The European Journal of Comparative Economics
Vol. 5, n. 1, pp. 33-56

ISSN 1722-4667

Comparing Direct and Indirect Taxation: The Influence of Framing on Tax Compliance

Christoph Watrin, Robert Ullmann

Abstract

Standard theory of the optimal mix of direct and indirect taxation implicitly assumes that compliance is not influenced by the framing of the taxes. According to our findings, this is not the case. Using an experimental approach, we examine whether framing the tax payment decision as income tax or consumption tax influences compliance. We find that median compliance is 10.2 percentage points higher in the income tax framing. Further, we find that subjects' reaction to a change in tax rates is comparable, but reaction towards a change in detection rates is higher in the consumption tax scheme. We conclude that behavioral patterns should be taken into account when drawing conclusions about the direct-indirect tax mix.

JEL Classification: C91, H21, H24, H26

Keywords: comparative analysis of tax systems, behavioral public finance, optimal tax mix, noncompliance, framing

1. Introduction

The idea of a consumption-based tax system has gained ever greater attention in tax policy debates of recent years. In the European Union (EU), most of the ten states that became EU members in May 2004 rely more on a value-added tax than on a corporate income tax. Furthermore, these states generally have lower corporate income tax rates than do the old EU member states. This new tax structure within the EU has given rise to much competition among the old and new EU member states with regard to both tax rates and structure of the tax system. In Germany, for example, the Tax Reform Act 2008 reduced the overall corporate tax rate from about 40 to about 30 percent. Almost at the same time, the government has raised the value-added tax rate from 16 to 19 percent, thereby increasing the importance of indirect taxes in the overall tax system. The United States Congress also is currently debating the Fair Tax Act of 2007 (H.R.25), which is a proposal to repeal the income tax and other current federal taxes, such as the estate and gift tax, and implement a national sales tax instead.

In public finance theory, the efficiency effects of a consumption tax in comparison to an income tax have been discussed for many years. Primary arguments for a general consumption tax are that it has less negative effects on labor supply than the income tax and that it interferes less with the choice between present and future consumption. A consumption tax further increases the part of the national income saved, and thus leads to more capital formation and higher economic growth (Musgrave and Musgrave (1989)). The major argument in favor of income taxation is generally that it is more adequate for the redistribution of income between income groups (Saez (2004)).

However, the design of the optimal tax system depends not only on considerations about efficiency and equity, but also on fiscal administration and enforcement costs. In this paper we present a new insight in compliance differences between divergent framings of the tax payment decision. Our research is rooted in the notion of behavioral public finance, which implies that behavioral effects must be considered in a well designed tax system (McCaffery and Slemrod (2006a)). In particular,

we compare the tax payment decision between an income tax declaration scheme that has the possibility to underdeclare income and a consumption tax scheme with the possibility of buying a certain good either on the legal market, or on the black market where taxes on consumption are not levied.

For fiscal authorities, noncompliance, both in consumption and income tax, is a great concern. However, experimental research on the behavioral aspects of noncompliance has focused on income tax evasion, as summarized for instance by Alm (1991) and Andreoni *et al.* (1998). There is still little behavioral work on noncompliance in other areas of taxation (Webley *et al.* (2006)). Given the broad utilization of indirect taxation by governments around the world, this discrepancy is remarkable. To our knowledge, our study is the first work to explicitly focus on the behavioral differences in noncompliance between income and consumption tax.

In this study we develop an experimental design for the comparative analysis of compliance in consumption and income tax settings. The advantages of experimental work in the field of noncompliance are evident by the very nature of the problem of tax evasion, since noncompliance can only be observed with a high probability of error. Furthermore, our design allows us to explicitly control for the key determinants of compliance, namely the tax rate, detection probability, effectiveness of government spending, and penalty fees. It further enables us to control for omitted variables that are constant between the two tax framings. Hence, in our experimental design any differences in compliance can be attributed to a potential behavioral effect, that is, any combination of omitted variables inconsistent between the two tax payment framings.

We utilize this experimental setup to identify differences in the level of noncompliance for an income or a consumption tax-based pay scheme when there is an equal effective tax burden and equal enforcement. We further analyze, whether compliance in both tax payment framings is influenced differently by changes in tax rates and tax enforcement efforts, hence the relative weight of the two tax schemes in the overall tax system. Thus, to the current political debate about consumption versus income taxes we add the important aspect of noncompliance that has not yet been incorporated.

We report significantly lower compliance when the tax payment decision is framed as a consumption tax rather than an income tax. Even though we do not find different reactions towards an increase in tax rates, we do find significant differences in reaction towards a change in detection probabilities. Particularly we observe that individuals react more strongly when detection probabilities are increased in the consumption tax scheme than when they are increased in the income tax scheme.

The paper is organized as follows: In Section two we review noncompliance models and related previous work on behavioral public finance. In Section three, we first develop a theoretical framework that allows us to predict the theoretical behavior of a fully rational taxpayer. We then derive our research hypothesis, and further describe the details of the experimental design. We discuss our findings in Section four. In Section five, we conclude by discussing possible future research on the issues raised by our experimental results.

2. Motivation and Prior Research

2.1 Comparing Direct and Indirect Taxation

The first researcher to discuss the problem of optimal taxation was Ramsey (1927). In his pioneering work, he assumes a model in which the government wants to raise a fixed amount of revenue merely by consumption taxation, and he analyzes the tax rates that should be applied to a number of different goods. He assumes that all goods can be taxed at different rates, which does not seem feasible in a market system without a social planner. His basic conclusion is that the percentage of reduction in demand for each good should be the same after introducing a consumption tax to the tax system. This model is known as the Ramsey optimum tax rule.

Numerous other scholars have also analyzed the basic question of the usefulness of a consumption tax in general. Hotelling (1938) finds that because of the distortion of optimal consumption caused by consumption taxes, welfare provisions are maximized by an income tax rather than a consumption tax. The models by Lerner (1970) and Dixit (1970) explicitly take into account goods that cannot be taxed with a consumption tax, which is an implicit assumption of Ramsey (1927). The Lerner and Dixit models both show that in the case of a non-taxable goods sector, if taxes on certain goods are increased, then policy design must incorporate the possibilities of a shift of consumption towards the untaxed sector, and vice versa. Thus, policy makers must be concerned about cross-elasticity between the taxable and the non-taxable goods or sectors. A summary of the earlier models can be found in Sandmo (1976); a more recent summary appears in Slemrod (1990). However, we note that so far, none of the models have allowed for the possibility of tax evasion.

The economics of crime argument, which was developed by Becker (1968), was introduced to noncompliance research by Allingham and Sandmo (1972). These authors analyze how rational individuals should react in an income tax-based system in which evasion is possible and compliance cannot be fully enforced by the authorities.

In the following years, some papers analyze the question of evasion only in consumption taxes. Gordon (1990), Cremer and Gahvari (1993), and McLaren (1998) are some examples of this theme of research. The Cremer/Gahvari and McLaren models both investigate how a firm would react if it had the possibility to evade sales taxes. Both models find that this would lead to a modified Ramsey optimum tax rule. Gordon additionally includes the demand side for "under the counter" transactions into his model. He particularly analyzes the effect of a noncompliance penalty on consumers compared to a penalty on firms that engage in such transactions.

More recent models incorporate a consumption tax into the economics of crime approach in taxation, such as Boadway *et al.* (1994) and Richter and Boadway (2005). Unfortunately, those models continue to allow for noncompliance only in income taxes, even though the authors explicitly model consumption taxes. The basic assumption of this approach is that a tax on consumption can largely be enforced, so evasion is highly unlikely. To the astute reader, the assumption of a full enforcement of the consumption tax seems a rather strong one. To our knowledge, there has never been a model that explicitly allowed for income and consumption tax evasion at the same time.

2.2 Framing in Tax Decision Making

The issue of comparing income and consumption tax ultimately breaks down to a question of framing the tax payment decision. To a rational individual, only the expected amount of payment should matter. Kahneman and Tversky (1981) introduced the concept of framing into economics, which basically states that if the same problem is presented in two different ways, responses can be different. Also, the application of framing within the research of optimal taxation is not completely novel and has been previously utilized by, for instance, McCaffery (2000), Sausgruber (2002), and Krishna and Slemrod (2003). Most of the research so far evolves around the question of how taxpayers perceive the burden that the government has placed on them. In his experiments, Sausgruber (2002) finds that the mere question of who has to transfer the taxes to the tax authorities (the company or the consumer) leads to a significant difference in the perceived tax burden. When individuals do not have to transfer the taxes, the perception of tax burden is significantly lower, even though economic burden is the same.

From a policy perspective, Krishna and Slemrod (2003) argue that policy makers might take into account the perception of a certain increase in taxes, and thus are probing for the lowest perceived burden at equal tax revenues in order to assure reelection. McCaffery (2000) takes this idea one step further by stating that cognitively favored tax systems could be more stable, and thus are more likely to survive in the long run.

McCaffery and Slemrod (2006b) subsume all the above in the novel notion of behavioral public finance. They also provide a thorough review of the topics that are currently debated in this area of research. Within this area, it is generally implied that "form matters" when government revenues are to be levied from the taxpayers. Research in behavioral public finance brings forward biases that are known from psychology or behavioral finance in the area of public finances and thrives to discover biases that are yet unknown. We add to previous research by investigating how framing of the tax payment decision influences individual compliance. Such differences in compliance merely caused by differently framing the tax payment decision are of great important to policy makers in designing optimal tax systems.

Despite its enormous importance for public finance, research on framing and noncompliance hardly exists. The first paper indicating that differences in tax compliance could occur conditional on the framing of the tax-payment decision is a working paper by Gueth and Sausgruber (2004). They design an experiment in which subjects vote for either a mixture of income and consumption taxes, in which the consumption tax is fully enforced and the income tax is not enforced at all, or an income tax scheme that cannot be enforced. Even though the authors are not particularly interested in the direct-indirect tax mix, they find that the subjects' tax morale is significantly higher in the latter scheme. Regrettably, the Gueth and Sausgruber (2004) experiments do not allow for direct comparison between the two schemes. The schemes differ significantly in complexity and risk of detection, two factors that other researchers believe might be key in the analysis of noncompliance (see for instance Milliron (1985); Forest and Sheffrin (2002) for issues of complexity and Fischer *et al.* (1992) for a review about research on detection probability). Further, implementing a voting pattern does not allow for comparing the entire subject pool in both tax schemes, but only for comparing the experimental sessions that voted for each tax scheme at least once, what might result in a self-selection of data points.

3. Experimental Design

3.1 Analytical Background of the Experiment

3.1.1 Multiple-Stage Game

The main difficulty for the theoretical modeling is to set government parameters such that we can directly compare noncompliance choices in the income tax decision and consumption tax decision. Our solution is to construct an experimental setup in which the compliance behavior of rational individuals should be indifferent between the two tax schemes, unless there is a nonconstant omitted factor, for example a behavioral influence, between the two tax payment decisions. To predict individuals' behavior in these settings, we must derive the following determinants of behavior:

1. The effect of the compliance decision in either the income tax or the consumption tax scheme on an individual's expected revenue.

2. The effect of changes in tax rates for either of the two types of tax framings on individual behavior.

3. The effect of changes in enforcement effort for either of the two types of tax framings on individual behavior.

We design our experiment as a public-goods game in which individuals pay for a public good by the means of taxation with the possibility of evasion. At no point do we attempt to find the optimal solution of this model, because we are only interested in the relative comparison of tax schemes.

For the theoretical model we set up a multiple-stage game with a sequence similar to the model of Richter and Boadway (2005):

Stage 1; **Government policies:** The government sets tax and detection rates. In the course of the experiment, we alter tax rates and detection rates for both income and consumption taxes separately.

Stage 2; **Decision on income tax reporting:** The individual takes tax and detection rates as given, and chooses the income to report for income taxation. Income taxes are then deducted based on the reported income. The amount of the reported income left after income taxes are levied and the income not reported for income taxation are available for consumption. We do not allow for changes in labor supply, wages and abilities of the individuals, but we do assign a random income from a known distribution to each individual to illustrate different income levels.

Stage 3; **Household budget allocation**: The individual may choose the part of the household budget (disposable income after income taxes) to spend on two private goods, A and B. If the individual decides to purchase good A, a consumption tax will be levied, if the individual decides to purchase good B there will be no consumption tax.

After the compliance decisions in stage 2 and stage 3 the government conducts random audits for both payments with a priori known detection rates. Penalties will

arise from these audits if an individual is found to be noncompliant in either of the two stages. Taxes, but not penalties, are then used to finance a public good which benefits all individuals in equal shares.

Using a multiple-stage game instead of comparing multiple one-stage games has numerous advantages: First, playing both different tax schemes in each treatment enables us not only to observe individuals' reaction in the stage in which we alter parameters, but also to control for the reaction in the stage in which we do not alter parameters. Second, using a multiple-stage game allows us to control for errors caused by certain omitted variables, such as fatigue or excitement, because both tax schemes are played in turns. Thus, any change in the experimental environment that we do not model explicitly should be equally present in the income and consumption tax scheme. Finally, a multiple-stage game allows us to control for signaling and reputation building with respect to compliance in a particular tax payment scheme: When contributing to the public good, an individual sends two interacted signals. The first is the level of overall compliance and the second is the level of income earned. Even though in a public-goods game we cannot prevent inferences about overall compliance of other group members completely, using a multiple-stage game allows us to prevent inferences about the compliance in each separate stage when the public good is financed by the total amount of taxes paid. We diminish the chance of robust inferences about the overall level of compliance further by assigning income at random to each individual. Ensuring that signaling and reputation building is negligible, and thereby ensuring that the relevant decisions between subjects are independent, is necessary if we are to apply powerful statistical tests of inference.

Throughout the remainder of the paper we use the tilde (~) to indicate stochastic variables. Subscripts I and C shall indicate a variable regarding the income or consumption tax scheme, respectively. At the beginning of stage 2, we assign a stochastic income to each individual:

$$\tilde{Y}_I \sim U(\alpha, \beta) \qquad [1]$$

Gross income from labor $\tilde{Y}_I$ thus is uniformly distributed with lower and upper bounds α and β, respectively. Income taxes are then levied based on the declared income as shown in equation (2):

$$\tilde{Y}_C = \tilde{Y}_I (1 - r_I \tau_I) \qquad [2]$$

The variable r_I is the fraction of gross income $\tilde{Y}_I$ reported to the income tax authorities, therefore it denotes the compliance rate in the income tax scheme. τ_I is the linear income tax rate. We note that a linear income tax simplifies the model significantly, but does so without any loss of generalization.

The individual then may choose in stage 3 to apportion the household budget $\tilde{Y}_C$ between two private goods. Good A carries a consumption tax and good B is free of any tax. We denote the fractions of $\tilde{Y}_C$ spent in gross value of the taxed good A as r_C. Hence, in analogy to the income tax stage, r_C represents the compliance rate in stage 3. We denote net value of the goods as v_A^{net} and v_B^{net}, respectively. With τ_C we denote the proportional consumption tax rate:

$$v_A^{net} = \tilde{Y}_C\left(r_C \frac{1}{1+\tau_C}\right) \quad [3]$$

$$v_B^{net} = \tilde{Y}_C(1-r_C) \quad [4]$$

We do not allow for saving, as a result household budget not spent in good A must be spent in good B. Equations (3) and (4) therefore show that consuming an equal fraction of household budget in the gross value of good A, yields less net value than consumption in good B. We assume that to the individual, the benefit from consuming the net value of the private goods is equal for the two goods. We think of goods A and B, as the same consumption bundle, one that is produced solely in the legal economy (good A) and one that is produced solely in the illegal economy, and therefore, not burdened with a consumption tax (good B). We thus avoid any distortional effects caused by a change in consumption tax, which may lead to the purchase of a different consumption bundle.[1] This approach allows us to isolate the noncompliance decision.

$$\lambda_{A,B} v_A^{net} = \lambda_{A,B}\tilde{Y}_C\left(r_C \frac{1}{1+\tau_C}\right) \quad [5]$$

$$\lambda_{A,B} v_B^{net} = \lambda_{A,B}\tilde{Y}_C(1-r_C) \quad [6]$$

$\lambda_{A,B}$ denotes the multiplier that represents the value of the private goods A and B to each individual. Hence, $\lambda_{A,B}$ introduces consumer surplus into the reward function, and it is reasonable to assume $\lambda_{A,B} > 1$. Diminishing marginal utility is generally not a problem if the reward is solely monetary and if we can assume that, ceteris paribus, all individuals would prefer more money to less money (monotonicity or nonsatiation).[2] We note that we could alternatively set interest rates at zero percent, which would make consumption the dominant decision, but doing so does not influence the outcomes of the model.

The public good is modeled as a group fund to which each individual can contribute $\tilde{P}_I$ from the gross income in the income tax stage and $\tilde{P}_C$ from the household budget in the consumption tax stage, respectively.

$$\tilde{P} = \tilde{P}_I + \tilde{P}_C \qquad = \tilde{Y}_I(r_I \tau_I) + \tilde{Y}_C\left(r_C \frac{\tau_C}{1+\tau_C}\right) \quad [7]$$

Thus, the payment into the group fund depends on compliance rates, tax rates, and income in both stages. We do now need to consider, which expected revenue function $E(\tilde{R})$ the individual faces in each stage. When deciding about income tax compliance in stage 2 the individual faces the following expected revenue:

[1] For some discussion about the strength of this assumption, see Gordon (1989).

[2] For discussion of induced values in experiments refer to Smith (1976) or Wartick *et al.* (1999).

$$E(\widetilde{R}_I^k) = \widetilde{Y}_I^k\left[\lambda_{A,B}(1-r_I^k\tau_I)-(1-r_I^k)\rho_I\delta_I\right]+\frac{\lambda_P}{N}\sum_{j=1}^{N}\widetilde{P}_I^j$$
$$= \widetilde{Y}_I^k\left[\lambda_{A,B}(1-r_I^k\tau_I)-(1-r_I^k)\rho_I\delta_I\right]+\frac{\lambda_P}{N}\sum_{j=1}^{N}\widetilde{Y}_I^j\tau_I r_I^j \quad [8]$$

Superscripts k and j denote variables unique to the individual k and j, respectively. We explicitly model the rational behavior of individual k and assume that all other individuals j face the exact same expected revenue functions.

We multiply all payments into the group fund by a public goods multiplier $\lambda_P > 1$, indicating the social welfare created by the public good. The public good then is divided in equal shares among the N individuals in the group. We denote the probability of detection of tax evasion set by the local authorities in the income tax scheme as δ_I and the penalty rate when audited and found noncompliant as ρ_I. The first term in brackets in equation (8) is the expected revenue that results from the income that remains after income taxes, and the second term is the expected value of a penalty. We note that equation (8) assumes that the household budget after income taxes will be fully consumed in the net value of the private goods. Hence, in stage 2, we model the income tax phase as if there is no consumption tax phase in which noncompliance could occur. The sum in equation (8) is the payment from the public good that depends on the decisions of individual k and all other individuals j $\forall$ j ≠ k.

In stage 3, where the individual chooses between good A from which consumption taxes will be levied and the untaxed good B, the individual faces the expected revenue function:

$$E(\widetilde{R}_C^k) = \widetilde{Y}_C^k\left[\lambda_{A,B}\left(\left(r_C^k\frac{1}{1+\tau_C}\right)+(1-r_C^k)\right)-(1-r_C^k)\rho_C\delta_C\right]+\frac{\lambda_P}{N}\sum_{j=1}^{N}\widetilde{P}_C^j$$
$$= \widetilde{Y}_C^k\left[\lambda_{A,B}\left(1-r_C^k\left(\frac{\tau_C}{1+\tau_C}\right)\right)-(1-r_C^k)\rho_C\delta_C\right]+\frac{\lambda_P}{N}\sum_{j=1}^{N}\widetilde{Y}_C^j\left(\frac{\tau_C}{1+\tau_C}\right)r_C^j \quad [9]$$

In equation (9), we again model consumer surplus by the private goods multiplier $\lambda_{A,B}$ on the net value of good A or good B, respectively. The parameters δ_C and ρ_C denote the probability of detection and the penalty rate with respect to the noncompliance rate r_C in the consumption tax stage. In analogy to the income tax stage, the penalty is paid based on the household budget that is not consumed in the taxed private good A.

We have now set up the theoretical model that we use to make predictions about the issues outlined above, in which we assume a risk neutral individual who maximizes expected value. We recall that we do not intend to find the optimal solution to the problem stated above, because we may have omitted relevant nonmonetary variables, such as conditional cooperation, as discussed, for instance, by Fischbacher *et al.* (2001) and Frey and Torgler (2007), or matters of fairness, as proposed, for instance, by Rabin (1993). However, if we can assume that these omitted variables are not different between the income tax scheme and the consumption tax scheme, then comparing compliance decisions between those schemes rather than solving for the optimum will difference out all constant omitted variables.

3.1.2 Rational Behavior in the Income and Consumption Tax Frame

When contributing to the public good in either the income tax stage or the consumption tax stage, a rational taxpayer should only be concerned about the effects on the individual expected revenue. For our analysis, we compare the two tax schemes to make predictions about compliance levels. If we take for granted that the individual demands an equal rate of return in both framings of the tax payment decision, we must analyze partial derivatives of the standardized expected revenues. In the income tax stage, we use equation (8) to derive:

$$\frac{\partial \dfrac{E(\widetilde{R}_I^k)}{\widetilde{Y}_I^k}}{\partial r_I^k} = -\lambda_{A,B}\tau_I + \rho_I\delta_I + \frac{\lambda_P}{N}\tau_I \qquad [10]$$

For the consumption tax stage, from equation (9) we derive:

$$\frac{\partial \dfrac{E(\widetilde{R}_C^k)}{Y_C^k}}{\partial r_C^k} = -\lambda_{A,B}\left(\frac{\tau_C}{1+\tau_C}\right) + \rho_C\delta_C + \frac{\lambda_P}{N}\left(\frac{\tau_C}{1+\tau_C}\right) \qquad [11]$$

If the omitted variables are constant between both tax schemes, then the following must hold:

$$\frac{\partial \dfrac{E(\widetilde{R}_I^k)}{Y_I^k}}{\partial r_I^k} = \frac{\partial \dfrac{E(\widetilde{R}_C^k)}{Y_C^k}}{\partial r_C^k} \qquad [12]$$

Equation (12) implicitly assumes that compliance does not depend on parameters of the distribution for the random variables $\tilde{Y}_I$ and $\tilde{Y}_C$, particularly the mean value. If we set $\delta_I = \delta_C$ and $\rho_I = \rho_C$, then risk of detection and the penalties the individual will face if detected are exactly the same in both tax schemes, relative to $\tilde{Y}_I$ and $\tilde{Y}_C$. We set $\tau_I = \tau_C/(1 + \tau_C)$. Thus, the effective tax burden is the same in both stages. Equation (12) then yields:

$$r_I^k = r_C^k \qquad [13]$$

Thus, if a rational individual is concerned solely about his or her individual revenue and not about the framing of the tax-paying decision, particularly if there is no systematic dislike for either of the two tax schemes, then we must find $r_I^k = r_C^k$ on average.

3.1.3 Reaction to a Change in Tax Rates

We are further interested in the effect of a change in tax rates on the revenue of individual k. With all other factors constant, an implicit differentiation of equation (8) yields:

$$\frac{\partial r_I^k}{\partial \tau_I} = \left[\frac{\lambda_P - \lambda_{A,B} N}{\lambda_{A,B} \tau_I N - \rho_I \delta_I N - \lambda_P \tau_I} \right] r_I^k \qquad [14]$$

For a comparable change in tax burden in the consumption tax stage, an implicit differentiation of equation (9) gives:

$$\frac{\partial r_C^k}{\partial \left(\frac{\tau_C}{1+\tau_C} \right)} = \left[\frac{\lambda_P - \lambda_{A,B} N}{\lambda_{A,B} \left(\frac{\tau_C}{1+\tau_C} \right) N - \rho_C \delta_C N - \lambda_P \left(\frac{\tau_C}{1+\tau_C} \right)} \right] r_C^k \qquad [15]$$

If we again set $\tau_I = \tau_C/(1 + \tau_C)$, $\delta_I = \delta_C$, and $\rho_I = \rho_C$ as a starting point, the terms in brackets in equations (14) and (15) are equal. Thus, the compliance elasticities of taxes for a comparable change in tax rate are equal for the income and consumption tax-payment decisions.

$$\frac{\partial r_I^k}{\partial \tau_I} \cdot \frac{\tau_I}{r_I^k} = \frac{\partial r_C^k}{\partial \left(\frac{\tau_C}{1+\tau_C} \right)} \cdot \frac{\left(\frac{\tau_C}{1+\tau_C} \right)}{r_C^k} \qquad [16]$$

If we ensure that a change in the tax burden is the same in the income and consumption tax payment decision, it follows from equation (16):

$$\frac{\partial \tau_I}{\tau_I} = \frac{\partial \left(\frac{\tau_C}{1+\tau_C} \right)}{\left(\frac{\tau_C}{1+\tau_C} \right)} \Leftrightarrow \frac{\partial r_I^k}{r_I^k} = \frac{\partial r_C^k}{r_C^k} \qquad [17]$$

The intuition behind the equality of the elasticities is that regardless of the starting values r_I^k and r_C^k in our experimental environment, the relative reaction in compliance should not be different when we alter the tax burden in either the income tax or the consumption tax stage by the same amount.

3.1.4 Reaction to a Change in Detection Probability

Here, we consider the effect of a change in detection probability on the expected revenue, and therefore on the compliance decision of the individual. We decide to alter detection probabilities instead of penalty rates. An equivalent change in penalty rates would have the exact same effect on the expected revenue.[3] Again, we use implicit differentiation to derive from equations (8) and (9):

$$\frac{\partial r_I^k}{\partial \delta_I} = \left[\frac{-\rho_I N}{\lambda_{A,B}\tau_I N - \rho_I \delta_I N - \lambda_P \tau_I} \right](1 - r_I^k) \qquad [18]$$

$$\frac{\partial r_C^k}{\partial \delta_C} = \left[\frac{-\rho_C N}{\lambda_{A,B}\left(\frac{\tau_C}{1+\tau_C}\right) N - \rho_C \delta_C N - \lambda_P \left(\frac{\tau_C}{1+\tau_C}\right)} \right](1 - r_C^k) \qquad [19]$$

In analogy to equations (14) and (15), with $\tau_I = \tau_C/(1 + \tau_C)$, $\delta_I = \delta_C$, and $\rho_I = \rho_C$ as starting points, the terms in brackets are equivalent in equations (18) and (19). Thus, we compute the elasticities:

$$\frac{\partial r_I^k}{\partial \delta_I} \cdot \frac{\delta_I}{r_I^k} = \frac{\partial r_C^k}{\partial \delta_C} \cdot \frac{\delta_C}{r_C^k} \qquad [20]$$

Consistent with the argument developed in the previous section, when we alter detection probabilities for the exact same amount in both stages, it must hold:

$$\frac{\partial \delta_I}{\delta_I} = \frac{\partial \delta_C}{\delta_C} \Leftrightarrow \frac{\partial r_I^k}{r_I^k} = \frac{\partial r_C^k}{r_C^k} \qquad [21]$$

Again, regardless of the starting values r_I^k and r_C^k, a rational taxpayer should not react differently to an equal change in detection probabilities in either of the two schemes.

3.2 Hypotheses Development

We have taken great care to ensure that monetary variables are modeled explicitly and that constant omitted variables are treated exactly alike in both tax stages of our model. Therefore, differences between the two payment schemes should not be expected. However, as is known from other areas of behavioral research, individuals are strongly influenced by the way a problem is presented. Given our previous rationale about influences of framing in behavioral public finance, we ultimately conjecture that subjects will not behave as predicted by the purely monetary theoretical model even

[3] For some experimental evidence on differences between altering detection probability and penalty rates, see Alm *et al.* (1992) and Hessing *et al.* (1992).

though effective tax burden, penalty rates, detection probability, and public goods multiplier are exactly the same in both tax schemes. Since previous research is, to our knowledge, not available on the subject at hand, we cannot draw on earlier results. Theoretically, if we find a significant deviation from the predictions made by our theoretical model, we can conclude that there is a nonconstant omitted variable which we have not incorporated, but which should be considered by policy makers. Given the fundamental objective of our analysis, we do not thrive to investigate the cause for differences between the two tax schemes. However, in order to increase robustness of our findings and to increase understanding of potential differences, we also include additional analyses. We first ask, whether a change in effective tax burden, which would influence relative weight between the two schemes, has different effects in the tax schemes. Second, we compare the effects when detection probability is altered. Our overall hypothesis is that individuals do not behave in the way predicted by the theoretical model in section 3.1.

We first analyze if we are able to find evidence for differences in the general level of compliance between the two tax payment framings when tax burden and detection rates are equal in both schemes. A behavioral factor would be indicated in this case if compliance rates r_I and r_C were different. Therefore, we conjecture:

H1: With equal tax burden and enforcement, compliance rates are significantly different when we frame the compliance decision as an income declaration decision or a consumption decision.

Our model also predicts that a change in the income tax rate τ_I should have the exact same effect on individual compliance as would an identical change in the term $\tau_C/(1 + \tau_C)$. Equation (17) shows that compliance elasticities for comparable tax rate changes are equal. Thus, framing would affect compliance if we observe:

H2: Individuals' reactions to a change in tax burden in the income tax are significantly different from the reaction to a comparable change in consumption tax.

Using the elasticities allows us to analyze H2 independent of the findings in H1. Thus, even if we find different compliance rates at equal tax burden in H1, we might find an additional behavioral interaction factor if adjustment towards a change in tax rates were different. A change in tax rates in either of the tax schemes shifts relative weight of the respective tax in the multiple-stage game, thus allows us to draw inferences about the behavioral effects of the relative importance of the two tax schemes in the overall tax system.

We can use a similar line of argument with respect to equation (21), when we alter detection probabilities in the income or the consumption tax scheme. Again, we might find an additional nonconstant omitted variable, if we were able to find different elasticities of compliance for a change in detection probabilities.

H3: Individuals' reactions to a change in detection probabilities in the income tax are significantly different from the reaction to a comparable change in consumption tax.

Again, since we use elasticities, we already control for any findings that might occur in H1.

3.3 Experimental Design Details

To differentiate between the theoretical model and the question of framing in the experimental setup, we label the tax compliance decision in the income tax scheme as phase1, and the decision in the consumption tax scheme as phase2. After phase2, each individual is informed about audits in phase1 and phase2, and penalties are deducted from the overall revenue. The random gross income at the beginning of each round is uniformly distributed with $\alpha = 1000$ and $\beta = 1600$:

$$\widetilde{Y}_I \sim U(1000, 1600) \qquad [22]$$

After income is randomly assigned, phase1 and phase2 contain the exact same decision, especially the decision on what proportion of $\tilde{Y}_I$ or $\tilde{Y}_C$, respectively, to enter, on which contributions to the public good will be based. The only difference is that we frame the decision in phase1 as declaring income, and the decision in phase2 as buying goods. Further, we must present tax rates with different numerical values, even when the rates imply the exact same effective burden, since income tax and consumption tax are calculated differently. Audits were conducted after each phase but results are revealed only after both phases are completed to ensure that penalty payments do not affect decisions during a round. Subjects were informed about their results at the end of each round.

We designed the instructions[4] and the experimental tool with great care to ensure that they included only tax neutral terminology, and that the decision screens in which the decisions had to be entered were identical for both phases. Subjects were provided with an onscreen calculator to circumvent errors caused by miscalculations. To avoid end-game effects, we informed the subjects that each treatment would run for at least ten rounds and that further rounds would be played based on the results of a random draw. The probability of another round (the end of treatment) was set at 70 percent (30 percent).

Most of the parameters in the model above were kept constant throughout all treatments. The parameter values were available to the subjects at all times on their written instructions. Table 1 shows the constant parameters:

Table 1: Parameters Held Constant Between Treatments

Parameter	Value
λ_P	2.0
$\lambda_{A,B}$	1.1
ρ_I	1.1
ρ_C	1.1

λ_P is the group fund multiplier, $\lambda_{A,B}$ is the multiplier for the value of the private goods A and B, ρ_I and ρ_C are the penalty rates for the income and consumption tax scheme, respectively.

Setting $\lambda_P = 2.0$ and $\lambda_{A,B} = 1.1$ makes it mutually beneficial to contribute the maximum amount to the group fund (thus, a social planner would want to make sure that this is the case). We use the argument of a social planner to justify the introduction

[4] A copy of the instructions can be made available by the authors upon request.

of an audit into the system. Doing so allows us to avoid loaded terminology (Wartick *et al.* (1999)) and also to emphasize the importance of the public good to the subjects. The chosen parameters further yield $\partial E(\tilde{R}_I^k)/\partial r_I^k < 0$ and $\partial E(\tilde{R}_C^k)/\partial r_C^k < 0$, hence from a game theory perspective, it would be individually beneficial to pay no taxes at all.

A significant incentive to comply with the public-goods provision was created by setting a penalty rate of $\rho_I = \rho_C = 1.1$ on the amount of money not declared/not consumed in good A. We note that a factor of 1.1 means that all income not reported will be seized and that there will be an additional fine of ten percent on this amount. Hence, penalties were significant compared to the taxes that would have been paid if the amount was rightfully declared in phase1 or consumed in goods of the legal economy in phase2.

Table 2: Parameters Altered Between Treatments

Treatment	Parameters			
	τ_I	τ_C	δ_I	δ_C
T1: base case	20.00 %	25.00 %	10.00 %	10.00 %
T2: high income tax rate	33.33 %	25.00 %	10.00 %	10.00 %
T3: high consumption tax rate	20.00 %	50.00 %	10.00 %	10.00 %
T4: high income detection probability	20.00 %	25.00 %	20.00 %	10.00 %
T5: high consumption detection probability	20.00 %	25.00 %	10.00 %	20.00 %

τ_I and τ_C are the tax rates; δ_I and δ_C are the detection rates for the income and consumption tax scheme, respectively.

We basically use a 4-factor 2×2×2×2 (income tax rate, consumption tax rate, income tax detection probability, and consumption tax detection probability) within-subject design, but omit all the interaction effects in which we are not particularly interested. Table 2 shows that subjects face a higher numerical value for the consumption tax rate than for the income tax rate. However, since we compute the income tax payment based on the gross income, but compute the consumption tax payment based on the net value of the good, the burden from income (20 percent) and consumption (25 percent) tax in T1 is effectively equal. The effective income tax burden in T2 (33 percent) is also equal to the effective consumption tax burden in T3 (50 percent). Further, the probability of detection is equal in the income tax of T4 and in the consumption tax of T5.

Since all treatments incorporate a significant amount of both types of taxes, there is little difference in the complexity of the overall tax system between treatments that might disturb comparative analysis. All treatments were played in each experimental session in a random sequence. To capture learning effects at the beginning of sessions, subjects played several trial rounds directly after the instructor read the instructions aloud while subjects read along. The parameters of the trial rounds corresponded to the parameters of the first treatment in the particular session.

To increase robustness of our results, we further asked the subjects to complete a short questionnaire at the end of each treatment with the intention to elicit subjective perceptions. We first asked the subjects about their perception of tax burden in the income tax and the consumption tax framing. We called these multiple-choice variables PERC_INC and PERC_CON, with a minimum value of one and a maximum value of six, where a high value indicates a high perceived tax burden. We further elicited an indirect measure of the perceived tax burden, namely, asking for the minimum group

fund multiplier λ_{Pmin} at which subjects would be willing to comply fully in each phase. The variables MULT_INC and MULT_CON indicate the values for phase1 and phase2, respectively. We note that for $\lambda_{Pmin} = 4 \cdot \lambda_{A,B}$, full compliance dominates all other alternatives, regardless of any other group member's decisions, tax rates, and detection probabilities. For the differences between PERC_INC and PERC_CON as well as MULT_INC and MULT_CON we expect similar results as for r_I and r_C.

We used N = 8 subjects in each of ten computer-based experimental sessions. Each subject drew a random number from a bingo cage prior to the start of the session. We assured subjects that their decisions could only be tracked by the system based on their subject number, which was not known to the experimenters at any time. To provide a maximum amount of anonymity, even within their peers, subjects were then randomly assigned to two groups of four subjects each. Subjects had no information about group composition at any time. Both groups had an own group fund, so the two groups played completely independent from one another. Further, anew random assignment of groups at the beginning of each treatment allowed us to analyze treatments independently from one another. The undergraduate and graduate student subjects (average age: 23.35 years; average number of semesters: 6.2) had previously registered on a website that is regularly used for recruiting subjects to economic experiments. About two weeks prior to the experimental sessions, 598 potential subjects received an email informing them about the experiments and asking them to register for the experiments online. Subjects were then automatically scheduled to sessions at their own convenience at a first come, first served basis. Sessions lasted approximately 1.5 hours and the 80 subjects received an average pay of € 19.32.

4. Data Analysis

4.1 Descriptive Statistics

The experimental sessions yield data on the subject level for 4,880 rounds. We analyze the average treatment level data per individual to avoid any problems that might be caused by different numbers of rounds in a treatment, and also to increase the robustness of our statistical results. We use nonparametric statistical methods to avoid assumptions about our experimental data (e.g., normally distributed variables). However, we find that using parametric statistical tests where they are feasible produces similar results. To distinguish between the variables in the theoretical section of this paper and the realizations of the variable, in the remainder of the paper we denote realizations with a preceding "_". When we discuss a specific value of a variable in a certain treatment, we indicate the treatment name at the beginning of the variable name.

Table 3 shows the average compliance rates for all treatments in the income tax (phase1 with variable $\_r_I$) and consumption tax (phase2 with variable $\_r_C$). As a test for robustness of results, we further utilize the answers given in the after treatment questionnaires. Those variables provide support for the experimental results if a relatively low (high) average compliance rate $\_r_I$ and $\_r_C$ correlates with a high (low) perception of tax burden (_PERC_INC and _PERC_CON). Further, a relatively low (high) average compliance rate should correlate with a high (low) minimum multiplier for the group fund (_MULT_INC and _MULT_CON), at which point subjects would be willing to comply fully.

Table 3: Descriptive Statistics for Experiment and Questionnaire Variables

Variable	Quartiles			Mean	STD
	25 %	50 %	75 %		
Experiment					
T1_r_I	0.350	0.682	0.851	0.604	0.301
T1_r_C	0.322	0.580	0.799	0.560	0.292
T2_r_I	0.206	0.583	0.811	0.524	0.329
T2_r_C	0.300	0.624	0.831	0.573	0.320
T3_r_I	0.369	0.745	0.911	0.630	0.334
T3_r_C	0.194	0.478	0.768	0.473	0.332
T4_r_I	0.695	0.911	1.000	0.803	0.243
T4_r_C	0.247	0.565	0.754	0.519	0.315
T5_r_I	0.319	0.656	0.897	0.576	0.346
T5_r_C	0.633	0.855	0.974	0.780	0.243
Questionnaire					
T1_PERC_INC	2.000	3.000	3.000	2.900	1.109
T1_PERC_CON	2.000	3.000	4.000	3.150	1.223
T1_MULT_INC	3.000	3.000	4.000	4.043	4.231
T1_MULT_CON	3.000	3.000	4.000	4.099	4.147
T2_PERC_INC	3.000	4.000	5.000	3.638	1.343
T2_PERC_CON	2.000	3.000	4.000	3.012	1.049
T2_MULT_INC	3.000	4.000	4.000	4.856	6.641
T2_MULT_CON	2.500	3.000	4.000	4.412	6.556
T3_PERC_INC	2.000	2.000	3.000	2.362	1.046
T3_PERC_CON	3.000	5.000	6.000	4.375	1.578
T3_MULT_INC	2.000	3.000	4.000	3.849	4.259
T3_MULT_CON	3.000	3.500	5.000	4.401	4.153
T4_PERC_INC	2.000	3.000	3.000	2.725	0.954
T4_PERC_CON	2.000	3.000	3.750	2.950	1.101
T4_MULT_INC	2.000	3.000	4.000	3.657	4.085
T4_MULT_CON	3.000	3.000	4.000	4.027	4.003
T5_PERC_INC	2.000	3.000	3.000	2.712	1.058
T5_PERC_CON	2.000	3.000	3.000	2.812	1.068
T5_MULT_INC	3.000	3.000	4.000	3.954	4.059
T5_MULT_CON	2.000	3.000	4.000	3.616	4.119

One subject, who reported 9,999,999 as _MULT_INC and _MULT_CON in all treatments, was excluded for all analyses using those variables.
Preceding labels T1 – T5 indicate the treatment in which variables have been measured. The variables r_i and r_c are the compliance rates in phase1 and phase2, respectively. _PERC indicates perceived tax burden, _MULT indicates the minimum multiplier at which subjects would have been willing to comply fully. _INC and _CON indicate variables for phase1 and phase2, respectively.

One of the underlying assumptions for the statistical inference tests is that the decision of a particular subject in phase1 (_r_I) and phase2 (_r_C) does not depend on the decisions of the other subjects. As discussed above, we deliberately designed the experiment to avoid the effect of possible communication via signaling. Therefore, our main concern in this area is that the possible communication about individuals' overall level of compliance between subjects might have an effect on the particular

$\_r_I/\_r_C$-combination. Since communication is only possible in case of complete noncompliance, the independence assumption would be reasonable if the ratio between $\_r_I$ and $\_r_C$ does not change when the overall level of compliance changes. The data support our assumption. A comparative analysis of $\_r_I/\_r_C$-combinations chosen by subjects whose overall compliance was lower than the 25th percentile and subjects whose overall compliance was higher than the 75th percentile yields no significant differences ($p = 0.836$, Mann-Whitney-U, two-tailed).

For the questionnaire variables, all of the median values of the _MULT_INC and _MULT_CON are significantly lower than 4.4, the group fund multiplier at which full compliance would dominate in both phases over all other alternatives (all at the 1 percent-level, Wilcoxon signed rank test, two tailed). Thus, we find evidence that subjects had a good understanding of the interaction between the two private goods and the public good. We further conclude that subjects would fully comply in treatments when there was a group fund multiplier lower than the absolute dominance multiplier. This result indicates that detection probabilities and penalty rates are significant enough to deter full noncompliance.

We further utilize elasticities in our theoretical model to predict subjects' behavior when tax rates and detection probabilities are altered. We calculate estimates for the elasticities from the experimentally observed changes in compliance rates when we alter tax burden and detection probabilities, respectively. We compute the values for equations (16) and (20) between values in T1 and the respective values in the other treatments that are relevant for our analysis. We note that we compute the values for the perception of tax burden (_PERC) and the minimum multiplier, at which subjects would comply fully (_MULT), similar to the way in which we computed those in equations (16) and (20).

Table 4: Descriptive Statistics for the Elasticity Estimates

Variable	Quartiles			Mean	STD
	25 %	50 %	75 %		
Increasing Tax Burden (H2)					
T1T2_ELAST_r_I	-0.672	-0.097	0.126	-0.131	0.941
T1T3_ELAST_r_C	-0.835	-0.141	0.250	-0.088	1.233
T1T2_ELAST_PERC_INC	0.000	0.375	1.375	0.612	1.071
T1T3_ELAST_PERC_CON	0.000	0.750	1.500	0.922	1.389
T1T2_ELAST_MULT_INC	0.000	0.000	0.500	0.327	1.576
T1T3_ELAST_MULT_CON	0.000	0.000	0.500	0.239	0.827
Increasing Detection Probability (H3)					
T1T4_ELAST_r_I	0.090	0.436	1.210	1.584	3.790
T1T5_ELAST_r_C	0.052	0.869	2.301	1.861	4.100
T1T4_ELAST_PERC_INC	-0.500	0.000	0.000	0.115	1.096
T1T5_ELAST_PERC_CON	-0.667	0.000	0.000	-0.025	0.913
T1T4_ELAST_MULT_INC	-0.609	-0.286	0.000	-0.126	0.717
T1T5_ELAST_MULT_CON	-0.571	-0.333	0.000	-0.240	0.496

Preceding labels T1 – T5 indicate the treatments between which elasticities are computed. The variables $\_r_I$ and $\_r_C$ are the compliance rates in phase1 and phase2, respectively. _PERC indicates perceived tax burden, _MULT indicates the minimum multiplier at which subjects would have been willing to comply fully. _INC and _CON indicate variables for phase1 and phase2, respectively.

When looking at Table 4, we note that the elasticity with respect to tax burden (H2) seems to be higher for the variable _PERC than for the variable _MULT. Differences are highly significant ($p = 0.000$ for income tax and $p = 0.001$ for consumption tax, Wilcoxon, two-tailed). For the elasticity with respect to detection probability (H3), this effect is exactly opposite, but the differences are nonsignificant ($p = 0.139$ for income tax and $p = 0.152$ for consumption tax, Wilcoxon, two-tailed). Both observations are reasonable, since a change in tax rate should have a stronger effect on the perceived tax burden than on the minimum multiplier at which full compliance would occur, and vice versa for a change in detection probabilities. These results serve as further evidence that subjects understood the interaction between variables in the experiment, and support the credibility of subjects' reactions towards a change in the experimental environment.

We note that the main assumption of the elasticity measure is that changes are incremental. However, if it is true that the real relationship between change in compliance (and questionnaire variables) towards changes in tax burden and detection probabilities, respectively, is the same for income and consumption taxes, we should not find differences.

4.2 Results

Table 5 shows the comparative results of the variables that are either measured in the experiment or obtained through the questionnaire.

Table 5: Comparative Results Between the Two Tax Payment Framings

Income Tax Variable	Consumption Tax Variable	Wilcoxon	
		Δ Sum of Ranks	p-value
Direct Comparison (H1)			
T1_r_I	T1_r_C	+	0.014**
T1_PERC_INC	T1_PERC_CON	-	0.042**
T1_MULT_INC	T1_MULT_CON	-	0.390
Increasing Tax Burden (H2)			
T1T2_ELAST_r_I	T1T3_ELAST_r_C	+	0.883
T1T2_ELAST_PERC_INC	T1T3_ELAST_PERC_CON	-	0.071*
T1T2_ELAST_MULT_INC	T1T3_ELAST_MULT_CON	-	0.877
Increasing Detection Probability (H3)			
T1T4_ELAST_r_I	T1T5_ELAST_r_C	-	0.021**
T1T4_ELAST_PERC_INC	T1T5_ELAST_PERC_CON	+	0.453
T1T4_ELAST_MULT_INC	T1T5_ELAST_MULT_CON	+	0.105

*The Wilcoxon-Test is based on the general form: income tax variable – consumption tax variable. For Δ Sum of Ranks a "+" indicates a higher absolute value of positive sum of ranks and "-" indicates a higher absolute value of negative sum of ranks. Thus "+" ("-") indicates a higher (lower) income tax variable compared to the corresponding consumption tax variable. *, **, and *** highlight significant differences at level 10 %, 5 %, and 1 %, the Wilcoxon-Test is carried out two-tailed.*

Since our main goal is to analyze the comparative results between the different tax payment decisions, we investigate whether subjects' compliance behavior is different between the income reporting framing and the consumption framing, respectively (H1).

Therefore, we compare experimentally observed compliance rates $T1\_r_I$ and $T1\_r_C$. Table 3 shows that the observed compliance rates differ by 10.2 percentage points in the median, and Table 5 yields evidence that these differences are significant ($p = 0.014$). This finding supports H1. Those results further are strongly supported by significant differences between T1_PERC_INC and T1_PERC_CON ($p = 0.042$). Differences in the desired multiplier for full compliance point in the expected direction, but are nonsignificant (T1_MULT_INC < T1_MULT_CON, $p = 0.390$).

Table 6: Spearman Correlations for Variables in Treatment 1 (T1)

Income Tax Framing (phase1)		
	T1_PERC_INC	**T1_MULT_INC**
$T1\_r_I$	-0.375***	-0.342***
	(0.001)	(0.002)
T1_MULT_INC	-0.034	--
	(0.767)	--
Consumption Tax Framing (phase2)		
	T1_PERC_CON	**T1_MULT_CON**
$T1\_r_C$	-0.163	-0.282**
	(0.148)	(0.012)
T1_MULT_CON	0.094	--
	(0.410)	--

*Two-tailed p-values in parentheses. *, **, and *** highlight significance at level 10 %, 5 %, and 1 %.*

Because H1 is the main hypothesis of this paper, in Table 6 we also analyze the correlation between the relevant variables. All correlations point in the same direction, as we expected from the findings in Table 5 and significance of correlation between the variables indicates the robustness of the reported findings.

We are further interested in the analysis of differences in elasticities when we alter the effective tax burden (H2). The absolute median value of the tax burden elasticities of compliance is higher in the consumption tax phase than in the income tax phase but differences are nonsignificant ($p = 0.883$). The findings from the questionnaire variables _PERC and _MULT both point in the same direction as the experimentally observed compliance rates. However, the differences are only marginally significant for the perception of tax burden ($p = 0.071$) and nonsignificant for the minimum multiplier at which subjects would comply fully ($p = 0.877$).

We next consider whether we find differences in elasticities when detection probabilities are increased (H3). To draw our conclusions, we compare the detection probability elasticities of compliance from Table 5. We generally find positive elasticities in our data. Experimentally observed elasticities of compliance rates are higher in the consumption tax phase than in the income tax phase and the differences are significant ($p = 0.021$). The differences in the questionnaire variables _PERC and _MULT both point in the direction that supports the findings in experimentally observed compliance rates, but both differences are nonsignificant ($p = 0.453$ for _PERC and $p = 0.105$ for _MULT).

4.3 Discussion

Our analysis of the compliance rates and questionnaire responses in T1 indicates strong support for H1, which posits that compliance is significantly different between the consumption tax phase and the income tax phase. In T1, effective tax burden and detection rates are equivalent, and all constant omitted variables difference out. Therefore our results indicate a strong effect of framing on tax compliance. When looking at the questionnaire variables we find that subjects' perceived tax burden is higher in the consumption tax scheme, which could be one determinant explaining the differences in compliance.

For H2 we note that in experimental noncompliance research it is well established that compliance decreases when tax rates increase, as can be seen, for instance, in the literature reviews from Jackson and Milliron (1986) or Alm (1991). Thus, the negative elasticities for the reaction in $\_r_I$ and $\_r_C$ are what previous research indicated. Nevertheless, we cannot reject the Null hypothesis that elasticities are equal in both phases. We have in consequence not been able to identify an additional behavioral factor that indicates individuals' reactions toward a comparable change in tax rates (H2). Results therefore indicate that the relative importance of direct and indirect taxation in the overall tax system does not have an influence on compliance.

From previous experimental research, we would then expect a positive elasticity of compliance when detection probability is increased, as can, for instance, be found in Fischer *et al.* (1992) and is also found in our analysis. For H3, we find evidence that individuals react stronger when enforcement efforts are increased in the consumption tax than in the income tax. However, this evidence is not as robust as the findings from H1, since it cannot be found in two of the three relevant variables.

As with any applied research, our approach has some caveats concerning internal and external validity which we need to draw attention to. Internal validity refers to the ability of the experiment to identify the main variables that influence the outcome. To allow for clear-cut predictions of individual behavior of a rational taxpayer, our theoretical model must rely on assumptions: First, the experimental design permits two decisions in two different phases. The sequence of those phases is fixed, i.e., the income tax payment decision is always prior to the consumption tax payment decision, which could lead to unmeasured sequential effects. Hence a possible alternative explanation for our findings could be that with several consecutively levied tax payments, individuals generally comply more during the earlier tax payment decisions. Thus, if we were to switch the sequence of the tax payment decisions, we should then expect to see compliance shift more towards the consumption tax scheme. We could further consider running recurring tax payment decisions, for instance two or more income tax schemes, and might then observe compliance decreasing in the latter stages of the multiple-stage game, even though there are no differences in the framing of the tax payment decision. In our experimental setup we are not able to control for such a case. However, the sequence we choose for our experiment seems to be the most realistic representation of tax systems around the world. Many other theoretical models either explicitly or implicitly rely on the same assumption, such as those of Atkinson and Stiglitz (1976), Atkinson and Stern (1980), Hines (2004), and Richter and Boadway (2005). Second, our theoretical model relies strongly on the assumptions that the individual will regard the decisions in the two tax schemes as independent, as it is conjectured in equations (8) and (9), and further, that the individual decides based solely on the expected value of the monetary variables, given that all omitted variables are

constant. We can neither verify nor falsify these assumptions and thus they remain weaknesses of the investigation. Future research should investigate whether omitted variables can be identified that can explain the behavior observed in our experiments.

External validity of our experiment could be reduced by the utilization of students as subjects. Students might have different attitudes or real world perceptions towards noncompliance than would other subjects outside the academic world. Further, due to low income of students and therefore low income taxes, consumption taxes have a stronger impact on students' after tax budget and, for the same reason, students might also be more familiar with consumption tax noncompliance situations. However, 53 of the 80 subjects were in the second half of their studies and 46 had handed in at least one income tax return. Thus, we conclude that numerous subjects have experience outside academia. Further, it does not seem feasible to believe that the cognitive processes of students regarding tax noncompliance are significantly different from those of non-students when tax neutral terminology is used, as the research by Wartick *et al.* (1999) indicates. In fact, using student subjects allows us to take advantage of the presumably low income outside the laboratory, so in the experimental environment, monetary incentives are more likely to induce necessary reward dominance. A further external validity aspect is that we do not account for the different costs of audits, which certainly should be a decision variable for a policy maker. Also, external validity is limited by the fact that the opportunity to evade taxes might be different for both tax schemes, for instance that the consumption tax scheme needs a black market seller of goods to be able to evade at all. Finally, it must generally be noted that experimental research can hardly incorporate all factors that might affect tax compliance behavior. In our experimental design, omitted factors could include the notion of fairness, fear of social proscription, fear of imprisonment or even physical punishment, subjective perception about the spending of taxes, acceptance of a nation's political system, and many others. If, on the one hand, it could be assumed that these factors affect direct taxation and indirect taxation similarly, then the effects on our comparative analysis regarding noncompliance would be negligible. If, on the other hand, it must be suspected that the factors which are not incorporated have a different effect on direct taxation and indirect taxation, this would in fact yield additional support for our general notion that the framing of the tax payment decision has an effect on tax compliance.

With these caveats in mind, we find different behavioral effects in the data. First, we find highly significant differences in compliance rates and perceived tax burden between the income tax and consumption tax phase, when effective tax rates are equal (H1). This finding is the single most important result of our study. It is strongly supported by the data for all treatments. Second, even though we find decreasing compliance rates when tax rates are increased, when taxes are altered we do not find significant differences in the elasticities of compliance between the income and consumption tax scheme (H2). Last, we find marginal support for H3, that subjects react differently towards a change in detection probabilities under different framing. Overall, we find that nonconstant omitted variables do influence compliance decisions in the income and consumption tax scheme differently. Behavioral differences therefore should be considered when designing a tax system.

5. Conclusion

In our paper we use data from laboratory experiments to analyze different compliance decisions when policy makers structure the tax payment decision either as

an income tax or a consumption tax. Our main finding is that individuals are significantly less compliant under the consumption tax structure than under the income tax structure. Although we do not find differences in reaction towards a change in tax rates, we do find significant differences towards a change in detection probabilities between the two framings of the tax payment decision. Individuals seem to react more strongly towards a change in enforcement in the consumption than in the income tax scheme.

As is true with any applied work, this data should be viewed with caution. On the one hand, though the experiments are carefully designed to capture the essential parameters of the real environment (see Plott (1987) for the notion of parallelism), the external validity of the results is uncertain. On the other hand, field data are not available now, and are very unlikely to become available in the future.

Overall, we do find that individuals react differently to an income or a consumption tax in our experiments, and that those differences are significant. Therefore, a policy maker should consider the findings of this paper, bearing in mind the caveats of experimental research as a whole and of our experiments in particular. Although an indirect consumption tax might increase the employment level and might be easier to enforce, those benefits possibly will be outweighed by a higher level of noncompliance. Further, decreasing overall tax morale caused by a more consumption-tax-based system could lead to stronger feelings of inequity in the overall tax system. This effect might create a general increase in noncompliance, and also affect the income tax compliance in the long run.

Future research should increase robustness of our findings, either by additional experiments or through empirical analysis of archival data. It could also be interesting to examine the determinants that lead to the differences in noncompliance between different framings of taxation. Social welfare could be increased by levying public finances through a tax system in which the perceived tax burden of the overall tax system is minimal, all other factors – particularly government revenue – equal.

References

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

Alm, J. (1991), 'A Perspective on the Experimental Analysis of Taxpayer Reporting', *The Accounting Review*, 66, 577-593.

Alm, J., Jackson, B. R., McKee, M. (1992), 'Estimating the Determinants of Taxpayer Compliance with Experimental Data', *National Tax Journal*, 45, 107-114.

Andreoni, J., Erard, B., Feinstein, J. S. (1998), 'Tax Compliance', *Journal of Economic Literature*, 36, 818-860.

Atkinson, A. B., Stiglitz, J. E. (1976), 'The Design of Tax Structure: Direct Versus Indirect Taxation', *Journal of Public Economics*, 6, 55-75.

Atkinson, A. B., Stern, N. H. (1980), 'On the Switch from Direct to Indirect Taxation', *Journal of Public Economics*, 14, 195-224.

Becker, G. S. (1968), 'Crime and Punishment: An Economic Approach', *The Journal of Political Economy*, 76, 169-217.

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

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

Dixit, A. K. (1970), 'On the Optimum Structure of Commodity Taxes', *The American Economic Review*, 60, 295-301.

Fischbacher, U., Gaechter, S., Fehr, E. (2001), 'Are People Conditionally Cooperative? Evidence from a Public Goods Experiment', *Economics Letters*, 71, 397-404.

Fischer, C. M., Wartick, M. L., Mark, M. M. (1992), 'Detection Probability and Taxpayer Compliance: A Review of the Literature', *Journal of Accounting Literature*, 11, 1-46.

Forest, A., Sheffrin, S. M. (2002), 'Complexity and Compliance: An Empirical Investigation', *National Tax Journal*, 55, 75-88.

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

Gordon, J. P. F. (1989), 'Individual Morality and Reputation Costs as Deterrents to Tax Evasion', *European Economic Review*, 33, 797-805.

Gordon, J. P. F. (1990), 'Evading Taxes by Selling for Cash', *Oxford Economic Papers*, 42, 244-255.

Gueth, W., Sausgruber, R. (2004), 'Tax Morale and Optimal Taxation', *Max Planck Institute of Economics*, Strategic Interaction Group.

Hessing, D. J., Elffers, H., Robben, H. S. J., Webley, P. (1992), 'Does Deterrence Deter? Measuring the Effect of Deterrence on Tax Compliance in Field Studies and Experimental Studies', in Slemrod, J. (ed), *Why People Pay Taxes. Tax Compliance and Enforcement*, Ann Arbor, The University of Michigan Press.

Hines, J. R. (2004), 'Might Fundamental Tax Reform Increase Criminal Activity?' *Economica*, 71, 483-492.

Hotelling, H. (1938), 'The General Welfare in Relation to Problems of Taxation and of Railway and Utility Rates', *Econometrica*, 6, 242-269.

Jackson, B. R., Milliron, V. C. (1986), 'Tax Compliance Research: Findings, Problems, and Prospects', *Journal of Accounting Literature*, 5, 125-165.

Kahneman, D., Tversky, A. (1981), 'The Framing of Decisions and the Psychology of Choice', *Science*, 211, 453-458.

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

Lerner, A. P. (1970), 'On Optimal Taxes with an Untaxable Sector', *American Economic Review*, 60, 284-294.

McCaffery, E. J. (2000), 'Cognitive Theory and Tax', in Sunstein, C. R. (ed), *Behavioral Law & Economics*, Cambridge, Cambridge University Press.

McCaffery, E. J., Slemrod, J. (2006a), *Behavioral Public Finance*, New York, Russel Sage Foundation.

McCaffery, E. J., Slemrod, J. (2006b), 'Toward an Agenda for Behavioral Public Finance', in McCaffery, E. J. and Slemrod, J. (eds), *Behavioral Public Finance*, New York, Russel Sage Foundation.

McLaren, J. (1998), 'Black Markets and Optimal Evadable Taxation', *The Economic Journal*, 108, 665-679.

Milliron, V. C. (1985), 'A Behavioral Study of the Meaning and Influence of Tax Complexity', *Journal of Accounting Research*, 23, 794-816.

Musgrave, R. A., Musgrave, P. B. (1989), *Public Finance in Theory and Practice*, New York et al., McGraw-Hill Book Company.

Plott, C. R. (1987), 'Dimensions of Parallelism: Some Policy Applications of Experimental Methods', in Roth, A. E. (ed), *Laboratory Experimentation in Economics: Six Points of View*, Cambridge et al., Cambridge University Press.

Rabin, M. (1993), 'Incorporating Fairness into Game Theory and Economics', *American Economic Review*, 83, 1281-1302.

Ramsey, F. P. (1927), 'A Contribution to the Theory of Taxation', *The Economic Journal*, 37, 47-61.

Richter, W. F., Boadway, R. (2005), 'Trading Off Tax Distortion and Tax Evasion', *Journal of Public Economic Theory*, 7, 361-381.

Saez, E. (2004), 'Direct or indirect tax instruments for redistribution: short-run versus long-run', *Journal of Public Economics*, 88, 503-518.

Sandmo, A. (1976), 'Optimal Taxation: An Introduction to the Literature', *Journal of Public Economics*, 6, 37-54.

Sausgruber, R. (2002), *Ursachen Und Wirkungen Von Steuerillusionen*, Innsbruck, Nomos.

Slemrod, J. (1990), 'Optimal Taxation and Optimal Tax Systems', *The Journal of Economic Perspectives*, 4, 157-178.

Smith, V. L. (1976), 'Experimental Economics: Induced Value Theory', *The American Economic Review*, 66, 274-279.

Wartick, M. L., Madeo, S. A., Vines, C. C. (1999), 'Reward Dominance in Tax Reporting Experiments: The Role of Context', *The Journal of the American Taxation Association*, 21, 20-31.

Webley, P., Adams, C., Elffers, H. (2006), 'Value Added Tax Compliance', in Slemrod, J. (ed), *Behavioral Public Finance*, New York, Russel Sage Foundation.