): Fiscal Anarchie in the UK: Modelling Poll Tax Noncompliance. *Journal of Public Economics* 64: 137-52.

Boadway, R., M. Marchand, and P. Pestieau (1994): Towards a Theory of Direct-Indirect Tax Mix. *Journal of Public Economics* 55: 71-88.

Camerer, C. (2003): *Behavioral Game Theory.* Experiments in Strategic Interaction. Princeton Univ Press: Princeton.

Creedy, J. (2000) Measuring Welfare Changes and the Excess Burden of Taxation. *Bulletin of Economic Research* 52: 1-47.

Cremer, H. and F. Gahvari (1993): Tax Evasion and Optimal Commodity Taxation. *Journal of Public Economics* 50: 261-275.

Fehr, E. and A. Falk (2002): Psychological Foundations of Incentives. *European Economic Review* 46: 687-724.

Fischbacher, U. (1999): *z-Tree. Toolbox for Readymade Economic Experiments.* IEW Working paper 21. University of Zurich.

Fullerton, D. and G.E. Metcalf (2002): Tax Incidence. In: A. J. Auerbach und M. Feldstein (Eds.). *Handbook of Public Economics.* Elsevier. Amsterdam: 1787-872.

Kahan, D.M. (1997): Social Influence, Social Meaning, and Deterrence. *Virginia Law Review* 83: 349-95.

Ledyard, J.O. (1995): Public Goods: A Survey of Experimental Research. In: J.H. Kagel and A.E. Roth (Eds.). *The Handbook of Experimental Economics.* Princeton Univ Press. Princeton: 111-94.

McCaffery, E.J. (2000): Cognitive Theory and Tax. In: C.R. Sunstein (Ed.): *Behavioral Law and Economics.* Cambridge Univ. Press. Cambridge: 398-421.

Quattrone, G.A. and A. Tversky (1988): Contrasting Rational and Psychological Analyses of Political Choice. *American Political Science Review* 82: 719-36.

Quirmbach, H.; C. Swenson, and C. Vines (1996): An Experimental Examination of Gerneral Equilibrium Tax Incidence. *Journal of Public Economics* 61: 337-358.

Rabin, M. (2002): A Perspective on Psychology and Economics. *European Economic Review* 46: 657-85.

Richter W.F. and R.W. Boadway (2003): Trading off Tax Distortion and Tax Evasion. *Journal of Public Economic Theory*. Forthcoming.

Saez, E. (2001): Using Elasticities to Derive Optimal Income Tax Rates. *Review of Economic Studies* 68: 205-29.

Schneider F. and D.H. Enste (2003): *The Shadow Economy* - An International Survey. Cambridge Univ. Press. Cambridge.

Slemrod, J. (1990): Optimal Taxation and Optimal Tax Systems. *Journal of Economic Perspectives* 4: 157-78.

Slemrod, J. and A. Krishna (2003): Behavioral Public Finance: Tax Design as Price Presentation. *International Tax and Public Finance* 10: 189-203.

Slemrod, J. und S. Yitzhaki (2002): Tax Avoidance, Evasion, and Administration. In: Auerbach, A. and Feldstein, M. (eds.): *Handbook of Public Economics* Volume 3. Elsevier. Amsterdam: 1423-1470.

Tyran, J.-R. and L. Feld (2001): *Why People Obey the Law. Experimental Evidence from the Provision of Public Goods.* Working paper. University of St. Gallen 2001-14.

Tyran, J.-R. and Sausgruber, R. (2002): *A Little Fairness may Induce a Lot or Redistribution in Democracy.* Discussion Paper 2002-30. University of St.Gallen.

Yitzhaki, S. (1987): On the Excess Burden of Tax Evasion. *Public Finance Quarterly* 15: 123-37.

Appendix. Sample instructions (originally in German)

General Instructions. Part I: Thank you for participating in the experiment. You receive Euro 4 for having shown up on time. If you read these instructions carefully and follow all the rules, you can earn more. The Euro 4 and all additional amount of money will be paid to you in cash immediately after the experiment. During the experiment we shall not speak of euros but rather of points. Points are converted to euros at the following exchange rate: 100 Points = 25 Cents (Euro 0.25).
It is prohibited to communicate with the other participants during the experiment. If you have any questions, please ask us. We will gladly answer your questions individually. It is very important that you follow this rule, otherwise we shall have to exclude you from the experiment and from all payments.

Detailed Instructions: The experiment is divided into periods. In every period you receive an amount of points, which we refer to as your endowment. Your endowment is a randomly generated integer number between 110 and 190 points. Here, every number between 110 and 190 is equally likely to occur. You will learn your endowment in every period.
In every period you make the following decisions: you have to decide how to use you endowment for the purchase of two commodities X and Y. The quantities you can buy depend on your endowment and in the prices of commodities X and Y:

The price of commodity X is 1.1 points.
The price of commodity Y is 50 points.

Example: Your endowment is 150 points. You buy X=68.18 units of commodity X and Y=1.5 units of commodity Y. With that purchase you exactly use up your endowment $[(68.18 \times 1.1) + (1.5 \times 50) = 150]$.
How your income is determined: Your income is the product of the units of X and Y you buy:

$$\text{Your income} = X \times Y$$

In the above example you have bought X=68.1 units of X and Y=1.5 units of Y. Thus, your income is 102.27 ($= 68.18 \times 1.5$) points.
You will take your decisions by computer. At the beginning of every period you will see the following input screen (*original instructions included a screen-figure here*). The number of the period appears in the top left corner of the screen. In the top right corner, you

can see how many seconds remain to take your decision. The first line shows "Your Endowment" in the current period. In the input fields below you must enter the units of commodities wish to buy. Confirm your choice by pressing the OK button.
Please note:

- If time expires before you have made a choice your income is zero.
- Your expenditures on commodity purchases may not exceed your endowment.
- Endowment not spent on the purchase of commodities result in zero income.

At the end of a period, a result screen will appear (*original instructions included a screen-figure here*). Here, your decisions are summarized again. In this example the participant has purchases 68.18 units of commodity X and 1.5 units of commodity Y. The last line shows your income in points (here: 102.27).

Instructions. Part II:

The experiment will now continue for additional 15 periods under the following rules: Participants are randomly assembled into groups of 5 persons; i.e. apart from you there will be 4 additional persons in your group. The composition of your group remains the same in each period. That is, your group members will be the same from one period to the next. The identity of your group members will not be revealed to you at any time. In this part of the experiment taxes are raised to fund a project. The project generates income for you and the other participants. There are two different kinds of taxes:

1. A tax of τ ("tau") percent on your endowment. You will shortly be informed about the percentage size of τ.
2. A tax of 27 percent imputed into the price of commodity X.

In every period you take two decisions:

1. First you decide how much of your endowment you declare for the purpose to be taxed. Of the endowment you declare for taxation τ percent will be deducted as tax. The endowment that remains to you after taxation is called your remaining endowment.

   Example: The tax rate tau is 50 percent ($\tau = 0.5$). Your endowment is 170 points. You declare 40 points. From these 40 points there will be 20 (40×0.5) points deducted as tax. Your remaining endowment is then 150 points ($170 - 20$).

2. After your remaining endowment is determined you decide again how to use it for the purchase of commodities X and Y. The quantities you can buy now also depend on whether good X is taxed.

   – If commodity X is taxed its price rises to 1.4 points instead of 1.1 points as previously; i.e. the price rises by 27 percent. If commodity X is not taxed its price remains at 1.1 points.

   – The price of commodity Y remains at 50 points.

   Example: Your remaining endowment is 150 points. Commodity X is taxed such that its price is 1.4 points. Commodity Y's price is 50 points. You buy X=56.42 units of commodity X and Y=1.42 units of commodity Y. With that purchase you exactly use up your remaining endowment $[(56.42 \times 1.4) + (1.42 \times 50) = 150]$.

How your income is determined: Your income consists of two parts:

(1) Your income from the project. This income is determined as follows:

$$\text{Income from the project} = 0.5 \times \text{sum of all individual tax payments}$$

The income from the project is determined in the same way for all other participants; i.e., they receive the same income from the project. For example, if you and the other participants pay in sum 200 points in taxes, you and the other participants will earn 100 ($= 0.5 \times 200$) points, each. If you and the other participants pay in sum, e.g., 20 points in taxes, you and the other participants will earn 10 ($= 0.5 \times 20$) points.

How the your individual tax payments is determined: Your tax payments are, first, τ percent of the endowment you declare for taxation. Second, if commodity X is taxed your expenditure for purchasing commodity X includes a tax of 27 percent. (To calculate the exact amount of commodity tax multiply the purchased units of X with the net price and the tax rate; i.e., $X \times 1.1 \times 0.27$.)

(2) Your income from purchasing commodities. As previously, this income is the product of units of X and Y you buy:

$$\text{Your income} = X \times Y$$

In the above example you have bought X=56.42 units of X and Y=1.42 units of Y. Thus, your income is 80.12 ($= 56.42 \times 1.42$) points.

Voting Proposal

Before the experiment will start under the new rules, you and the other 4 participants in your group vote between two alternatives how to fund the project. The alternative that is wins the majority of votes (i.e., 3 or more) is implemented for the next 15 periods.

Alternative 1:

- The tax τ on your declared endowment is $\tau = \mathbf{0.05}$ (= 5 percent)
- Commodity X is taxed; its price is **1.4** points.

Alternative 2:

- The tax τ on your declared endowment is $\tau = \mathbf{0.5}$ (= 50 percent)
- Commodity X is not taxed; its price is **1.1** points.

Please note:
Under alternative 1 participants pay taxes if they purchase commodity X. Since commodity X has become more expensive and you can buy less units from it, you will earn less income from commodity purchase. On the other hand, you will earn income from the project even if you and the other participants do not declare any endowment for the purpose to be taxed.

Under alternative 2 participants pay taxes only if they declare endowment for taxation. The more you declare, the less you can earn form commodity purchase. On the other hand, the more you declare the higher is the income from the project that you and the other participants receive.