No. 59 october 2010

Lisa Bruttel
Tim Friehe

On the path-dependence of tax compliance

Research Paper Series
Thurgau Institute of Economics and Department of Economics
at the University of Konstanz

Member of

thurgauwissenschaft

www.thurgau-wissenschaft.ch

THURGAU INSTITUTE
OF ECONOMICS
at the University of Konstanz

On the path-dependence of tax compliance*

Lisa Bruttel† Tim Friehe‡

October 20, 2010

Abstract

This paper presents experimental evidence that tax compliance is path dependent. We show that individuals faced with the same current tax enforcement parameters, will nevertheless choose different compliance if they have faced different tax enforcement parameters in the past. This finding has important policy implications. For instance, legal harmonization in the EU cannot be expected to reliably yield similar behavior in countries with different legal histories.

Keywords: tax compliance, path dependence, experiment

JEL-Classification: C91, H26, K42

*We would like to thank Kate Bendrick, Gerald Eisenkopf, Urs Fischbacher, Simeon Schudy, Pascal Sulser, Verena Utikal, Irenaeus Wolff, and Natalie Zimmer for very helpful suggestions.

†University of Konstanz, Department of Economics, Box 131, 78457 Konstanz, Germany. E-mail: lisa.bruttel@uni-konstanz.de. Phone: 0049-7531-88-3214. Fax: 0049-7531-88-2145.

‡University of Konstanz, Department of Economics, Box 136, 78457 Konstanz, Germany. E-mail: tim.friehe@uni-konstanz.de. Phone: 0049-7531-88-2534. Fax: 0049-7531-88-4135.

1 Introduction

The design of mechanisms to steer private decision-making in desired directions is an important aspect of public economics. The policy designer predicts how different institutions influence individual decision-making by considering their impact on individual payoffs. For instance, because governments need to raise funds to finance public goods, increasing the expected costs of tax evasion by raising the audit probability may be viewed as a good policy if seeking to increase tax declarations (e.g., Allingham and Sandmo 1972).

This paper argues that there are circumstances in which this focus on individual payoffs is not sufficient for deriving optimal institutions. In a tax-compliance setup, we establish that factors which influenced benefits and costs in the past still bear importance for tax compliance in the present, although benefits and costs in the present are no longer affected by these factors. Stated differently, if individual A and individual B have faced dissimilar institutions in the past, they may behave differently now even though their expected benefits and costs are the same. Given the difficulty of obtaining appropriate field data to study path dependence, we make use of an experiment to establish our findings. The tax-compliance context is special for several reasons. For instance, it may be argued that (i) this is an area of norm obedience where rational decision-making is most likely, (ii) there are no payoff interdependencies among participants, and (iii) tax enforcement is not as loaded with social meaning as compared to other incentive systems (see, e.g., Franzoni 2009).[1] These facets of the setup help us to keep other aspects out of the decision-making context in order to fully focus on the potential path dependence of compliance behavior.

In our laboratory setting, there are 20 rounds. In each round individuals first earn gross income in a real-effort task. Next, participants are asked to report their gross income, knowing that reported income is subject to an income tax and that any tax evasion may be detected and, in the case of detection, will be penalized. Our treatment variable is the penalty multiplier, which transforms evaded taxes into the penalty level payable upon detection. The level of the penalty multiplier may be either high or low. In terms of expected income it always pays off for the participants to declare an income of zero, irrespective of the treatment.

[1] It has been established that positive and negative incentives may have subtle, counterintuitive effects since they can, for instance, signal a lack of trust (see, e.g., Bowles 2008).

The results suggest that tax compliance is path dependent. Individuals who had a high penalty multiplier in rounds 1-10 and a low penalty multiplier in rounds 11-20, declare a higher share of income in rounds 11-20 than individuals having had a low penalty multiplier in all 20 rounds. Similarly, individuals who had a low penalty multiplier in rounds 1-10 and a high penalty multiplier in rounds 11-20, tend to declare a lower share of income in rounds 11-20 than individuals having been assigned a high penalty multiplier for all 20 rounds.

Our evidence suggests that experience with a law enforcement system affects compliance decisions. Therefore, personal history must be taken into account when determining optimal government policy. These different experiences imply heterogeneity among agents and thus tend to reduce attainable welfare levels if government policy cannot be made contingent on individual experiences. There are very important practical circumstances in which our results are relevant. For example, there are many areas of the law in which the European Union seeks to harmonize national legal frameworks. Environmental crimes are a case in point. There are large differences between the criminal sanctions provided for environmental offenses in the member states. Directive 2008/99/EC on the protection of the environment through criminal law seeks to redress this. Given, however, that compliance with the law may be path dependent, there is little reason to expect that individual behavior in countries whose legal rules were different before the harmonization will rapidly converge as a result of legal harmonization. In a similar vein, immigrants experience different institutions and regulations in their home country before migrating, and thus may respond differently to institutions than natives do. Finally, we refer to an example outside of the context of legal system incentives: the possibility of path dependence is relevant to incentive systems in firms, since the present employees of the firm often have differing previous employers, and different employers imply differing incentive systems.

The paper at hand analyzes the path dependence of the decision concerning compliance with a legal norm and for which some related work can be found in the literature. The decision to undertake a criminal act is often influenced by the level of human capital, and thus by decisions made in the past (see, e.g., Lochner 2010). Similarly, it may be that the privately optimal income declaration at present is dependent on declarations in the past, for example, because detection by tax authorities in the current round may uncover tax evasion in past rounds (see, e.g., Baumann and Friehe 2010). However, in our context, it is parameters of the past that influence the norm compliance decision in the present. There are also other studies which analyze individual tax compliance in the lab. For instance, Alm et al. (2009)

test repercussions of different types of enforcement information dissemination, Alm and McKee (2006) discuss the consequences of audit certainty, and Alm et al. (1995) are interested in social norm effects on individual compliance.[2] However, whether present individual compliance is affected by past enforcement has, to the best of our knowledge, not been dealt with as such in the literature.

The structure of the article is as follows. In Section 2, we present the experimental design and procedures. The behavioral hypotheses are laid out in Section 3, before discussing the experimental results in Section 4. Section 5 concludes the study.

2 Experimental design and procedures

2.1 Experimental design

The experiment lasted 20 rounds, $t = 1, ..., 20$.[3] There were three stages in a round. In Stage 1, participants could earn gross income I_t by performing a real-effort, computer-based task. The earnings task required subjects to find a number in a 12 by 12 matrix. Each cell of the matrix contained a number drawn randomly out of the set of natural numbers $\{1, ..., 9\}$. To earn points, participants had to state the number in row x and column y, x, $y \in \{1, ..., 12\}$. After they stated the correct number, participants obtained 10 points and got a new combination of row and column numbers. An incorrect number prompted an error message which called for another try. This earnings task lasted 45 seconds. In Stage 2, participants were informed via the computer of their gross round income. They were instructed to make an income declaration $S_t \leq I_t$, knowing that a tax rate of $\tau = .3$ was applied to declared income, implying a tax payment of τS_t. In Stage 3, the final stage in a round, it was randomly determined whether the income declaration was audited. The audit probability p was equal to .1. Detected tax evasion, i.e., if there was an audit and $I_t - S_t > 0$ held, led to a penalty of $\tau(I_t - S_t)\gamma_t$, where $\tau(I_t - S_t)$ is the tax evaded and γ_t is the penalty multiplier in round t (Yitzhaki 1974).[4] The

[2] For recent surveys on tax compliance, see Franzoni (2009) and Slemrod (2007).

[3] Our experimental design is similar to that used by Alm et al. (2009).

[4] The round income of a punished subject could be negative. In that case, losses were offset against gains in earlier rounds. Furthermore, participants were given an initial endowment of 240 points for compensation of potential losses in initial rounds.

written instructions and information presented on the computer screens during the experiment used a taxation vocabulary, with phrases such as "income", "tax declaration", "tax evasion", and "penalty payment". In this way, we made sure that individuals were aware of the fact that tax honesty, i.e., $I_t = S_t$, was expected. The penalty multiplier γ_t was the treatment variable and was equal to either 4 or 8. It stayed constant for rounds 1-10 and 11-20 in all treatments. Thus, we may denote γ_{1-10} (γ_{11-20}) as the penalty multiplier applying in the first (second) ten rounds. The experimental design consists of 4 treatments (see Table 1). Treatments 4-4 and 8-8 had the same penalty multiplier throughout all 20 rounds. In treatments 4-8 and 8-4, it changed after 10 rounds.

Treatment	γ_{1-10}	γ_{11-20}	Subjects
4-4	4	4	33
8-4	8	4	30
4-8	4	8	31
8-8	8	8	31

Table 1: Number of subjects (= number of independent observations) per treatment.

All participants were instructed at the beginning of the experiment about the possibility of a change in the rules after 10 rounds. Before the start of round 11, all subjects received a message on their computer screen, saying either that the penalty multiplier would change to another level or that the rules of the experiment, including the level of the penalty multiplier, would remain the same. Only after showing recognition of this information by clicking a button, was round 11 started. At the end of each round, participants were informed of their resultant income, whether an audit of their income declaration took place, and about the size of any taxes or penalties paid. Participants did not obtain information about other participants' behavior.

2.2 Experimental procedures

The experiment was computerized using z-Tree (Fischbacher 2007). A total of 125 students from various disciplines took part in at most one of the four treatments. They were recruited via ORSEE (Greiner 2004). The experiment took place in the *Lakelab*, the laboratory for experimental economics at the University of Konstanz, in May 2010. Sessions lasted less than 90 minutes. The experimental currency was points, with 120 points converted into 1 Euro after

the experiment. On average, participants earned 16.5 Euros in the experiment. The protocol before the start of the experiment was as follows: Subjects first received written instructions for participating in the experiment, and then had to answer control questions which were shown on their computer screen. The experiment started only after all subjects had answered the control questions correctly. After the main experiment, we elicited participants' attitudes toward risk using the Holt and Laury (2002) procedure. At the end of the session, the participants were asked to complete a questionnaire.

3 Behavioral hypotheses

Considering the subject's problem in Stage 2 of a given round t (i.e., taking income as given) we may argue that the participant seeks to

$$\max_{S_t} E_t = pU(A) + (1-p)U(N) \tag{1}$$

where p is the audit probability, U represents the utility function, $A = I_t - \tau S_t - \tau(I_t - S_t)\gamma_t$ is the income level in the audit state of the world, and $N = I_t - \tau S_t$ is the income level in the no audit state of the world. For a given level of income, expected utility responds to a change in S_t according to

$$\frac{dE_t}{dS_t} = \tau\left[(\gamma_t - 1)pU'(A) - (1-p)U'(N)\right]. \tag{2}$$

An increase in the level of declared income increases (decreases) utility in the (no) audit state of the world. Corner solutions may arise for specific combinations of tax enforcement parameters and the tax rate (see, e.g., Allingham and Sandmo 1972). Assuming an interior solution, i.e., that (2) is equal to zero, makes it possible to derive the way in which the privately optimal declaration S_t^* changes with respect to the penalty multiplier:

$$\frac{dS_t^*}{d\gamma_t} = -\frac{1}{\frac{d^2E_t}{dS_t^2}}\tau p\left\{U'(A) - \tau(\gamma_t - 1)(I_t - S_t)U''(A)\right\} \tag{3}$$

which is greater than zero as long as $U' > 0$ and $U'' < 0$ (i.e., for risk-averse participants). This directly leads to our first hypothesis:

H1: A higher penalty multiplier induces a higher level of declared income.

The above formal considerations make use of enforcement parameters of relevance in round

t and do not incorporate enforcement parameters which were applicable in past rounds but are no longer relevant for payoffs. This approach is in line with the intuitive idea that, to direct behavior, the policy maker needs to adequately manipulate individual marginal benefits and costs. However, observations in areas other than tax compliance lend support to our idea that past enforcement parameters are in fact not irrelevant for decision-making in the present (see, e.g., Liebowitz and Margolis 1995). This potential relevance is of central interest in this study. Participants invest conscious effort into solving the problem given the initial penalty multiplier. This basis is, we hypothesize, too internalized to lose its force once one parameter of the set of enforcement parameters is changed. Our design makes it possible to inquire into the consequences of enforcement information lingering in the heads of participants. It is reasonable to expect that present behavior will to some extent be guided by present enforcement parameters, but can also to some extent be driven by past enforcement parameters. Specifically, we construct the following hypotheses concerning responses to changes in the penalty multiplier:

H2: Individuals with penalty multipliers $\gamma_{1-10} = 8$ and $\gamma_{11-20} = 4$ declare more income in rounds 11-20 than individuals with penalty multipliers $\gamma_{1-10} = \gamma_{11-20} = 4$.

H3: Individuals with penalty multipliers $\gamma_{1-10} = 4$ and $\gamma_{11-20} = 8$ declare less income in rounds 11-20 than individuals with penalty multipliers $\gamma_{1-10} = \gamma_{11-20} = 8$.

4 Experimental results

Figure 1 summarizes our experimental evidence. It shows declared income divided gross income averaged over the subjects in each treatment in round t, $t = 1, ..., 20$. Tax declaration in rounds 1-10 is significantly lower if the penalty multiplier is $\gamma_{1-10} = 4$ instead of $\gamma_{1-10} = 8$ (Wilcoxon rank sum test, two-sided, p-value $< .01$). This suggests that subjects respond to the different penalty multipliers in the way predicted by our formal considerations in Section 3.[5] For the treatments with a change in the penalty multiplier, we find a strongly significant upward (downward) shift in declaration rates in 4-8 (8-4) between the first and the second

[5]Note that there is no statistically significant difference between the income declarations in the two treatments 4-4 and 4-8 (p-value = .5121) and between 8-4 and 8-8 (p-value = .6336) in the first 10 rounds.

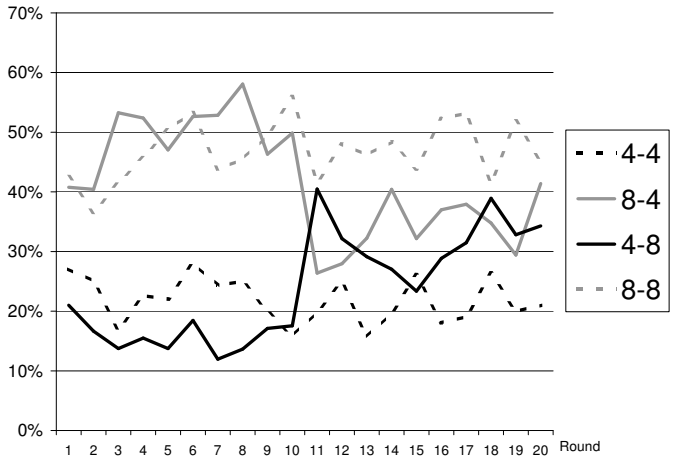


Figure 1: Average tax declaration

halves of the experiment.[6] For the treatments without a change in the penalty multiplier (i.e., treatments 4-4 and 8-8) we find that declaration rates remain constant over time.[7]

Result 1 *A higher penalty multiplier induces a higher level of declared income.*

While participants' income declaration rates respond strongly to the different levels of the penalty multiplier, there is no such variation in observed effort levels. In making this claim, we have taken the number of correct answers in a given round to be a good proxy for participants' effort (see Figure 2). There is a clear upward trend in the number of correct answers, indicating learning in identifying numbers in the matrix. However, the average number of correct answers is not statistically different across treatments with $\gamma_{1-10} = 4$ and $\gamma_{1-10} = 8$ (p-value = .3303).[8] Our conclusion is that effort incentives are not affected by the level of the penalty multiplier.

[6]We tested the average income declaration rate in rounds 1-10 against that in rounds 11-20 using a two-sided Wilcoxon signed rank test, and obtained p-values $< .01$ for both treatments.

[7]The p-values are equal to .5235 for 4-4 and .8746 for 8-8.

[8]In the interest of singling out the effect interest, we only included the values from rounds 1-10 in the comparison.

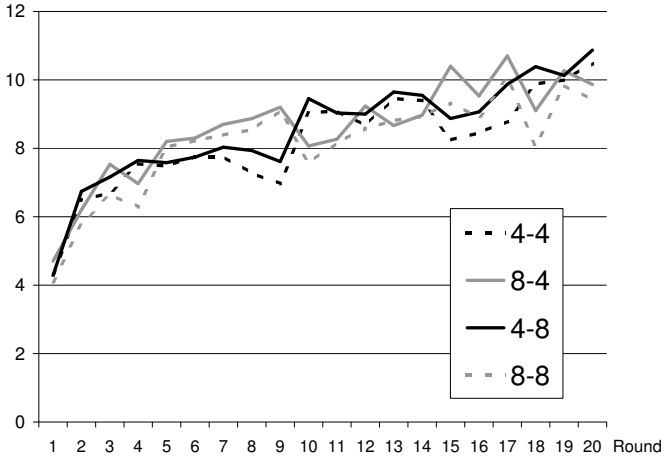


Figure 2: Average number of correct answers

Path dependence of declaration rates

Our central interest is with the potential path dependence of tax compliance. In exploring this issue, we have included in this study the treatments 4-8 and 8-4 in which the penalty multiplier changes after 10 rounds. Since participants in treatment 4-4 and participants in treatment 8-4 faced the same material payoffs in rounds 11-20, the standard argument would be that income declaration rates of the respective groups should be indistinguishable. Similarly, the fact that participants in treatment 8-8 and participants in treatment 4-8 faced the same enforcement parameters in rounds 11-20 should imply similar tax compliance choices by the respective subjects. However, Figure 3 illustrates that this is not the way decision-making actually took place. In rounds 11-20, participants in treatment 8-4 declare more than participants in treatment 4-4. Similarly, participants in treatment 4-8 declare less than participants in treatment 8-8. Actually, in treatment 8-4, income declaration rates in rounds 11-20 remained significantly above those in treatment 4-4 (Wilcoxon rank sum test, two-sided, p-value $< .1$). Comparing treatment 4-8 and treatment 8-8, we obtain exactly the same effect (p-value $< .1$). This supports the conclusion that current behavior is guided by current enforcement parameters, but is at the same time also driven by past enforcement parameters as hypothesized in Section 3. The

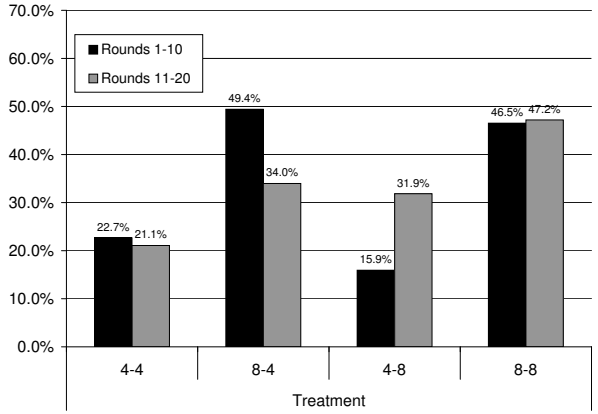


Figure 3: Average income declaration rate in the first and second halves of the experiment

behavior of participants whose penalty multiplier has been changed after 10 rounds falls into "middle ground" and can easily be differentiated from behavior of participants who had the same penalty multiplier in rounds 11-20 but had not experienced a change in its level.[9]

When comparing the income declarations of participants in treatments with the same initial penalty multiplier γ_{1-10}, we find further support for the path dependence of tax compliance. The difference between the income declarations of participants in treatment 4-4 (8-8) and subjects in treatment 4-8 (8-4) in rounds 11-20 is not statistically significant (two-sided tests, p-value (4-4 vs. 4-8) = .1323, p-value (8-4 vs. 8-8) = .1335). Stated differently, behavior by participants in treatment 4-8 in rounds 11-20 is more similar to behavior by participants in treatment 4-4 than to behavior by participants in treatment 8-8. Analogously, participants

[9]We carefully ensured that participants recognized the message informing about the change in the level of the penalty multiplier. First, we announced in the instructions that they would receive new information about the rules of the game after round 10. In addition, they had to click a button after reading this new information before round 11 started. In the data, it shows that these measures were sufficient to draw the subjects' attention to the change. Declaration rates indeed strongly react in round 11 in both treatments 8-4 and 4-8 moving towards the levels of their companion treatments 4-4 and 8-8, respectively. Comparing only the declaration rates in round 11, we find no difference in income declarations, neither between 8-4 and 4-4 (p-value = .3669) nor between 4-8 and 8-8 (p-value = .9761). Rather, path dependence appears to have manifested in later rounds.

in treatment 8-4 behave more similar to participants in treatment 8-8 than to participants in treatment 4-4 in rounds 11-20.

Result 2 *Individuals with penalty multipliers* $\gamma_{1-10} = 8$ *and* $\gamma_{11-20} = 4$ *declare more income in rounds 11-20 than individuals with penalty multipliers* $\gamma_{1-10} = \gamma_{11-20} = 4$.

Result 3 *Individuals with penalty multipliers* $\gamma_{1-10} = 4$ *and* $\gamma_{11-20} = 8$ *declare less income in rounds 11-20 than individuals with penalty multipliers* $\gamma_{1-10} = \gamma_{11-20} = 8$.

As a robustness check of our central finding, we regress the share of income S_t/I_t which is declared in round t on a set of explanatory variables (see Table 2). The explanatory variables include two dummy variables for the value of the penalty multipliers in the first and the second half of the 20 rounds, respectively, controlling for the main treatment effects. These dummy variables take a value one if the penalty multiplier is high ($\gamma = 8$), and zero if it is low ($\gamma = 4$). Furthermore, $punish_{t-1}$ is a dummy variable indicating whether or not a subject was punished in the previous round. We restrict our regression to rounds 12-20 where no more changes in the penalty multiplier occur. Round 11 is thus not included in the regression. The lagged dummy variable $punish_{t-1}$ would otherwise cause behavior from round 10 (before the change in the penalty multiplier) to enter the regression. The index *round* accordingly ranges from 12 to 20. As a control, Table 2 also contains the results of a regression including round 11. The variable $income_t$ measures the gross income generated in round t. We use gross income instead of net income as a control variable for potential wealth effects because gross income enters the tax declaration problem in Stage 2 as an exogenous variable. The variables *male* and *religious* use self-stated answers from the post-experimental questionnaire. The former is a gender dummy variable equal to one for men and zero for women. The latter can take the values 0, .25, .75, and 1, depending on whether participants categorize themselves as "not religious", "rather not religious", "rather religious" or "religious". Finally, $riskaversion$ represents the number of risk-averse choices in the Holt and Laury (2002) procedure.

Table 2 summarizes the regression results. They confirm the lasting effect of the penalty multiplier γ_{1-10}. Past experience with a high penalty multiplier, i.e., with $\gamma_{1-10} = 8$, significantly increases the average declaration, which is qualitatively similar to the effect of the present penalty multiplier $\gamma_{11-20} = 8$.

S_t/I_t	Rounds 12-20			Rounds 11-20		
$\gamma_{1-10} = 8$	.1525**	.1547**	.1316**	.1410**	.1431**	.1214**
	(.0624)	(.0606)	(.0583)	(.0615)	(.0599)	(.0577)
$\gamma_{11-20} = 8$	.1131*	.1139*	.1324**	.1197*	.1201**	.1382**
	(.0625)	(.0608)	(.0583)	(.0615)	(.0602)	(.0578)
$punish_{t-1}$		-.2628***	-.2550***		-.2396***	-.2369***
		(.0382)	(.0363)		(.0365)	(.0350)
$income_t$		.0008	.0011		.0007	.0010
		(.0011)	(.0011)		(.0011)	(.0010)
$round$		.0029	.0023		.0026	.0021
		(.0035)	(.0035)		(.0032)	(.0032)
$male$			-.0520			-.0552
			(.0558)			(.0559)
$religious$			-.0280			-.0314
			(.0784)			(.0776)
$riskaversion$			.0516***			.0479**
			(.0192)			(.0188)
$constant$	.2043***	.1025	-.1785	.2051***	.1174	-.1369
	(.0519)	(.0977)	(.1342)	(.0503)	(.0885)	(.1300)
R^2	.0524	.0843	.1295	.0501	.0770	.1175

Table 2: Regression coefficients: % income declaration. Standard errors in brackets. *** denotes significance at the 1% level, ** at the 5% level and * at the 10% level. Standard errors are clustered by subject.

Being punished has a negative and significant effect on future income declaration. This effect might be due to some participants' lack of understanding of the independence of control probabilities across rounds.[10] In fact, some participants stated in the post-experimental questionnaire that they perceived the risk of being controlled as smaller than usual after a control.[11]

The variables income and time have no significant effect on the level of income declaration while risk aversion significantly increases tax honesty in our experiment. The data does not confirm a statistically significant effect of gender or religiosity on declaration behavior. We, thus, cannot confirm previous findings of the empirical tax compliance literature (see, e.g., Torgler 2007) that people who call themselves religious are more tax honest than others.

Heterogeneity

Figure 4 gives information about the share of subjects who have an average declaration in rounds 1-10 of a given magnitude. For instance, gathering participants in treatments 4-4 and 4-8, we find that 24 out of 63 subjects (38%) declare no income, and only one participant declares total income in all ten rounds. In contrast, in rounds 1-10 only 7 out of 60 subjects (12%) in treatments 8-4 and 8-8 declare no income while 8 declare at least 90 and up to 100%.

Figure 5 shows how subjects with different declaration inclinations in rounds 1-10 change their behavior in rounds 11-20. To simplify the illustration, we cluster subjects into the three groups of low, medium, and high declarers, given their average income declaration in rounds 1-10. The groups are defined such that low declarers declare up to 10% of their income in rounds 1-10, whereas high declarers declare at least 90%. Subjects with declaration rates between 10 and 90% in rounds 1-10 are denominated medium declarers. Here we see that the treatment effects in 4-8 and 8-4 result from reactions of participants from different groups. In 8-4, the decrease in tax honesty is a consequence of a behavioral adaption of both medium and high declarers. In 4-8, some low declarers turn into medium, and some into high, declarers after the

[10] For similar evidence, see Bruttel and Kamecke (2010).

[11] We asked the question '*Did you change your behavior after being controlled and if so, how?*' to 81 of our 125 subjects. 9 participants out of these 81 responded something like "the probability of another control is smaller after being punished" while only 5 correctly noticed that the probability of control continues to be equal to 10%.

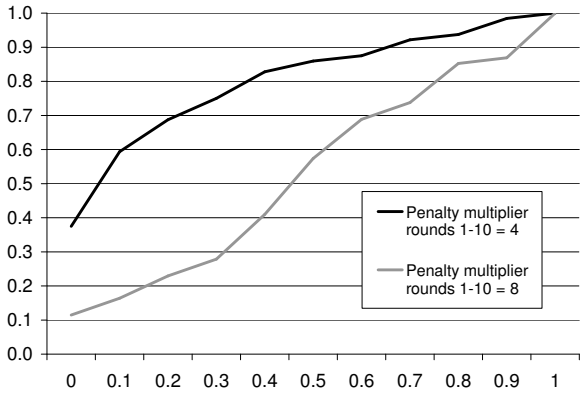


Figure 4: Distribution of income declaration rates in rounds 1-10

increase in the penalty multiplier. In addition, some medium declarers declare almost all of their income in rounds 11-20. The behavior in 8-8 is stable since only few subjects declare a notably different income share in rounds 11-20 compared to rounds 1-10. In 4-4, there are more changes in individual declaration behavior while the average declaration rate over all subjects in this treatment remains stable.

Discussion of results

There are different mechanisms which may be decisive for the observed path dependence. For instance, it has to be noted that path dependence of tax compliance is related to the concept labeled anchoring (see, e.g., Ariely et al. 2003, Stewart 2009). An anchor is commonly defined to be completely irrelevant to the task at hand, but may still influence judgment or behavior. Most prominently, Tversky and Kahneman (1974) spun a wheel of fortune with numbers ranging from 0 to 100, asked subjects whether the number of African nations in the UN was greater or less than that number, and then requested subjects to estimate the actual number of African nations. These estimates were significantly related to the number spun on the wheel (the anchor). Path dependence in tax compliance is created by factors which in fact used to influence benefits and costs in the past and thus used to be relevant to the task at hand. In that regard,

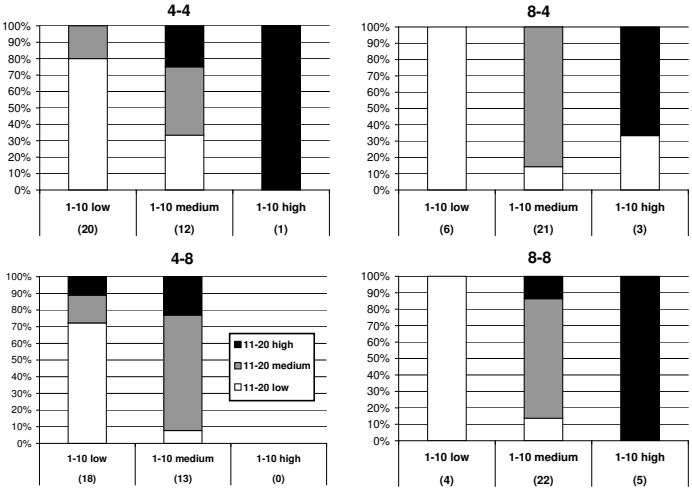


Figure 5: Declaration behavior of low, medium and high declarers from rounds 1-10 in rounds 11-20.

path dependence is different from anchoring. Still, it is possible that past decisions regarding tax compliance function as an anchor for present decisions.

The behavior may also be explained by taking conformity-seeking behavior or compliance to a social norm into account. Individuals may prefer to behave as they think others do (conformity) or they may feel obliged to behave in a certain way (social norm). With regard to the latter, it is conceivable that a norm is introduced simply by having either mild or drastic penalties for tax evasion (see, e.g., Cooter 1998). However, there are no payoff interdependencies in our setup, whereas this is usually the case in studies concerning social norms (see, e.g., Alm et al. 1999). Conformity-seeking behavior is complicated by the fact that individuals do not know how others behave, but they nonetheless form beliefs about others' decisions. The post-experimental questionnaire asked subjects to state the expected average declaration rate of other participants. Subjects with the low penalty multiplier in rounds 1-10 stated lower beliefs (35.61% in 4-4 and 35.48% in 4-8) than those with the high penalty multiplier (47.00% in 8-4 and 50.97% in 8-8). The difference between the expected declaration rate of subjects with $\gamma_{1-10} = 4$ and with $\gamma_{1-10} = 8$ is indeed highly significant (p-value $< .01$). In contrast, there is virtually

no difference between the expected declaration rates of subjects with $\gamma_{11-20} = 4$ and those with $\gamma_{11-20} = 8$ (p-value = .5317). Thus, beliefs about others' compliance behavior mirror the main treatment effect. Individual declaration rates and beliefs about others' declarations have a correlation coefficient of .5756. This correlation gives some support to the idea that subjects condition their tax declaration on their belief about the honesty of others. Such conditioning has been established in other studies of tax compliance (see Frey and Torgler 2007). In addition, however, to the discussed effects, we note that a consensus effect (see Ross et al. 1977) could also explain this similarity between behavior and beliefs stated ex post.

Finally, our data may be interpreted as resulting from the imperfect adaptation to changes in the decision problem. For participants in treatments 4-8 and 8-4, a key parameter changed after ten rounds. This requires that subjects consider how individually optimal behavior needs to adapt in response to this change. However, this re-optimization requires effort, in particular because the decision problem involves a choice in the presence of risk. Participants invested notable conscious effort into solving their initial tax compliance problem in the first ten rounds, based on the penalty multiplier γ_{1-10} contained in the instructions and in the control questions. In contrast, the change in the level of the penalty multiplier after round 10 was recognized "in passing", without much extra time for calculating optimal decisions anew. These combined aspects could lead to a strong argument for reverting to the level of the declaration rate which used to be perceived as optimal given the original statement of the problem, instead of undertaking a true re-optimization.

5 Conclusion

We establish that factors having no impact on current payoffs may still contribute to determining individual choices. In order to establish this point, we provide experimental evidence showing that past enforcement parameters affect present tax compliance behavior. Although the penalty for tax evasion applicable in past rounds is not relevant for current payoffs, individuals who were used to a high penalty declared a higher share of income, given a low level of the penalty, than individuals who continuously had a low penalty. The observed choices suggest that indeed both past and present enforcement parameters have an influence on present individual behavior. This strongly suggests that norm compliance in the tax realm is path dependent.

Our findings have important policy implications. Some of these, for example with respect to EU harmonization, were elaborated on before. The findings concerning round 11 of the experiment also merit discussion in this context. Observing something like the strong initial reaction to the change in the penalty level in our experiment may mislead policy makers evaluating the effectiveness of their policy changes. Behavioral adaptations in the short run do not necessarily provide convincing evidence for the effectiveness of a policy change. Our data shows that individuals may fall back into past decision patterns relatively quickly. In this context, our results suggest that policy makers need to invest in ensuring that individuals affected by the policy change actually concern themselves with its repercussions in determining their optimal behavior.

Appendix: Instructions for treatments with $\gamma_{1-10} = 4$

The following gives the translations of the German instructions for treatments 4-8 and 4-4. Instructions for the other treatments were identical except for the parts concerning the third stage.

General instructions:

Thank you for participating in this experiment.

From now on, please remain seated and do not talk to other participants. These instructions are identical for all participants. Please read them carefully. If you have a question regarding the experiment, please raise your hand. We will come to you to help.

This experiment will last 20 rounds. Each round comprises a sequence of decisions and events. There are three stages in each round, which will be described below. There may be a change in your decision problem after 10 rounds. This (possibly varied) decision problem will be valid for the remaining 10 rounds. Your payment in this experiment depends on your decisions and luck, but not on other participants' decisions.

Your gains and losses are counted in points during the experiment. After the experiment, all points will be added up and you will receive 1 Euro cash for each 120 points you scored in

the experiment. In addition, you receive an initial endowment of 240 points.

Before the experiment starts, we will ask you some control questions. This is to ensure that all participants understand the instructions. Your answers to these questions do not influence your final payment.

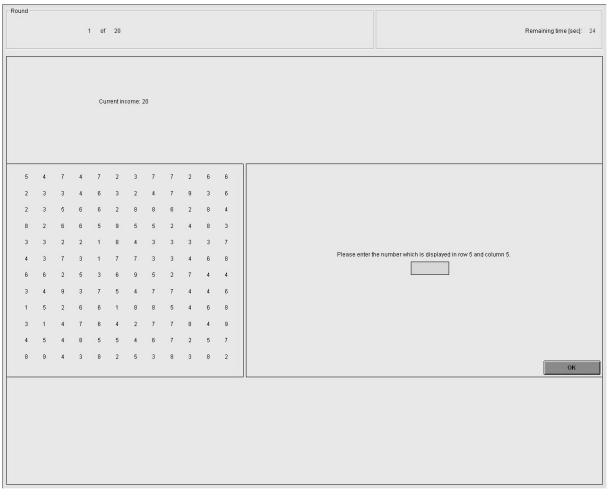
After the main experiment, there will be a lottery experiment. You will receive the instructions for this second part of the experiment on your computer screen after the end of the first part.

Detailed description of one round:

In this experiment, you will go through a series of decisions and events in each round. One round consists of three stages:

1) Earnings task:

In the first stage, you can earn points. You need to find numbers in a matrix with 12 columns and 12 rows. You have 45 seconds to find as many numbers as possible. You receive 10 points for each correct number. The screen of this stage looks as follows:



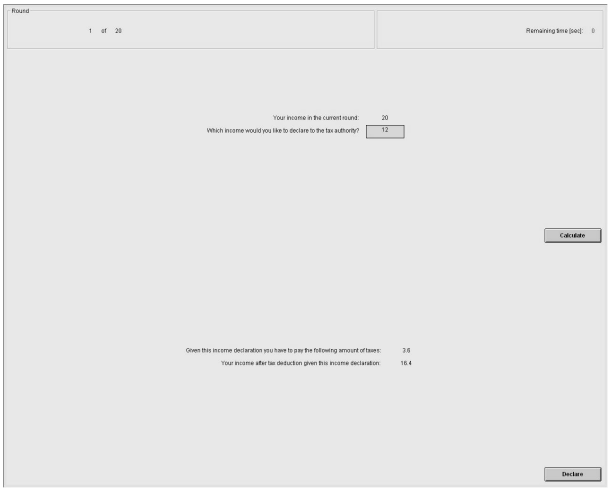
2) Tax declaration:

In the second stage, you must declare your income (in points). Your declared income may be smaller than your actual income, but not larger. Declared income is subject to an income tax of 30%. If you, for example, have a gross income of 20 points and declare an income of 12 points, you have to pay 3.6 points in taxes.

Gross income	Declared income	Taxes paid
20	12	3.6

Table 3: Example with gross income = 20 and declared income = 12.

The tax declaration screen is shown in the following figure:



Enter your declared income into the indicated field. Your declared income may be smaller than or equal to the actual income you earned in this round. Clicking on "calculate" shows you both how many points would be subtracted as taxes and what your net income would be in this round, if that amount were your declared income. Clicking on "calculate" has no other consequence besides providing you this information. Your entry is binding only after you click "submit income".

3) Tax audit:

In the third stage, your tax declaration is audited with a probability of 10%. If you did not declare all of your income in Stage 2 and are controlled, there will be a penalty amounting to the evaded taxes multiplied by the penalty multiplier four. The following tables show one example with and one without tax audit.

Gross income	Declared income	Taxes paid	Taxes not paid	Points subtracted	Net income
20	12	3.6	2.4	9.6	6.8

Table 4: Example with income > declaration and tax audit

Gross income	Declared income	Taxes paid	Taxes not paid	Points subtracted	Net income
20	12	3.6	2.4	0	16.4

Table 5: Example with income > declaration with no tax audit

If you earned 20 points in the first stage and declared 12 points in the second stage, you will receive a net income of 6.8 points if your tax declaration is controlled (10% probability) and 16.4 points in the case that it is not (90% probability).

At the end of each round you will be informed about:

- your round income in points,
- your taxes paid in points,
- whether you have been controlled or not, and
- your net income after deduction of taxes and a potential penalty due to tax evasion.

Please fill out the short questionnaire after the experiment. After you have done so, you will be paid your net income in cash (1 Euro per 120 points).

References

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

Alm, J., McKee, M., 2006. Audit Certainty, Audit Productivity, and Taxpayer Compliance.

National Tax Journal 59, 801-816.

Alm, J., Jackson, B.R., McKee, M., 2009. Getting the Word Out: Enforcement Information Dissemination and Compliance Behavior. Journal of Public Economics 93, 392-402.

Alm, J., McClelland, G.H., Schulze, W.D., 1999. Changing the Social Norm of Tax Compliance by Voting. Kyklos 52, 141-171.

Alm, J., Sanchez, I., de Juan, A., 1995. Economic and Non-Economic Factors in Tax Compliance. Kyklos 48, 3-18.

Ariely, D., Loewenstein, G., Prelec, D., 2003. "Coherent Arbitrariness": Stable Demand Curves Without Stable Preferences. Quarterly Journal of Economics 118, 73-105.

Baumann, F., Friehe, T., 2010. Tax Evasion, Investment, and Firm Activity. Finanzarchiv 66, 1-14.

Bowles, S., 2008. Policies Designed for Self-Interested Citizens May Undermine "The Moral Sentiments": Evidence from Economic Experiments. Science 320, 1605 - 1609.

Bruttel, L., Kamecke, U., 2010. Infinity in the Lab - The Role of the Continuation Rule for Cooperation in Repeated Games. Working Paper.

Cooter, R., 1998. Expressive Law and Economics. Journal of Legal Studies 27, 585-608.

Fischbacher, U., 2007. z-Tree: Zurich Toolbox for Ready-made Economic Experiments. Experimental Economics 10, 171-178.

Fischbacher, U., Heusi, F., 2008. Lies in Disguise. An Experimental Study on Cheating, TWI Working Paper 40, Thurgau Institute of Economics.

Franzoni, L.A., 2009. Tax Evasion and Avoidance. In: Garoupa, N., (ed.). The Encyclopedia of Law and Economics, 2nd edn. Cheltenham: Edward Elgar.

Frey, B.S., Torgler, B., 2007. Tax Morale and Conditional Cooperation. Journal of Comparative Economics 35, 136-159.

Greiner, B., 2004. The Online Recruitment System ORSEE 2.0 - A Guide for the Organization of Experiments in Economics. University of Cologne, Working Paper Series in Economics 10.

Holt, C.A., Laury, S. K., 2002. Risk Aversion and Incentive Effects. American Economic Review 92, 1644-1655.

Liebowitz, S.J., Margolis, S.E., 1995. Path Dependence, Lock-In, and History. Journal of Law, Economics, and Organization 11, 205-226.

Lochner, L., 2010. Education Policy and Crime. NBER Working Papers No. 15894.

Ross, L., Greene, D., House, P., 1977. The "False Consensus Effect": An Egocentric Bias in Social Perception and Attribution Processes. Journal of Experimental Social Psychology 13,

279-301.

Slemrod, J., 2007. Cheating Ourselves: The Economics of Tax Evasion. Journal of Economic Perspectives 21, 25-48.

Stewart, N., 2009. The Cost of Anchoring on Credit-card Minimum Repayments. Psychological Science 20, 39-41.

Torgler, B., 2007. Tax Compliance and Tax Morale: A Theoretical and Empirical Analysis. Cheltenham: Edward Elgar.

Tversky, A., Kahneman, D., 1974. Judgment under Uncertainty: Heuristics and Biases. Science 185, 1124-1131.

Yitzhaki, S., 1974. A Note on: Income Tax Evasion: A Theoretical Analysis. Journal of Public Economics 3, 201-202.

THURGAU INSTITUTE
OF ECONOMICS
at the University of Konstanz

Hauptstr. 90
CH-8280 Kreuzlingen 2

Telefon: +41 (0)71 677 05 10
Telefax: +41 (0)71 677 05 11

info@twi-kreuzlingen.ch
www.twi-kreuzlingen.ch