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

On the path dependence of tax compliance*

Lisa Bruttel[†] Tim Friehe[‡]

This version: April 2013 - First version: November 2010

Abstract

This paper presents experimental evidence that tax compliance is path dependent. For given values of the audit probability and the fine for tax evasion, we compare the income declaration of subjects who faced a change in one of the two parameters to that of subjects who experienced no such parameter change. We show that past tax enforcement regimes continue to have an impact on current income declarations. This finding may be explained by reference-dependent preferences and it has important policy implications. For instance, legal transplants cannot be expected to reliably yield similar behavior in countries with different legal histories.

Keywords: tax compliance, path dependence, reference dependence, experiment
JEL-Classification: C91, H26, K42

*We would like to thank Florian Baumann, Kate Bendrick, Gerald Eisenkopf, Dirk Engelmann, Ido Erev, Urs Fischbacher, Mario Mechtel, Karim Sadrieh, Simeon Schudy, Pascal Sulser, Verena Utikal, Irenaeus Wolff, participants in both the EEA annual meeting 2011 in Oslo and in a research seminar in Magdeburg, and three anonymous reviewers for very helpful suggestions. Alexandra Haas and Natalie Zimmer provided excellent research support.

[†]University of Konstanz, Department of Economics, Box 131, 78457 Konstanz, Germany. lisa.bruttel@uni-konstanz.de. Phone: 0049-7531-88-3214. Fax: 0049-7531-88-2145. (corresponding author)

[‡]University of Bonn, Center for Advanced Studies in Law and Economics, Adenauerallee 24-42, 53113 Bonn, Germany. CESifo, Munich, Germany. E-mail: tim.friehe@uni-bonn.de.

1 Introduction

1.1 Motivation and main results

Tax evasion has existed for as long as taxation has.[1] While it is a concern in every kind of tax system and for all nations around the world, the extent to which tax evasion poses a major problem varies from country to country. This variety exists both in comparisons of developed countries and even more so when developed countries are compared to transition and developing countries (see, e.g., Alm 2012, Andreoni et al. 1998, Fuest and Riedel 2009, Gërxhani and Schram 2006, Slemrod 2007).[2] Greece is a case in point because it seems particularly plagued by tax evasion, a fact that has been attributed to complex regulation and insufficient enforcement (Katsios 2006, Matsaganis et al. 2012). Such a country plagued by high tax evasion may seek to adopt the regulatory regime of a country with significantly less tax evasion, i.e., to opt for a legal transplant. For instance, Greece may implement tax enforcement similar to that of Northern European countries.[3] However, whether the transplanted legal regime will prove similarly effective in the adopting country is unclear. Garoupa (2012) argues that changes in enforcement due to the process of European harmonization may not quickly translate into changes in compliance behavior.[4] In other words, past experience with a different enforcement system may continue to influence tax compliance, that is, produce a kind of *path dependence* of tax evasion.

Such path dependence cannot be explained by a standard expected utility setup, as we show in our section on behavioral predictions. However, the expected utility theory has been criticized because, in many if not most circumstances, both the absolute level of the variable in question *and* how this level compares to a reference point are important for individual well-being and decision-making (Kahneman and Tversky 1979, Köszegi and Rabin 2006, Schmidt

[1]See, for example, Plato's *Politeia*: "whenever there are taxes to be paid, the just man pays more from an equal estate and the unjust less" [343d].

[2]The estimates in Schneider and Enste (2000) on the related concept of the shadow economy are similarly of interest in this context.

[3]In this vein, Germany offered to send tax specialists to Greece (see, e.g., the newspaper article *Germany offers to send tax men to Greece* at www.thelocal.de; site last visited March 15, 2013).

[4]Martinez-Vazquez and Torgler (2009) provide an empirical study for Spain, relying on European and World Values Survey data, that indicates that general institutional reforms, including tax policy and tax administration reforms, led to increases in reported tax morale over a long time horizon.

et al. 2008). With regard to the reference point, it is argued that the subject will use either the status quo, the lagged status quo or expectations (see, e.g., Abeler et al. 2011, Sugden 2003).[5] Transferred to our setting, subjects may use their experience with the old tax regime to benchmark the outcome under the new tax regime. Subjects with experience in different tax regimes will have diverging reference points, which may influence their marginal incentives to declare income. Indeed, using a framework with reference-dependent preferences, we derive our central hypothesis that there is path dependence of tax compliance behavior. Specifically, subjects who have experienced weak enforcement early on will tend to declare less income when faced with strict enforcement than subjects who have only experienced strict enforcement, whereas subjects who have experienced strict enforcement early on will tend to declare more income when faced with weak enforcement than subjects who have only experienced weak enforcement.

In order to empirically study whether or not tax compliance is indeed path dependent in the field, it would be necessary to have data from countries with the same current tax enforcement regime but dissimilar ones in the past, holding everything else constant. This is difficult because tax regimes are rarely easily compared and, most importantly, because tax evasion is by its very nature a hidden activity (Alm 2012).[6] To circumvent such problems, we use an experiment to study the potential path dependence of tax compliance. In our experiment, 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 income tax and that any tax evasion may be detected and, when detected, will be penalized. Our treatment variables are the level of the penalty multiplier and the audit probability, both of which may be either high or low. In terms of expected income, it pays off for a risk-neutral participant to declare an income of zero in all of our treatments. Our setup entails no behavioral interaction or payoff interdependence among participants, which helps us to keep other aspects out of the decision-making context, that is, to concentrate on the potential path dependence of compliance behavior.

Our empirical analysis of the experimental data yields the following results. Comparing groups with and without a parameter change after round 10, we find that past values of the

[5] A different source of reference points without bearing on our study is the outcome of others (e.g., Frank 1985).

[6] The case of East and West Germany presents a scenario in which comparability after the reunification and heterogeneity before the reunification apply. Feld et al. (2008) study the German case focusing on the evolution of reported tax morale.

penalty multiplier and the audit probability indeed have a persistent impact on declaration behavior after the parameter change, irrespective of whether the enforcement incentives became stricter or weaker. For example, individuals who experienced a low penalty multiplier in rounds 1-10 and a high penalty multiplier in rounds 11-20 declared less income in rounds 11-20 than individuals who had a high penalty multiplier in all 20 rounds.

Regarding the question raised earlier about whether a transplanted legal regime will prove similarly effective in the adopting country, our central result suggests a negative answer. Our findings dovetail nicely with results on legal transplants: The result that a legal transplant is more successful when the set of rules is adapted to the adopting country may be interpreted as implying that the transplanted set of rules should accommodate prevailing expectations about individual outcomes (see, e.g., Berkowitz et al. 2003).

1.2 Related literature

The use of experimental methods to study tax compliance has been well-established. Some examples include Becker et al. (1987), Alm, Jackson, and McKee (1992, 2009), Alm, McClelland, and Schulze (1992, 1999), Alm and McKee (2004, 2006), Coricelli et al. (2010), Cullis et al. (2012), and Gërxhani and Schram (2006). Regarding the comparative statics with respect to the level of the audit probability and the level of the fine that are the focus of our study, the experimental studies of tax compliance mostly find that tax compliance increases with both tax enforcement parameters (see, e.g., the recent survey by Alm 2012). However, to the best of our knowledge, the question of whether present individual compliance is affected by past tax enforcement parameters has not yet been addressed in the literature.

Spillover effects have been studied in other domains than taxation. These contributions consider the effects of incentives once they are removed, that is, the behavioral spillovers of incentives provided at an earlier point in time. For example, a field experiment conducted by Meier (2007) establishes that matching donations in one time period can dampen charity contributions after the matching is phased out. Famously, Gneezy and Rustichini (2000) show that the introduction of fines for tardy parents at a childcare center increases tardiness even after the fines are abolished. Gächter et al. (2008) establish that the use of explicit incentives in an employment relationship can crowd out voluntary cooperation after the reliance on such incentives is discontinued, whereas Brandts and Cooper (2006) and Hamman et al.

(2007) analyze behavioral repercussions of temporary strong incentives in the minimum-effort game. Having an interest in the long-lasting effects of the rules of the game instead of explicit incentives, Bohnet and Huck (2004) study whether or not having had a trust-supporting fixed-partner setup spills over to interactions with random strangers. Moreover, instead of focusing on behavioral changes over time in a given decision context (i.e., exploring whether there are spillovers across periods of time) as we do in the present study, it is also possible to consider behavioral spillovers across decision contexts. Cason et al. (2012) show that cooperation in a minimum-effort game can be achieved through a behavioral spillover from a median-effort game when the former is played after the latter. Similar results have been found by Ahn et al. (2001) and Knez and Camerer (2000), among others. In contrast to the aforementioned studies, we consider a setting without strategic effects, thereby ruling out this channel for potential path dependence.[7]

The explanation put forth in this paper for the observed path dependence of tax compliance relies on reference-dependent preferences. The literature on this topic spans important theoretical contributions such as Apesteguia and Ballester (2009), Köszegi and Rabin (2006, 2007), and Sugden (2003) as well as empirical analyses including Abeler et al. (2011) and Crawford and Meng (2011). Our derivation of hypotheses can be compared to Bernasconi and Zanardi (2004) and Dhami and al-Nowaihi (2007), given that their studies for the prospect theory framework similarly take a role of reference points in tax compliance for granted (see also the survey by Hashimzade et al. forthcoming). However, following Sugden (2003), we consider state-dependent reference points; we also abstain from assuming probability weighting and consider a dynamic setting with lagged reference points.

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; the experimental results are detailed in Section 4. Section 5 discusses our findings, and Section 6 concludes the study.

[7]Under the title path dependence, other contributions consider very different mechanisms (see, e.g., the survey by Liebowitz and Margolis 2000). As with most of the experimental studies, these contributions rely, in one way or another, on effects resulting from interactions between individuals or firms. For example, it is argued that increasing returns to scale can cause path dependence for competing technologies and the location of industrial centers (Arthur 1989, Krugman 1994). Such path dependence is thus at best loosely related to the within-subject behavioral spillovers in our study.

2 Experimental design and procedures

2.1 Experimental design

The experiment lasted for 20 rounds, $t = 1, ..., 20$.[8] There were three stages in a round. In Stage 1, participants could earn gross income I_t by performing a computer-based, real-effort task. This earnings task required subjects to find numbers in a 12 by 12 matrix. Each cell of the matrix contained a number drawn randomly out of the set of natural numbers $\{1, ..., 9\}$. Participants had to state the number in row x and column y, $x, y \in \{1, ..., 12\}$. For each correct number they received 10 points. An incorrect number prompted an error message that called for the subject to try again. After stating the correct number, participants received a new combination of row and column number. This effort task lasted for 45 seconds.[9] In Stage 2, participants were informed via the computer of their gross income I_t in that round. They were instructed to make an income declaration S_t ($\leq I_t$), knowing that a tax rate of $\tau = .3$ was applied to their declared income (implying a tax payment of τS_t). In Stage 3, the final stage in the round, the income declaration was controlled with the audit probability p_t. Detected tax evasion (i.e., when there was an audit and $S_t < I_t$) led to a penalty of $\tau(I_t - S_t)\gamma_t$, where $\tau(I_t - S_t)$ was the tax evaded and γ_t the penalty multiplier in round t (Yitzhaki 1974).[10] The written instructions and the 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., $S_t = I_t$, was expected.[11]

The treatment variables in our experiment were the penalty multiplier and the audit probability. Tax enforcement parameters stayed constant for rounds 1-10 and 11-20 in all treatments. The penalty multiplier γ_k ($k \in \{1 - 10; 11 - 20\}$) was equal to either 4 or 8. The audit probability p_k took values of either .1 or .2. Depending on the treatment, either one of the two

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

[9]Repeating the real-effort task at the beginning of each round instead of only once before the first round made our design more natural and was intended to strengthen the subjects' perception of having earned the income they would pay taxes on.

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

[11]See Appendix A for a translated version of our instructions.

parameters changed after ten rounds or both parameters remained the same in all 20 rounds. When the penalty multiplier was varied, the audit probability remained fixed at .1. When the audit probability was varied, the penalty multiplier remained fixed at 4 (see Table 1). This yielded seven treatments in total. Treatment BothLL is the baseline with low values for both parameters ($\gamma_{1-10} = \gamma_{11-20} = 4$ and $p_{1-10} = p_{11-20} = .1$) in all 20 rounds. Treatments FineLH, FineHL, and FineHH have a high penalty multiplier in either rounds 1-10 or rounds 11-20 or both. Treatments ProbLH, ProbHL, and ProbHH have a high audit probability in either rounds 1-10 or rounds 11-20 or both.

Treatment	γ_{1-10}	γ_{11-20}	p_{1-20}	Subjects	Treatment	p_{1-10}	p_{11-20}	γ_{1-20}	Subjects
BothLL	4	4	.1	33					
FineHL	8	4	.1	30	ProbHL	.2	.1	4	34
FineLH	4	8	.1	31	ProbLH	.1	.2	4	33
FineHH	8	8	.1	31	ProbHH	.2	.2	4	32

Table 1: Treatments.

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, stating either how the tax enforcement parameters would change or that the rules of the experiment would remain the same. Round 11 began only after the participants demonstrated their recognition of this new information by clicking a button. At the end of each round, participants were informed of their income, of whether an audit of their income declaration had taken place, and of the size of any taxes or penalties paid. Participants did not obtain any information about the behavior of other participants.

2.2 Experimental procedures

The experiment was computerized using z-Tree (Fischbacher 2007). A total of 224 students from various disciplines recruited via ORSEE (Greiner 2004) took part in one of the seven treatments at most. The experiment took place in the *Lakelab*, the laboratory for experimental economics at the University of Konstanz, in May 2010 and May 2012. 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 15.29 euros in the experiment, including a 2 euro show-up fee paid to cover potential losses in the first rounds. The protocol before the start of the experiment was as follows: Subjects first received written instructions for participating in the experiment, and were then asked to answer control questions 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. Finally, the participants were asked to complete a questionnaire. At the end of the session, participants were called one by one to the exit. They received their payment in cash outside the laboratory with sufficient time between two participants to ensure privacy with respect to the amount of money received.

3 Behavioral hypotheses

In each round of our experiment, the individuals determine their effort in the first stage and then decide what level of income to declare to the tax authority in the second stage. We detail the decision on declared income, considering the effort decision as given. In our empirical analysis, we will similarly condition the level of declared income on the level of gross income (i.e., on effort).[12]

In this section, we will consider the predictions that follow from a standard expected utility setup and contrast them with those derived from reference-dependent expected utility (Sugden 2003). In the theory of reference-dependent expected utility, the reference point is exogenous and not explained. In this regard, we make use of a lagged status quo, which has empirical support (e.g., Baucells et al. 2011, Genesove and Mayer 2001) and naturally allows for spillover effects from enforcement parameters relevant at earlier points in time. Specifically, we assume that participants do not bring a reference point to the lab. However, it is natural to conceptualize that the point in time at which the rules for the second part of the experiment are announced presents an opportunity to consider the experience up to this point as a reference point for the subsequent part. Individuals who optimize given the tax regime in part 1 (that is, rounds 1-10) will use this experience to benchmark the outcome in part 2 (that is, rounds 11-20). We will thus simplify by relying on a two-part framework, which admittedly does not capture all

[12]This does not challenge the idea that different levels of tax enforcement parameters and possible changes in their magnitudes can influence the private optimum regarding effort provision. Nevertheless, our focus is on tax evasion, and the experimental data suggests that treatment differences are mainly restricted to this stage of decision-making.

complexities of the decision-making situation at hand, but contains the elements central to our research. We are primarily interested in how the one-time change in tax enforcement influences subsequent declaration behavior.

3.1 Predictions for the standard expected utility setup

Firstly, consider a standard expected utility setup (e.g., Alm et al. 2009). In part t, $t = 1, 2$, participants select the level of income declaration S_t that maximizes their expected utility, taking into account the audit probability p_t, the tax rate τ, and the level of the sanction γ_t, where $\gamma_t > 1 > p_t, \tau > 0$. Expected utility in part t can be written as

$$E_t = p_t U(A_t) + (1 - p_t) U(N_t), \tag{1}$$

where $A_t = I_t - \tau S_t - \tau\gamma_t(I_t - S_t)$ is the net income level in the audit state of nature, and $N_t = I_t - \tau S_t$ is the net income level in the no-audit state of the world.[13] Note that A_t increases with the level of S_t, whereas N_t decreases with S_t. Intuitively, declaring all income maximizes net income in the audit state of the world, whereas declaring no income maximizes net income in the no-audit state.

The marginal influence of an increase in the level of declared income on expected utility is given by

$$\frac{dE_t}{dS_t} = \tau\left[(\gamma_t - 1)p_t U'(A_t) - (1 - p_t)U'(N_t)\right]. \tag{2}$$

An increase in the level of declared income increases utility in the audit state of the world, because a higher declaration increases net income in that state and $U' > 0$. An increase in the level of declared income decreases utility in the no-audit state of the world, because a higher declaration decreases net income in that state. The privately optimal level of income declaration $S^*(p_t, \gamma_t)$ balances out the marginal utility from, on the one hand, income after taxes and penalties in the audit state of nature and, on the other hand, from income after taxes

[13] This part-wise optimization is most natural, as the information provided in part 1 does not comprise information on the parameters relevant to the decision-making in part 2. However, given that there are no interdependencies between parts in this standard expected utility setup, note that the results are not affected if we substitute that the subject maximizes E_1 (E_2) with regard to the declaration in part 1 (2) for the assumption that the subject maximizes $E_1 + \delta E_2$ in part 1 with δ as discount factor when all parameters were known at the beginning of part 1.

in the no-audit state. All of the tax enforcement parameter combinations that we consider in our experiment imply that choosing $S_t = 0$ maximizes expected payoffs (due to $p_t \gamma_t < 1$). As a result, risk-neutral and risk-seeking individuals (i.e., individuals with $U'' = 0$ and $U'' > 0$, respectively) will not declare any income. In contrast, risk-averse subjects (i.e., individuals for whom $U'' < 0$ holds) will choose an interior declaration level (although corner solutions may also arise for risk-averse individuals under specific combinations of tax enforcement parameters, utility function, and the tax rate; see, e.g., Allingham and Sandmo 1972). Assuming an interior solution for $S^*(p_t, \gamma_t)$ (i.e., assuming that (2) is equal to zero), we derive how the privately optimal declaration of income changes with respect to the penalty multiplier and the level of the audit probability:

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

$$\frac{\partial S^*}{\partial p_t} = -\frac{1}{\frac{d^2 E_t}{dS_t^2}} \tau \left\{ (\gamma_t - 1) U'(A_t) + U'(N_t) \right\}. \tag{4}$$

Both are greater than zero for risk-averse individuals, using $d^2 E_t / dS_t^2 < 0$. We summarize as follows.

Proposition 1 *Suppose that all subjects maximize* (1) *in parts 1 and 2. In that case:*
(i) A higher penalty multiplier γ_t and a higher audit probability p_t induce a weakly higher level of declared income $S^*(p_t, \gamma_t)$.
(ii) The optimal declaration $S^*(p_2, \gamma_2)$ in part 2 is independent of the audit probability and the penalty multiplier in part 1, p_1 and γ_1, respectively.

Proof. Claim (i) uses (3)-(4) for interior $S^*(p_t, \gamma_t)$. The possibility that a higher penalty multiplier and a higher audit probability do not induce a higher level of declared income pertains to the scenario with boundary solutions. Risk-neutral and risk-seeking subjects will declare no income irrespective of which of our tax enforcement parameter combinations applies, such that their optimal declaration does not change when γ_t or p_t is increased. Claim (ii) makes use of the fact that the condition that is solved by $S^*(p_2, \gamma_2)$, that is, (2) for $t = 2$, is independent of p_1 and γ_1. ■

3.2 Predictions for the reference-dependent expected utility setup

Next, we consider a reference-dependent expected utility setup following Sugden (2003). We assume that the reference point is a lagged status quo. Furthermore, we suppose that subjects select the level of declared income in part 1 without regard to the fact that the outcome in part 1 (that is determined by $S^*(p_1, \gamma_1)$) acts as a reference point for the subjects in part 2.[14] A myopic decision by the participant seems reasonable in our context, particularly due to the fact that, in our design, participants lack information about the decision-making context in the second part (making it difficult to forecast the likely repercussions of the reference point). In the remaining section, we will first present the specification of reference-dependent expected utility used in our analysis and then describe how gains and losses relative to the part 1 outcome may change marginal incentives in part 2. We will establish that two thresholds emerge for the level of declared income in part 2, which are important since they determine whether or not the subject is in the loss or gain region in the audit and no-audit state, respectively. As a next step, we consider the way in which these thresholds are shaped by the tax enforcement parameters in part 1 to prepare our discussion of potential path dependence of tax compliance, that is, of potential differences in the level of income declared by subjects with a change in tax enforcement on the one hand and the level of income declared by subjects without a change in tax enforcement on the other.

Our specification features a role for both direct utility and gain-loss utility, and that gains and losses relative to the reference outcome are assessed as being contingent on the state of nature (Köszegi and Rabin 2006, 2007, Sugden 2003).[15] More specifically, we assume that expected utility in part 2 is given by

$$E_2^R = p_2\{U(A_2) + \mu_a[U(A_2) - U(A_1)]\} + (1 - p_2)\{U(N_2) + \mu_n[U(N_2) - U(N_1)]\}. \quad (5)$$

[14] In contrast, Koch and Nafziger (2011) and Falk and Knell (2004), for example, study settings in which decision-makers influence reference points directly.

[15] The main difference between the approaches by Köszegi and Rabin (2006, 2007) and Sugden (2003) lies in how an outcome is compared with the reference lottery. In the former framework, each outcome is compared with all outcomes in the support of the reference lottery. In the latter framework, an outcome is compared only with the outcome that would have resulted from the reference lottery in the same state. We follow the approach proposed by Sugden (2003), which is also used in Schmidt et al. (2008), for example, and judged to be more realistic in some cases even by Köszegi and Rabin (2007, Fn. 4).

The respective net income levels in the 'audit' and 'no audit' states that result in part 1, A_1 and N_1, represent the reference points in the determination of the level of declared income in part 2, denoted S_2, thereby influencing gain-loss utility. In order to arrive at gain-loss utility, gains and losses are weighted by the coefficient of loss aversion μ_j, $j = a, n$, where

$$\mu_j = \begin{cases} \lambda\eta & \text{if } J_2 \leq J_1 \\ \eta & \text{if } J_2 > J_1 \end{cases} \tag{6}$$

with $\eta > 0$, and $J_t = A_t, N_t$. The piece-wise linear specification of gain-loss utility is common in the literature (see, e.g., Abeler et al. 2011, Gill and Stone 2010, Köszegi and Rabin 2006). We will assume $\lambda > 1$ in the following argumentation, reflecting loss aversion.[16]

The marginal influence of an increase in the level of declared income on reference-dependent expected utility is given by

$$\begin{aligned} \frac{dE_2^R}{dS_2} = & \underbrace{\tau\left[(\gamma_2 - 1)p_2 U'(A_2) - (1-p_2)U'(N_2)\right]}_{\alpha} \\ & + \underbrace{\tau\left[(\gamma_2 - 1)p_2 U'(A_2)\mu_a - (1-p_2)U'(N_2)\mu_n\right]}_{\beta}, \end{aligned} \tag{7}$$

and comprises an effect on direct utility (Term α) and on gain-loss utility (Term β). With respect to the latter, the relative importance of influencing the utility comparison in the audit and the no-audit state of nature is illustrated by the coefficients of loss aversion, μ_a and μ_n. These coefficients take one of two possible values, depending on the level of income declared in part 2 as is described more precisely below. The definition of μ_j in (6) clearly shows that the marginal influence of an increase in the level of declared income is not continuous for all S_2. We will denote the level of declared income that maximizes reference-dependent expected utility in part 2 by $S^R(p_1, p_2, \gamma_1, \gamma_2)$.

Along the lines of the exposition in Abeler et al. (2011), we now seek to describe the levels of the part 2 income declaration S_2 for which the marginal benefits of an increase in S_2 are inflated relative to marginal costs (when compared to the standard incentives that flow from (2)) and under what circumstances the reverse holds. Relatively higher marginal benefits result when, for a given level of S_2, a subject is in the loss region in the audit state of the world and in the gain region in the no-audit state of the world, such that $\mu_a > \mu_n$. This case would arise when the net income in the audit state of nature in part 2 falls short of that in part 1 (i.e.,

[16] On the loss aversion coefficient, see, for instance, Köbberling and Wakker (2005).

$A_2 \leq A_1$), whereas the net income in the no-audit state of the world in part 2 exceeds the net income from part 1 (i.e., $N_2 > N_1$). Such a combination results, for example, when S_2 is less than the first period declaration $S^*(p_1, \gamma_1)$, the gross income in part 2 is not less than that in part 1, and the penalty multiplier is constant. More precisely, from $A_2 \leq A_1$, we find that a subject is in the loss region in the audit state when

$$S_2 \leq \frac{I_2(\tau\gamma_2 - 1) - I_1(\tau\gamma_1 - 1)}{\tau(\gamma_2 - 1)} + \frac{\gamma_1 - 1}{\gamma_2 - 1} S^*(p_1, \gamma_1) = C_a(p_1, \gamma_1, \gamma_2), \tag{8}$$

such that $\mu_a = \lambda\eta$. Thus, a declaration $S_2 \leq C_a$ will cause the net income in part 2 to compare unfavorably with the net income from part 1. This is intuitive given that A_t increases with S_t. When considering the no-audit state of the world, we deduce from $N_2 \leq N_1$ that a subject is in the loss region when

$$S_2 \geq \frac{I_2 - I_1}{\tau} + S^*(p_1, \gamma_1) = C_n(p_1, \gamma_1), \tag{9}$$

such that $\mu_n = \lambda\eta$. Thus, a declaration $S_2 \geq C_n$ will cause the net income in part 2 to compare unfavorably with the net income from part 1. The intuition is that N_t decreases with S_t. The comparison of net income levels from the no-audit state of nature, N_2 and N_1, is influenced by the tax enforcement parameters applicable in part 1 only because the level of declared income $S^*(p_1, \gamma_1)$ is shaped by these, whereas there is also a direct effect of the penalty multiplier regarding the comparison of net income levels in the audit state.

The two thresholds C_a and C_n cannot be ranked at a general level. Using them, however, we are in the position to state how marginal benefits and marginal costs are affected by gain-loss utility.[17]

Lemma 1 *(i) When* $0 \leq S_2 < \min\{C_a, C_n\}$*, then the ratio of marginal benefits from increasing the level of declared income to marginal costs is higher in* (7) *than in* (2).
(ii) When $S_2 \in (\min\{C_a, C_n\}, \max\{C_a, C_n\})$*, then the ratio of marginal benefits from increasing the level of declared income to marginal costs is the same in* (7) *as in* (2).
(iii) When $I_2 \geq S_2 > \max\{C_a, C_n\}$*, then the ratio of marginal benefits from increasing the level of declared income to marginal costs is lower in* (7) *than in* (2).

Proof. When comparing the ratios of marginal benefits to marginal costs that follow from (7) and (2), the applicable levels of μ_a and μ_n (as defined in (6)) are of critical importance, because

[17] The ranges discussed in Lemma 1 do not necessarily exist all at the same time, as the thresholds may be negative or greater than I_2 for some parameter configurations and $S^*(p_1, \gamma_1)$.

how the ratio $(1+\mu_a)/(1+\mu_n)$ compares to one is decisive regarding whether or not the ratio differs between the two settings. In order to determine which of the two possible levels μ_j takes, we rely on the definitions (8) and (9). ■

Importantly, when the level of declared income in part 2 is below (above) the minimum (maximum) of the two thresholds, the marginal benefit of increasing S_2 is higher (lower) than in the standard model.[18] Note that both thresholds, C_a and C_n, increase in the level of part 2 gross income, where $I_2 > I_1$ may result from learning of the subjects in our experiment. This connotes that the range described in Lemma 1 (i) will be relatively more important, implying a tendency for relatively higher income declarations due to being in the loss region in the audit state of the world and in the gain region in the no-audit state of the world.

Before we turn to how the tax enforcement parameters applicable in part 1 influence the levels of the thresholds, we would like to mention the case of risk-neutral subjects briefly. While the prediction in the standard expected utility framework for these individuals is that they do not declare any income for the tax enforcement parameters considered in our experiment, this may change when preferences are reference-dependent and the range $0 \leq S_2 < \min\{C_a, C_n\}$ exists. However, given that risk-neutral subjects have $S^*(p_1, \gamma_1) = 0$ in (8) and (9), the existence of that range is not guaranteed. Furthermore, the effect due to reference-dependent preferences would have to be sufficiently strong to justify opting against the boundary solution. In summary, it may be expected that risk-neutral subjects declare no income in part 2, and thereby do not change declaration behavior subsequent to the change in tax enforcement.

The tax enforcement parameters applicable in the first part influence the levels of the thresholds that determine whether or not the subject is in the loss or the gain region in a specific state of the world in part 2. In this way, experience with a different tax regime in part 1 may imprint on the declaration behavior of subjects with reference-dependent preferences in part 2. The influence on the thresholds can be explicated as follows:

[18] From Lemma 1, marginal benefits may be relatively inflated only for small levels of S_2, while marginal costs may be relatively inflated only for high levels of S_2. As a result, we still have a concave programming problem, complicated by the discontinuities. This is a characteristic that argues for a unique maximum and produces a tendency against boundary solutions.

$$\begin{aligned}
\frac{\partial C_a}{\partial p_1} =& \frac{\gamma_1 - 1}{\gamma_2 - 1}\frac{\partial S^*(p_1,\gamma_1)}{\partial p_1} > 0 && (10)\\
\frac{\partial C_n}{\partial p_1} =& \frac{\partial S^*(p_1,\gamma_1)}{\partial p_1} > 0 && (11)\\
\frac{\partial C_a}{\partial \gamma_1} =& \frac{S^*(p_1,\gamma_1) - I_1 + (\gamma_1 - 1)\partial S^*(p_1,\gamma_1)/\partial\gamma_1}{\gamma_2 - 1} && (12)\\
\frac{\partial C_n}{\partial \gamma_1} =& \frac{\partial S^*(p_1,\gamma_1)}{\partial \gamma_1} > 0. && (13)
\end{aligned}$$

Thus, a higher audit probability in part 1 increases the range of levels of S_2 that imply a loss in the audit state of the world and increases the range of levels of S_2 that imply a gain in the no-audit state of the world, as follows from (10) and (11). In this way, a higher audit probability makes the range described in case (i) in Lemma 1 applicable to more S_2, that is, the case with relatively higher marginal benefits of declaring an additional unit of income. In contrast, a higher audit probability makes the range described in case (iii) in Lemma 1 applicable to fewer levels of S_2, that is, the case with relatively lower marginal benefits of declaring an additional unit of income. When the penalty multiplier is increased, the effect is not clear with respect to the audit state of nature. Firstly, a higher γ_1 increases the penalty payment in part 1, which lowers C_a. Secondly, a higher γ_1 induces an increase in the level of declared income in part 1, which increases C_a. In contrast, a higher penalty multiplier in the first part clearly increases C_n, such that more levels of S_2 imply a gain in the no-audit state of the world, as follows from (13).

We now turn to the possibility of finding *path dependence* of tax evasion. For this, we discuss how subjects with a change in tax enforcement may be differently affected by gain-loss utility considerations when compared to participants who have not experienced such a change. For this discussion, we will assume that the level of income I_t does not vary across treatments, and denote the critical values for subjects with experience in only one tax regime by K_j instead of C_j to avoid misunderstandings.[19] From (8) and (9), it is clear that $K_a \leq K_n$ since $\gamma_1 = \gamma_2$ and $\tau < 1$. Note that the marginal incentives of subjects with and without changes in the tax enforcement parameters (detailed in equation (7)) are the same except for the levels of μ_a and

[19] In our framework, the assumption that gross income levels (i.e., effort) do not vary across treatments is sensible, because subjects lack interesting uses of time that compete with the real-effort task. Indeed, the alternative to performing the real-effort task is to stare at a grey computer screen for 45 seconds until the next screen appears. Thus, we have reason to believe that subjects tend to invest maximum effort in the task, independently of the finer details of our incentives provided.

μ_n. The latter differ between the two groups only because of the different regime experiences (which are manifested in the different levels for the critical thresholds).

Regarding the possibility of path dependence of tax evasion, we arrive at the following proposition.

Proposition 2 *Suppose that $S^R(p_1, \gamma_1, p_2, \gamma_2) \in (0, I_2)$.*
(a) Subjects in ProbLH declare weakly less income in part 2 than individuals in ProbHH.
(b) Subjects in ProbHL declare weakly more income in part 2 than individuals in BothLL.
(c) Subjects in FineLH declare weakly less income in part 2 than individuals in FineHH when (i) $\partial C_a/\partial\gamma_1 > 0$, and when (ii) $\partial C_a/\partial\gamma_1 < 0$ unless the declaration of subjects in FineHH falls within the range $(K_a, \min\{C_a, C_n\})$ or $(K_n, \max\{C_a, C_n\})$, should these exist.
(d) Subjects in FineHL declare weakly more income in part 2 than individuals in BothLL when (i) $\partial C_a/\partial\gamma_1 > 0$, and when (ii) $\partial C_a/\partial\gamma_1 < 0$ unless the declaration of subjects in BothLL falls within the range $(C_a, K_a]$.

Proof. See Appendix B. ■

Experience with a tax regime influences the reference point in a way that reflects in the declaration behavior in part 2. More precisely, the tax enforcement in part 1 co-determines whether or not the subject is in the loss or the gain region in a specific state of the world for some S_2 in part 2 via its impact on the thresholds C_a and C_n. Two subjects with different experiences from part 1 may then be distinguished by the fact that one perceives higher marginal benefits from an increase in the level of declared income at some S_2, because the outcome in the audit state of nature is in the loss range for one subject but not the other, for instance. Take the contrast between subjects in ProbLH and ProbHH as an example, which we will illustrate based on Figure 1.[20]

In Figure 1, we use (i) as a shortcut to indicate case (i) detailed in Lemma 1, that is, the range of income declarations in part 2 that are associated with inflated marginal benefits, and accordingly for (ii) and (iii). It follows that the range of S_2 with inflated marginal benefits is strictly smaller for subjects in ProbLH than for subjects in ProbHH, while the range of S_2 with inflated marginal costs is strictly larger in ProbLH. This ensures that the income declaration

[20] The ranking $C_a < K_a < C_n < K_n$ is an example. Generally, we know that $\min\{C_a, C_n\} < K_a$ and $\max\{C_a, C_n\} < K_n$, such that $C_n < K_a$ is also possible.

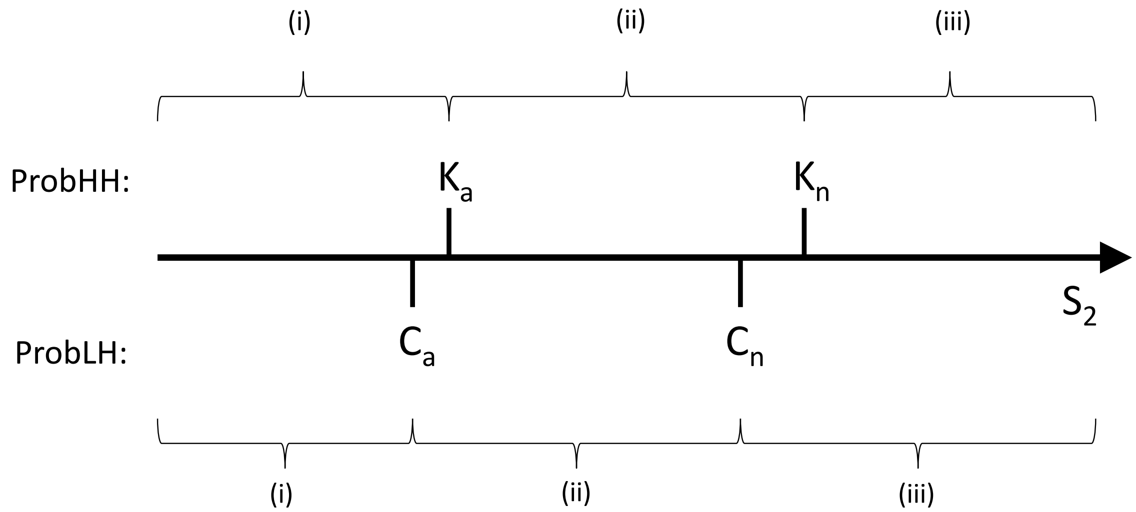


Figure 1: Exemplary configuration for ProbLH and ProbHH.

of subjects in ProbLH is weakly less than that of subjects in ProbHH, as the ratio of marginal benefits to costs is weakly smaller for the former participants than that for the latter. In other words, a subject in ProbLH will exit the loss range in the audit state of nature at a smaller level of S_2 than a subject in ProbHH (as $C_a < K_a$). The subject in ProbLH chose a smaller first-part income declaration in view of the weak enforcement in part 1. This smaller first-part income declaration translates into a lower net income in the audit state of nature. As a result, the net income in part 2 exceeds that of part 1 at a lower level of S_2 for subjects in ProbLH than is true for subjects in ProbHH. Similarly, a subject in ProbLH will enter the loss range in the no-audit state of nature at a lower level of S_2 than a subject in ProbHH (as $C_n < K_n$). This is precisely because the former chose a smaller income declaration in part 1, which translates into a higher net income in the no-audit state of nature. As a result, the net income in part 2 falls short of that of part 1 at a lower level of S_2 for subjects in ProbLH when compared to subjects in ProbHH. In other words, subjects in ProbLH perceive a relatively stronger bias in favor of a lower declaration than subjects in ProbHH. It is forces such as these that may materialize in a path dependence of tax evasion.

4 Experimental results

Our experiment has two decision stages, the real-effort task and the income declaration. In the following, we will first describe the development of gross income I_t over time and consider possible treatment differences in this stage. Next, we will turn to our main interest, the impact of tax enforcement on declaration behavior in the second stage. We will divide this latter discussion into a descriptive analysis and a regression. Finally, we will briefly discuss additional insights from our regression regarding the effect of other explanatory variables on declaration behavior and relate them to findings in the literature.

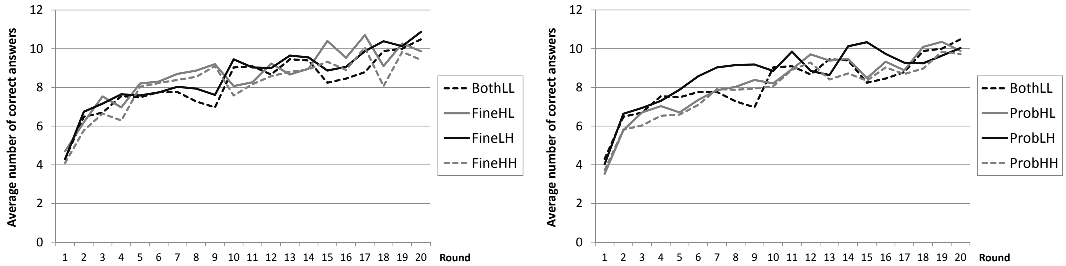


Figure 2: Average number of correct answers

Figure 2 illustrates the number of correct answers per round in the real-effort task, which we interpret as a good proxy for participants' effort in the first stage. There is a clear upward trend in the number of correct answers, indicating learning in the identification of numbers in the matrix. Comparing average gross income in rounds 1-10 and in rounds 11-20 using the two-sided Wilcoxon signed ranks test, we obtain p-values $< .01$ for all treatments.[21]

Result 1 *The average income in rounds 11-20 is larger than in rounds 1-10, irrespective of the treatment.*

The average number of correct answers is not statistically different across treatments with $\gamma_{1-10} = 4$ and $\gamma_{1-10} = 8$ (p-value $= .3303$).[22] In contrast, there is a weakly significant difference between the average number of correct answers in treatments with $p_{1-10} = .1$ and $p_{1-10} = .2$ (p-value $< .1$). This latter finding is intuitive, because the lower audit probability makes a unit of gross income more valuable in terms of expected utility when the audit probability is low.

Result 2 *Effort does not vary with the penalty multiplier. A higher audit probability tends to reduce effort.*

[21]We use averages at the subject level as our unit of observation.

[22]To isolate the effect of the current parameters on effort, we have only included the values from rounds 1-10 in the comparison.

Let us now turn to behavior in the second stage, the declaration stage. The average declaration is illustrated in Figure 3. In order to normalize declarations with respect to potentially differing gross income levels, the figure illustrates declared income divided by gross income, averaged over the subjects in each treatment.[23] Let us first focus on the darker bars for rounds 1-10. The figure demonstrates that stricter tax enforcement induces a higher share of income declarations. Indeed, tax declarations in rounds 1-10 are significantly lower when the penalty multiplier is lower and when the audit probability is lower (Wilcoxon rank sum tests, two-sided, p-values $< .01$). This indicates that subjects responded to the two parameters as predicted by our formal considerations in Section 3.[24]

Result 3 *Both a higher penalty multiplier γ_{1-10} and a higher audit probability p_{1-10}, on average, induce a higher share of declared income.*

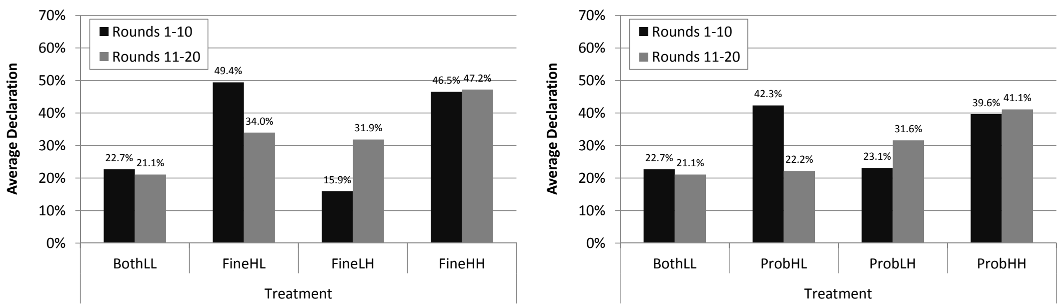


Figure 3: Average income declaration rate in the first and second halves of the experiment

Next, we concentrate on the declaration rates in the four treatments with a parameter change, FineHL, FineLH, ProbHL, and ProbLH. Figure 3 indicates that average income declarations reflect the change in tax enforcement parameters. Subjects increase the share of declared income when tax enforcement becomes stricter and declare less income when tax enforcement becomes less strict. This impression is confirmed by statistical analysis. For the treatments with a change in the penalty multiplier, we find a strongly significant upward (downward) shift in declaration rates in FineLH (FineHL) between the first and second halves of the experiment.

[23]Focusing on the ratio of declared income is standard in the literature and, given our assumption of a linear income tax, is comparable to taxes paid relative to taxes owed or to taxes evaded relative to taxes owed (see, e.g., Alm et al. 2009).

[24]Note that there is no statistically significant difference between the income declarations in the two treatments BothLL and FineLH (p-value = .5121) and between FineHL and FineHH (p-value = .6336) in the first 10 rounds. The same holds for BothLL vs. ProbLH (p-value = .5401) and for ProbHL vs. ProbHH (p-value = .7917).

Testing the average income declaration rate in rounds 1-10 against that in rounds 11-20 using a two-sided Wilcoxon signed ranks test we obtain p-values $< .01$ for both treatments. For the treatment in which the audit probability decreases (ProbHL), the downward shift in declaration rates is similarly significant with p-value $< .01$. For the upward movement in ProbLH, we do not obtain a difference in declarations that is statistically significant at conventional levels (p-value = .2380). In our treatments without a parameter change (i.e., BothLL, FineHH, and ProbHH), declaration rates remain constant over time.[25]

Result 4 *Income declarations in rounds 11-20 are higher than in rounds 1-10 in FineLH, whereas they are lower in FineHL and ProbHL. Income declarations tend to increase in rounds 11-20 compared to rounds 1-10 in ProbLH.*

Our central research interest is the potential path dependence of tax compliance. We now compare average declarations in rounds 11-20 across treatments with the same combination of audit probability and penalty multiplier in rounds 11-20, but different ones in rounds 1-10. More specifically, we compare FineHL to BothLL, FineLH to FineHH, ProbHL to BothLL, and ProbLH to ProbHH. Since the participants in treatments BothLL and FineHL (ProbHL) face the same material payoffs in rounds 11-20, the standard argument would be that the income declaration rates of the respective groups should be indistinguishable (see Section 3.1). Similarly, the fact that participants in treatments FineHH (ProbHH) and FineLH (ProbLH) face the same enforcement parameters in rounds 11-20 suggests similar tax compliance choices from all of these subjects. However, Figure 3 illustrates that this is not how decisions were actually made. Participants in treatment FineHL declared more than participants in treatment BothLL. Similarly, participants in treatment FineLH declared less than those in treatment FineHH (Wilcoxon rank sum tests, two-sided, both p-values $< .1$). When the audit probability is varied, there is no effect that it as easily observable. Comparing ProbLH and ProbHH, the difference in average declaration amounts to roughly 10 percentage points, but the difference is not statistically significant (p-value = .1718). For the comparison between ProbHL and BothLL, the difference in declarations in rounds 11-20 even appears negligible at first (p-value = .7044). However, as we have already seen that gross income varies across treatments – as do other important determinants of declaration to be detailed below – a simple comparison of mean income declarations is likely to miss important effects. We therefore continue with a regression analysis of declaration behavior.

[25]The p-values are equal to .5235 for BothLL, .8746 for FineHH, and .8220 for ProbHH.

We regress subjects' income declaration in round t on a set of explanatory variables. The regression analysis uses the absolute level of the income declaration in part 2 as the dependent variable, which is consistent with our model in Section 3. The empirical model is a GLS model with subject random effects, including period dummies for all rounds. The explanatory variables include two dummy variables for both the value of the penalty multiplier γ and the detection probability p in the first and second parts of our experiment respectively. The variables $\gamma_{current} = 8$ and $p_{current} = .2$ take the value one when the currently applicable tax enforcement parameter is high, and zero when it is low, where the situation with $\gamma = 4$ and $p = .1$ is the base. The variables $\gamma_{past} = 8$ and $p_{past} = .2$ are included to address our main interest, i.e., the potential path dependence of tax compliance. The dummy variable $\gamma_{past} = 8$ is equal to one in rounds 11-20 when the penalty multiplier for the subject was high in rounds 1-10 and is otherwise equal to zero, and likewise for $p_{past} = .2$. In rounds 1-10, both variables are set to zero as there is no history with different parameters yet. Thus, these two variables capture the lasting effect of a high penalty multiplier $\gamma_{1-10} = 8$ or a high audit probability $p_{1-10} = .2$ in rounds 1-10 on income declaration in rounds 11-20.

The regression includes several control variables. $punish_{t-1}$ is a dummy variable indicating whether or not a subject was punished in the previous round. Controlling for potential differences in first-stage behavior, the variable $income_t$ measures the gross income generated in round t, while $income_t^2$ controls for a possible non-linear influence. 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", "somewhat not religious", "somewhat religious" or "religious". Religiosity has been established as a variable with explanatory power regarding tax evasion in earlier studies (see, e.g., Torgler 2007). Finally, $risk\_aversion$ represents the number of risk-averse choices in the Holt and Laury (2002) procedure.[26]

Table 2 summarizes our regression results. Column (1) presents the regression results without further controls. Starting with the regression presented in column (2), we include our control variables. Our results confirm a lasting effect of the parameter set in rounds 1-10 across all treatments. Past experience with a high penalty multiplier, i.e., with $\gamma_{1-10} = 8$, significantly

[26] As suggested by Holt and Laury (2002), we use the absolute number of risk-averse choices as the measure for risk aversion, thereby ignoring multiple switches.

S_t	1	2	3	4
			(Excl. zero declarers)	(Only risk averse)
$\gamma_{current} = 8$	16.09***	16.00***	19.09***	14.00***
	(1.771)	(1.675)	(1.925)	(1.915)
$p_{current} = 0.2$	14.35***	16.57***	19.44***	16.58***
	(1.727)	(1.638)	(1.905)	(1.957)
$\gamma_{past} = 8$	7.762***	9.732***	9.459***	9.276***
	(2.404)	(2.278)	(2.611)	(2.638)
$p_{past} = 0.2$	4.850**	5.314**	4.315*	6.512**
	(2.389)	(2.264)	(2.598)	(2.774)
$punish_{t-1}$		-10.50***	-13.59***	-11.37***
		(1.265)	(1.546)	(1.457)
$income_t$		0.210***	0.178**	0.218**
		(0.0706)	(0.0853)	(0.0859)
$income_t^2$		0.00136***	0.00217***	0.00141***
		(0.000407)	(0.000495)	(0.000489)
male		-3.550	-0.931	-3.696
		(3.284)	(3.405)	(3.890)
religious		3.032	2.448	0.606
		(4.343)	(4.430)	(5.134)
risk_aversion		2.830***	2.653***	6.965***
		(1.011)	(1.005)	(1.869)
constant	3.483	-24.10***	-23.82***	-51.63***
	(2.555)	(7.157)	(7.456)	(13.09)
Observations	4,480	4,480	3,660	3,300
Subjects	224	224	183	165

Table 2: Regression coefficients: GLS with subject random effects. Period dummies were included in all regressions. Standard errors in brackets. *** denotes significance at the 1% level, ** at the 5% level and * at the 10% level.

increases the average declaration. Similarly, having had a high audit probability $p_{1-10} = .2$ increases the average declaration in rounds 11-20. Columns (3) and (4) are discussed in the next section.

Result 5 *Individuals with penalty multiplier* $\gamma_{1-10} = 8$ *declare on average more income in rounds 11-20 than individuals with penalty multiplier* $\gamma_{1-10} = 4$. *Individuals with audit probability* $p_{1-10} = 0.2$ *declare on average more income in rounds 11-20 than individuals with audit probability* $p_{1-10} = 0.1$.

In addition to the coefficients of the variables related to our main research question, the coefficients of other controls are interesting as well. For example, 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.[27] 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.[28] This negative influence of an audit in the previous period has been found in other studies as well (see, e.g., Alm and McKee 2004, Coricelli et al. 2010, Kastlunger et al. 2009). Gross income and also squared gross income have a positive influence on declared income. Men seem to declare less income than women, but not significantly so.[29] An intuitive finding is that risk aversion significantly increases tax honesty in our experiment. In contrast to previous findings of the empirical tax compliance literature (see, e.g., Torgler 2007), our data analysis does not confirm a statistically significant effect of religiosity on declaration behavior. We thus cannot confirm that people who call themselves religious are more honest with respect to their taxes than others.

[27] For similar evidence, see Bruttel and Kamecke (2012).

[28] We asked the question '*Did you change your behavior after being controlled and if so, how?*' to 180 of our 224 subjects. 15 participants out of these 180 responded with answers like "the probability of another control is smaller after being punished" while only 6 explicitly noticed that the probability of control continues to be equal to 10% or 20%, respectively.

[29] Gërxhani (2007) discusses gender effects in tax honesty in greater detail.

5 Discussion

In our discussion of the experimental results, we firstly address potential differences between the path dependence of tax compliance when the penalty multiplier is varied and the path dependence when the audit probability changes. Next, we discuss the heterogeneity of individual income declaration behavior. Finally, we comment on issues related to the identification of path dependence in tax compliance.

Penalty multiplier vs. audit probability

Figure 3 indicates that the path dependence of tax compliance may be a more significant phenomenon when it is the penalty that is varied. However, note that the gross income levels are somewhat lower in the case of a high audit probability than in the case of a low one, whereas there is no such effect on effort in treatments with different levels for the penalty multiplier. Such differences in the levels of gross income will generally cause different shares of declared income to be privately optimal. In our econometric analysis in which we include gross income as a control variable, we also find different coefficients for $\gamma_{past} = 8$ and $p_{past} = 0.2$ (see, for example, regression 2 in Table 2). One reason for such observations might be that the difference in the shares of declared income in rounds 1-10 between FineLH and FineHH is already more pronounced than the difference in the shares of declared income between ProbLH and ProbHH (see Figure 3); this makes it relatively more difficult to establish a significant difference regarding the level of declared income of subjects in ProbLH and that of subjects in ProbHH in rounds 11-20. In Figure 3, we also observe that when starting from the relatively higher share of declared income, subjects in ProbLH increase their declaration in view of the increase in the audit probability by less than subjects in FineLH do in view of the increase in the penalty multiplier. This may be understood since, for risk-averse subjects, the privately optimal adaptation of the level of declared income to the increase in the audit probability in part 2 will tend to be smaller than the adaptation to a change in the penalty multiplier when the change in the latter affects risk-bearing costs more via its impact on the variance of the lottery given by $V = \tau^2 \gamma_t^2 (I_t - S_t)^2 (p_t - p_t^2)$.[30] In summary, while our data points to path

[30]For example, doubling the penalty multiplier starting from its initial value increases the variance by more than doubling the audit probability starting from its initial value, while both changes imply the same expected penalty.

dependence of tax evasion being slightly more important for changes in the penalty multiplier than for changes in the audit probability, the present results are not sufficient to settle the matter.[31]

Heterogeneity

We will now discuss the heterogeneity of individual behavior observed in our experiment. The first matter of interest is the variety in income declarations for the initial tax enforcement parameters. Figure 4 illustrates the proportion of subjects who make an average declaration of a given magnitude in rounds 1-10. For instance, among participants in treatments BothLL and FineLH, 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 FineHL and FineHH declare no income, while 8 declare 90 to 100%. In addition to illustrating the extreme scenario of no declared income, the figure also illustrates (in the sense of first-order stochastic dominance) that the majority of subjects have declaration rates below 50% when the penalty multiplier is low. For the treatments varying the audit probability, the differences are similar, albeit smaller.

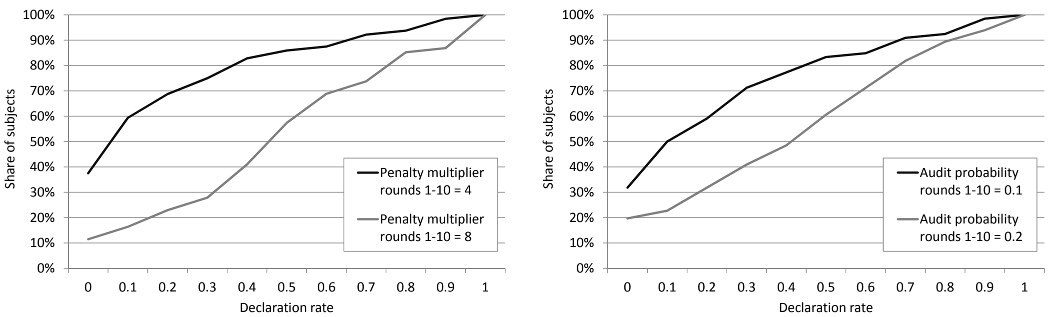


Figure 4: Distribution of income declaration rates in rounds 1-10

The next matter of interest is the variety in the reaction to a change in the levels of tax enforcement parameters. Figure 5 shows how subjects change their behavior in rounds 11-20 compared to rounds 1-10. More precisely, the figure illustrates by how many percentage points the average declaration in rounds 11-20 is smaller or larger than in rounds 1-10. In all

[31] Interestingly, taking into account probability weighting is not useful in explaining this potential difference in observed path dependence, because the overweighting of small probabilities would actually strengthen path dependence in treatment ProbHL rather than weakening it (see, e.g., Hashimzade et al. (forthcoming) on probability weighting models).

treatments, there is a large proportion of subjects who do not adjust their behavior at all or adjust it by less than 10 percentage points (61% in FineLH, 43% in FineHL, 39% in ProbLH, and 32% in ProbHL), including a subgroup of about one-quarter of our subjects who declare no income throughout (26% of all subjects in FineLH, 10% in FineHL, 21% in ProbLH, and 24% in ProbHL). Regression (3) in Table 2 excludes subjects who declare an income of zero in all rounds in order to test whether our results regarding path dependence are attributable solely to these participants. The difference in the regression results compared to our main regression (2) is quite small, so we can rule out this explanation.

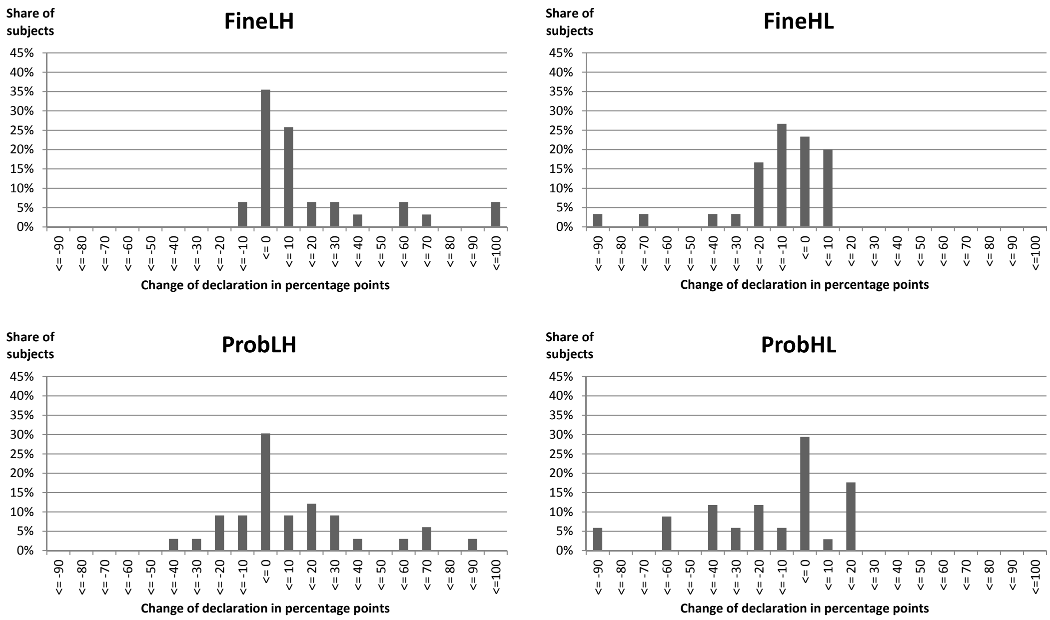


Figure 5: Change of average declaration from rounds 1-10 to rounds 11-20 in percentage points.

About half of all subjects adjust a little, i.e., their average declaration in rounds 11-20 is between 10 and 50 percentage points above or below their declaration in the first ten rounds, but adjusted in the direction expected for risk-averse individuals. A relatively small share of subjects react strongly to the parameter change by adjusting their average declaration by more than 50 percentage points (15% in FineLH, 6% in FineHL, 12% in ProbLH, and 15% in ProbHL). There are also some subjects who adjust their behavior in the "wrong" direction. These individuals may have been guided by aspects not controlled for in this analysis, such as the experience of an audit.

Identification of path dependence in tax compliance

We propose to identify path dependence of tax compliance. In the following section, we comment on possible objections, such as the possibility that subjects did not understand the parameter change.

To avoid the situation in which path dependence is merely a consequence of some participants disregarding the information about the parameter change, we carefully ensured that participants recognized the message informing them of the change. First of all, we announced in the instructions that they would receive new information about the rules of the game after round 10. In addition, before round 11 started, the participants had to click a button after reading this new information. To test whether these measures were sufficient to draw the subjects' attention to the change, we consider the time they spent viewing the information screen before clicking the button, assuming that those who clicked the button faster potentially paid less attention to the new information. More precisely, we divide subjects in our treatments with a parameter change into two groups, depending on whether or not they spent more time than the median on the information screen, and compare by how many percentage points (in absolute terms) they change their declarations in rounds 11-20 as compared to rounds 1-10. We find no effect. The fast subjects actually react more to the parameter change in three out of the four treatments. This strongly indicates that all subjects understood the parameter change.

After ruling out ignorance of the parameter change as an explanation for behavioral spillovers from the first to the second part of the experiment, we can look for other explanations for the observed path dependence. Is path dependence possibly attributable only to the subset of risk-neutral and risk-seeking subjects? To test for this possibility, we repeat the full regression presented above (see column (4) in Table 2) excluding 22 subjects classified as risk-seeking and 37 subjects categorized as risk-neutral by their decisions in the Holt and Laury lotteries. The regression yields the same effects as it did for the full sample, allowing us to exclude risk-seeking behavior as an explanation for behavioral spillovers in declaration decisions.

6 Conclusion

We show that factors having no impact on current material payoffs may still influence individual choices. In order to establish this point, we provide experimental evidence showing that present tax compliance behavior is affected by past enforcement parameters. For example, individuals who had grown accustomed to a low penalty declared less income when making their declaration decision under new higher levels of the penalty than individuals who had continuously experienced the high level of the penalty, even though the penalty for tax evasion applicable in past rounds had no material impact on current payoffs. The observed choices suggest that indeed both past and present enforcement parameters have an influence on present individual behavior.

The path dependence of tax compliance can be explained by individuals having reference-dependent preferences with lagged reference points. The reference point is shaped by the past tax regime, and the reference point is critical to the privately optimal tax compliance in the current tax regime. Thus, the reference point introduces a relationship of the past tax regime and behavior in the current one. Other concepts may also help in understanding our experimental data. For instance, the path dependence of tax compliance could be related to the concept referred to as anchoring (see, e.g., Ariely et al. 2003, Stewart 2009, Tversky and Kahneman 1974). Following Kahneman (1992), a reference point determines a benchmark, in comparison to which an outcome is evaluated as a gain or loss. In this sense, reference points influence behavior in our experiment because of their impact on whether or not some payoff will represent a gain or a loss for the subject. In contrast, anchoring directly biases decisions towards the anchor, such that the decision variable should not diverge too strongly from the anchor. It thus directly impacts decision-making by shifting the perception of what a "normal" declaration is. In our context, the level of the income declaration in part 1 of the experiment may serve as an anchor for the declaration in part 2. Both effects, reference dependence and anchoring, manipulate behavior in a very similar way and cannot be disentangled in this experiment.

The findings from our study may have important policy implications. When thinking about transplanting tax enforcement systems, for instance, when using regimes from Northern Europe in the south, it is important to bear in mind that the behavioral adaptation to the new system by southerners may take much more time than the pure legal implementation. This lag can be attributed to being accustomed to the old regime in terms of expectations, for instance. The fact that the old tax enforcement regime continues to shape behavior may also be of relevance for the behavior of tax inspectors, such that professional guidance from outsiders along the process of adjustment may be valuable. Furthermore, it may be worth trying to accelerate the adaptation by explicitly informing the population about the new rules of enforcement and their consequences for possible payoff outcomes. The preliminary results regarding the extent of path dependence being contingent on the tax enforcement instrument that is varied also seems important for the decision about how to devise a new enforcement system.

Tentatively, it may be conjectured that our results generalize to other domains outside the tax compliance context. There are many areas of the law in which the European Union is seeking to harmonize national legal frameworks. It seems likely that the process of behavioral adaptation follows a similar path as in the narrow field of taxation, such that rules applicable before the harmonization continue to co-determine decisions for quite some time. In a similar vein, immigrants experience different institutions and regulations in their home country which influence their reference points before migrating, and thus may respond differently to their new country's institutions than natives do. Finally, path dependence could be relevant to incentive systems in firms, since employees in a firm have often had different previous employers, and different employers imply heterogeneous incentive systems.

Our contribution opens up many avenues for future research. For example, research is needed to answer whether there will be full behavioral convergence at some point in time and how precisely the speed of convergence differs between policy instruments. Related issues are also of interest when subjects have experienced several regimes in sequence instead of only two. In the example regarding employees switching employers, for example, in many circumstances subjects will have experienced many different regimes. Another issue is whether or not it is possible to actively deal with differences in reference points in some way, for example, by explicitly addressing the realism of expectations individuals have.

Appendix

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

(The following section provides the translations of the German instructions for treatments FineLH and BothLL. 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 in cash for each 120 points you scored in the experiment. In addition, you will 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.

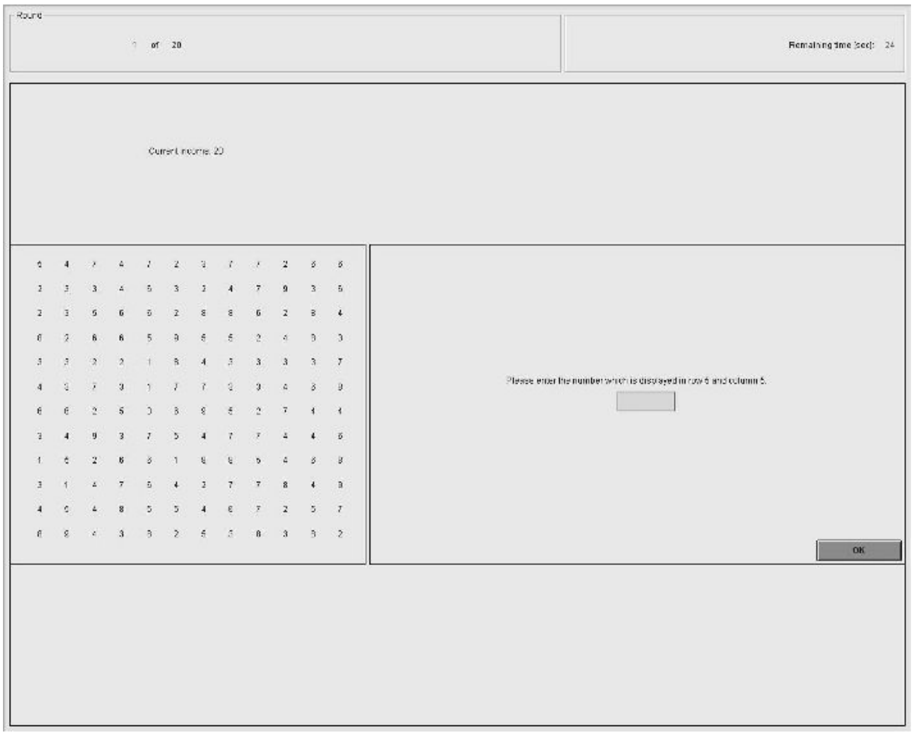
After the main experiment, there will be a lottery experiment. You will receive the instructions for this second part of the experiment on your computer screen after the end of the first part.

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 for this stage appears as follows:



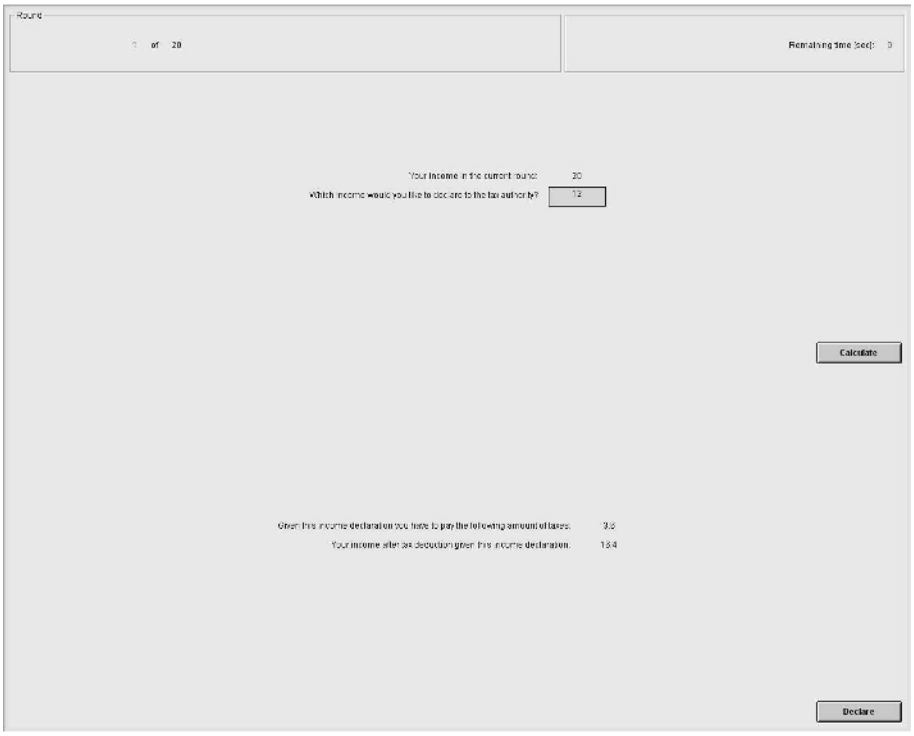
2) Tax declaration:

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 will have to pay 3.6 points in taxes.

Gross income	Declared income	Taxes paid
20	12	3.6

Example with gross income = 20 and declared income = 12.

The tax declaration screen is shown in the following figure:



Enter your declared income in the field indicated. 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 with 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 a 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).

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

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

Example with income > declaration with no tax audit

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).

Appendix B: Proof of Proposition 2

Proof. Claim (a): From (10) and (11), it is clear that $C_a < K_a$ and $C_n < K_n$. Subjects in ProbLH are in the loss region in the audit state of nature for strictly fewer levels of S_2. Similarly, subjects in ProbLH are in the loss region in the no-audit state of nature for strictly more levels of S_2. This ensures that $\min\{C_a, C_n\} < K_a$ and $\max\{C_a, C_n\} < K_n$. Thus, the range of S_2 with inflated marginal benefits (costs) is strictly smaller (larger) for subjects in ProbLH than for subjects in ProbHH, as follows from reference to Lemma 1. This allows for subjects in ProbLH declaring strictly less income in part 2 than subjects in ProbHH, because there are ranges of S_2 in which subjects in ProbHH perceive a higher marginal benefits to costs ratio than subjects in ProbLH. At the same time, this rules out that subjects in ProbLH declare more income than subjects in ProbHH.

Claim (b): From (10) and (11), it is clear that $C_a > K_a$ and $C_n > K_n$. Subjects in ProbHL are in the loss region in the audit state of nature for strictly more levels of S_2. Similarly, subjects in ProbHL are in the loss region in the no-audit state of nature for strictly fewer levels of S_2. This ensures that $\min\{C_a, C_n\} > K_a$ and $\max\{C_a, C_n\} > K_n$. In other words, the range of S_2 with inflated marginal benefits (costs) is strictly larger (smaller) for subjects in ProbHL than for subjects in BothLL, as follows from reference to Lemma 1. This allows for subjects in ProbHL declaring strictly more than subjects in BothLL, because there are ranges of S_2 in which subjects in BothLL perceive a lower marginal benefits to costs ratio than subjects in ProbHL. At the same time, this rules out that subjects in ProbHL declare less income than subjects in BothLL.

Claim (c): (i) Suppose that $\partial C_a/\partial\gamma_1 > 0$. In this case, it follows from (12) and (13) that $C_a < K_a$ and $C_n < K_n$. This ensures that $\min\{C_a, C_n\} < K_a$ and $\max\{C_a, C_n\} < K_n$. In other words, the range of S_2 with inflated marginal benefits (costs) is strictly smaller for subjects in FineLH than for subjects in FineHH, as follows from Lemma 1. This allows for subjects in FineLH declaring less than subjects in FineHH, and rules out that subjects in FineLH declare more income than subjects in FineHH. (ii) Suppose that instead $\partial C_a/\partial\gamma_1 < 0$, such that $C_a > K_a$ and $C_n < K_n$. At a general level, different rankings of the thresholds are possible. For example, the case in which $C_n < K_a < C_a < K_n$ would entail that subjects in FineLH have inflated marginal benefits for strictly fewer levels of S_2 and inflated marginal costs for strictly more levels. Out of the possible configurations, the only possibility to have subjects in FineLH declaring more than subjects in FineHH would arise when the declaration of the latter subjects in part 2 falls within the range $(K_a, \min\{C_a, C_n\})$ or $(K_n, \max\{C_a, C_n\})$, should these exist. If these ranges exist, they signify levels of S_2 for which the ratio of marginal benefits to costs is higher for subjects in FineLH.

Claim (d): (i) Suppose that $\partial C_a/\partial\gamma_1 > 0$. In this case, it follows from (12) and (13) that $C_a > K_a$ and $C_n > K_n$. This ensures that $\min\{C_a, C_n\} > K_a$ and $\max\{C_a, C_n\} > K_n$. In other words, the range of S_2 with inflated marginal benefits (costs) is strictly larger (smaller) for subjects in FineHL than for subjects in BothLL, as follows from reference to Lemma 1. This allows for subjects in FineHL declaring more than subjects in BothLL, and rules out that subjects in FineHL declare less income than subjects in BothLL. (ii) Suppose that instead $\partial C_a/\partial\gamma_1 < 0$, such that $C_a < K_a$ and $C_n > K_n$. In this scenario, the range of both inflated marginal benefits and the one of inflated marginal costs are smaller for subjects in FineHL. When the declaration of subjects in BothLL falls within the range $[K_n, C_n)$, then marginal costs are relatively inflated only for these subjects whereas subjects in FineHL face equally inflated marginal benefits and marginal costs, such that the latter will choose a higher level of declared income. When the declaration of subjects in BothLL falls within the range of either $(0, C_a]$ or (K_a, K_n), then subjects in FineHL face the same marginal incentives and will thus declare the same level of income in part 2. In contrast, subjects in FineHL would declare less than subjects in BothLL should the declaration of the latter subjects fall within the range $(C_a, K_a]$. ■

References

Abeler, J., Falk, A., Goette, L., and D. Huffman, 2011. Reference points and effort provision. American Economic Review 101, 470-92.

Ahn, T.K., Ostrom, E., Schmidt, D., Shupp, R., and J. Walker, 2001. Cooperation in PD games: Fear, greed, and history of play. Public Choice 106, 137-155.

Allingham, M.G., Sandmo, A., 1972. Income tax evasion: A theoretical analysis. Journal of Public Economics 1, 323-338.

Alm, J., 2012. Measuring, explaining, and controlling tax evasion: Lessons from theory, experiments, and field studies. International Tax and Public Finance 19, 54-77.

Alm, J., and M. McKee, 2004. Tax compliance as a coordination game. Journal of Economic Behavior and Organization 54, 297-312.

Alm, J., McKee, M., 2006. Audit certainty, audit productivity, and taxpayer compliance. National Tax Journal 59, 801-816.

Alm, J., Jackson, B.R., and M. McKee, 1992. Institutional uncertainty and taxpayer compliance. American Economic Review 82, 1018-1026.

Alm, J., Jackson, B.R., and M. McKee, 2009. Getting the word out: Enforcement information dissemination and compliance behavior. Journal of Public Economics 93, 392-402.

Alm, J., McClelland, G.H., and W.D. Schulze, 1992. Why do people pay taxes? Journal of Public Economics 48, 2138.

Alm, J., McClelland, G.H., and W.D. Schulze, 1999. Changing the social norm of tax compliance by voting. Kyklos 52, 141-171.

Andreoni, J., Erard, B., and J. Feinstein, 1998. Tax compliance. Journal of Economic Literature 36, 818-860.

Apesteguia, J., and M.A. Ballester, 2009. A theory of reference-dependent behavior. Economic Theory 40, 427-455.

Ariely, D., Loewenstein, G., and D. Prelec, 2003. "Coherent arbitrariness": Stable demand curves without stable preferences. Quarterly Journal of Economics 118, 73-105.

Arthur, W.B., 1989. Competing Technologies, Increasing Returns, and Lock-In by Historical Events. Economic Journal 97, 642-665.

Baucells, M., Weber, M., and F. Welfens, 2011. Reference-Point Formation and Updating. Management Science 57, 506-519.

Becker, W., Buchner, H.J., and S. Sleeking, 1987. The impact of public transfer expenditures on tax evasion: An experimental approach. Journal of Public Economics 34, 243-252.

Berkowitz, D., Pistor, K., and J.F. Richard, 2003. Economic development, legality, and the transplant effect. European Economic Review 47, 165-195.

Bernasconi, M., and A. Zanardi, 2004. Tax evasion, tax rates, and reference dependence. FinanzArchiv 60, 422-445.

Bohnet, I., and S. Huck, 2004. Repetition and reputation: Implications for trust and trustworthiness when institutions change. American Economic Review Papers & Proceedings 94, 362-366.

Brandts, J., and D.J. Cooper, 2006. A change would do you good ... An experimental study on how to overcome coordination failure in organizations. American Economic Review 96, 669-693.

Bruttel, L., and U. Kamecke, 2012. Infinity in the lab - How do people play repeated games? Theory and Decision 72, 205-219.

Cason, T.N., Savikhin, A., and R.M. Sheremeta, 2012. Behavioral spillovers in coordination games. European Economic Review 56, 233-245.

Coricelli, G., Joffily, M., Montmarquette, C., and M.C. Villeval, 2010. Cheating, emotions, and rationality: an experiment on tax evasion. Experimental Economics 13, 226-247.

Crawford, V.P., and J. Meng, 2011. New York City cab drivers' labor supply revisited: Reference-dependent preferences with rational-expectations targets for hours and income. American Economic Review 101, 1912-32.

Cullis, J., Jones, P., and A. Soliman, 2012. 'Spite effects' in tax evasion experiments. Journal of Socio-Economics 41, 418-423.

Dhami, S., and A. al-Nowaihi, 2007. Why do people pay taxes? Prospect theory versus expected utility theory. Journal of Economic Behavior and Organization 64, 171-192.

Falk, A., and M. Knell, 2004. Choosing the Joneses: endogenous goals and reference standards. Scandinavian Journal of Economics 106, 417-435.

Feld, L.P., Torgler, B., and B. Dong, 2008. Coming closer? Tax morale, deterrence and social learning after German Unification. Working Paper 232, QUT School of Economics and Finance.

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

Frank, R.H., 1985. The demand for unobservable and other nonpositional goods. American Economic Review 75, 101-116.

Fuest, C., and N. Riedel, 2009. Tax evasion, tax avoidance, and tax expenditure in developing countries: A review of the literature. Report prepared for the UK Department of International Development, Oxford.

Gächter, S., Kessler, E., and M. Königstein, 2008. Performance incentives and the dynamics of voluntary cooperation. Unpublished working paper.

Garoupa, N., 2012. An economic analysis of legal harmonization: The case of law enforcement within the European Union. In: Eger, T., and H.B. Schäfer (eds.). Research Handbook on the Economics of European Union Law. Cheltenham: Edward Elgar.

Genesove, D., and C. Mayer, 2001. Loss Aversion and Seller Behavior: Evidence from the Housing Market. Quarterly Journal of Economics 116, 1233-1260.

Gërxhani, K., 2007. Explaining gender differences in tax evasion: The case of Tirana, Albania. Feminist Economics 12, 119-155.

Gërxhani, K., and A. Schram, 2006. Tax evasion and the source of income. Journal of Economic Psychology 27, 402-422.

Gill, D. and R. Stone, 2010. Fairness and desert in tournaments. Games and Economic Behavior 69, 346-364.

Gneezy, U., and A. Rustichini, 2000. A fine is a price. Journal of Legal Studies 29, 1-17.

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.

Hamman, J., Rick, S., and R.A. Weber, 2007. Solving coordination failure with "all-or-none" group-level incentives. Experimental Economics 10, 285-303.

Hashimzade, N., Myles, G.D., and B. Tran-Nam, forthcoming. Applications of behavioral economics to tax evasion. Journal of Economic Surveys.

Holt, C.A., and S.K. Laury, 2002. Risk aversion and incentive effects. American Economic Review 92, 1644-1655.

Kahneman, D., 1992. Reference points, anchors, norms, and mixed feelings. Organizational Behavior and Human Decision Processes 51, 296-312.

Kahneman, D., and A. Tversky, 1979. Prospect theory: an analysis of decision under risk. Econometrica 47, 263-292.

Kastlunger, B., Kirchler, E., Mittone, L., and J. Pitters, 2009. Sequences of audits, tax compliance, and taxpaying strategies. Journal of Economic Psychology 30, 405-418.

Katsios, S., 2006. The shadow economy and corruption in Greece. South-Eastern Europe Journal of Economics 1, 61-80.

Koch, A.K., and J. Nafziger, 2011. Self-regulation through goal setting. Scandinavian Journal of Economics 113, 212-227.

Köbberling, V., and P.P. Wakker, 2005. An index of loss aversion. Journal of Economic Theory 122, 119-131.

Köszegi, B., and M. Rabin, 2006. A model of reference-dependent preferences. Quarterly Journal of Economics 121, 1133-1165.

Köszegi, B., and M. Rabin, 2007. Reference-dependent risk attitudes. American Economic Review 97, 1047-1073.

Krugman, P.R., 1994. Peddling Prosperity: Economic Sense and Nonsense in the Age of Diminished Expectations. New York, Norton.

Liebowitz, S.J., and S.E. Margolis, 2000. Path dependence. In: Bouckaert, B. and G. De Geest (eds.), The encyclopedia of law and economics, Volume I. The History and Methodology of Law and Economics, Cheltenham, Edward Elgar.

Martinez-Vazquez, J. and B. Torgler, 2009. The evolution of tax morale in modern Spain. Journal of Economic Issues 18(1), 1-28.

Matsaganis, M., Leventi, C., and M. Flevotomou, 2012. The crisis and tax evasion in Greece: what are the distributional implications? CESifo Forum 13, 26-32.

Meier, S., 2007. Do subsidies increase charitable giving in the long run? Matching donations in a field experiment. Journal of the European Economic Association 5, 1203-1222.

Schmidt, U., Starmer, C., and R. Sugden, 2008. Third-generation Prospect Theory. Journal of Risk and Uncertainty 36, 203-223.

Schneider, F., and D.H. Enste, 2000. Shadow economies: size, causes, and consequences. Journal of Economic Literature 38, 77-114.

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.

Sugden, R., 2003. Reference-dependent subjective expected utility. Journal of Economic Theory 111, 172-191.

Torgler, B., 2007. Tax compliance and tax morale: A theoretical and empirical analysis. Cheltenham: Edward Elgar.

Tversky, A., and D. Kahneman, 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