Essays on behavioral tax research and tax accounting

INAUGURALDISSERTATION

zur Erlangung des akademischen Grades eines Doktors
der Wirtschaftswissenschaft des Fachbereichs Wirtschaftswissenschaft
der Freien Universität Berlin

Vorgelegt von
Matthias Sünwoldt, M.Sc.
aus Perleberg
2016

Erstgutachter: Prof. Dr. Jochen Hundsdoerfer

Zweitgutachter: Prof. Dr. Martin Fochmann

Drittgutachter: Prof. Dr. Frank Hechtner

Tag der Disputation: 23. Juli 2016

Publikationsverzeichnis und Angabe der Koautoren

Diese Dissertation besteht aus insgesamt fünf Beiträgen:

Erster Beitrag: Does legality matter? The case of tax avoidance and evasion

- Dieser Beitrag wurde gemeinsam mit Herrn Prof. Dr. Kay Blaufus (Leibniz Universität Hannover), Herrn Prof. Dr. Jochen Hundsdoerfer (Freie Universität Berlin) und Herrn Prof. Dr. Martin Jacob (WHU Otto Beisheim School of Management Vallendar) erstellt. Der Anteil der vier Autoren beträgt jeweils ein Viertel.
- Der Beitrag wurde im Rahmen von Seminaren an der NHH Bergen, an der WU Wien, an der Friedrich Schiller Universität Jena und an der Otto von Guericke Universität Magdeburg sowie im Rahmen des ZEW Workshops „Empirische Steuerforschung“, des Berlin-Köln Doktorandenworkshops und der Tagung des VfS „Ausschuss Unternehmensrechnung“ vorgestellt.
- Eine Vorversion wurde 2015 unter gleichlautendem Titel unter meinem Geburtsnamen *Braune* als FACTS Working Paper 2015/23 des Fachbereichs Wirtschaftswissenschaft der Freien Universität Berlin veröffentlicht und ist abrufbar unter http://hdl.handle.net/10419/114185.
- Der Beitrag wurde im *Journal of Economic Behavior & Organization* Vol. 127, S. 182-206 veröffentlicht.

Zweiter Beitrag: Self-serving bias and tax morale

- Dieser Beitrag wurde gemeinsam mit Herrn Prof. Dr. Kay Blaufus (Leibniz Universität Hannover), Herrn Prof. Dr. Jochen Hundsdoerfer (Freie Universität Berlin) und Herrn Prof. Dr. Martin Jacob (WHU Otto Beisheim School of Management Vallendar) erstellt. Der Anteil der vier Autoren beträgt jeweils ein Viertel.
- Der Beitrag wurde im Jahr 2015 unter meinem Geburtsnamen *Braune* in der Fachzeitschrift *Economics Letters* Vol. 131(C) S. 91–93 veröffentlicht.

Dritter Beitrag: **Mental accounting and the timing of pension taxation**

- Dieser Beitrag wurde gemeinsam mit Herrn Prof. Dr. Kay Blaufus (Leibniz Universität Hannover), Herrn Prof. Dr. Jochen Hundsdoerfer (Freie Universität Berlin) und Frau Nadja Wolf (Leibniz Universität Hannover) erstellt. Der Anteil der vier Autoren beträgt jeweils ein Viertel.
- Der Beitrag wurde im Rahmen der Jahrestagung 2015 der GfEW e. V., des Workshops „Coping with Difficult Decisions – An Experimental Economics Perspective" in Düsseldorf, des WHU „Accounting & Tax Research" Seminars und im Rahmen des ZEW Workshops „Lab Experiments in Public Economics" vorgestellt und wurde für eine Präsentation im Rahmen der Pfingsttagung 2016 des VHB e. V. angenommen.

Vierter Beitrag: **Do tax earmarking and voting on the tax rate influence tax avoidance and labor supply in the lab?**

- Dieser Beitrag wurde gemeinsam mit Herrn Prof. Dr. Kay Blaufus (Leibniz Universität Hannover) und Herrn Prof. Dr. Jochen Hundsdoerfer (Freie Universität Berlin) erstellt. Der Anteil der drei Autoren beträgt jeweils ein Drittel.
- Der Beitrag wurde im Rahmen des Workshops „Accounting & Taxation Research" in Mannheim, des arqus e. V. Doktorandenworkshops in Bochum, des Workshops „Experimentelle Steuerforschung" in Paderborn, des WHU „Accounting & Control Research" Seminars und im Rahmen der Pfingsttagung 2013 des VHB e. V. vorgestellt.

Fünfter Beitrag: **The information content of large book-tax differences – Empirical evidence from Germany**

- Dieser Beitrag wurde ohne Koautoren erstellt.
- Der Beitrag wurde im Rahmen des ZEW Workshops „Empirical Taxation Research" und des Doktorandenworkshops 2015 an der WU Wien vorgestellt.

Weitere Publikationen, die nicht Bestandteil der Dissertation sind, bestehen nicht.

Contents

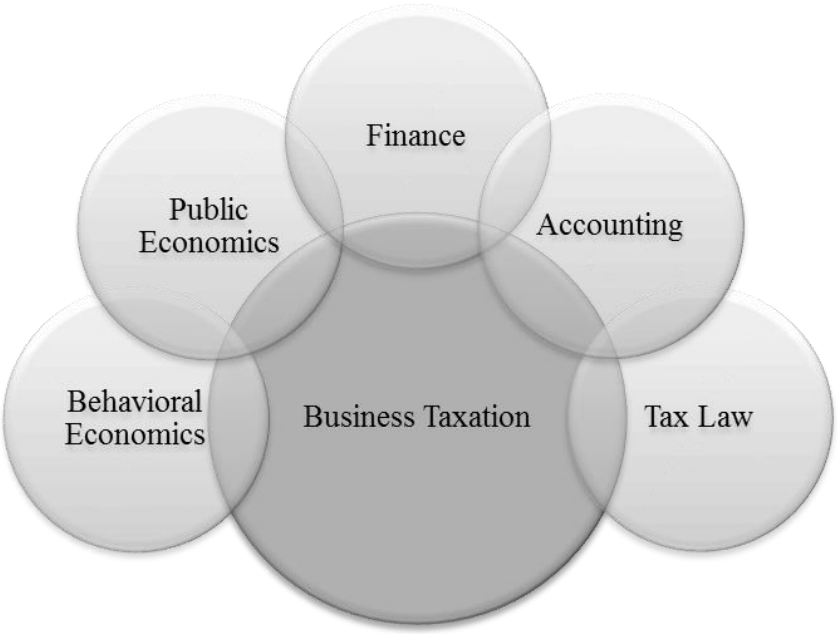
List of tables

List of figures

1 General introduction

Over the past ten years, research in business taxation has changed, especially from a German perspective. Hundsdoerfer et al. (2008) already illustrate the importance of empirical research to observe whether taxes unfold real tax burden effects, distributive effects, and economic decision effects. Nevertheless, related empirical evidence is rather rare at this time. Almost ten years later, empirical research is the state of the art in business taxation.[1] Moreover, the focus changed from a rather national and normative view to a diversified international view. Figure 1 displays the current development of empirical research in business taxation. It is related to a variety of different fields of research ranging from behavioral economics to tax law. Moreover, the different fields of research are interdisciplinary as well (see for instance Maydew 2001 for the interdisciplinary of tax accounting).

Figure 1: Empirical research in business taxation and interfaces



This thesis focuses on research questions related to the interface between business taxation and behavioral/public economics (first four studies) and with accounting (last study). Thus, the thesis is divided into two, rather different, parts. The research fields related to finance and tax law are not considered.[2] In the following, a short overview on behavioral/public economic and business taxation is presented and it is shown how the first four thesis studies are connected, which make up the main part of the thesis. Afterwards, the second part of this thesis, which is related to research in accounting, is introduced and an overall connection is presented.

[1] This development is controversial. For instance, Schneider et al. 2013 discuss the importance of non-empirical, thus, normative and analytical research in business taxation.

[2] The area taxation and (corporate) finance deals with research questions related to, for instance, cost of capital, capital structure and payout policy (Graham 2006). Tax law research, or legal tax research, strongly focuses on the tax law and the tax system itself to identify occurring problems, to develop possible solutions, and to derive recommendations for tax authorities and tax payers. Thus, this field of research covers both empirical and theoretical approaches, and it generally does not consider possible effects on observable behavior.

Behavioral economics could either be seen as a separate field of research besides public economics or it could be seen as a subfield (Chetty 2015). Whereas typical research questions are often similar (e.g., effects of taxation on labor supply, consumption and saving, welfare), the main difference between these two fields is the assumption on whether subjects behave rationally. Relaxing the rationality assumption allows researchers to explain subjects' real behavior and it leads to different or even new policy implications. For instance, Chetty et al. (2013) observe that the impacts of the U.S. Earned Income Tax Credit (EITC) on labor supply depend on individuals' knowledge about the EITC schedule. Differences with respect to knowledge are not considered in a neoclassical standard model. Thus, relaxing the rationality assumption helps to explain individuals' behavior and leads to new policy implications (e.g., considering knowledge as an additional determinant). Another example is the "slippery slope" framework in the context of tax compliance (Kirchler et al. 2008). Based on power of and trust in tax authorities as dimensions for compliance and as an alternative to a neoclassical standard model of tax evasion (Allingham and Sandmo 1972, Srinivasan 1973), this framework is widely used to explain real compliance behavior (e.g., Kastlunger et al. 2013, Kogler et al. 2013, Kogler et al. 2015). Moreover, Prinz et al. (2014) formalized this framework and linked it to the standard economic model of tax evasion.

The first four of the five studies of this thesis can be referred to the business taxation and behavioral/public economics interface. The first study examines whether the legality (legal tax avoidance versus illegal tax evasion) affects individuals' tax minimization behavior under different conditions (risk-free, risk and penalties, priming). This study could be seen as a prime example for the interface between behavioral and public economics. Based on a neoclassical model, legality simply does not matter. Including moral costs into this model leads to different predictions. The results confirm these predictions: legality is important under certain conditions. The second study reveals that moral evaluation of tax evasion is subject to a self-serving bias, which could explain why previous studies find lower tax morale among self-employed individuals (Alm and Torgler 2006). Again, this effect is not addressed by a neoclassical approach. The third study discloses that the timing of pension taxation (immediate versus deferred) influences individuals' perceived fairness of work payment due to mental accounting and that this influences individuals' work effort. Moreover, deferred taxation decreases risk taking. From a neoclassical perspective, the timing of pension taxation should not influence the work effort and risk taking in this setting because the timing does not affect individuals' wealth. The fourth study examines whether democratic participation and earmarking a tax influences real effort and legal tax avoidance. Both, participation and earmarking do not influence individuals' payouts and thus, are again wealth-independent. Nevertheless, earmarking and democratic participation lowers legal tax avoidance but does not affect labor supply.

Summing up, the four studies reveal that different tax policies affect individuals' behavior beyond the rationality assumption. Table 1 summarizes these four studies in terms of tax policy, response variables and observed individuals' responses. The first two studies and, at least partly, the fourth study are related to the broad area of tax compliance research whereas the last two studies are related to the research on labor supply. All four studies focus on

individual behavior in experimental settings, in which hypothetical effects of various instruments of tax policies on the individual behavior are examined.

Table 1: Summary of studies related to the behavioral economics research

Study	Tax policy	Response variables	Individuals' responses
First study *Legality*	Defining the borderline between legality and illegality	Tax Minimization	Depending on financial consequences and priming: strong response in the absence of negative financial consequences, weaker response when negative financial consequences are present and priming is used
Second study *Self-serving bias*	Access to tax evasion	Moral evaluation of illegal tax minimization	Tax evasion is judged less unethical when individuals have the opportunity to evade
Third study *Pension taxation*	Timing of taxation	Perceived fairness, real effort, risk taking	Deferred taxation increases perceived fairness of work payment, which indirectly affects work effort. Deferred taxation decreases risk taking
Fourth study *Voting and earmarking*	Earmarking and democratic participation	Legal tax minimization, real effort	Earmarking and democratic participation lowers tax avoidance, real effort is not affected

The fifth study, which forms the second part of this thesis, is related to the interface of business taxation and accounting. This interface has become increasingly important in recent years (Graham et al. 2012). Based on tax related financial statement positions (e.g., deferred taxes, income taxes, uncertain tax positions), researchers try to gain empirical evidence for a variety of research questions. One of the main research area in this context deals with the question of whether and how managers use tax accounts for earnings management (e.g., in terms of income increasing management, smoothing, meeting analysts' forecast, or avoiding losses). Thus, tax accounts may exhibit incremental information content on the earnings quality for stakeholders, primarily investors. Another important area is research on corporate tax avoidance. Since tax return information is rare, the measurement of tax avoidance is often based on financial statements. Research in tax avoidance is concerned with the measurement, the determinants, and the consequences of corporate tax avoidance (Hanlon and Heitzman 2010, p. 137). The study presented in this thesis examines whether large book-tax differences, which are calculated on the basis of deferred taxes, exhibit incremental information content regarding the persistence of pre-tax earnings, which is a common proxy for earnings quality. Moreover, this study also analyses whether large book-tax differences likely arise due to earnings management, tax avoidance or changes in tax loss carryforwards and how these different sources affect the earnings quality proxy.

The overall link between the first four studies and the fifth study is, rather generally spoken, empirical tax research. Taxes influence a variety of decisions ranging from individual real effort to corporate earnings management. Nevertheless, a less general link is also present, and that is tax avoidance. The last study as well as the majority of the other studies consider tax

avoidance. Whereas the behavioral economics studies examine determinants (e.g, legality, voting and earmarking) of individual tax avoidance, the last study deals, inter alia, with the identification of corporate tax avoidance. Factors that determine individual tax avoidance often apply to corporate tax avoidance as well (Hanlon and Heitzman 2010, p. 138). For instance, the *legality* of a potential tax avoidance strategy might affect manager's decision on implementing such a strategy. Nevertheless, likely arising agency problems complicate a direct attribution of the behavioral economics results, which could be addressed by future research. In the following, the studies of this thesis are presented in more detail.

Does legality matter? The case of tax avoidance and evasion

The legality or illegality of an action is an explicit way for policy makers to affect the social acceptance of this action (expressive function of law, Sunstein 1996). Law expresses social values and legality may act as a reference point when individuals rationalize their decisions (Cooter 1998, 2000). The first study contribute to this theoretical approach and empirically test whether declaring tax minimization as legal or illegal affects individual decisions. Several prior studies compare tax with non-tax situations (e.g., Alm et al. 1992, Durham et al. 2014) but there is very little empirical research on the effect of legality on tax minimization behavior (Kirchler et al. 2003, Bobek and Hatfield 2003, Dwenger et al. 2015). However, there is no empirical evidence on the effect of legality on real tax minimization decisions. This study contribute to the literature on the role of legality in individual decision making processes by examining empirically how and when legality affects an individual's tax minimization decisions. Therefore, three real effort experiments are conducted.

In the first experiment, tax minimization behavior is compared in the absence of any detection or penalty risk. Theoretically, the illegal opportunity should causes higher moral costs than the legal opportunity does. In line with that theory, the study reveal that labeling a tax minimization opportunity as unambiguously illegal results in significantly less tax minimization compared to labeling tax minimization as unambiguously legal. More generally, this finding is consistent with the expressive function of law. Declaring an action as illegal affects behavior even if the illegal action is not penalized.

Outside the lab, tax evasion is typically associated with positive detection and penalty risk. Moreover, due to tax law ambiguity, tax avoidance also bears the risk that the revenue agency will assess an additional income tax payment and corresponding interest charges upon audit. Therefore, a second experiment is conducted, in which detection risk, negative detection consequences (penalties in the case of evasion and interest charges in the case of avoidance) and implicit monitoring are considered. The prior difference between legal and illegal tax minimization is no longer observable. There are four possible explanations for this effect:

- Participants in the evasion treatment could decide to use only a small fraction of the maximum income concealment to maintain their positive self-concept (Mazar et al. 2008). However, the vast majority of participants in the first experiment evaded either nothing or they evaded the full amount of six sheets. This still holds when penalties and detection risk are introduced.
- The line between legal and illegal behavior could be blurred due to detection risk and penalties. An additional survey reveal that the difference in subjects' moral evaluations

between illegal and legal tax minimization is not much affected by the introduction of risk.

- Risk and penalties could increase the cognitive load of taxpayers (Dohmen et al. 2010). This could reduce the importance of intrinsic preferences for obeying the law and reduce the effectiveness of injunctive norms (Kredentser et al. 2012, Dwenger et al. 2015).
- Penalties and the implicit introduction of monitoring could undermine intrinsic motivation (Gneezy and Rustichini 2000, Fehr and Falk 2002, Falk and Kosfeld 2006).

The third experiment addresses the crowding out and cognitive load explanations by introducing moral priming. Consistent with the argument that moral priming reduces crowding out and reinforces the legality effect, a legality effect is observable in the third experiment.

Taken together, the series of experiments shows that legality can have strong effects on individuals' behavior. In line with the expressive law approach, defining the borderline between legality and illegality can be used to affect moral costs. In recent years, the tax authorities seem to have been trying to shift the classical line between avoidance and evasion (Friese et al. 2008). However, the risk of negative financial consequences could suppress the legality effect especially when subjects have low tax morale.

Self-serving bias and tax morale

The second study analyze whether moral evaluations of tax evasion are egoistically biased. Several researchers have emphasized the importance of tax morale in explaining observed tax evasion (e.g., Baldry 1986, Cummings et al. 2009). Tax morale is a special form of solidarity behavior with a low level of social interaction and interpersonal coordination (Brosig-Koch et al. 2011). It is often described as the intrinsic willingness to pay taxes (Alm and Torgler 2006, Russo 2013). The study empirically tests whether tax morale arises independently of individuals' economic situation or whether there is self-serving bias. In a real-effort experiment, subjects are randomized to treatments with and without tax minimization opportunities. It is shown that individuals' tax morale is subject to a self-serving bias. Individuals with the opportunity to evade taxes consider tax evasion less unethical compared to those without this opportunity. The results are robust to different detection probabilities and economically similar but legal avoidance opportunities. Spillover effects of evasion opportunities on other moral evaluations (e.g., bribery or lying) are not observable.

Mental accounting and the timing of pension taxation

Does the timing of taxation affect the behavior of individuals even if it does not affect their wealth? While the question is of general interest, it is especially important in the field of pension taxation. Most OECD countries use tax incentives to encourage private retirement savings plans. Some countries do not tax interest earned in savings plans funded with after-tax contributions (e.g., Hungary and Luxembourg), whereas other countries defer payroll taxation on the contributions until the pensions are paid out (e.g., Austria, the Netherlands and the UK). The study concentrates on the question of whether the timing of taxation (immediate taxation versus deferred taxation) affects individual labor supply and investment risk decisions. From a neoclassical view, given a time-constant tax rate, there is no reason why the

timing of taxation should affect the behavior of individuals as both alternatives are equivalent in present value terms. The equivalence of deferred and immediate taxation is based on the assumption that subjects perceive the total tax burden accurately. However, there is increasing evidence that many subjects misperceive taxes (e.g., Sausgruber and Tyran 2005, Chetty et al. 2009, Blumkin et al. 2012, Fochmann and Weimann 2013, Fochmann et al. 2013, Blaufus et al. 2013).

With respect to the equivalence between immediate and deferred taxation, the third study proposes that mental accounting (Thaler 1985, 1990, 1999) leads to a deviation from the "neutrality result." Based on this concept, some subjects should assign their wealth decision to two separate mental accounts (a work account and an investment account) rather than determining their wealth based on one aggregated decision. Thus, work and investment decisions should differ between immediate and deferred taxation.

A real-effort experiment is conducted to test the prediction. In a first step, subjects perform a work task lasting one hour. In a second step, subjects invest their earned money. They make five lottery decisions, and in each decision, they choose between two lotteries that differ in risk but not in expected returns. As a result, mental accounting seems to be important. Subjects in the deferred tax treatment perceive their wage as significantly more fair and this perception (indirectly) increases working effort. Moreover, subjects in the deferred tax treatment make less risky investment. Thus, the presumed neutrality regarding the timing of taxation does not hold. This is of relevance to the current debate in some countries such as the United Kingdom over a change from deferred pension to immediate pension taxation.

Do tax earmarking and voting on the tax rate influence tax avoidance and labor supply in the lab?

In November 2012, voters in California passed the ballot measure Proposition 30 by a 54% to 46% vote.[3] From a tax research standpoint, it would be interesting to know whether these additional taxes caused behavioral responses similar to those caused by other taxes, taking into account that the tax revenues are earmarked and that taxpayers had a voice regarding their implementation.

The fourth study addresses this question by conducting a real-effort experiment. The tax revenue is transferred to the faculty library (the earmarking treatment). Democratic participation rights are modelled as a voting on the tax rate. The results reveal that earmarking significantly lowers tax avoidance. Democratic participation reduces tax avoidance even further. Furthermore, the individual tax rate vote affects the subsequent avoidance decision. Those who voted for a high tax rate avoid significantly fewer taxes. With respect to labor supply, any significant impact of earmarking and participation rights is not observable. This result suggests that a feature of a tax could influence different forms of taxpayers' behavioral responses in different ways. Taxpayers' avoidance responses seem to be much more elastic than their real responses.

[3] This measure temporarily raised the state sales tax and income tax rates and earmarked the resulting revenues for education. These revenues have subsequently allowed the California State University network to avoid a 9% tuition hike.

The information content of large book-tax differences – Empirical evidence from Germany

Previous empirical research dealing with the information content of large book-tax differences (BTDs), which include primary temporary and permanent differences between book income and taxable income, shows mixed results. On the one hand, a relatively clear negative relation between large temporary book-tax differences (tBTDs) and earnings persistence is revealed (for instance, Hanlon 2005; Blaylock et al. 2012; Tang and Firth 2012; Krummet 2011). On the other hand, Jackson (2015) finds little evidence that large positive tBTDs are associated with upward earnings management. This contradicts the results presented by Blaylock et al. (2012), who find that large positive tBTDs are positively correlated with upward earnings management. Moreover, the low earnings quality is likely influenced by factors other than large tBTDs, including profitability, size, age, and large transitory items (Graham et al. 2012), and that the calculation of tBTDs is distorted by the recognition of net operating loss carryforwards (NOLCs), even when controlling for NOLCs (Guenther 2011). Furthermore, it is theoretically unclear why firm-years with large negative tBTDs contain lower earnings persistence, as shown by Hanlon (2005) and Blaylock et al. (2012).

The purpose of the fifth study is to examine the association between large positive and large negative tBTDs and earnings persistence for German public companies. Prior research is extended by precisely controlling for the deferred tax expense that is caused by NOLCs. The NOLCs-portion of deferred tax expense does not reflect income measurement differences between book and taxable income, which are the basis for this and also the related information content studies.

As a result, lower earnings persistence when firms exhibit large positive tBTDs is not observable. This also holds for firms with large positive tBTDs that are identified as (presumed) upward earnings managing firms according to several approaches of the Jones model. Thus, the results differ from the results presented by Hanlon (2005) and Blaylock et al. (2012) and confirm rather critical prior literature, e.g., Guenther (2011). With respect to tax avoidance as a potential explanation for large positive tBTDs, it is shown that 27.9% of the firm-years that exhibit large positive tBTDs are categorized as tax avoiders compared to 16.8% (18.6%) of firm-years that exhibit large negative tBTDs (small tBTDs). This indicates that large positive tBTDs can be used as a signal of tax avoidance.

In contrast, earnings persistence is relatively low when firms exhibit large negative tBTDs. Although this is in line with prior literature, it is unclear what causes this relation. Hand-collected disclosure information for influential observations reveals that various accounts affect the level of tBTDs. Thus, for the majority of cases, a plain or obvious relation between these high negative tBTDs and earnings persistence is not observable. The low earnings persistence mainly results from direct and indirect consequences of the 2008 crisis, changing market conditions, and a database misspecification. Because these are rather external effects, the study concludes that the observed negative relation can be explained by internal management activities only for a small part of the observations.

2 Does legality matter? The case of tax avoidance and evasion

Kay Blaufus, Jochen Hundsdoerfer, Martin Jacob, Matthias Sünwoldt

Die Studie wurde 2016 im *Journal of Economic Behavior & Organization* (Vol. 127, S. 182-206) veröffentlicht.

doi:10.1016/j.jebo.2016.04.002

Der Artikel ist in der Online-Version aus Gründen des Urheberrechts nicht enthalten.

3 Self-serving bias and tax morale

Kay Blaufus, Jochen Hundsdoerfer, Martin Jacob, Matthias Sünwoldt

Die Studie wurde 2015 in der Fachzeitschrift *Economics Letters* (Vol. 131 (C), S. 91-93) veröffentlicht.

doi:10.1016/j.econlet.2015.03.041

Der Artikel ist in der Online-Version aus Gründen des Urheberrechts nicht enthalten.

.

4 Mental accounting and the timing of pension taxation

Kay Blaufus, Jochen Hundsdoerfer, Matthias Sünwoldt, Nadja Wolf

Abstract

We study whether the timing of pension taxation influences work effort and risk taking decisions. In a real effort experiment, participants first earn money within one hour and then invest the money earned by making five independent portfolio decisions. Participants in the immediate taxation treatment must pay taxes on their wages, but their invested pension capital and the respective returns are tax-exempt, whereas participants in the deferred taxation treatment do not need to pay taxes on their wages but pay taxes on their withdrawal of the invested pension capital and returns. After-tax payoffs are equal in the two tax systems. However, as a result of mental accounting, we expect and find that participants in the deferred taxation treatment perceive their work payment to be significantly fairer and that the perceived fairness significantly influences work effort. Moreover, we find that deferred taxation decreases risk taking.

Keywords

Mental Accounting · Tax Perception · Fairness · Behavioral Taxation · Deferred Taxation

JEL Classification

H24 · H31

We thank Ido Erev, Martin Fochmann, Martin Jacob, Florian Morath, and the participants at the 2015 Conference of the German Association of Experimental Economics, the 2015 Closing Workshop on "Coping with Difficult Decisions - An Experimental Economics Perspective," the 2016 WHU Accounting & Tax Research Seminar, and the 2016 ZEW Workshop on "Lab Experiments in Public Economics." Matthias Sünwoldt gratefully acknowledges financial support from KPMG.

4.1 Introduction

Does the timing of taxation affect the behavior of individuals even if it does not affect their wealth? We answer this question affirmatively. As a result of mental accounting, the timing of taxation is important even in the absence of wealth differences between alternatives.

While this question is of general interest, it is especially important in the field of pension taxation. Most OECD countries use tax incentives to encourage private retirement savings plans (Yoo and de Serres 2005). Some countries do not tax interests earned in savings plans funded with after-tax contributions, whereas other countries defer payroll taxation on contributions until pensions are paid out. The first method implies immediate taxation of wages but no taxation of saving returns (TEE, i.e., taxable contributions, exempt accumulations and exempt withdrawals). This method has been implemented in countries such as Hungary, Luxembourg, and the U.S. (Roth individual retirement arrangements and Roth 401(k) plans) and has recently been proposed in the UK (HM Treasury 2015). The second method implies deferred taxation of wages and savings returns (EET, i.e., exempt contributions, exempt accumulations and taxable withdrawals), and it is currently offered in countries such as Austria, the Netherlands, Norway, the UK, and the U.S. (401(k) plans).

We concentrate on the question of whether the timing of taxation (immediate taxation or TEE vs. deferred taxation or EET) affects individual labor supply and investment risk decisions. Suppose that a subject earns € 100 and wants to invest in a one-year pension plan. Under an immediate taxation scheme (TEE), the subject pays taxes on her wage income but does not pay taxes on future withdrawals, including interest income. Assuming an interest rate of 10% and a tax rate of 60%, a total of € 40 remains for her investment, and she receives € 44 after one year. Under deferred taxation (EET), wage income is tax free (i.e., contributions are tax deductible), but withdrawals from the pension plan are fully taxed. In this case, investing € 100 in the pension plan leads to a withdrawal of € 110 after one year and € 44 after taxes. Thus, the timing of taxation (immediate or deferred) does not affect her wealth. From a neoclassical perspective, given a time-constant tax rate, the timing of taxation should not affect individual behavior because both alternatives are equivalent in present value terms.

This "neutrality result" has significantly influenced tax research and the evaluation of tax policy options. This view has been considered in the discussion of different pension tax systems (e.g., Yoo and de Serres 2005, Huang 2008). Furthermore, it is used to study tax effects in a variety of contexts, such as the choice between present and future consumption, between lifetime and testamentary gifts, between retention and distribution of corporate earnings, between receiving or deferring income, and between different forms of doing business (Warren 1986). Despite the importance of this equivalence between immediate and deferred taxation for tax policy and research, its empirical validity is an open issue. Only recently has a direct test been offered by Beshears et al. (2015). However, their results contradict the neoclassical neutrality assumption regarding the timing of taxation, as they find that future after-tax pensions will differ between immediate and deferred pension tax systems. Their findings suggest that taxpayers generally neglect taxes and thus do not adjust their contributions to pension plans according to different tax treatments. This behavior in turn leads to different expected pensions in immediate and deferred tax systems. The current paper

adds to the discussion of the economic equivalence of immediate and deferred pension taxation by studying whether the timing of taxation influences subjects' work effort and risk taking decisions.

The equivalence of deferred and immediate taxation is based on the assumption that subjects accurately estimate their total tax burden. However, increasing evidence suggests that many subjects misperceive taxes because they do not pay attention to less salient taxes (e.g., Sausgruber and Tyran 2005, Chetty et al. 2009, Blumkin et al. 2012, Fochmann and Weimann 2013), because they focus on pre-tax-values instead of after-tax returns (Fochmann et al. 2013 and Weber and Schramm 2016), because they use simple decision heuristics (Blaufus et al. 2013) or because taxes induce negative emotions that affect decision making (Blaufus and Möhlmann 2014).

With respect to the equivalence between immediate and deferred taxation, we propose that mental accounting (Thaler 1985) leads to a deviation from the "neutrality result." Prior research provides evidence that mental accounting is of relevance in the tax context: a tax refund delivered in monthly amounts increases current spending more than the same yearly total tax reduction delivered in one lump sum (Chambers and Spencer 2008), and a tax decrease implemented gradually over several years leads to a greater increase in risky investment than a tax change implemented all at once (Falsetta et al. 2013). In the context of life-cycle consumption, Thaler (1990) argues that individuals tend to use mental accounts when evaluating savings and consumptions. He considers three broad accounts when categorizing types of wealth: a current income account, an asset account and a future income account. Based on the fungibility of money, the marginal propensity to consume should be equal for all three accounts. These mental accounts could explain some of the observed saving and consumption anomalies, such as the tendency for consumption to respond too sensitively to current income. Individuals seem to evaluate the different mental accounts separately. Considering this concept, we expect that some subjects assign their wealth decisions to two separate mental accounts (a work account and an investment account) rather than determining their wealth based on one aggregated decision. If these individuals are subject to deferred (immediate) taxation, they will make work decisions as if wages are tax free (subject to tax) and make investment decisions as if returns are fully taxable (tax exempt). Thus, we expect that work and investment decisions differ between immediate and deferred taxation.

Despite the usual concerns about external validity, a laboratory experiment has obvious advantages in answering our research question. First, the neutrality between deferred and immediate pension taxation requires that present and future tax rates are equal and that future tax rates are known with certainty. This requirement can easily be met only in an experimental situation. Second, productivity and work performance often may be only roughly estimated using archival data. Hence, to test the hypotheses, we conduct a real-effort laboratory experiment with subjects randomly assigned to either an immediate or deferred taxation treatment. In the first step, subjects perform a work task lasting one hour, and we measure their work effort. In the second step, subjects invest their earned money. They make five lottery decisions, and in each decision, they choose between two lotteries that differ in risk but not in expected returns.

We find that mental accounting is important. In line with the mental accounting hypothesis, subjects in the deferred taxation treatment perceive their wage as significantly more fair, and this perception (indirectly) increases work effort. Moreover, subjects in the deferred taxation treatment make less risky investments. Thus, the presumed neutrality regarding the timing of taxation does not hold. This finding is of relevance to the current debate in countries such as the United Kingdom regarding a change from deferred pension to immediate pension taxation. Policy makers who decide between deferred and immediate taxation should consider that neoclassical predictions could be misleading. Whereas neoclassical economics would predict the same tax revenues (in present value terms), our results suggest the tax revenues differ because of the different behavioral effects of immediate and deferred taxation. In addition to observing direct effects on after-tax pensions (Beshears et al. 2015), we demonstrate that a change may cause unexpected effects on work effort and risk taking.

4.2 Hypotheses development

4.2.1 Work effort

Given a world where W denotes the wage that is used for saving purposes, i represents the market interest rate that applies to the savings over period n; τ denotes the investor's tax rate, which is constant over time; and the investor's wealth WE_i under immediate taxation is given by the following equation:

$$WE_i = W(1-\tau)(1+i)^n. \quad (1)$$

Under immediate taxation, the wage income that is used for savings is subject to taxes. However, the return from these savings is tax exempt. In contrast, under deferred taxation, the wage income that is used for savings remains tax free, but the return from savings is fully taxable. Thus, deferred taxation leads to the following equation for the investor's wealth:

$$WE_d = W(1+i)^n(1-\tau). \quad (2)$$

Both tax systems lead to the same wealth if we assume equal wages, the same time horizon, and the same tax and interest rates ($WE_i = WE_d$). Thus, according to neoclassical economics, working and savings behavior are unaffected by the timing of taxation. An overview of the two tax systems is presented in Table 9.

Table 2: Overview of the taxation system

	Deferred Taxation	**Immediate Taxation**
Work stage	Work: *no tax*	Work & *taxation*
Investment stage	Investment & *taxation*	Investment: *no tax*
	Equal present values	

Mental accounting could distort this wealth neutrality. The concept of mental accounts is based on the idea that individuals evaluate financial activities comparable to the managerial and financial accounting that is used by firms and other organizations (Thaler 1985 and Thaler 1999). Mental accounting consists of three main components (Thaler 1999): perception of outcomes and evaluation of decisions, frequency of evaluation or "balancing," and assignment of financial activities. The last component is particularly relevant for our

experimental setting. In mental accounting, similar to financial accounting, individuals label expenditures and funds as different accounts based on sources and uses (e.g., regular income vs. windfall). As a consequence, the fungibility of money, introduced by the account-based budgeting, could be violated (for experimental evidence, see Heath and Soll 1996). In our experimental design, we do not directly refer to the fungibility problem, but we argue that subjects in our experiment use two different accounts (a work income account and an investment income account) because we assume that subjects separate their income based on income sources (work income vs. investment income). If we consider that subjects use these different mental accounts, the above derived "neutrality result" may no longer hold. Under deferred taxation, subjects might perceive their wages as tax free because they neglect the indirect, less salient deferred taxation. In contrast, under immediate taxation, subjects might perceive the wage income as fully taxable, neglecting the tax advantage of tax-exempt investment income. Thus, subjects who use different mental accounts for work and investment income may perceive the wage rate to be higher under deferred taxation than under immediate taxation.

A higher perceived wage rate should lead to an increase in work effort[4] (see, for example, Fehr and Goette 2007) as leisure becomes less valuable (*substitution effect*). Thus, we would expect participants in the deferred taxation treatment to work harder than those in the immediate taxation treatment. However, if subjects use mental accounting, an *income effect* might also occur, and reduced effort will be the result. This *income effect* will arise if subjects have a certain amount in mind that they are striving to earn. When they fulfill their goal, subjects will exert less work effort. Given the perceived higher wage rate, subjects in the deferred taxation treatment should exert less work effort. Hence, either the opposing effects, the *income* and *substitution effects*, offset one another, or one predominates the other. Depending on which effect predominates, we formulate our first hypothesis as follows:

Hypothesis 1: Work effort differs between deferred and immediate taxation.

If we actually find that work effort differs between treatments, this result would contradict the neoclassical neutrality assumption and support the mental accounting hypothesis. By contrast, if we do not find any difference in work effort, this does not necessarily imply that subjects make decisions in line with neoclassical assumptions. Alternatively, subjects could use mental accounting although income and substitution effects offset one another.

To directly test the mental accounting hypothesis, we formulate a second hypothesis: subjects who actually use different mental accounts for wage and investment income should perceive the wage rate in the immediate tax system as less fair than the wage rate under deferred taxation. This difference results from the work mental accounting that occurs during the work task. If subjects think only of their work mental account, they encounter a wage rate of € 0.80 (after taxes) in the immediate taxation treatment but a rate of € 2.00 (tax-free) in the deferred taxation treatment. Thus, subjects should perceive the wage rate to be ever higher under deferred taxation if mental accounting is used. We formulate Hypothesis 2a accordingly:

[4] In our experiment, subjects decide not on the hours to work but on the effort to expend. As subjects decide between work effort and on-the-job leisure (Dickinson 1999), we may only measure work effort within that time span.

Hypothesis 2a: *The perceived fairness of wages is higher under the deferred taxation of wages.*

The impact of perceptions of a payment's fairness on work effort has been broadly examined in the literature. In the context of gift exchanges and reciprocity, Akerlof (1982), Akerlof (1984), Fehr et al. (1998), Riedl and Tyran (2005) and Dohmen et al. (2009) find that a payment's fairness is positively correlated with work effort. Hence, the fairer subjects feel a payment is, the more effort they will exert in their labor. Thus, we formulate the second hypothesis as follows:

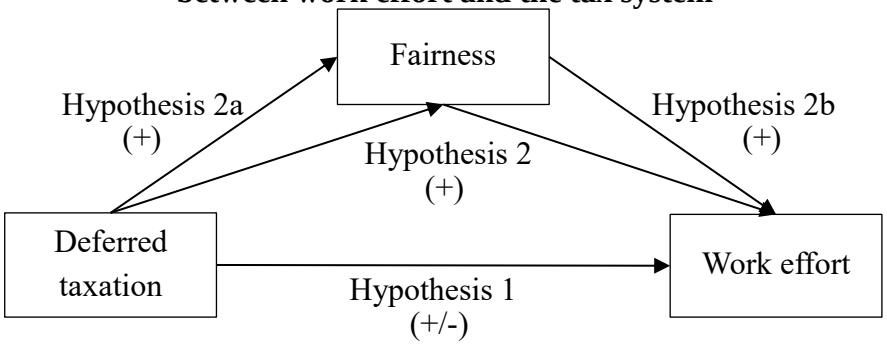
Hypothesis 2b: *The higher the perceived fairness of wages is, the greater the work effort is.*

Summarizing Hypotheses 2a and 2b, we expect participants in the deferred taxation treatment to engage in mental accounting and perceive their work task payment fairer and therefore offer greater work effort. Thus, if we consider perceived fairness to be a mediator variable, we can specify our first hypothesis with regard to an indirect effect:

Hypothesis 2: *Deferred taxation indirectly increases work effort because the perceived fairness of wages operates as a mediator variable.*

The general theoretical model underpinning the testing of these hypotheses is illustrated in Figure 9. We expect to find a direct effect of the deferred tax system on work effort (Hypothesis 1). Moreover, fairness mediates the relationship between the tax system and work effort. Under the deferred taxation system, the perceived fairness of a payment is higher than the perceived fairness under the immediate tax system (Hypothesis 2a). Fairness and work effort are positively correlated: an increase in perceived fairness will improve work effort (Hypothesis 2b). Thus, we expect to find a positive indirect effect of deferred taxation on work effort (Hypothesis 2).

Figure 2: Theoretical model of fairness mediating the relationship between work effort and the tax system



4.2.2 Risk taking

Rational choice theory would also predict no differences in subjects' investment decisions.[5] However, mental accounting entails a perception of full taxation (tax exemption) of investment returns under deferred (immediate) taxation, as subjects consider only their mental investment account. Thus, we again expect differences between the two tax systems due to mental accounting. First, subjects in the deferred tax system have more to invest because their income has not yet been taxed. This perceived *income effect* may lead to higher or lower levels of risk taking or may not affect risk taking at all depending on the subjects' risk preferences. If one assumes constant relative risk aversion on average (in line with the experimental results of Chiappori and Paiella 2011), one would expect no effect. However, if one assumes decreasing relative risk aversion (in line with many experimental and empirical studies, e.g., Levy 1994, Calvet and Sodini 2014), one would expect greater risk taking under deferred taxation.

Second, both investment alternatives differ solely in the variance of returns. Whereas immediate taxes have already reduced the investment amount, taxes in the deferred taxation treatment reduce the variance of returns. Although after-tax payoffs are equal, the risk in the deferred tax treatment might be perceived as lower, which could increase risk taking under deferred taxation.

Third, prior behavioral research finds that taxation results in lower risk taking because of the additional complexity induced by taxes (Ackermann et al. 2013). If subjects engage in mental accounting, they will consider only their investment account when making portfolio decisions. Regarding the mental investment account, only subjects under deferred taxation face taxes at the investment stage, as the returns and invested capital for subjects under immediate taxation are tax exempt. Thus, the taxation of returns and invested capital in the deferred taxation treatment might cause an increase in complexity, which would lower risk taking behavior.

In sum, mental accounting leads to a difference in risk taking behavior between deferred and immediate taxation. However, the theoretical direction is unclear and is thus an empirical question.[6] We therefore test the following hypothesis:

Hypothesis 3: Levels of risk taking differ between deferred and immediate taxation.

4.3 Experimental protocol and sample

4.3.1 Experimental protocol

We use a between-subjects design with the timing of taxation (immediate versus deferred) as the treatment variable.[7] Subjects are randomly assigned to the two treatment groups. The

[5] Note that the taxation of investment returns and the amount invested in the deferred tax treatment are comparable with a wealth tax in the classical framework of Atkinson and Stiglitz (1980): future wealth is reduced because of taxation, and this leads to less risk taking than in a situation without taxation (assuming decreasing relative risk aversion). However, in contrast to Atkinson and Stiglitz (1980), we compare the deferred tax system with an immediate tax system, both of which offer identical future wealth.

[6] Note that a rational individual will not change her risk-taking behavior between the treatments.

[7] Our immediate versus deferred taxation design does not consider the concept of the "time value of money TVM" because this concept does not fit our experimental approach. The participants receive one net payment at the end of the experiment. Thus, there is no time effect at all. For a similar approach, see Falsetta et al.

experiment was conducted in eight sessions at the computerized experimental laboratory of the Leibniz University of Hanover. The z-Tree software was used (Fischbacher 2007). We present a translation of the instructions and screenshots of the experiment in Appendices A1 and A2.

The experiment is divided into two parts: a work task and portfolio decisions. To keep our results general, we do not frame the experiment as a pension tax task. We use a work task for two reasons. First, we are able to examine the influence of the taxation system on work effort. Second, we prevent the occurrence of the *house money effect* that might increase the participants' risk taking if they do not invest money earned previously that occurred as a windfall.[8] The work task lasts one hour. During this time, the participants are asked to digitize answer sheets for a multiple-choice exam. Each answer sheet consists of 40 rows (questions) with 6 possible columns (answers). The participants are asked to correctly transfer each checked box by clicking the respective check box on the computer screen. This work task offers two advantages: On the one hand, it is largely independent of the participants' education and abilities; thus, all participants have the opportunity to earn the same money. On the other hand, the correctness of the digitized sheets is automatically controlled, thus enabling the payment to be dependent on the participants' accuracy. The participants are granted one practicing period to familiarize themselves with the work's design and task. However, the practicing period is not relevant to the payout.

Before each digitalization, the participants enter a four-digit number that identifies a certain answering scheme. We use twelve different answering schemes, each comprising 60 different four-digit numbers. Based on the answering sheet's number, the computer checks the correctness of the work task. Only accurately transferred answering sheets are paid. After each sheet, the participant is told about the correctness of her work and the amount of money she has already earned. For each correct sheet, the participant earns € 2 (before taxes).[9] A countdown projected on the front wall informs the participants how much time is left for work. During the work task, the participants are allowed to surf the internet. Thus, we offer them an on-the-job leisure alternative if they want to pause or end their work task (Dickinson 1999).

The two treatments differ only regarding the time of taxation.[10] In the immediate taxation treatment, the participants are told that their gross wage is taxed at a rate of 60%.[11] Thus, they earn a net wage of € 0.80 per correctly transferred sheet. However, their subsequent portfolio decisions have tax-free returns. In the deferred taxation treatment, the participants are told that their wage is tax free but that they must pay taxes at a rate of 60% on both the returns of their portfolio decisions and the invested capital. This summarizing taxation information is presented in the first paragraph of the second part of the instructions (see Appendix A1.2).

(2013). We argue that the time effect, which occurs in practice, would rather strengthen our result because the probability of different mental accounts would increase because of time effects.

8 For a literature overview of the house money effect, see Clark (2002) as well as Weber and Zuchel (2005).

9 Tax revenues from this experiment are not distributed among participants. Instead, they are used for further experimental research at the experimental laboratory of the Leibniz University of Hanover.

10 Note that all subjects are informed about both parts of the experiment (work and investment task) and the respecting tax treatments before the experiment starts (see Figure 10 and Appendix A1 for the instructions).

11 We decided for this rather high tax rate to increase the tax salience.

The taxation information is then given again for the working and investment periods separately. To ensure that the salience regarding taxation is the same between treatments, we include the same type of information. More precisely, we inform the participants on the taxation of their wage (tax free versus 60% tax rate on gross wages) and their investment (tax free versus 60% tax rate on invested capital and returns). Thus, both treatments lead to the same tax burden and after-tax payoff, but they differ in the timing of taxation. After the work task ends, the investment phase starts. At the beginning of the investment phase, the participants are given an overview of their work results (i.e., how many correct sheets they have entered). The immediate taxation treatment group is also informed about the gross and net wage (in euro) and about withholding taxes, whereas the deferred taxation treatment group only receives additional information on their wage, as taxes are not withheld at that time.

In the investment phase, the participants successively confront five portfolio decisions. In each of the decisions, the participant is presented with two alternative independent lotteries, with one lottery always riskier than the other, but both have the same expected value. The participants in the immediate taxation treatment (deferred taxation treatment) are asked to distribute their entire net wage (wage) between the two lotteries in each portfolio decision. Within each lottery, three different states may occur with a probability of one-third. The lotteries are presented in Table 10 with the riskier lottery on the left side.

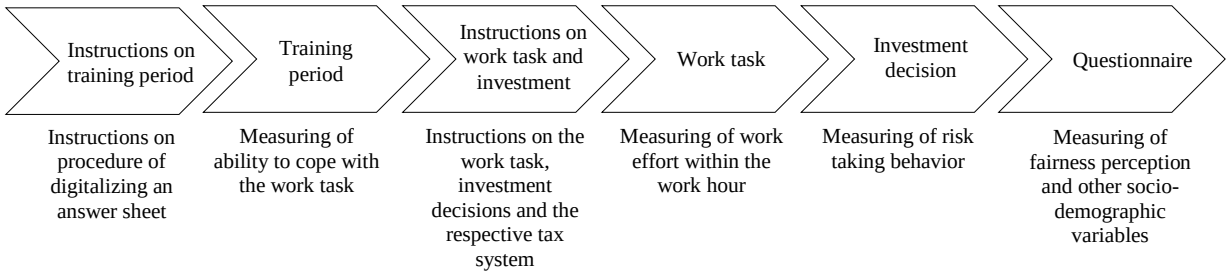
Table 3: Overview of portfolio decision lotteries

Decision period		Riskier lottery			Less risky lottery		
	State	1	2	3	1	2	3
	Probability	1/3	1/3	1/3	1/3	1/3	1/3
1	Rate of return	60%	30%	0%	40%	30%	20%
2		70%	40%	10%	50%	40%	30%
3		50%	30%	10%	30%	30%	30%
4		40%	20%	0%	30%	20%	10%
5		60%	30%	0%	40%	30%	20%

The portfolio decisions occur in this fixed order. However, the presentation of the riskier lottery and less risky lottery varies. Thus, in the second and fifth decision periods, the riskier lottery is presented on the monitor's right-hand side, but in all other periods, it is presented on the left-hand side. The participants in the immediate taxation treatment are informed that the return of their portfolio decision is tax free. In contrast, the participants in the deferred taxation treatment are informed that a tax with a rate of 60% on the return and invested amount is withheld after their decisions.

After the investment phase, the participants are asked to answer a questionnaire that collects socio-demographic data. A translated version of the questionnaire is presented in Appendix A3. For payout purposes, only one portfolio decision is relevant. To determine the relevant return rate, the participant must throw the dice twice. The first throw decides on the decision period, and the second determines the state of environment. The participants are successively and separately paid out cash, whereby the payment is rounded up to the next ten cents. A timeline of the experimental procedure is illustrated in Figure 10.

Figure 3: Timeline of experimental protocol



4.3.2 Sample

A total of 121 students (49 females and 72 males) participated in the experiment. The subjects were 23.4 years on average, and 46.3% studied either in the Faculty of Economics and Management or in the Faculty of Philosophy.[12] The subjects earned € 19.78 on average in approximately 120 minutes (approximately € 9.89 per hour), with a range from € 9.80 to € 46.40. Table 11 provides an overview of the main characteristics. The results reveal no significant differences in the individual characteristics between the treatments.

Table 4: Descriptive statistics for individual characteristics
SOCIAL SCIENCE displays whether a subject either studies in the Faculty of Economics and Management or in the Faculty of Philosophy.

	Mean	Median	Standard deviation
MALE	59.50%		
SOCIAL SCIENCE	46.28%		
AGE (years)	23.42	23.00	3.38

4.4 Results

4.4.1 Work effort

4.4.1.1 Bivariate analyses

We start our analysis by examining the impact of the tax system on work effort (Hypothesis 1). The variable WORK EFFORT is measured by the number of answer sheets that the subjects correctly transfer to the computer. The descriptive statistics for this variable are displayed separately for the immediate taxation and deferred taxation treatments in Table 12. On average, 19.3 correct sheets were digitized in the immediate taxation treatment, whereas 18.6 correct sheets were transferred in the deferred taxation treatment. This difference, however, is not significant (Mann-Whitney U, $p = 0.714$). Thus, we cannot confirm the first hypothesis that deferred taxation affects work effort. However, as noted previously, this result does not inevitably mean that subjects do not use mental accounting. Rather, subjects may use mental accounting, while income, substitution, and fairness effects offset one another.

[12] All other subjects studied in the Faculty of Mathematics and Physics, the Faculty of Natural Sciences, the Faculty of Engineering, the Faculty of Electrical Engineering and Informatics, the Faculty of Architecture, or the Faculty of Law.

Table 5: Overview of work effort categorized by treatment

The table presents key figures of the variable WORK EFFORT. It depicts the number of correctly digitized answer sheets within the work hour and is displayed for each taxation treatment separately. We use a two-tailed Mann-Whitney U test to examine whether there is a significant difference between the treatments.

Treatment	WORK EFFORT	
	Immediate Taxation	Deferred Taxation
Mean	19.30	18.60
Median	19.00	19.00
Standard deviation	5.66	4.51
Minimum	11	10
Maximum	33	29
No. of subjects	61	60
Mann-Whitney U test	p = 0.714	

To test the use of mental accounts, we collect data on perceptions of payment fairness in the post-experimental questionnaire (Hypothesis 2a).[13] Using a 10-point scale, subjects were asked to report the perceived fairness of the payment (1 = not fair at all and 10 = totally fair). Table 13 reveals a significant difference in the perceptions of payment fairness between the immediate and deferred taxation treatments. The mean level of perceived fairness is 5.02 in the immediate taxation treatment and 6.52 in the deferred taxation treatment. By using a two-tailed Mann-Whitney U test, we find that this difference is strongly significant at a 1% level. Thus, the participants who must pay taxes on their wages perceive the payment as less fair than those who pay taxes at a later point (i.e., when deciding on their investment portfolio). This result demonstrates that the subjects use different mental accounts for their wage income and their investment income. By ignoring the deferred taxation of their wage income in the investment phase, the subjects in the deferred tax treatment seem to evaluate the fairness of their payment using the pre-tax payment (€ 2), whereas the subjects in the immediate tax treatment use the after-tax payment (€ 0.80). Hence, we can confirm Hypothesis 2a, which states that mental accounting causes the perceived fairness of wages to be higher under deferred taxation of wages.

[13] We openly asked the participants at the end of our experiment (after they decided on the work effort and risky investment). Thus, the participants were aware of the taxation system not only because of the given instructions but also because of their experiences during the actual experiment.

Table 6: The impact of deferred taxation on fairness perception

The table presents key figures of the perceptions of payment fairness. FAIRNESS displays the individual's self-reported satisfaction of the work task's payment measured on a 10-point scale where 1 = not fair at all and 10 = totally fair. We use a two-tailed Mann-Whitney U test to examine whether there is a significant difference between the treatments.

Treatment	FAIRNESS	
	Immediate taxation	Deferred taxation
Mean	5.02	6.52
Median	5	7
Standard deviation	2.19	2.41
Minimum	1	1
Maximum	10	10
No. of subjects	61	60
Mann-Whitney U test	p = 0.0008	

Next, we analyze how perceived fairness affects work effort (Hypothesis 2b) in order to investigate the effect of the tax system on work effort by examining the indirect effects via perceived fairness. Regarding the fairness perception, we use a median split to divide subjects into two groups. All subjects who report the perceived fairness of the work's payment to be below the median of the total sample (i.e., less than or equal to 5, measured on a 10-point scale where 1 = not fair at all and 10 = totally fair) are categorized as LOW FAIRNESS. All subjects who report at least the median level of perceived fairness (i.e., at least 6 points) are categorized as HIGH FAIRNESS. Using these two groups, we can analyze whether fairness perceptions significantly influence work effort. An overview of the work effort within fairness groups and the respective bivariate test is given in Table 14.

Table 14 depicts work effort measured as the number of correctly transferred sheets and categorized according to perceived fairness. Whereas the LOW FAIRNESS group digitizes 18.14 sheets on average within the work hour, the HIGH FAIRNESS group digitizes 19.70 sheets on average. The difference is significantly different from zero ($p = 0.061$). Hence, we can conclude that participants who perceive a payment as less fair tend to work less. Conversely, we can confirm Hypothesis 2b, which states that the greater the perceived fairness is, the greater the work effort is. The results regarding Hypotheses 2a and 2b point to an indirect effect of the timing of taxation (Hypothesis 2). Under deferred taxation, subjects perceive their payment to be fairer, and this increased fairness perception leads to greater work effort. In Section 4.4.1.2, we examine whether this indirect effect is significant.

Table 7: The impact of perceived fairness on work effort

WORK EFFORT depicts the number of correctly digitalized sheets within the work hour. It is analyzed for two binary categories of FAIRNESS. Subjects are allocated to the group LOW FAIRNESS if they report the perceived fairness to be less or equal than 5 and to the group HIGH FAIRNESS in all other cases. We use a two-tailed Mann-Whitney U test to examine whether there is a significant difference between the low and high fairness group.

	WORK EFFORT	
Fairness perception	LOW FAIRNESS	HIGH FAIRNESS
Mean	18.14	19.70
Median	17	20
Standard deviation	5.20	4.95
Minimum	11	10
Maximum	33	32
No. of subjects	58	63
Mann-Whitney U test	p = 0.061	

4.4.1.2 Mediation analysis

In the previous section, we demonstrate how the timing of taxation affects perceived fairness and show that stronger perceptions of fairness increase work effort (confirming Hypotheses 2a and 2b). However, in contrast to Hypothesis 1, we do not yet find a direct effect of the tax system on work effort. Using a structural equation model (SEM), we test whether the tax system indirectly affects work effort if perceived fairness operates as a mediator variable, apart from a possible direct effect of the tax system on work effort. Assuming sequential ignorability (Imai et al. 2013), we estimate the following two equations:

$$EFFORT = \alpha_{01} + \alpha_{11} DEFERRED + \alpha_{21} FAIRNESS + \varepsilon_{01} \quad (3)$$

$$FAIRNESS = \alpha_{02} + \alpha_{12} DEFERRED + \varepsilon_{02} \quad (4)$$

The widely used traditional, but rather conceptual, mediation analysis is based on Baron and Kenny (1986). A formal test related to Alwin and Hauser (1975) calculates the mediated or indirect effect by multiplying both direct effects: the direct effect of the tax treatment on fairness perception (coefficient α_{12}) and the direct effect of fairness perception on work effort (coefficient α_{21}). The calculation of asymptotic standard errors of the indirect effect is based on the multivariate delta method (Bishop et al. 1975, Sobel 1982, 1986). Statistical significance is derived through a comparison of the indirect effect divided by the asymptotic standard errors to a standard normal distribution (see MacKinnon et al. 2007 for a review).

To estimate the simple linear SEM in one analysis, we use Stata 14 based on maximum likelihood and the Stata command “estat teffects” to decompose into direct, indirect, and total effects, again based on the delta method (Stata Corp 2015 p. 157, Sobel 1987). The results are identical to the traditional approach.

Figure 4: A structural equation model linking the tax system treatment, perceived fairness, and work effort

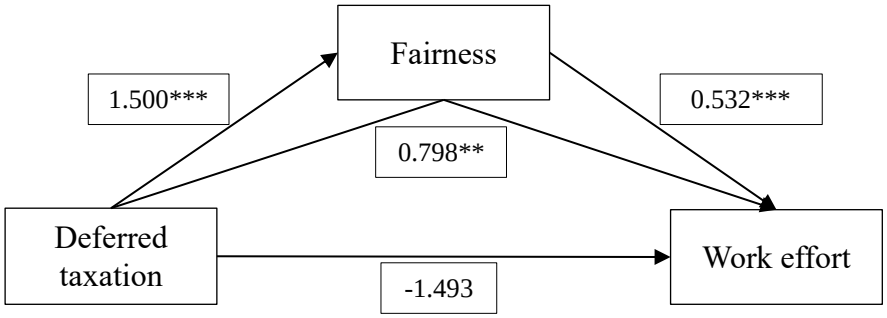


Table 8: Mediation analysis

WORK EFFORT displays the number of correctly digitized answer sheets within the work hour. FAIRNESS displays the individual's self-reported satisfaction of the work task's payment measured on a 10-point scale where 1 = not fair at all and 10 = totally fair. DEFERRED TAXATION is a dummy variable equal to one if the observation belongs to the deferred taxation treatment. The structural equation model includes equations (3) and (4). The indirect effect is calculated by multiplying both direct effects, i.e., the direct effect of the tax treatment on the fairness perception and the direct effect of the fairness perception on the work effort. The total effect is the sum of the direct and the indirect effect. Standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1.

	FAIRNESS	WORK EFFORT		
		Direct	Indirect	Total
DEFERRED TAXATION	1.500*** (0.416)	-1.493 (0.944)	0.798** (0.368)	-0.695 (0.924)
FAIRNESS		0.532*** (0.196)		
Constant	3.516*** (0.655)	18.120*** (1.574)		
No. of observations	121	121	121	121
R-squared	0.116		0.062	
Wald Chi2	13.04***		7.95**	

Figure 11 and Table 15 display the results. As already presented in the previous section, the tax system has a strongly significant influence on perceived fairness ($p < 0.001$): as a result of mental accounting, subjects perceive the payment to be fairer than under deferred taxation than under immediate taxation (Hypothesis 2a). Fairness itself has a significant impact ($p = 0.007$) on work effort, as greater fairness increases effort (Hypothesis 2b). We observe a significant indirect effect of the timing of taxation on work effort, with fairness as a mediator ($p = 0.030$). Hence, we can confirm Hypothesis 2, which states that the deferred taxation treatment indirectly increases work effort as a result of mental accounting because the perceived fairness of wages operates as a mediator variable. We find neither a significant direct effect of deferred taxation on work effort ($p = 0.114$) nor a total effect as the sum of direct and indirect effects ($p = 0.452$). Thus, the SEM also fails to confirm Hypothesis 1.

Table 9: Mediation analysis including socio-demographic variables and ability

WORK EFFORT displays the number of correctly digitized answer sheets within the work hour. FAIRNESS displays the individual's self-reported satisfaction of the work task's payment measured on a 10-point scale where 1 = not fair at all and 10 = totally fair. DEFERRED TAXATION is a dummy variable equal to one if the observation belongs to the deferred taxation treatment. AGE displays participants' age measured in years. MALE is a dummy variable equal to one if the participant is male. SOCIAL SCIENCE displays whether a subject either studies at the faculty of economics or at the faculty of philosophy. ABILITY displays participants' required time to pass the practicing period. The structural equation model includes equations (3) and (4). The indirect effect is calculated by multiplying both direct effects, i.e., the direct effect of the tax treatment on the fairness perception and the direct effect of the fairness perception on the work effort. The total effect is the sum of the direct and the indirect effect. Standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1.

	FAIRNESS	WORK EFFORT		
		Direct	Indirect	Total
DEFERRED TAXATION	1.700*** (0.423)	-0.595 (0.926)	0.672* (0.359)	0.077 (0.886)
FAIRNESS		0.396** (0.187)		
AGE	0.035 (0.062)	0.003 (0.127)	0.014 (0.025)	0.017 (0.129)
GENDER	0.300 (0.477)	-0.312 (0.870)	0.119 (0.176)	-0.194 (0.868)
SOCIAL SCIENCE	0.101 (0.005)	-0.864 (0.852)	0.040 (0.165)	-0.824 (0.868)
ABILITY	0.009** (0.005)	0.041*** (0.010)	0.004 (0.003)	0.045*** (0.010)
Constant	4.211** (1.891)	26.513*** (3.966		
No. of observations	121	121	121	121
R-squared	0.266		0.186	
Wald Chi2	27.63***		19.04***	

To test whether these results are robust, we include socio-demographic control variables such as age (AGE), gender (MALE), faculty (SOCIAL SCIENCE), and ABILITY. ABILITY controls for the inherent and undistorted ability of the participant to perform the work task. ABILITY is measured as the time that participants need to finish the practicing period, multiplied by -1. Thus, a high value of ABILITY indicates a greater subject-specific ability to cope with the work task. Because the practicing period was conducted before the detailed treatment instructions were distributed (see Section 4.3.1), we ensure that the tax system does not affect the control variable ABILITY. We extend equations (3) and (4) and include AGE, MALE, SOCIAL SCIENCE and ABILITY as observed exogenous variables for our endogenous variables WORK EFFORT and FAIRNESS. Table 16 displays the results, which show that the previously reported effects remain the same. Additionally, we observe significant and plausible positive effects of ability on work effort (p-value < 0.001) as well as on perceived fairness (p-value = 0.041). None of the other control variables significantly affect perceived fairness or work effort.

4.4.2 Risk taking

4.4.2.1 Bivariate analyses

In the following section, we examine whether the timing of taxation influences investment decisions (Hypothesis 3). As shown in Table 17, the results reveal no linear relation between the tax system and risk taking. We observe a significant effect only if we consider decisions with high risk (i.e., decisions in which the percentage invested in the riskier asset exceeds a certain threshold, e.g., 2/3 or 3/4). In the deferred tax system, the subjects make significantly less decisions with high risk (Mann-Whitney U, p = 0.002).

Table 10: The impact of deferred taxation on risk taking

The table presents key figures of the risk taking variable under different conditions. The number of subjects in the Immediate (Deferred) Taxation Treatment amounts to 61 (60). Risk taking (continuous) denotes the average percentage invested in the riskier lottery, whereas risk taking (risky share > 50% [60%] {75%}) takes the value 1 if the subject invests at least 50% [60%] {75%} into the riskier lottery and zero otherwise. The minimum (maximum) of all risk taking variables amounts to zero (one). We use two-tailed Mann-Whitney U tests to examine whether there is a significant difference between the treatments.

Risk taking variable	Treatment	Mean	Median	SD	p-value
Risk taking (continuous)	Immediate	0.47	0.30	0.30	0.871
	Deferred	0.46	0.5	0.24	
Risk taking (risky share > 50%)	Immediate	0.35	0	0.48	0.372
	Deferred	0.31	0	0.46	
Risk taking (risky share > 66%)	Immediate	0.22	0	0.42	0.002
	Deferred	0.13	0	0.33	
Risk taking (risky share > 75%)	Immediate	0.19	0	0.39	0.002
	Deferred	0.10	0	0.30	

4.4.2.2 Panel analysis

Participants made five independent investment decisions in which they were asked to distribute their earned money into two lotteries with different levels of risk. To exploit the panel structure of our data and to meet the requirement of our left- and right-censored variable, we use the continuous RISK TAKING variable[14] and run a random-effects tobit panel regression. We do not find a significant treatment effect (p = 0.633) independent of whether we control for socio-demographic variables such as age, gender, and faculty. Thus, we cannot confirm Hypothesis 3 when analyzing investment behavior with the continuous variable.[15]

Therefore, we again examine risk taking behavior by analyzing decisions with high risk based on the variable HIGH RISK TAKING, which indicated whether a subject invests at least 75% of the income earned to the riskier lottery.[16] We use the variable HIGH RISK TAKING as a

[14] By dividing the amount invested in the riskier lottery by the total earned amount, we calculate the share of income that was invested more riskily.

[15] However, if we consider only risk-taking decisions with positive contributions in the riskier lottery, we find a significant treatment effect on risk taking.

[16] We also created the HIGH RISK TAKING variable with a 66% investment in the riskier lottery as a threshold. All results remain qualitatively unchanged but are not reported here.

dependent variable and run random-effects logit panel regressions.[17] The variable HIGH RISK TAKING is a dummy variable equal to one if the subject invests at least 75% of her income into the riskier asset in period t (with t ranging from 1 to 5).

Table 18 displays two logit regressions run separately that include an increasing number of independent variables. Model 1 tests the influence of the tax system on the subjects' risk taking behavior. Model 2 additionally controls for the socio-demographic variables age (AGE), gender (MALE) and faculty (SOCIAL SCIENCE). Additionally, we include the time needed for the investment decision for each period, DECISION TIME$_t$, as well as whether the riskier lottery was displayed on the monitor's left-hand side LEFT SIDE$_t$ as control variables.

Table 11: Random-effects logit regressions (dependent variable: HIGH RISK TAKING$_t$)

HIGH RISK-TAKING$_t$ is a dummy variable equal to one if the subject invests at least 75% of her income into the riskier asset in period t (with t = 1 to 5). DEFERRED TAXATION is a dummy variable equal to one if the observation belongs to the deferred taxation treatment. AGE displays participants' age measured in years. MALE is a dummy variable equal to one if the participant is male. SOCIAL SCIENCE denotes whether a subject either studies at the faculty of economics or at the faculty of philosophy. DECISION TIME$_t$ displays the required time for the investment decision in period t. LEFT SIDE$_t$ is a dummy variable and takes the value 1 if the riskier lottery is presented on the monitors' left-hand side. Standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1.

	HIGH RISK TAKING$_t$ (1)	HIGH RISK TAKING$_t$ (2)
DEFERRED TAXATION	-1.245**	-1.246**
	(0.568)	(0.541)
AGE		-0.026
		(0.080)
MALE		1.593***
		(0.572)
SOCIAL SCIENCE		-0.707
		(0.535)
DECISION TIME$_t$		0.001
		(0.002)
LEFT SIDE$_t$		0.272
		(0.308)
CONSTANT	-1.197	-1.354
	(0.836)	(2.177)
No. of observations	605	605
No. of subjects	121	121
Prob > chi2	0.028	0.034

The logit regressions in Table 18 demonstrate that the tax system significantly influences risk taking behavior in both models (p = 0.028 and p = 0.023, respectively). Under deferred taxation, participants take less risk in investing (fewer decisions are made to invest at least 75% in the riskier lottery) than they do under the immediate tax treatment. Thus, when the tax is already levied directly on the income earned, people make more risky investments, whereas their investments are less risky when the income earned is tax free but the return and invested

[17] We find the same results when using a random-effects probit regression.

amount are taxed. This result supports Hypothesis 3. Model 2 reveals that, in line with prior research (Croson and Gneezy 2009), we find a significant influence of MALE on risk taking ($p = 0.005$). On average, male subjects make one more risky investment (i.e., investing at least 75% in the riskier lottery) than females in all five lotteries.

4.4.2.3 Diversification heuristics

The result that deferred taxation leads to lower risk taking corresponds to the behavioral argument that taxation leads to lower risk taking because of the additional complexity induced by taxes (Ackermann et al. 2013). If subjects use mental accounting, they will notice that their returns and the invested amount are taxed in the deferred taxation treatment. In the immediate taxation treatment, however, the taxation occurs only in the work stage, such that subjects are not confronted with any further taxes in the investment stage. Thus, the investment decision is more complex under deferred taxation because taxes are levied there.

Higher complexity could also lead to an increase in the amount of time needed to make an investment decision. However, our results show that deferred taxation does not significantly extend the decision time.[18] This surprising result suggests that subjects might use decision heuristics to reduce the complexity in the environment. In line with the complexity argument, we expect that subjects more frequently use simple heuristics in the deferred tax treatment. We present the distribution of RISK TAKING in Figure 12.

As shown in Figure 12, risk taking behavior differs between the two treatments. In the immediate tax treatment, the share of total investment decisions that is invested fully in the less risky lottery or fully in the riskier lottery (RISK TAKING = 0 or RISK TAKING = 1) is significantly higher than that in the deferred taxation treatment ($p < 0.001$, chi2 test). Accordingly, the subjects in the deferred taxation treatment diversify more between the two lotteries than the subjects in the immediate taxation treatment. Note that in the used lottery setting it is not possible to minimize risk with diversification.[19] Nevertheless, the subjects seem to prefer diversification. This result is consistent with prior findings on the use of a diversification heuristic. The subjects display a desire for variety that leads to more diversity than they actually want (Read and Loewenstein 1995, Hedesstrom et al. 2004). In particular, the number of decisions that distribute the earned income evenly between both lotteries is higher in the deferred taxation treatment.

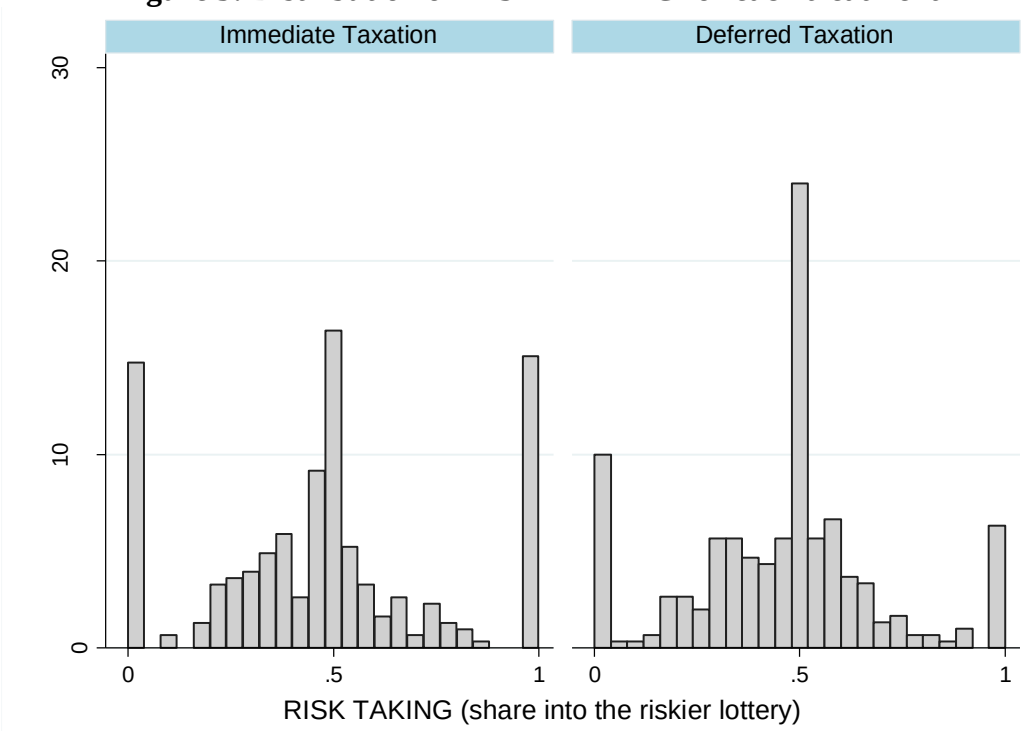
We observe that 12.5% of the investment decisions made in the immediate taxation treatment are based on an even split of the income between the two assets. In contrast, 24% of the investment decisions in the deferred taxation treatment are based on this even distribution

[18] In the immediate taxation treatment, the subjects need 53.7 seconds on average to make a single investment decision, whereas the subjects need 55.9 seconds on average in the deferred taxation treatment. We perform a two-tailed Mann-Whitney U test to examine whether the decision time between the tax treatments differs significantly ($p = 0.614$).

[19] Assuming either a concave or convex expected utility function, risk seeking subjects should always invest 100% in the riskier lottery, whereas risk averse subjects should invest 100% in the less risky lottery. If a subject rather shows a risk utility function that is increasing in small variances but decreasing in larger variances, an interior solution may be optimal. (For a justification of Friedman-Savage utility functions even in perfect capital markets, see Hartley and Farrell 2002). An interior solution can further be optimal if a subject demands a minimum rate of return above the worst outcome of the riskier lottery but, apart from that, seeks risk. However, these interior solutions do not inhibit diversification characteristics.

($p < 0.001$, chi2 test and Fisher's exact test). We use the aggregated application of the even split heuristic to test whether we observe statistically significant differences between the two tax system treatments. HEURISTIC represents the individual's heuristic behavior measured on a 6-point scale, where 0 = the subject never uses the even split heuristic and 5 = the subject uses this heuristic in all 5 periods. A two-tailed Mann-Whitney U test reveals a significant difference between the two tax system groups (p-value = 0.021). Moreover, we use the panel structure and run random-effects logit regressions with $HEURISTIC_t$ as an independent variable (a dummy variable equal to one if the subject uses the even split heuristic in period t, with t ranging from 1 to 5; p-value = 0.013). The higher tax complexity in the deferred taxation treatment appears to be accompanied by a significant increase in the use of the diversification heuristic, especially the even split heuristic (see Benartzi and Thaler 2001 for the widespread use of the 1/n heuristic). We interpret the greater application of these special heuristics as a proxy for higher complexity and argue that this increased complexity leads to a lower propensity for risk taking in the deferred taxation treatment.

Figure 5: Distribution of RISK TAKING for each treatment



4.4.2.4 Panel regressions with regard to diversification heuristics

In the previous section, we demonstrated that subjects in the immediate taxation system invest their total earned money in the less risky lottery significantly more often because they do not use the diversification heuristic as frequently as subjects in the deferred tax treatment. If we exclude all decisions in which the share invested in the riskier lottery is zero (RISK TAKING = 0) and run the random-effects tobit panel regression again, we find that risk taking differs significantly between the two treatments. These results are illustrated in Table 19. We find that risk taking is significantly greater in the immediate taxation treatment. Again, we additionally find that male subjects show a higher propensity for risk taking than females do.

Table 12: Random-effects tobit regressions (dependent variable: RISK TAKING$_t$)

RISK TAKING$_t$ is the percentage invested in the riskier lottery in period t (with t = 1 to 5). Only observation where RISK TAKING$_t$ >0 are included. DEFERRED TAXATION is a dummy variable equal to one if the observation belongs to the deferred taxation treatment. AGE displays participants' age measured in years. MALE is a dummy variable equal to one if the participant is male. SOCIAL SCIENCE denotes whether a subject either studies at the faculty of economics or at the faculty of philosophy. DECISION TIME$_t$ displays the required time for the investment decision in period t. LEFT SIDE$_t$ is a dummy variable and takes the value 1 if the riskier lottery is presented on the monitors left-hand side. Standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1.

	RISK TAKING$_t$ (1)	RISK TAKING$_t$ (2)
DEFERRED	-0.055*	-0.067**
TAXATION	(0.033)	(0.032)
AGE		-0.005
		(0.005)
MALE		0.095***
		(0.032)
SOCIAL SCIENCE		-0.038
		(0.032)
DECISION TIME$_t$		0.000
		(0.000)
LEFT SIDE$_t$		0.015
		(0.015)
CONSTANT	0.620***	0.715***
	(0.518)	(0.131)
No. of observations	530	530
No. of subjects	118	118
Prob > chi2	0.094	0.023

4.5 Robustness checks

In this section, we subject our results to a series of robustness checks. In particular, we study the development of work effort over time, examine whether DEFERRED TAXATION indirectly affects RISK TAKING (moderated by EFFORT), and investigate whether our results are affected by subjects misunderstanding the instructions.

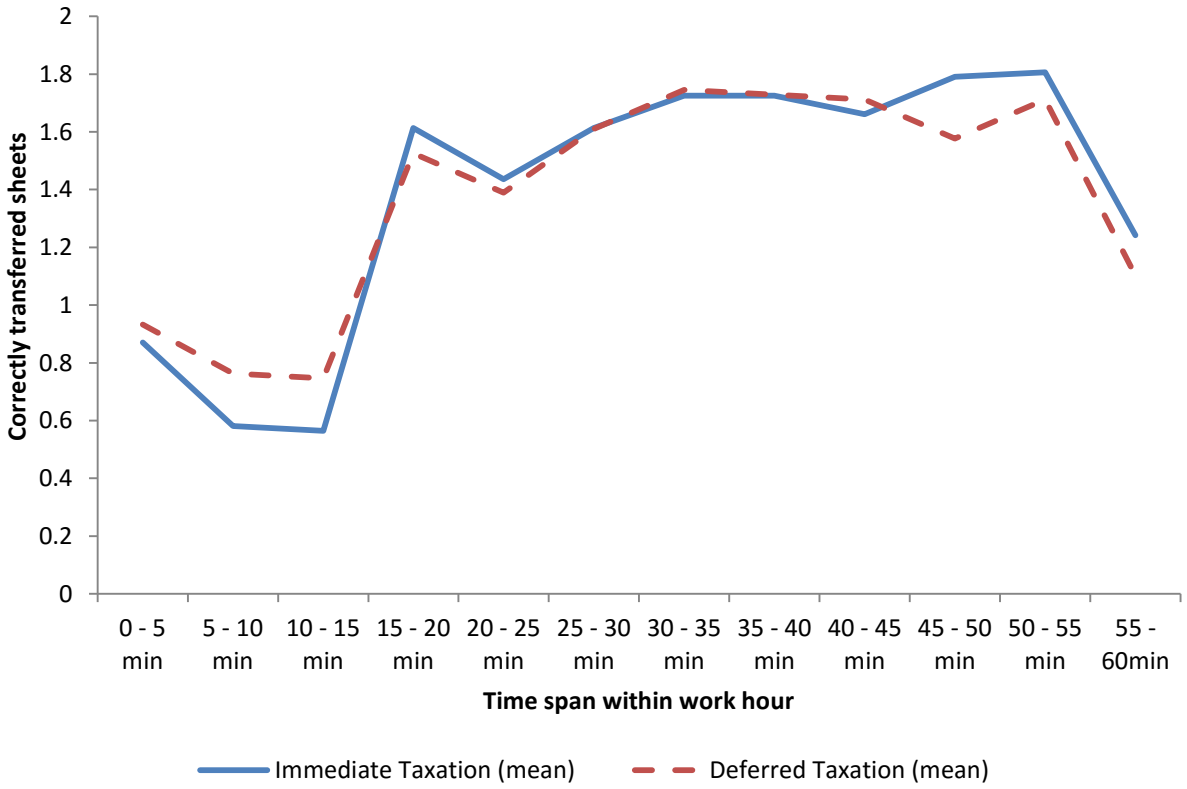
4.5.1 Work effort over time

Thus far, we have analyzed the impact of deferred taxation on work effort by examining the number of correctly transferred sheets within one work hour. However, the tax system might not only influence the output as total work effort but also trigger productivity effects over time. Thus, subjects might start working hard much earlier under one tax system, whereas they might work longer under the other system. Therefore, to control for those effects, we analyze productivity over time.

First, we cluster work time into time spans of 5 minutes, thus generating twelve points of time at which we measure the number of correctly transferred sheets. Second, we determine the number of sheets per time span. The results are illustrated in Figure 13 and separated by tax system. The findings reveal only minor differences in the number of digitized sheets between the treatments that are never significant (analyzed with a Mann-Whitney U test). Thus, we can

conclude that deferred taxation does not directly influence productivity at a certain point within the work hour.

Figure 6: Number of correctly transferred sheets over time



4.5.2 Structural equation modelling including risk taking

In the next step, we integrate the investment decision into the SEM to account for work effort, which might also influence risk taking beyond the timing of taxation. The existing literature reveals that risk taking behavior is likely influenced by a house money effect. In this context, effort could negatively influence risk taking (Arkes et al. 1994). The main results of the SEM are presented in Table 20. To integrate high risk taking in the SEM, we generate a binary variable that takes the value 1 if the subject invests at least 75% of her income in the riskier lottery and 0 otherwise for all five decision periods. Subsequently, we sum these variables over all periods to create the variable HIGH RISK TAKING. Depending on the single decision periods, HIGH RISK TAKING can take values from 0 (i.e., a subject never invests at least 75% of her income in the riskier asset) to 5 (i.e., a subject invests at least 75% of her income in the riskier asset in all five decision periods). Note that we do not use ABILITY as control variable for HIGH RISK TAKING, as the ability to cope with the work task correlates with work effort. We observe a significant negative effect of the tax system treatment on high risk taking ($p = 0.014$). Hence, we again confirm Hypothesis 3, as subjects under deferred taxation make less risky investments because of the tax-induced complexity. Moreover, we find that work effort negatively influences risk taking behavior (p-value = 0.073), which is

consistent with prior findings: greater effort leads to an increase in risk aversion.[20] The results of the SEM's work effort analysis remain unchanged.

Table 13: Mediation analysis including socio-demographic variables, ability, and risk taking behavior

WORK EFFORT displays the number of correctly digitized answer sheets within the work hour. FAIRNESS displays the individual's self-reported satisfaction of the work task's payment measured on a 10-point scale where 1 = not fair at all and 10 = totally fair. HIGH RISK-TAKING$_t$ is a dummy variable equal to one if the subject invests at least 75% of her income into the riskier asset in period t (with t = 1 to 5). DEFERRED TAXATION is a dummy variable equal to one if the observation belongs to the deferred taxation treatment. AGE displays participants' age measured in years. MALE is a dummy variable equal to one if the participant is male. SOCIAL SCIENCE displays whether a subject either studies at the faculty of economics or at the faculty of philosophy. ABILITY displays participants' required time to pass the practicing period. The structural equation model includes equations (3) and (4). The indirect effect is calculated by multiplying both direct effects, i.e., the direct effect of the tax treatment on the fairness perception and the direct effect of the fairness perception on the work effort. The total effect is the sum of the direct and the indirect effect. Standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1.

	FAIRNESS	WORK EFFORT			HIGH RISK TAKING
		Direct	Indirect	Total	
DEFERRED	1.700***	-0.595	0.672*	0.077	-0.525**
TAXATION	(0.423)	(0.926)	(0.359)	(0.886)	(0.214)
FAIRNESS		0.396**			
		(0.187)			
WORK EFFORT					-0.037*
					(0.021)
AGE	0.035	0.003	0.014	0.017	-0.009
	(0.062)	(0.127)	(0.025)	(0.129)	(0.032)
GENDER	0.300	-0.312	0.119	-0.194	0.679***
	(0.477)	(0.870)	(0.176)	(0.868)	(0.214)
SOCIAL SCIENCE	0.101	-0.864	0.040	-0.824	-0.303
	(0.005)	(0.852)	(0.165)	(0.868)	(0.986)
ABILITY	0.009**	0.041***	0.004	0.045***	
	(0.005)	(0.010)	(0.003)	(0.010)	
Constant	4.211**	26.513***			2.154**
	(1.891)	(3.966			(0.986)
No. of observations	121	121	121	121	121
R-squared	0.266		0.186		0.130
Wald Chi2	27.63***		19.04***		18.10***

4.5.3 Instructions check

In the post-experimental questionnaire, we ask the participants in the deferred (immediate) taxation treatment to state the amount of their wage after (before) taxation per sheet, and we ask those in the deferred (immediate) taxation treatment to report the tax rate with respect to the taxation of the invested amount and the return (wage). Eleven participants fail to answer both questions correctly. To conduct a robustness test, we run SEM regressions and random-effects logit regressions that include a further control variable named CHECK. This variable is a dummy variable equal to one if the subject answers both questions correctly and zero

[20] Considering the context of tax evasion decisions, Kirchler et al. (2009) find that subjects evade more money if only low effort was required to earn it. Hence, as a tax evasion decision can be tantamount to a risk-taking decision, we should expect subjects who work harder to take less risk.

otherwise. Additionally, we exclude the participants who fail to answer the questions correctly and re-run all regression models. The results remain qualitatively unchanged. We observe a highly significant effect of the tax system treatment on perceived fairness, a significant direct effect of fairness on work effort and a significant indirect effect of the tax system treatment on work effort via fairness as a mediator. The direct effect of deferred taxation on work effort remains non-significant. Moreover, we find that deferred taxation leads to a significantly lower amount of risky investments (measured as 75% dummy variable and 66% dummy variable).

4.6 Conclusion

Using a real-effort laboratory experiment, we study whether the immediate or deferred taxation of compulsory pension contributions affects labor and investment decisions. In contrast to neoclassical predictions but in line with the assumption that individuals use different mental accounts for their work income and their investment income, we find that deferred taxation of wage income results in greater perceptions of fairness of the wage payment. This perception in turn increases participants' work effort. Thus, the timing of taxation (immediate versus deferred taxation) indirectly affects work effort. Furthermore, also consistent with the mental accounting hypothesis, the results show that risk taking is lower under deferred taxation.

Our study contributes to the behavioral tax research that emphasizes the importance of incorporating psychological insights into the economic analysis of tax policy. Our study reveals that the assumption of neutrality regarding the timing of taxation does not hold. This finding has consequences for many tax policy issues. In the field of pension taxation, deferred and immediate taxation compete. Under immediate taxation (TEE), all income earned is taxed in the respective year, even if it is paid into a pension plan. However, the withdrawal is untaxed. In contrast, under deferred taxation (EET), contributions are deductible, but withdrawals are fully taxed. Although prior (neoclassical) research regards both pension tax systems as equivalent, our findings suggest that policy makers should consider that both pension tax systems might lead to different tax revenues because behavioral responses differ between the two systems as a result of mental accounting. While our study investigates only the effect on work effort and risk taking, one might expect also differences concerning the effect on subjects' savings. If subjects use different mental accounts for their current and future income, they may not pay sufficient attention to the full taxation of their future income under a deferred tax system. This oversight could lead to less savings (and thus lower after-tax pensions) in a deferred tax system compared with an immediate tax system, as observed by Beshears et al. (2015). Moreover, future research could examine the effects of mental accounting in other tax policy fields, such as dividend and capital gain taxation, where neoclassical predictions also use the neutrality concept in considering the timing of taxation.

References

Ackermann, H., Fochmann, M., & Mihm, B. 2013. Biased effects of taxes and subsidies on portfolio choices. *Economics Letters*, 120(1), 23-26.

Akerlof, G. A. 1982. Labor Contracts as Partial Gift Exchange. *Quarterly Journal of Economics*, 97(4), 543-569.

Akerlof, G. A. 1984. Gift Exchange and Efficiency-Wage Theory: Four Views. *American Economic Review*, 74(2), 79-83.

Alwin, D. F., & Hauser, R. M. 1975. The Decomposition of Effects in Path Analysis. *American Sociological Review*, 40(1), 37-47.

Arkes, H. R., Joyner, C. A., Pezzo, M. V., Nash, J. G., Siegel-Jacobs, K., & Stone, E. 1994. The Psychology of Windfall Gains. *Organizational Behavior and Human Decision Processes*, 59(3), 331-347.

Atkinson, A. A., & Stiglitz, J. E. 1980. Lectures on Public Economics. London: McGraw-Hill.

Baron, R. M., & Kenny, D. A. 1986. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173-1182.

Benartzi, S., & Thaler, R. H. 2001. Naive diversification strategies in defined contribution saving plans. *American Economic Review*, 91(1) 79-98.

Beshears, J., Choi, J. J., Laibson, D., & Madrian, B. C. 2015. Does front-loading taxation increase savings? Evidence from Roth 401 (k) introductions. *Journal of Public Economics*, forthcoming.

Bishop, Y. M. M., Fienberg, S. E., & Holland, P. W. 1975. Discrete Multivariate Analysis: Theory and Practice. Cambridge.

Blaufus, K., Bob, J., Hundsdoerfer, J., Kiesewetter, D., & Weimann, J. 2013. Decision heuristics and tax perception – An analysis of a tax-cut-cum-base-broadening policy. *Journal of Economic Psychology*, 35, 1-16.

Blaufus, K., & Möhlmann, A. 2014. Security Returns and Tax Aversion Bias: Behavioral Responses to Tax Labels. *Journal of Behavioral Finance*, 15(1), 56-69.

Blumkin, T., Ruffle, B. J., & Ganun, Y. 2012. Are income and consumption taxes ever really equivalent? Evidence from a real-effort experiment with real goods. *European Economic Review*, 56(6), 1200-1219.

Calvet, L. E., & Sodini, P. 2014. Twin Picks: Disentangling the Determinants of Risk-Taking in Household Portfolios, *Journal of Finance*, 69(2), 867-906.

Chambers, V., & Spencer, M. 2008. Does changing the timing of a yearly individual tax refund change the amount spent vs. saved? *Journal of Economic Psychology*, 29(6), 856-862.

Chetty, R., Looney, A., & Kroft, K. 2009. Salience and Taxation: Theory and Evidence. *American Economic Review*, 99(4), 1145-1177.

Chiappori, P., & Paiella, M. 2011. Relative Risk Aversion Is Constant: Evidence from Panel Data. *Journal of the European Economic Association*, 9(6), 1021-1052.

Clark, J. 2002. House money effects in public good experiments. *Experimental Economics*, 5(3), 223-231.

Croson, R., & Gneezy, U. 2009. Gender differences in preferences. *Journal of Economic Literature*, 47(2), 448-474.

Dickinson, D. L. 1999. An experimental examination of labor supply and work intensities. *Journal of Labor Economics*, 17(4,1), 638-670.

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. 2009. Homo Reciprocans: Survey Evidence on Behavioural Outcomes. *The Economic Journal*, 119(536), 592-612.

Falsetta, D., Rupert, T. J., & Wright, A. M. 2013. The effect of the timing and direction of capital gain tax changes on investment in risky assets. *The Accounting Review*, 88(2), 499-520.

Fehr, E., & Goette, L. 2007. Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment. *American Economic Review*, 97(1), 298-317.

Fehr, E., Kirchsteiger, G., & Riedl, A. 1998. Gift exchange and reciprocity in competitive experimental markets. *European Economic Review*, 42(1), 1-34.

Fischbacher, U. 2007. z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2), 171-178.

Fochmann, M., & Weimann, J. 2013. The effects of tax salience and tax experience on individual work efforts in a framed field experiment. *FinanzArchiv: Public Finance Analysis*, 69(4), 511-542.

Fochmann, M., Weimann, J., Blaufus, K., Hundsdoerfer, J., & Kiesewetter, D. 2013. Net Wage Illusion in a Real-Effort Experiment. *The Scandinavian Journal of Economics*, 115(2), 476-484.

Hartley, R., & Farrell, L. 2002. Can Expected Utility Theory Explain Gambling? *American Economic Review*, 92(3), 613-24.

Heath, C., & Soll, J. B. 1996. Mental Budgeting and Consumer Decisions. *Journal of Consumer Research*, 23(1), 40-52

Hedesstrom, T. M., Svedsater, H., & Garling, T. 2004. Identifying heuristic choice rules in the Swedish premium pension scheme. *Journal of Behavioral Finance*, 5(1), 32-42.

HM Treasury. 2015. Strengthening the incentive to save: a consultation on pensions tax relief (July 2015), Web. 03. Mar. 2016. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/442160/Strengthening_the_incentive_to_save_consultation__web_.pdf.

Huang, J. 2008. Taxable and tax-deferred investing: a tax-arbitrage approach. *Review of Financial Studies*, 21(5), 2173-2207.

Imai, K., Tingley, D., & Yamamoto, T. 2013. Experimental designs for identifying causal mechanisms. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176(1), 5-51.

Kirchler, E., Muehlbacher, S., Hoelzl, E., & Webley, P. 2009. Effort and Aspirations in Tax Evasion: Experimental Evidence. *Applied Psychology*, 58(3), 488-507.

Levy, H. 1994. Absolute and Relative Risk Aversion: An Experimental Study, *Journal of Risk and Uncertainty*, 8(3), 289-307.

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. 2007. Mediation Analysis. *Annual Review of Psychology*, 58, 593-614.

Read, D., & Loewenstein, G. 1995. Diversification bias: Explaining the discrepancy in variety seeking between combined and separated choices. *Journal of Experimental Psychology: Applied*, 1(1), 34-49.

Riedl, A., & Tyran, J. R. 2005. Tax liability side equivalence in gift-exchange labor markets. *Journal of public Economics*, 89(11/12), 2369-2382.

Sausgruber, R., & Tyran, J. R. 2005. Testing the Mill hypothesis of fiscal illusion. *Public Choice*, 122(1-2), 39-68.

Sobel, M. E. 1982. Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, 13, 290-312.

Sobel M. E. 1986. Some New Results on Indirect Effects and Their Standard Errors in Covariance Structure Models. *Sociological Methodology*, 16, 159-186.

Sobel, M. E. 1987. Direct and Indirect Effects in Linear Structural Equation Models. *Sociological Methods & Research*, 16(1), 155-176.

Stata Corp. 2015. Stata Structural Equation Modeling Reference Manual Release 14. Texas: Stata Press, Stata Corp. LP.

Thaler, R. H. 1985. Mental accounting and consumer choice. *Marketing Science*, 4(3), 199-214.

Thaler, R. H. 1990. Anomalies: Saving, fungibility, and mental accounts. *Journal of Economic Perspectives*, 4(1), 193-205.

Thaler, R. H. 1999. Mental Accounting Matters. *Journal of Behavioral Decision Making*, 12(3), 183-206.

Warren Jr, A. C. 1986. The Timing of Taxes. *National Tax Journal*, 39(4), 499-505.

Weber, M., & Schram, A. J. 2016. The Non-Equivalence of Labor Market Taxes: A Real-Effort Experiment. *The Economic Journal*, forthcoming.

Weber, M., & Zuchel, H. 2005. How do prior outcomes affect risk attitude? Comparing escalation of commitment and the house-money effect. *Decision Analysis*, 2(1), 30-43.

Yoo, K. Y., & De Serres, A. 2005. Tax treatment of private pension savings in OECD countries. *OECD Economic Studies*, 39, 73-110.

Appendix

A1 Instructions

We divided the instructions into two parts. The first part is identical for both treatments and describes the procedure and training periods. The second part partly differs between the treatments. In the following, the instructions (originally written in German) are presented.

A1.1 Part 1 of the instructions

Hello, and welcome to our experiment!

By participating in this experiment, you have the possibility of earning money being provided by the University of Hanover. The aim of the experiment is to analyse economic decision making. The payment you receive at the end of the experiment depends on your effort as well as on chance. Please read the instructions carefully and attentively.

If you have further questions, please contact the experimenter.

1. Procedure

We would like to point out that you are not allowed to talk to the other participant or to leave your seat during the experiment.

You received a table tennis ball with an identification number to start the experiment. Please carefully keep the ball with you. You need the ball to identify yourself as soon as the compensation is paid. The identification number enables you to hide your true identity.

2. Training period

The experiment starts with a training period in which you get to know the design and functionality of the experiment. This period is designed to help you in the experiment.

Your task is to digitize the answers marked on the sheets in front of you into an entry form on the computer. The sheets contain the answers from a multiple-choice exam. First, we ask you to enter the number of the sheet, which can be found at the top left corner of the page, into the field provided for it and press “Next.” Afterward, you will see the entry form for the sheet. It will be set up similar to the hard copy of the sheet in front of you. Please translate the marked answers for all 60 questions into the entry form on the computer. When you have finished translating the sheet, please press “Next.” Then, you will be given information on whether or not you correctly digitized on the sheet.

After every participant has finished the training period, you will be given a sign by the instructor. At that point, please open the envelope on your desk. Do not open the envelope before that.

A1.2 Part 2 of the instructions

1. Design of the experiment

[Immediate Taxation Treatment only: The experiment consists of two parts: You start with a one-hour working period in which you earn your wage. This wage is taxed immediately. Therefore, the experimenter will withhold the tax. The tax rate is 60% (for further information, please refer to point 3). In the second part of the experiment, you decide how you intend to invest your net wage (your wage after tax). The return on any investment (interest earned) is tax free.]

[Deferred Taxation Treatment only: The experiment consists of two parts: You start with a one-hour working period in which you earn your wage. In the second part of the experiment, you decide how you intend to invest your wage. The return on any investment and the invested amount must be taxed. The tax rate is 60% (for further information, please refer to point 3).]

As a last step, you will be asked to answer a short questionnaire, which is needed to evaluate the experimental results. We would like to emphasize that all answers will be kept anonymous. You will receive your payment after all participants have finished the experiment.

2. Working period

You can see a paper pile in front of you. These are answer sheets to a multiple-choice exam that you have to digitize. This work task is exactly the same as during the previous training period.

Each answer sheet is numbered. You can find the sheet number at the top left corner. This sheet number is unique. Therefore, it is possible to identify the exam participant as well as the corresponding answer mask. As a consequence, it is extremely important that you type in the sheet number correctly!

First, we ask you to enter the number of the sheet, which can be found at the top left corner of the page, into the field provided for it and press "Next." Afterward, you will see the entry form for the sheet. It will be set up similar to the hard copy of the sheet in front of you. Please translate the marked answers for all 60 questions into the entry form on the computer. When you have finished translating the sheet, please press "Next." Then, you will be given information on whether or not you correctly typed in the sheet.

You get one hour to digitize as many sheets as you want or are able to do. During the whole period, a countdown will be projected onto the front wall of the room, telling you how much time is left. The experimenter will inform you once the hour has elapsed. We ask you to immediately stop digitizing the sheets at that moment, to look at the experimenter, and to press "Cancel" at his/her notice. Please stop digitizing sheets as soon as the working period is over!

If you want to take a break during the working period, you can use the internet, but you are not allowed to exit the room during the working period. You can access the internet by pressing the "Windows button" and clicking on the "Internet Explorer" symbol in the toolbar. Alternatively, you can access the internet by pressing "tab" and "alt" at the same time. If you want to continue your work after your break, you have to press the "sheet" symbol in the bottom toolbar. Please do not press "Cancel" during the working period. Please wait until the experimenter tells you to do so. If you press "Cancel" prematurely, you are not able to continue the work task.

Please enter every sheet number just once. If you enter the same sheet number more than once, this sheet cannot be taken into consideration, and as a consequence, you will receive no

compensation for this sheet. After the working period, you have the opportunity to make an investment decision.

3. Wage and taxation of wages

[Immediate Taxation Treatment only: You receive € 2.00 before taxation for each correctly digitized sheet; **60%** of this payment is withheld as tax, resulting in a payment of € 0.80 after taxation. After each digitized sheet, the computer informs you whether you entered the last sheet correctly and tells you the number of sheets you have entered correctly so far. At the beginning of the investment period, the number of sheets you entered, the number of sheets you entered correctly, the wage before taxation, the amount of taxes and the resulting wage after taxation are displayed.]

[Deferred Taxation Treatment only: You receive € 2.00 for each correctly digitized sheet. After each digitized sheet, the computer informs you whether you entered the last sheet correctly and tells you the number of sheets you entered correctly so far. In addition, you can see how much money you have earned so far. Your wage is not subject to any taxation.]

4. Investment period

[Immediate Taxation Treatment only: During this period of the experiment, you decide how to invest your wage after taxation.]

[Deferred Taxation Treatment only: During this period of the experiment, you decide how to invest your wage.]

Therefore, 2 different investment alternatives are displayed over 5 rounds. The return on the investment (interest) depends on the chosen investment alternative and the occurring state. Three possible states are given, with different impacts on your realized return. The probability for each state is equal to 1/3. For both investment alternatives, the same state occurs.

The investment alternatives are presented this way:

Investment A	State 1	State 2	State 3
Return	A%	B%	C%
Probability	1/3	1/3	1/3

Investment B	State 1	State 2	State 3
Return	X%	Y%	Z%
Probability	1/3	1/3	1/3

[Immediate Taxation Treatment only: You are asked to split up the after-tax wage between the two investment alternatives. Therefore, please enter the amount you are willing to invest in investment alternative A and in investment alternative B (amounts in euro and cent). You can also invest your after-tax wage solely in one of the two alternatives. You have to invest all of your after-tax wage. The amount you have to invest remains constant over the 5 rounds. After the fifth round, one round is chosen for your payment. This decision will be made randomly by throwing a dice (for further information, please refer to point 6). Each payment you receive from the investment is subject to no taxation (for further information, please refer to point 5).

[Deferred Taxation Treatment only: You are asked to split up the wage between the two investment alternatives. Therefore, please enter the amount you are willing to invest in investment alternative A and in investment alternative B (amounts in euro and cent). You can also invest your wage solely in one of the two alternatives. You have to invest all of your wage. The amount you have to invest remains constant over the 5 rounds. After the fifth round, one round is chosen for your payment. This decision will be made randomly by throwing a dice (for further information, please refer to point 6). Each payment you receive from the investment is subject to taxation (for further information, please refer to point 5).

5. Taxation of investment

[Immediate Taxation Treatment only: The payment you receive from the investment is not subject to any taxation.]

[Deferred Taxation Treatment only: The payment you receive from the investment is subject to taxation. The tax rate is **60%**, meaning that the return on the investment and the invested amount is taxed at a rate of 60%. You will receive the after-tax payment (the tax is withheld by the experimenter).

6. Payment

Please answer the questionnaire as soon as you have finished your investment decisions. As mentioned above, we need this information to evaluate the results. All answers remain anonymous!

After each participant has completed the questionnaire, you will be asked to come to the front desk to receive your payment. Please use the dice to determine your return on investment.

The first roll of the dice determines the round of the investment decision. If 1 turns up, the investment decision made in the first round is taken into consideration, and if 2 turns up, the investment decision made in round two is taken into consideration, and so forth. If 6 turns up, you have to roll the dice again.

The second roll of the dice determines the occurring state. If 1 or 2 turns up, state 1 occurs. If 3 or 4 turns up, state 2 occurs. If 5 or 6 turns up, scenario 3 occurs.

If no further questions remain, please wait until the countdown for the working period to start. After that, please begin to work.

Thank you for participating!

A2 Screenshots

Figure 7: The work task (z-Tree screenshot)

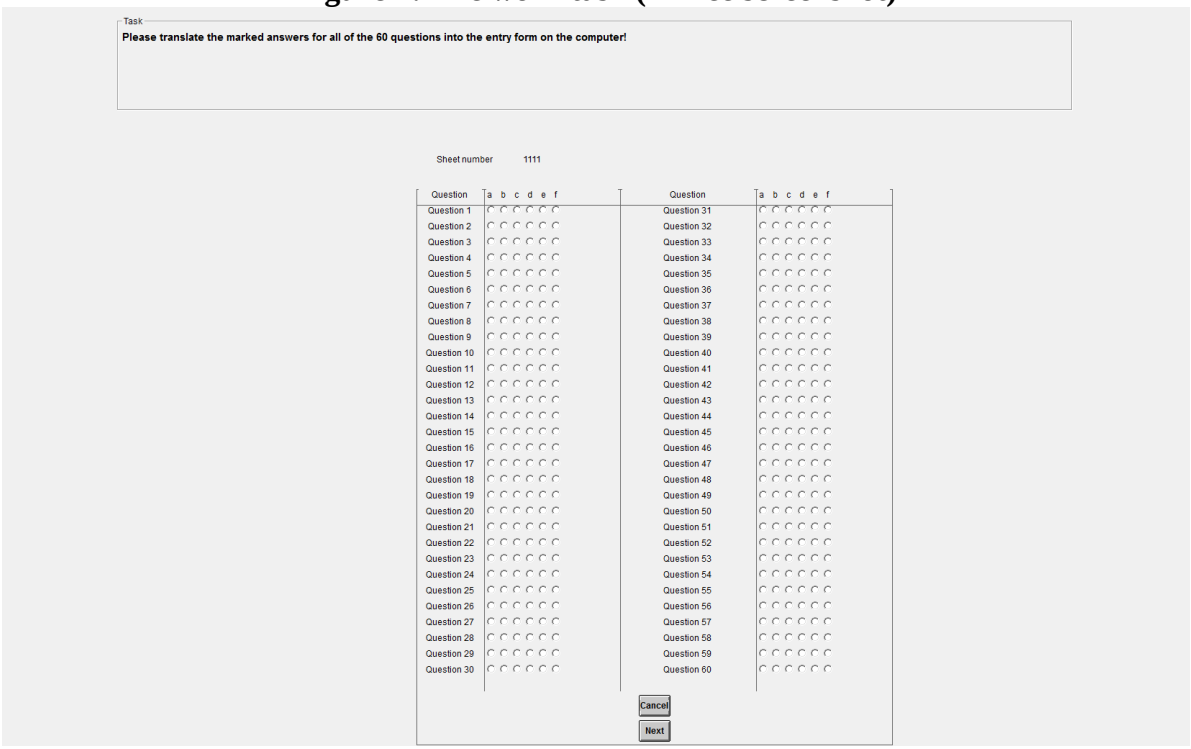


Figure 8: Information on actual work progress (z-Tree screenshot)

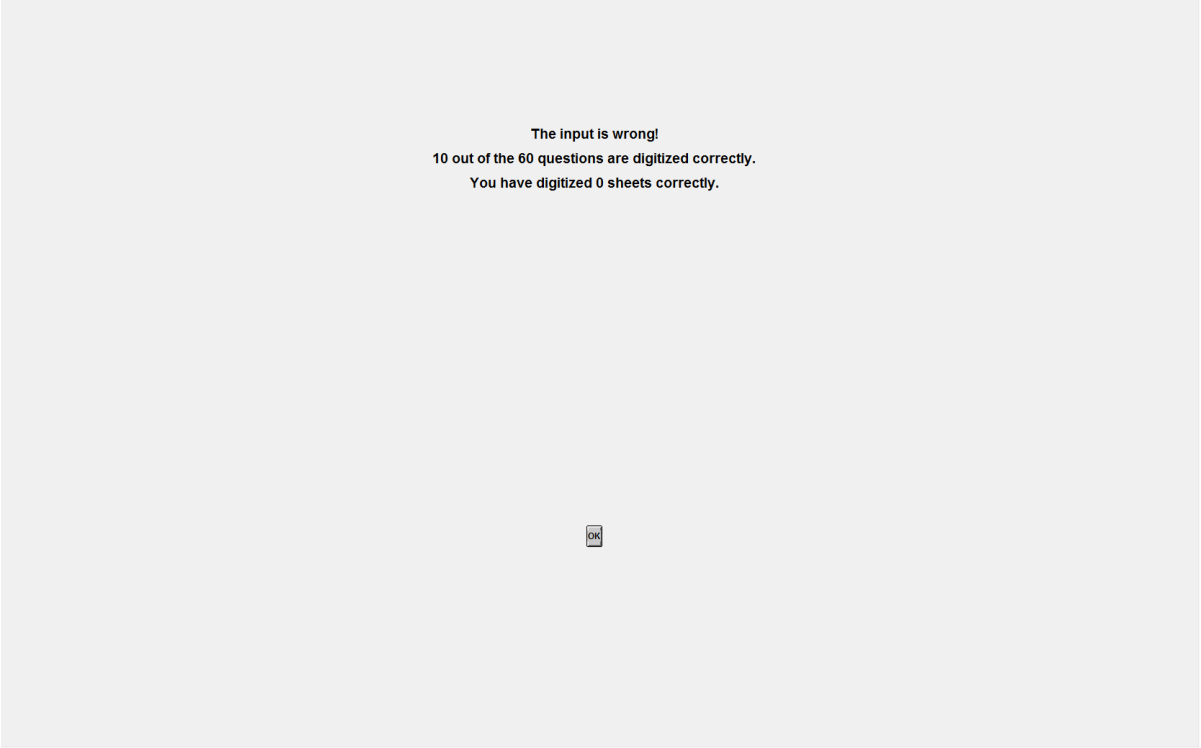


Figure 9: Entering a new sheet number and information on work progress, actual wage before and after tax, and the tax rate for the immediate taxation treatment (z-Tree screenshot)

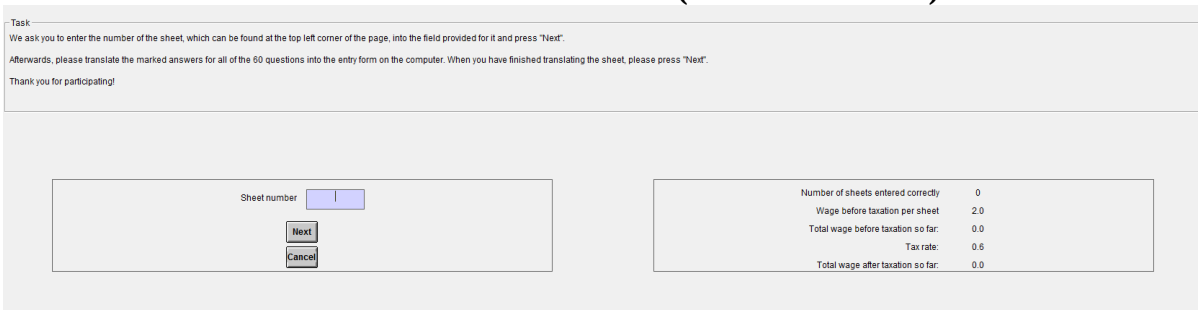


Figure 10: Entering a new sheet number and information on work progress and actual wage for the deferred taxation treatment (z-Tree screenshot)

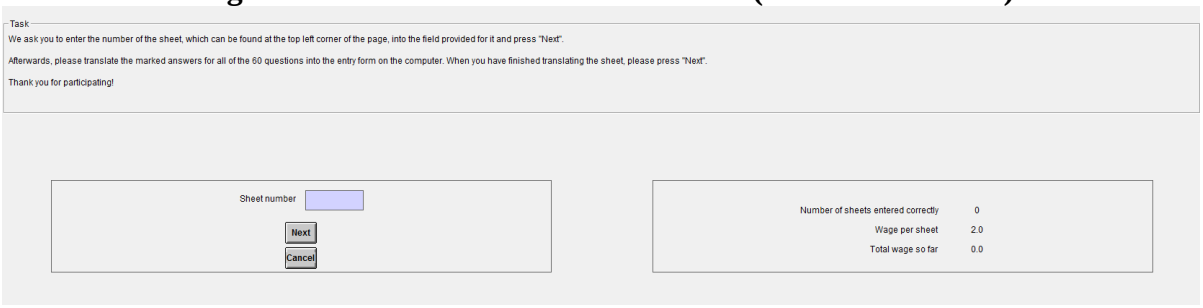


Figure 11: Investment decision stage period 1 for the immediate taxation treatment (z-Tree screenshot)

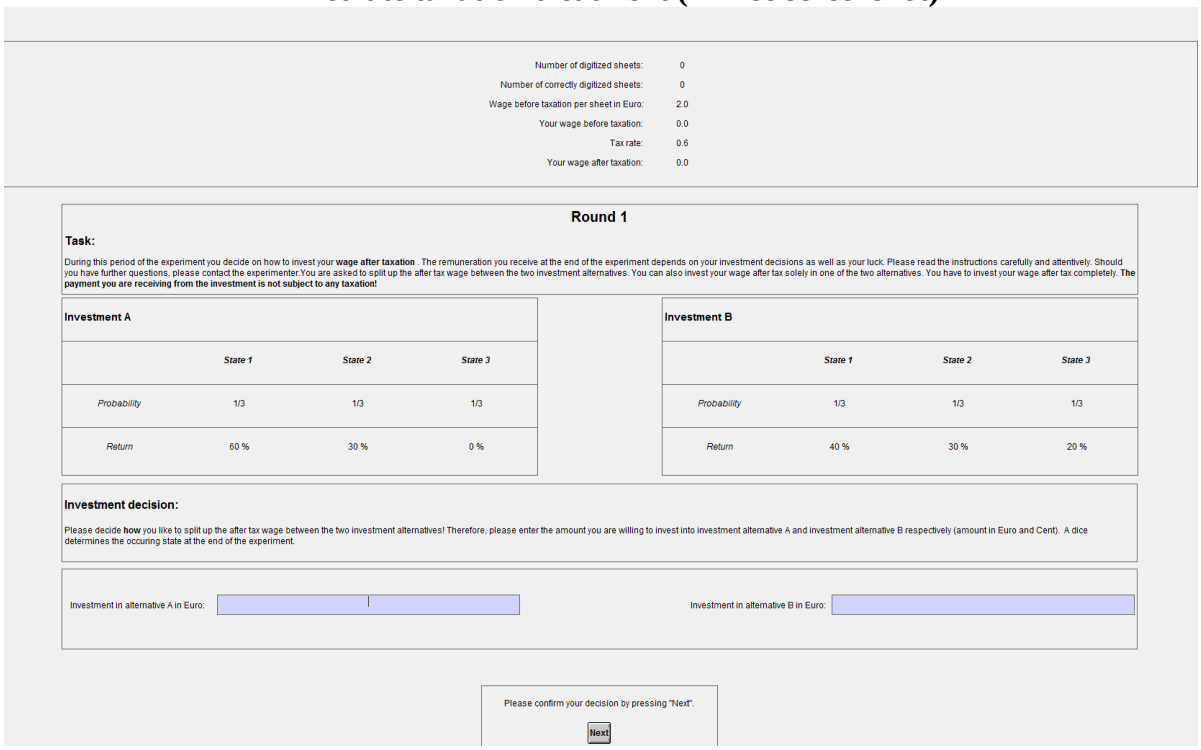
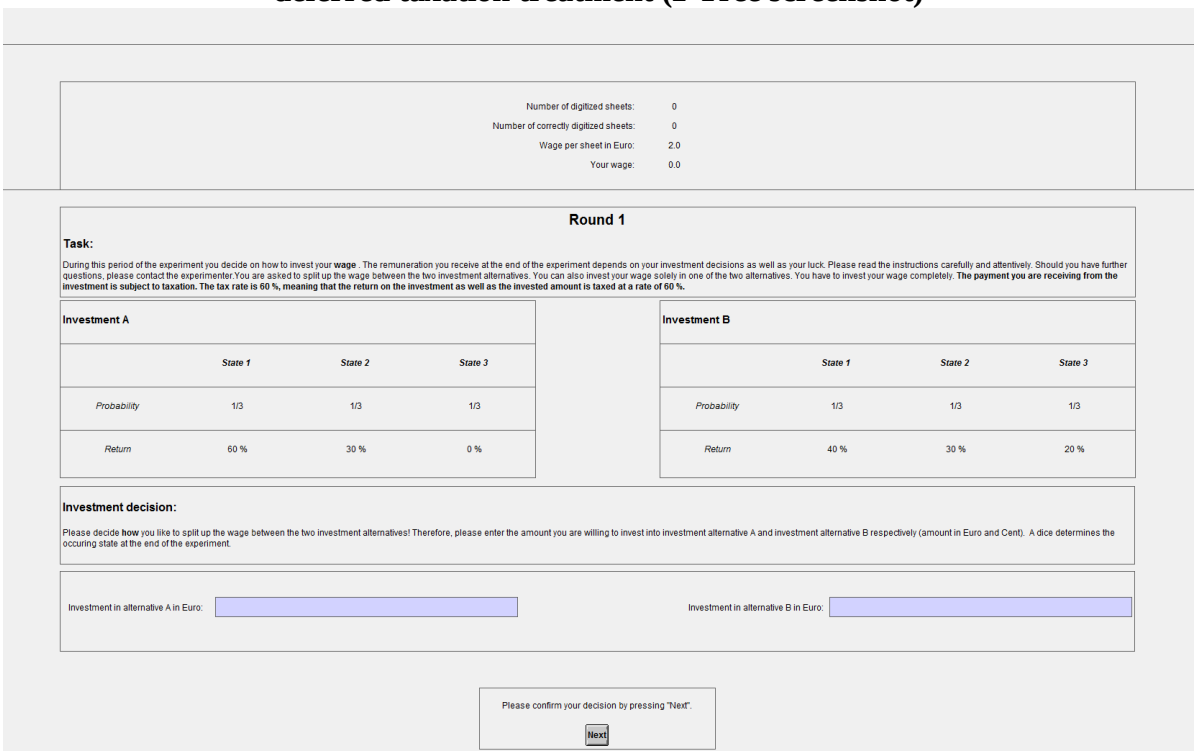


Figure 12: Investment decision stage period 1 for the deferred taxation treatment (z-Tree screenshot)



A3 Questionnaire

How did you perceive the work task during the experiment on a scale of 1 = very unpleasant to 10 = very pleasant?

Very unpleasant ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very pleasant
1 2 3 4 5 6 7 8 9 10

If you took a break during the work task, please tell us why!

- ☐ The task was unpleasant
- ☐ The wage was too low
- ☐ I had no more desire
- ☐ I did not take a break
- ☐ Others ☐

How important was the wage before taxation by determining the work effort on a scale of 1 = very unimportant to 10 = very important?

Very unimportant ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very important
1 2 3 4 5 6 7 8 9 10

Immediate *Taxation Treatment only:* Your wage was subject to taxation. What was the tax rate in percent?

☐ %

Deferred Taxation Treatment only: The return on the investment and the invested amount were subject to taxation. What was the tax rate in percent?

☐ %

How important was the taxation by determining the work effort on a scale of 1 = very unimportant to 10 = very important?

Very unimportant ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very important
1 2 3 4 5 6 7 8 9 10

Immediate *Taxation Treatment only:* Please tell us the amount of the wage after taxation per sheet in euro.

☐ euro

Deferred Taxation Treatment only: Please tell us the amount of the wage before taxation per sheet in euro.

☐ euro

Immediate *Taxation Treatment only:* How important was the wage after taxation by determining the work effort on a scale of 1 = very unimportant to 10 = very important?

Very unimportant ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very important
1 2 3 4 5 6 7 8 9 10

Deferred Taxation Treatment only: Did you recognize that the wage was effectively subject to taxation because of the taxation of the return on investment and the taxation of the invested amount?

☐ Yes ☐ No ☐ I don’ t know

How did you perceive the fairness of the work task wage on a scale of 1 = not fair at all to 10 = totally fair?

Not fair at all ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Totally fair
1 2 3 4 5 6 7 8 9 10

How would you rate your tax law knowledge on a scale of 1 = no knowledge to 10 = exceptionally knowledge?

No knowledge ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Exceptionally knowledge
1 2 3 4 5 6 7 8 9 10

How would you rate your investment knowledge (e.g., investing in securities, bonds, fixed deposit accounts) on a scale of 1 = no knowledge to 10 = exceptionally knowledge?

No knowledge ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Exceptionally knowledge
1 2 3 4 5 6 7 8 9 10

Was the mean return the same for the two investments alternatives in each round?

☐ Yes ☐ No ☐ I don' t know

How old are you?

☐

Are you female or male?

☐ Female ☐ Male

Which faculty are you enrolled in?

- ☐ Architecture and landscape
- ☐ Construction engineering and geodesy
- ☐ Electrical engineering and computer science
- ☐ Law
- ☐ Mechanical engineering
- ☐ Mathematics and physics
- ☐ Natural sciences
- ☐ Philosophy, business and economics (social science)
- ☐ Others ☐
- ☐ I am not a student

What qualification are you aiming at right now?

- ☐ Bachelor
- ☐ Master
- ☐ Diploma
- ☐ Magister [comparable to Master of Arts]
- ☐ 1st state examination
- ☐ 2nd state examination
- ☐ Doctoral degree
- ☐ Other ☐

Which academic semester are you in?

☐

What's your marital status?

- ☐ Marriage/registered partnership
- ☐ Unmarried
- ☐ Divorced/widowed

Do you have children?

☐ Yes ☐ No

What is your monthly disposable income (after rent; approximately)?

- ☐ < € 500
- ☐ € 501 - € 1.000
- ☐ € 1.001 - € 1.500
- ☐ € 1.501 - € 2.000
- ☐ > € 2.001

5 Do tax earmarking and voting on the tax rate influence tax avoidance and labor supply in the lab?

Kay Blaufus, Jochen Hundsdoerfer, Matthias Sünwoldt

Abstract

We conduct a laboratory experiment to analyze how earmarking of tax revenues and voting on the rate of the earmarked tax affect taxpayers' tax avoidance and labor supply behavior. We find that while earmarking tax revenues reduces tax avoidance, the opportunity to vote on the rate of the earmarked tax reduces tax avoidance even further. In contrast, earmarking tax revenues and voting on the tax rate do not affect real labor supply. This indicates that a feature of a tax could influence different forms of taxpayers' behavioral responses in different ways. In our experiment, taxpayers' avoidance responses clearly dominate their real responses.

Keywords

Behavioral Taxation · Tax Earmarking · Democratic Participation · Tax Avoidance · Substitution Effects · Real Effort Experiment

JEL Classification

M41 · M48 · D72 · H20 · H41 · Z18

We thank Martin Fochmann, Martin Jacob, Ulrike Stefani, Pia Vollert, the participants of the Mannheim Accounting & Taxation Research Workshop, the participants of the WHU Research Seminar in Accounting & Control, the participants of the session 5b during the 75th Annual Meeting of the VHB and the participants of the Paderborn Workshop for Experimental Research in Taxation for their helpful comments. Matthias Sünwoldt gratefully acknowledges financial support from KPMG.

I paid my income tax today.
I never cared what Congress spent.
But now I'll watch over ev'ry cent,
Examine ev'ry bill they pay,
They'll have to let me have my say.
(From the song: "I paid my income tax today", by Irving Berlin, 1942)

5.1 Introduction

In November 2012, voters in California passed ballot measure Proposition 30 by a 54% to 46% vote. This measure temporarily raised the state sales tax and income tax rates and earmarked the resulting revenues for education. These revenues have subsequently allowed the California State University network to avoid a 9% tuition hike. From a tax research standpoint, it would be interesting to know whether these additional taxes caused behavioral responses similar to those caused by other taxes, taking into account that the tax revenues are earmarked and that taxpayers had a voice regarding their implementation.

Taxpayers may respond to taxation by changing their real behavior, by legally avoiding taxes or even by illegally evading taxes. Previous accounting research has examined taxpayer and tax system characteristics to explain differences in taxpayers' compliance behavior (e.g., Richardson et al. 2013, Atwood et al. 2012) and real responses as investment or labor/leisure decisions (e.g., Swenson 1988, Sillamaa 1999b, Falsetta et al. 2013). The present study complements this research by addressing the question of whether two important tax system characteristics—earmarking the revenues of a tax and voting rights in setting the tax rate—have an effect on tax avoidance and real labor supply decisions.

Viewing the tax rate as the single determinant of taxpayers' reactions is clearly a heroic simplification (Saez et al. 2012). If the earmarking of tax revenues and the extent of democratic participation drive tax avoidance and labor supply, those additional tax characteristics may explain differences between countries or between pre-reforms and post-reforms in tax avoidance and in tax elasticities. The policy implications are obvious. In designing tax systems, earmarking approaches and democratic participation designs could be important options.

By earmarking tax revenues, the appropriation of the collected funds to the taxpayers is disclosed. Direct democratic participation in tax system features (e.g., tax rates) gives the taxpayers the impression of control over their contributions. Both elements highlight features of the tax system, which may not only enhance tax salience but also change taxpayers' deliberate reactions.

We contribute to previous research in the following ways. First, we conduct a real-effort laboratory experiment with a data input task. The participants are paid in real money, and the earmarked tax revenue is transferred to the announced institution (faculty library). This setting provides the participants with realistic financial incentives and avoids cheap-talk effects. Second, we model democratic participation rights as voting on the tax rate. We are able to do this because the tax revenue is earmarked for an external public good. Third, the setting allows us to test whether the tax rate a participant votes for explains her subsequent tax

avoidance decision. Fourth, we distinguish between tax avoidance and real effects of taxation, which opens a new perspective for empirical tax research.

We find that earmarking significantly lowers tax avoidance. Democratic participation, i.e., voting on the rate of the earmarked tax, reduces tax avoidance even further. Furthermore, we show that the individual tax-rate vote affects the subsequent avoidance decision. Those who vote for a high tax rate avoid significantly fewer taxes. With respect to labor supply, we do not observe any significant effect of earmarking or participation rights. This result suggests that a feature of a tax could influence different forms of taxpayers' behavioral responses in different ways. Taxpayers' avoidance responses seem to be much more elastic than their real responses.

The paper is organized as follows. An overview of the literature is given in Section 5.2. The research hypotheses are developed in detail in Section 5.3. Section 5.4 explains the experimental research design. Section 5.5 presents the main results, and the final section concludes by summarizing and discussing key findings.

5.2 Literature review

Prior accounting research has extensively examined how certain tax characteristics, such as tax rates, audit schemes, tax complexity or tax equity, affect tax evasion and taxpayers' real decisions (e.g., Kim et al. 2005, De Simone et al. 2013, Falsetta et al. 2013). However, surprisingly, accounting research is almost silent on two important tax characteristics: earmarking and participation rights. Considering fiscal exchange theory and the psychological tax-contract approach, we expect that both earmarking and participation rights may have a significant effect on taxpayers' behavioral responses.

Fiscal exchange theory highlights the importance of voice, legitimacy and consensus in the public decision process on tax features (Alm and Torgler 2006) and emphasizes the relevance of tax revenue allocation for tax compliance. Taxpayers may rank different government expenditures according to their personal preferences. Tax revenues may be spent on public goods or on redistribution, or they may be extracted by the government or other agents. Among the public goods that can be purchased with tax revenues, taxpayers develop preferences according to their personal demand for the public goods or their political opinions. In addition to fiscal exchange theory, Feld and Frey (2007) argue that tax compliance is based on a psychological tax contract between taxpayers and the government that establishes a fair, reciprocal exchange. One of the government's duties in this contractual relationship consists of providing fiscally equivalent public goods and services in exchange for tax payments. According to this approach, however, taxpayers comply even if they do not receive a full public good equivalent as long as they perceive the political process as fair (Feld and Frey 2007).

Although accounting research has not previously examined earmarking or participation rights, there are a number of related public economics studies. Earmarking taxes and its influence on taxpayers' behavior are analyzed theoretically in Buchanan (1963) and Goetz (1968). Several environmental economics studies find that the acceptance of ecological taxes can be raised by informing on how the funds are used (Beuermann and Santarius 2006, Kallbekken and Aasen

2010, Sælen and Kallbekken 2011). Moreover, there is experimental evidence that earmarking taxes reduces tax evasion (e.g., Alm et al. 1993).

Whether democratic participation rights influence tax resistance has been primarily analyzed in the context of tax evasion. Experimental (e.g., Alm et al. 1993, Alm et al. 1999, Feld and Tyran 2002, Tyran and Feld 2006) and empirical studies based on survey evidence (e.g., Torgler 2005) show that evasion decreases if subjects are granted participation rights. Hug and Spörri (2011) conduct a cross-country study and find that referendum institutions have almost no direct effect on tax morale, but they provide evidence for an indirect effect—that is, referendums strengthen the link between trust and tax morale.

The cited studies show that earmarking and participation rights may be useful instruments in reducing tax evasion. However, the influence of voting on tax rates and earmarking taxes on taxpayers' licit reactions, to our knowledge, has not been analyzed. A priori, it is unclear whether the results for tax evasion also hold for licit tax avoidance. Kirchler et al. (2003) show that legality influences taxpayers' fairness perceptions of different tax responses. In particular, they find that tax avoidance is perceived positively, whereas tax evasion is perceived negatively. Thus, the question of whether earmarking and participation are also helpful in reducing tax avoidance remains open.

Moreover, it is unclear how earmarking and voting rights affect real decisions such as labor supply.[21] Although previous accounting and public economics studies have used real-effort experiments to study the effect of taxes on labor supply, these studies focus on other tax system features, such as tax rates (Swenson 1988, Collins et al. 1992, Sillamaa 1999a, 1999b, Sutter and Weck-Hannemann 2003, Djanali and Sheehan-Connor 2012, Fochmann et al. 2012), tax fairness (Lévy-Garboua et al. 2009), tax salience (Fochmann and Weimann 2013), evasion opportunities (Collins and Plumlee 1991, Collins et al. 1992), and tax framing (Blumkin et al. 2012, Gamage et al. 2010).

Our contributions are twofold. First, we show how earmarking and participation in setting a tax rate affect licit tax avoidance. Second, we show how earmarking and participation in setting a tax rate affect labor supply. By comparing the effect of participation in setting a tax rate with the effect of earmarking the tax revenues, we can measure which effect is stronger in mitigating taxpayers' reactions. We can also observe whether the individual voting decision influences the subsequent labor supply and avoidance decision.

For our research question, an experiment has several obvious advantages. For example, financial incentives for the participants are set, and we can model a voting mechanism and the labor supply decision. Furthermore, our setting allows us to differentiate between labor supply decisions and tax avoidance by offering a costly tax loophole.

[21] For voting and real effort in a non-tax context, see, for instance, Mellizo et al. 2014.

5.3 Hypotheses

5.3.1 Tax avoidance

We concentrate on two characteristics of a tax system that may influence the perceived burden. First, we ask whether earmarking the revenues of a tax for a public good drive tax-avoidance behavior. Considering fiscal exchange theory, the psychological tax-contract approach and the empirical findings that tax compliance increases with the perceived adequacy of public services, we hypothesize that earmarking the revenues of a tax for a public good lowers tax avoidance more than a non-earmarked tax does. In a tax avoidance decision, taxes are explicit and salient. Because individuals may feel badly about explicitly avoiding a tax that has been earmarked for a positive purpose, a tax earmarked for a purpose that is perceived as positive should reduce tax avoidance.

H_{1a}: *Earmarking reduces tax avoidance.*

Second, an important property of a tax system is how tax rates or tax scales are set. Fiscal exchange theory highlights the importance of voice, legitimacy and consensus in the public decision process on tax features. In reality, we find a continuum between the direct participation of the taxpayers in setting tax rates and no participation whatsoever.[22] Most countries are in between; the elected government or parliament in a democratic process may decide tax rates. We are interested in the effects of democratic participation. According to the literature on tax evasion, we hypothesize that the democratic right to participate in setting the rate of the earmarked tax reduces tax avoidance.

H_{1b}: *The opportunity to vote on the tax rate of an earmarked tax reduces tax avoidance.*

5.3.2 Labor supply

It is well documented (e.g., Keane 2011) that taxpayers tend to react to a tax on wages by reducing their labor supply. However, recent evidence indicates that the standard assumption of zero utility from tax paying may not hold, thus implying that individuals work more than predicted by the standard labor/leisure models (Djanali and Sheehan-Connor 2012). Applying the framework stated herein (fiscal exchange theory, psychological tax-contract approach) and the empirical findings for tax compliance, we expect that earmarking the tax revenues increases labor supply. If the tax revenues are earmarked, taxpayers may view part of their labor supply as a charitable giving-in-kind. Psychological research suggests that this view could lead to a utility increasing "warm glow" effect due to, e.g., guilt reduction, a negative state relief, or self-reward (Djanali and Sheehan-Connor 2012).

H_{2a}: *Earmarking increases labor supply.*

Moreover, we suppose that democratic participation in setting the rate of the earmarked tax could increase individuals' utility perceived from paying taxes due to the rules of reciprocity implied by the psychological tax-contract approach. Thus, participation should further increase labor supply:

H_{2b}: *The opportunity to vote on the tax rate of an earmarked tax increases labor supply.*

[22] See Torgler (2005) for differences in direct democratic rights between Swiss cantons.

5.4 Experimental design and protocol

The experiment was conducted in the laboratory of the Freie Universität Berlin. We prevented communication and visual contact between participants with installed screen walls. The participants included 134 students from business administration and economics programs at the university. They were recruited by posters and by personal promotions in large undergraduate lectures.[23] Most participants were undergraduate students, though there was a very small number of business administration master's students. We used a between-subject design with two dependent variables, legal tax avoidance (*avoidance*) and labor supply (as *minutes*), and two independent variables, *earmarking* and *participation*. Subjects were randomly assigned to each treatment.

5.4.1 Tax avoidance and labor supply

As in previous accounting research, we conducted a real-effort experiment to measure labor supply. However, previous studies have often used rather artificial operationalization for leisure. In Djanali and Sheehan-Connor (2012), subjects can choose to work or to do nothing and wait for the experiment's continuation. In other studies, such as Swenson (1988) and Sillamaa (1999a, 1999b), subjects can read magazines or play games. In contrast, to achieve higher external validity in our experiment, subjects could choose not only their work effort but also their working time—i.e., they were allowed to finish the task and to quit the experiment whenever they wanted (with an upper limit of 185 minutes).

The real-effort task was a data input task such that paper test sheets from a modified multiple choice test had to be keyed into an Excel sheet. The gross wage was € 0.38 per *accurately* recorded sheet.[24] This wage and our estimation for the gross wage rate (€ 14 per hour) were common knowledge. For each participant, we counted the number of sheets recorded (*sheets*), the number of sheets accurately recorded (*correct sheets*) and the minutes she spent working (*minutes*). We use *minutes* to measure labor supply in the subsequent analyses. However, we will also conduct robustness checks with *effort* (correct sheets/minutes) as the dependent variable.

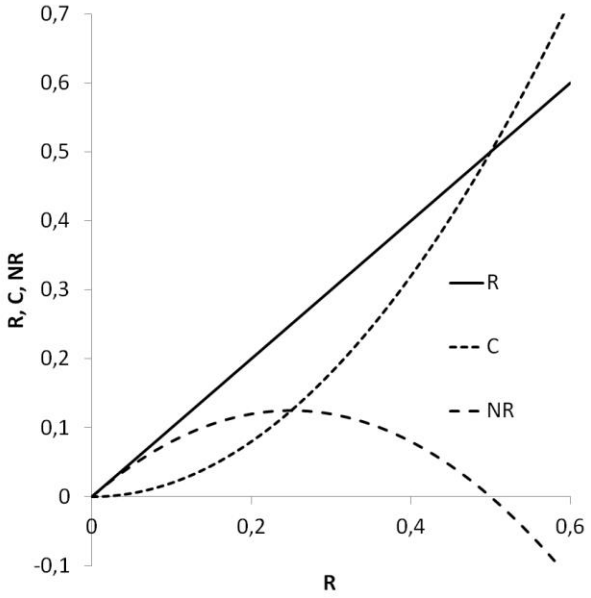
When a participant finished the real effort task, we calculated and announced her gross income and the tax bill. Subsequently, the participants were offered a tax-avoidance investment. The tax-avoidance investment allowed a net reduction of the tax bill. The participants were asked to set a percentage reduction of their tax bill. This reduction came with cost, and the marginal costs were increasing. Let R be the pre-cost tax-avoidance rate in percentage of the tax bill and C(R) be the tax avoidance cost in percentage of the tax bill with C' > 0, C`` > 0. The net tax avoidance rate is $NR = R - C(\mathrm{R})$ in percentage of the tax bill (see Figure 20). We present this tax avoidance investment in a very simple table using the wording "legal tax loophole" (see Appendix I). It is immediately comprehensible that—given the assumed cost function—a tax avoidance of 24% or 26% of the tax bill (each resulting in a

[23] To participate, the students had to apply by mail. Thus, we were able to prevent subjects' repeated participation.

[24] There was a limited number of possible box-ticking schemes. Therefore, we were able to control for correct and incorrect keyed sheets using a special VBA code.

12.5% net tax avoidance rate) maximizes the post-tax profit. We use the individual's chosen pre-cost tax avoidance rate to measure tax avoidance in all following analyses.

Figure 13: Tax avoidance investment: revenue (R), cost (C), and net revenue (NR) per tax euro



5.4.2 Earmarking and participation

Our two independent variables are *earmarking* and *participation*. In our setting, the dummy variable *earmarking* equaled one if the tax revenues were used to purchase textbooks in paper and/or online versions for the faculty's library.[25] *Earmarking* equaled zero if the tax revenues were not assigned to a specific use; rather, subjects were informed that their taxes were to be paid to the German Treasury. Obviously, we were unable to levy real taxes in the lab. To obtain as much external validity as possible, we donated the non-earmarked tax revenues to a special account of the German Central Bank that accepts gift contributions to reduce public debt (Möhlmann 2013).

The dummy variable *participation* takes a value of one if subjects had the opportunity to vote on the tax rate that was applied to their income of the real-effort task and zero otherwise. Voters were informed that the median tax rate would be set as the tax rate for every participant in the group. Secret paper ballots were used. The ballots contained a reminder of how the tax revenues were earmarked and a fixed scale (7%, 9%, …, 55%). The median voting design had two advantages. First, we could divide the treatment into two subsamples of (exactly or nearly) the same size. Second, a single vote could change the tax rate by 2 percentage points, which made each vote important. The voting was anonymous, but the participants had to note their participant ID number on the paper ballot. The participants noted

[25] To control whether the purchase of textbooks in paper and/or online versions is a sufficient proxy for a demanded public good, we used the answers from the post-experimental questionnaire. We asked the participants about their frequency of using the library in general and textbooks in paper and online versions from the library in particular. Moreover, we asked whether the library's future purchase of e-books was necessary. The answer scale ranged from 0 (seldom/not necessary) to 10 (often/highly necessary). The range of the mean answers was between 5 and 6. Thus, we assume that the purchase of textbooks is a sufficient proxy.

the tax rate they voted for on a second sheet for their own use. Subsequently, the tax rate outcome of the voting process was announced and noted on a flip chart.

Because our treatment variables are dichotomous, a full 2x2 design encompasses four treatments: *earmarking+participation*, *earmarking-only*, *participation-only* and *control*. A comparison of the treatments would require that the tax rate be identical in each treatment. The tax rate, however, was the outcome of the voting procedure in the participation treatments. Thus, we abstained from conducting a *participation-only* treatment because it was impossible to ensure that this treatment would lead to the same tax rate as the *earmarking+participation* treatment. In the following, we refer to the remaining three groups as follows:

- *earmarking+participation* (E+P): *earmarking* = 1 and *participation* = 1,
- *earmarking only* (E): e*armarking* = 1 and p*articipation* = 0, and
- *control*: *earmarking* = 0 and *participation* = 0.

Participants were randomly assigned to these treatment groups. To ensure that we used the same tax rate in all treatments, we started with the *earmarking+participation* group and used the result of the tax-rate vote as an (exogenous) tax rate for the two other groups. Therefore, we did not randomly assign the different treatments to the laboratory sessions. Instead, we first solely started with the *earmarking+participation* group in one session. After the *earmarking+participation* group was conducted, we started with the *earmarking-only* and *control* groups in separate sessions. We randomly assigned the participants to one of our sessions.[26] We describe the details of the experimental protocol in the next subsection.

5.4.3 Experimental protocol

In chronological order, our experiment consists of the following (for all instructions given to participants, see Appendix II):

- informing the participants of the earmarking of tax revenues (*earmarking-only* and *earmarking+participation*),
- setting the endogenous tax rate by median voting (*earmarking+participation*),
- real-effort task (all groups),
- tax avoidance investment decision (all groups), and
- post-experimental questionnaire (all groups).

Note that all subjects were informed about the tax avoidance opportunity after they had completed their real-effort task. This order was necessary to disentangle the direct effect of earmarking (participation) on real effort and tax avoidance. If subjects had known about the tax avoidance opportunity before they started working, it would likely have affected their labor supply decision.

We conducted two runs. The first run took place in June 2012, and the second occurred in May 2013. In the first run, we did not offer participants in the *earmarking+participation*

[26] In detail, before the students confirmed their participation, we informed them about the different session dates and told them that we would randomly assign the participants to one session. After the students confirmed their participation, we informed them about the assignment to one of our sessions.

group a tax-avoidance opportunity. To test the hypothesized effect of participation on tax avoidance, we conducted a second run in May 2013 with the *earmarking+participation* group and the *earmarking-only* group.[27]

In detail, the experiment was conducted in the following way. The participants were briefed that their earnings would be taxed (all groups), and they were told how the tax revenues were earmarked, i.e., purchase of textbooks in paper and/or online versions (*earmarking+participation* and *earmarking-only*). The *control* group was briefed on the appropriation of income taxes in Germany. Additionally, they were briefed on the voting process (*earmarking+participation*) and the real-effort task (all groups). As stated in the instructions, we paid the earmarked tax to the faculty library and the non-earmarked tax to the revenue office to increase external validity.[28]

The participants in the *earmarking+participation* group subsequently voted on the tax rate according to the median voting regime described in the previous subsection. The resulting tax rates were 13% (first run) and 16% (second run), which were then used in the other groups to ensure comparability between treatments.

Finally, the participants filled out a post-experimental questionnaire that included questions regarding demographic information (gender, age, net income, university courses). After completing the study, the participants were remunerated in cash. Subjects' reward was calculated as gross wage (€ 0.38 * number of accurately recorded sheets) minus taxes.

5.5 Results

5.5.1 Summary statistics

Table 21 lists summary statistics on tax avoidance, measured as the percentage by which the tax liability was reduced pre-cost (variable *avoidance*), and on labor supply, measured as working time in minutes (variable *minutes*)[29], separately for the three groups. Additionally, we list summary statistics on the earned gross wage (variable *gross wage*) and control variables *gender* and *age*, which we obtained from the post-experimental questionnaire.

In all, 134 students from business administration and economics participated in the experiment. Due to missing values regarding our control variables, however, we remove ten cases. Furthermore, we skip four observations because of concerns that these participants were unable to cope with the tax avoidance decision. These subjects chose a tax avoidance level with a clearly negative return. Our final sample, therefore, consists of 120 students.

[27] Because of the one-year gap, we do not compare the 2012 results with the 2013 results.

[28] Because voluntary tax paying is not directly possible, we transferred the tax amount to the revenue office using a special bank account and "debt settlement" as a reference.

[29] Alternatively, one could use *sheets* as a measure of labor supply instead of *minutes*. We opt for *minutes* because it is likely that *sheets* is more affected by individual learning effects and individual skills (eye-hand coordination). However, the results are the same when we use *sheets.*

Table 14: Summary statistics

The values shown are means with standard deviations in parentheses. *Control*, *E*, and *E+P* are group dummy variables that equal one if the observation belongs to the *control*, *earmarking-only* or *earmarking+participation* groups; *avoidance* is the percentage by which the tax liability is reduced pre-cost (our measure of tax avoidance); *minutes* is worked minutes before quitting, unsolicited, the real-effort task (our measure of labor supply); *gross wage* as total earned wage before tax in the experiment; *gender* is a dummy variable that equals one if the participant is male; *age* is the participant's age in years.

	Control	E		E+P	
	1st run only	1st run	2nd run	1st run	2nd run
Observations	22	24	23	23	28
Avoidance	0.19 (0.09)	0.13 (0.08)	0.14 (0.10)	n.a.	0.10 (0.09)
Minutes	101.69 (52.31)	93.67 (48.97)	96.00 (51.03)	110.73 (42.65)	79.86 (44.03)
Gross wage	13.61 (8.64)	13.97 (11.25)	14.65 (9.83)	17.35 (11.21)	12.65 (8.42)
Gender	64%	49%	48%	48%	43%
Age	21.81 (2.79)	21.63 (1.84)	22.48 (3.93)	21.65 (2.14)	22.75 (3.47)

5.5.2 Tax avoidance

Despite the relatively low tax rate, amounting to 13% (16%) in the first (second) run, we observe substantial tax avoidance in each of the three treatments. However, Table 21 shows that average tax avoidance in the *earmarking-only* group is lower than that in the *control* group, where taxes were not earmarked, but subjects were informed that their taxes were to be paid to the German Treasury. An (unsigned) Mann-Whitney U test shows that the difference between these two groups is significant ($p = 0.021$). Moreover, it appears that in the *earmarking+participation* group, where subjects voted on the tax rate of the earmarked tax, the degree of tax avoidance is even lower than that in the *earmarking-only* group. Nevertheless, the difference between these two groups is insignificant ($p = 0.198$). Together, these results are consistent with H_{1a}, but we are not able to find evidence for H_{1b} at this stage. However, to control for individual differences, we test H_{1a} and H_{1b} using multivariate OLS regression analysis.[30] Our results are presented in Table 22.

Panel A includes the regression results regarding H_{1a} (lower tax avoidance within the *earmarking-only* group than the *control* group), and Panel B includes the regression results regarding H_{1b} (lower tax avoidance within the *earmarking+participation* group than the *earmarking-only* group). We use the pre-cost tax avoidance rate (*avoidance*) as the dependent variable and the corresponding group dummy as the independent variable. *E* equals one if the observation belongs to the *earmarking-only* group and zero otherwise, and *E+P* equals one if the observation belongs to the *earmarking+participation* group and zero otherwise. We include *gross wage* as a control variable. Participants who worked a relatively long time could

[30] As a robustness test, we used tobit regression instead of OLS because we found a relatively high number of zero tax avoidance votes (6 in Panel A and 14 in Panel B). The results remain qualitatively unchanged.

have made essential use of tax avoidance to retain a higher net amount of compensation.[31] Furthermore, a high level of *gross wage* could lead to aggressive tax avoidance due to a higher gross advantage of tax avoidance. Additionally, we add *gender* (one if the gender is male) into our regression equation. It is possible that male participants behaved differently from female participants regarding tax avoidance because previous experimental literature provides evidence that male subjects evade more taxes (e.g., Kastlunger et al. 2010).

Table 15: Estimation results regarding tax avoidance

Treatment as *E (E+P)* as a group dummy variable that equals one if the observation belongs to the *earmarking-only* group in Panel A (*earmarking+participation* group in Panel B); *avoidance* is the percentage by which the tax liability is reduced pre-cost (our measure of tax avoidance); *gross wage* is the total earned wage before tax; *gender* is a dummy variable that equals one if the participant is male; e*xperience* is a dummy variable that equals one if the participant is older than 21 years. Standard errors in parentheses; ***p<0.01, **p<0.05, *p<0.1.

	Panel A earmarking-only vs. control group (1st run), treatment = E		Panel B earmarking+particip. vs. carmarking-only (2st run), treatment = E+P	
	(1)	(2)	(1)	(2)
VARIABLES	Avoidance	Avoidance	Avoidance	Avoidance
Treatment	-0.0488*	-0.0682*	-0.0293	-0.0706*
	(0.0271)	(0.0360)	(0.0258)	(0.0359)
Gross wage	0.000855	0.000712	0.00291**	0.00240*
	(0.00134)	(0.00135)	(0.00143)	(0.00140)
Gender	0.0235	0.0154	0.0571**	0.0763***
	(0.0275)	(0.0282)	(0.0256)	(0.0266)
Experience		0.0153		-0.0932**
		(0.0392)		(0.0383)
Experience*treatment		0.0348		0.0793
		(0.0543)		(0.0492)
Constant	0.159***	0.160***	0.0709**	0.118***
	(0.0296)	(0.0334)	(0.0311)	(0.0356)
Observations	46	46	51	51
Adj. R-squared	0.058	0.057	0.138	0.205

In model 2, we include the control variable *experience* and interact this variable with treatment variable *E* in Panel A and *E+P* in Panel B. *Experience* equals one if the participant was older than 21 years (approximately the mean age of the participants; see above). It is likely that older participants behaved differently from younger participants with respect to tax avoidance. Because our participants were economics students, at least two behavioral differences between younger and older students could have arisen. First, older students could have made relatively essential use of tax avoidance in every treatment because they had learned how to behave rationally (in terms of profit maximizing) either in the past while at the university or in the workforce before entering the university. Previous research indicates that students of economics assume the behavior of *homo economicus* to be the social norm (Cullis et al. 2012). Thus, earmarking the tax or participation as voting for the tax rate would reduce

[31] Alternatively, the individual effort could lead to aggressive tax avoidance due to a higher net amount of compensation. Consequently, we substitute *gross wage* with *gross wage per hour* in all of the models as a measure of individual effort. The results remain qualitatively the same.

their tax avoidance to a lesser extent than inexperienced participants'. Second, older participants may have behaved rationally in the control setting but reacted to earmarking and/or participation to a greater extent than inexperienced participants'. Thus, it is conceivable that due to earmarking the tax for the faculty library, older participants developed a stronger need to give something back to the faculty after studying a comparatively long time at the university. Thus, our treatment would reduce their tax avoidance to a greater extent than inexperienced participants'. To control for these different responses, we include the dummy variable *experience,* and we interact experience with the treatment variables *E* and *E+P* in Panels A and B.

In Panel A of Table 22, we compare the *earmarking-only* group with the *control* group, all else being equal. Earmarking taxes for a public good reduces the pre-cost tax avoidance rate significantly in both of our models (p-values < 0.10). The coefficients of the control variables are not significant. Thus, we find evidence to confirm H_{1a}. Earmarking the tax for a purpose that is perceived as positive decreases the amount of legal tax avoidance investment in comparison to a non-earmarked tax.

Panel B of Table 22, shows that participation in an earmarked tax further reduces tax avoidance. However, this effect is significant only if we control for *experience*. As predicted, we find significant positive signs for the coefficients of the control variables *gross wage* and *gender*. The coefficient of *experience* is negative in model 2 (p-value < 0.05), which clearly contradicts the assumption that more experienced economics students behave more in the manner of *homo economicus*. Overall, we find evidence to confirm H_{1b}.

It is possible that the *earmarking-only* and *control* participants' expectations of the "optimal" or "fair" tax rate for the earmarked tax differed from the exogenous valid tax rate. A different expectation could lead to different tax avoidance levels. For instance, if the individual expectation of the "fair" tax rate is higher than the exogenous and actually valid tax rate, a lower level of tax avoidance is imaginable. Therefore, we asked the participants of the *earmarking-only* and *control* groups about their hypothetical individual tax-rate vote for the earmarked tax. We do not find significant group differences with respect to the individual (hypothetical) tax-rate votes (p-values between 0.21 and 0.94).

To study the relation between the individual tax-rate vote and the individual amount of tax avoidance, we conduct an additional analysis. If a participant voted for low taxes, but the tax rate outcome of the voting process was higher, we could expect her to subsequently avoid the higher tax. A participant voting for a higher tax rate than the majority could subsequently feel obliged not to avoid the low tax. Therefore, we add the dummy variable *individual vote* in both of our models. The variable *individual vote* reflects the individual tax-rate vote. Adding *individual vote* to our model leads to a decreased sample size because we analyze only the *earmarking+participation* group. The results are presented in Table 23.

Table 16: Estimation results regarding tax avoidance and the individual tax-rate vote

Avoidance is the percentage by which the tax liability is reduced pre-cost (our measure of tax avoidance); *individual vote* is the tax rate the participants voted for; *gross wage* is the total earned wage before tax; *gender* is a dummy variable that equals one if the participant is male; *experience* is a dummy variable that equals one if the participant is older than 21 years. Standard errors in parentheses, ***p<0.01, **p<0.05, *p<0.1.

	(1)	(2)
VARIABLES	Avoidance (2nd run)	Avoidance (2nd run)
Individual vote	-0.238**	-0.238*
	(0.113)	(0.115)
Gross wage	0.00471***	0.00468**
	(0.00167)	(0.00177)
Gender	0.0641**	0.0649**
	(0.0279)	(0.0310)
Experience		-0.00209
		(0.0315)
Constant	0.0601*	0.0613
	(0.0345)	(0.0393)
Observations	28	28
Adj. R-squared	0.365	0.338

We find a negative and significant influence of *individual vote* on the pre-cost tax-avoidance rate in model 1 (p-value < 0.05). That is, participants who voted for a relatively high tax avoided the earmarked tax to a lesser extent. Furthermore, we are able to confirm significant positive effects of *gross wage* and *gender*. In model 2, we include the control variable *experience*. The results are qualitatively the same.

In untabulated robustness tests, we use the hypothetical individual tax-rate vote from the post-experiment questionnaire and test the same regression equations for the *earmarking-only* and *control* groups. We do not find a significant influence of the hypothetical individual tax-rate vote on the pre-cost tax avoidance rate. Thus, it seems that our findings are primarily attributable to the voting process itself. Summing up, we find evidence that the individual tax-rate vote influences the amount of tax avoidance if subjects are granted real voting rights.

5.5.3 Labor supply

Table 21 shows that subjects work, on average, longer (shorter) in the *earmarking-only* group (*earmarking+participation* group) than in the *control* group. However, these differences are insignificant, as confirmed in a nonparametric test (unsigned Mann-Whitney U).[32]

To control for individual differences, we use OLS regression analysis. Table 24 provides the results regarding H_{2a} (Panel A) and H_{2b} (Panel B first run and Panel C second run). We use the same model setting as in the tax avoidance chapter but with one change. We exclude the control variable *gross wage* because of the change in the dependent variable (*minutes* instead of *avoidance*). As before, we test our base model 1 without the control variables *experience* and *experience*treatment* and include these control variables in model 2.

[32] The corresponding p-values amount to 0.652 for *earmarking-only* vs. *control* (1st run), 0.12 for *earmarking+participation* vs. *earmarking-only* (1st run) and 0.192 for *earmarking+participation* vs. *earmarking-only* (2nd run).

The results in Panel A reject H_{2a}. When comparing the labor supply between the *earmarking-only* and *control* groups, we do not find an increased labor supply within the *earmarking-only* group. Furthermore, the sign of the coefficient of our group dummy variable *E* is negative, though not significant. With respect to the control variables, we find no significant effects.

The results in Panel B and Panel C reject H_{2b}. That is, we do not find an increased labor supply when we add participation into our tax system, all else being equal. The coefficient of the group dummy variable *E+P* in Panel B (first run) is positive in both of our models—not on a significant level in model 1 but on a significant level in model 2 (p-value < 0.10). Moreover, the coefficient of *E+P* in Panel C (second run) is negative but not significant. Thus, we fail to replicate the significant effect.

In untabulated robustness tests, we use *sheets*, *correct sheets* and *effort* (correct sheets/minutes) as alternative specifications to measure labor supply. However, the results remain qualitatively unchanged.

Moreover, we add *individual vote* into our models to test whether the individual tax-rate vote affects the labor supply within the *earmarking+participation* group (the same procedure as in the tax avoidance chapter). We fail to find a significant influence.

Table 17: Estimation results regarding labor supply

Treatment as *E (E+P)* as a group dummy variable that equals one if the observation belongs to the *earmarking-only* group in Panel A (*earmarking+participation* group in Panels B and C); *minutes* is the worked minutes before, unsolicited, quitting the real-effort task (our measure of labor supply); *gender* is a dummy variable that equals one if the participant is male; *experience* is a dummy variable that equals one if the participant is older than 21 years. Standard errors in parentheses, ***p<0.01, **p<0.05, *p<0.1.

	Panel A earmarking-only vs. control (1st run), treatment = E		Panel B earmarking+particip. vs. earmarking-only (1st run), treatment = P+E		Panel C earmarking+particip. vs. earmarking-only (2nd run), treatment = P+E	
	(1)	(2)	(1)	(2)	(1)	(2)
VARIABLES	minutes	minutes	minutes	minutes	Minutes	minutes
Treatment	-4.614	-21.25	14.58	29.67*	-16.52	-15.30
	(15.52)	(20.73)	(13.16)	(17.49)	(13.42)	(19.44)
Gender	13.01	9.884	24.10*	23.49*	-7.644	0.903
	(15.51)	(16.03)	(13.31)	(13.73)	(13.42)	(14.46)
Experience		-22.31	24.10*	12.53		-21.79
		(22.61)		(18.88)		(20.60)
Experience*		39.11	(13.31)	-36.25		-0.925
treatment		(31.25)		(26.68)		(26.74)
Constant		103.5***		79.12***		106.9***
		(17.09)		(12.84)		(14.84)
Observations	46	46	47	47	51	51
Adj. R-squared	-	-	0.068	0.067	-	0.004

Summing up, neither the treatment variable nor the control variables are able to explain the labor supply on a significant level. Two exceptions with weak significance are treatment variable *P+E* in Panel B model 2 and control variable *gender* in Panel B. It seems that the labor supply decision is different from the tax avoidance decision. Although subjects respond

elastically regarding tax avoidance and participation, their real responses are rather inelastic. One reason could be that tax avoidance decisions are, compared to labor supply decisions, explicit and salient.

5.6 Conclusion

We complement previous accounting research by analyzing how earmarking tax revenues and democratic participation in setting the rate of the earmarked tax affect tax avoidance and labor supply behavior. To this aim, we conducted a real-effort experiment in which the participants were paid real money. We randomly assigned the participants to two treatment groups (*earmarking-only* and *earmarking+participation*) and one *control* group (neither participation nor earmarking). The earmarked tax revenue was transferred to an announced institution (faculty library). The endogenous tax rate resulting from the vote in the *earmarking+participation* group was used in all groups to ensure comparability between treatments.

The first interesting result is that subjects are indeed willing to pay a positive amount of (earmarked) taxes. However, the vote in the *earmarking+participation* group leads to a rather low tax rate (13% in the first run, 16% in the second run), thus showing the limits of voluntary taxation.

To answer the question of whether subjects' real and tax avoidance responses were affected by earmarking and participation, our subjects could respond to taxes by reducing labor supply and/or by using a costly tax loophole (tax avoidance). Our findings show that tax avoidance is significantly reduced by earmarking the tax for a positive purpose. Adding participation in the form of voting on the rate of the earmarked tax further lowers tax avoidance. Furthermore, we show that the individual tax-rate vote affects the subsequent avoidance decision. Those who voted for a high tax rate avoided significantly fewer taxes. These results show that both earmarking approaches and democratic participation designs are important options in designing tax systems. By reducing tax avoidance, earmarking tax revenues and democratic participation directly increase tax revenues for given tax rates. Furthermore, earmarking the tax revenues and democratic participation may reduce taxpayers' perceived level of tax avoidance by other taxpayers. In this way, these instruments may raise the perceived degree of fairness of a tax system.

In contrast, we find that neither earmarking nor participation affects labor supply, thus suggesting that one feature of a tax could influence different forms of tax responses in a different way. Subjects seem to be more reluctant to change their real behavior than to change their tax avoidance behavior. Not only is this finding important for interpreting previous research, but it could also open several avenues for further research. Taxpayers have several options to react to elements of a tax design, and different elements may cause different types of reactions.

References

Atwood, T. J., Drake, M. S., Myers, J. N., Myers, L. A. 2012. Home Country Tax System Characteristics and Corporate Tax Avoidance International Evidence. *The Accounting Review* 87(6), 1831-1860.

Alm, J., Jackson, B. R., McKee, M. 1993. Fiscal exchange, collective decision institutions, and tax compliance. *Journal of Economic Behavior and Organization* 22(3), 285-303.

Alm, J., McClelland, G. H., Schulze, W. D. 1999. Changing the social norm of compliance by voting. *Kyklos* 52(2), 141-171.

Alm, J., Torgler, B., 2006. Culture differences and tax morale in the United States and Europe. *Journal of Economic Psychology* 27(2), 224-246.

Beuermann, C., Sanatarius, T. 2006. Ecological tax reform in Germany: handling two hot potatoes at the same time. *Energy Policy* 34(8), 917-929.

Blumkin, T., Ruffle, B. J., Ganun, Y. 2012. Are income and consumption taxes ever really equivalent? Evidence from a real-effort experiment with real goods. *European Economic Review* 56(6), 1200-1219.

Buchanan, J. M. 1963. The economics of earmarked taxes. *Journal of Political Economy* 71(5), 467-469.

Collins, Julie B., Murphy, D. P. and Plumlee, R. D. 1992. The Taxpayer's Dilemma: How Hard to Work and What to Report? Advances in Taxation. Volume 4, 31-53, edited by J.J. Stern, JAI Press, Greenwich, Connecticut.

Collins, J. B., Plumlee, R. D. 1991. The Taxpayer's Labor and Reporting Decision: The Effect of Audit Schemes. *The Accounting Review* 66(3), 559-576.

Cullis, J., Jones, P., Savoia, A. 2012. Social norms and tax compliance: Framing the decision to pay tax. *Journal of Socio-Economics* 41(2), 159-168.

De Simone, L., Sansing, R. C., Seidman, J. K. 2013. When are Enhanced Relationship Tax Compliance Programs Mutually Beneficial? *The Accounting Review* 88(6), 1971-1991.

Djanali, I., Sheehan-Connor, D. 2012. Tax affinity hypothesis: Do we really hate paying taxes? *Journal of Economic Psychology* 33(4), 758-775.

Feld, L. P., Frey, B. S. 2007. Tax Compliance as the Result of a Psychological Tax Contract: The Role of Incentives and Responsive Regulation. *Law & Policy* 29(1), 102-120.

Feld, L. P., Tyran, J.-R. 2002. Tax evasion and voting: An experimental analysis. *Kyklos* 55(2), 197-221.

Falsetta, D., Rupert, T. J., Wright, A. M. 2013. The Effect of the Timing and Direction of Capital Gain Tax Changes on Investment in Risky Assets. *The Accounting Review* 88(2), 499-520.

Fochmann, M., Weimann, J. 2013. The effects of tax salience and tax experience on individual work efforts in a framed field experiment. *FinanzArchiv: Public Finance Analysis* 69(4), 511-542.

Fochmann, M., Weimann, J., Blaufus, K., Hundsdoerfer, J., Kiesewetter, D. 2012. Net Wage Illusion in a Real Effort Experiment. *Scandinavian Journal of Economics* 115(2), 476-484.

Gamage, D., Hayashi, A., Nakamura, B. K. 2010. Experimental Evidence of Tax Framing Effects on the Work/Leisure Decision. SSRN Working Paper. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1629919

Goetz, C. J. 1968. Earmarked taxes and majority rule budgetary processes. *American Economic Review* 58(1), 128-136.

Hug, S., Spörri, F. 2011. Referendums, trust, and tax evasion. *European Journal of Political Economy* 27(1), 120-131.

Kallbekken, S., Aasen, M. 2010. The demand for earmarking: Results from a focus group study. *Ecological Economics* 69(11), 2183-2190.

Kastlunger, B., Dressler, S. G., Kirchler, E., Mittone, L., Voracek M. 2010. Sex differences in tax compliance: Differentiating between demographic sex, gender-role orientation, and prenatal masculinization (2D:4D). *Journal of Economic Psychology* 31(4), 542-552.

Keane, M. P. 2011. Labor Supply and Taxes: A Survey. *Journal of Economic Literature* 49(4), 961-1075.

Kim, C. K., Evans III, J. H., Moser, D. V, 2005. Economic and equity effects on tax reporting decisions. *Accounting, Organizations and Society* 30(7-8), 609-625.

Kirchler, E., Maciejovsky, B., Schneider, F. 2003. Everyday representations of tax avoidance, tax evasion, and tax flight: Do legal differences matter? *Journal of Economic Psychology* 24, 535-553.

Lévy-Garboua, L., Masclet, D., Montmarquette, C. A behavioral Laffer curve: Emergence of a social norm of fairness in a real effort experiment. *Journal of Economic Psychology* 30, 147-161.

Mellizo, P., Carpenter, J., Matthews, P. H. 2014. Workplace Democracy in the Lab. *Industrial Relations Journal* 45(4), 313-328.

Möhlmann, A. 2013. Investor home bias and sentiment about the country benefiting from the tax revenue. *Journal of Economic Psychology* 35(4), 31-46.

Richardson, G., Taylor, G., Lanis, R. 2013. The impact of board of director oversight characteristics on corporate tax aggressiveness: An empirical analysis. *Journal of Accounting and Public Policy* 32(3), 68-88.

Sælen, H., Kallbekken, S. 2011. A choice experiment on fuel taxation and earmarking in Norway. *Ecological Economics* 70(11), 2181-2190.

Saez, E., Slemrod, J., Giertz, S. 2012. The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review. *Journal of Economic Literature* 50(1), 3-50.

Sillamaa, M. A. 1999a. Taxpayer behavior in response to taxation: comment and new experimental evidence. *Journal of Accounting and Public Policy* 18(2), 165-177.

Sillamaa, M. A. 1999b. How work effort responds to wage taxation: An experimental test of a zero top marginal tax rate. *Journal of Public Economics* 73(1), 125-134.

Sutter, M., Weck-Hannemann, H. 2003. Taxation and the veil of ignorance – A real effort experiment on the Laffer curve. *Public Choice* 115(1-2), 217-240.

Swenson, C. 1988. Taxpayer behavior in response to taxation: An experimental analysis. *Journal of Accounting and Public Policy* 7(1), 1-28.

Torgler, B. 2005. Tax morale and direct democracy. *European Journal of Political Economy* 21(2), 525-31.

Tyran, J.-R., Feld, L. P. 2006. Achieving Compliance when Legal Sanctions are Non-deterrent. *Scandinavian Journal of Economics* 108(1), 135-156.

Appendix

Appendix I: Instructions work task

You have the opportunity to earn money by fulfilling a task. The wage depends on your personal effort. You can quit the task at any time. Immediately after quitting the task you receive your wage.

The task: In the beginning, you get a number of sheets from a multiple choice test. On the sheet, you will find an identification number and 6 answer options for each of the 60 questions. Your task is to key the identification number and the marked answer option for each question into an already arranged Excel sheet in front of you. Please remember: you can quit whenever you want.

Remuneration: Depending on the number of **error-free** keyed sheets, you receive your wage. For every error-free keyed sheet, you receive a gross wage of **EUR 0.38**. Assuming an average processing time in the range from 1.5 to 2 minutes per sheet, you can achieve a gross wage per hour in a range from EUR 11.40 to EUR 15.20.

Please note: You have to pay a tax on the gross wage. You only receive the net wage (wage after tax).

[Treatment *E, E+P* The tax is earmarked. That means that the tax revenues will be transferred to the economics library and will be spent to purchase textbooks in paper and/or online versions.]

[Treatment *Control*: Your tax will be transferred to the revenue office.]

[Treatment *E+P*: You can influence the tax rate. Therefore, you will receive a questionnaire from the experiment leader. Please mark on the scale how high you feel the tax rate of the earmarked should be. Please return the completed questionnaire to the experiment leader. **Attention**: The overall valid tax rate is calculated as the average (median) tax rate of all questionnaires. Thus, this median tax rate is valid for every participant. In other words: If all participants mark a very low tax rate, the valid tax will also be low. If a high tax rate is chosen by participants, a high tax rate will be applied accordingly. Consequently, you influence the level of the tax rate by your personal decision.]

The valid tax rate for all participants will be announced before you begin your task. Remember, you can quit the task at any time. Immediately after quitting the task you receive your wage after tax.

Based on the gross wage of EUR 0.38 Euro cents per sheet, an average processing time of 1.5 and a valid tax rate of, for example, 40% you would achieve an hourly wage after tax of EUR 9.12. Accordingly, with a tax rate of, for example, 20% you would achieve an hourly wage after tax of EUR 12.16.

Appendix II: Instructions for tax avoidance decision

You have the opportunity to reduce your tax liability through legal tax avoidance. For the tax advisory services you have to pay advisory costs. *[Treatments E, E+P: Please remember that your tax liability will be transferred to the economics library of the Freie Universität Berlin.]*

Which proportion of your tax burden would you like to avoid through tax planning?

The table below shows you which proportion of your taxes you can avoid in percentage of your primary tax liability. Additionally, you can see your net tax savings after tax advisory costs. For example, if you choose tax savings in the amount of 12 % (A), a net tax saving in the amount of 9.1% after tax advisory costs of your primary tax liability will be reached.

A Tax savings in % of your primarily tax liability (gross)	B Tax advisory costs in % of your primarily tax liability	**A - B = net tax savings = total tax savings after tax advisory costs in % of your primarily tax liability**	Please mark as appropriate
0%	0.0%	**0.0%**	
2%	0.1%	**1.9%**	
4%	0.3%	**3.7%**	
6%	0.7%	**5.3%**	
8%	1.3%	**6.7%**	
10%	2.0%	**8.0%**	
12%	2.9%	**9.1%**	
14%	3.9%	**10.1%**	
16%	5.1%	**10.9%**	
18%	6.5%	**11.5%**	
20%	8.0%	**12.0%**	
22%	9.7%	**12.3%**	
24%	11.5%	**12.5%**	
26%	13.5%	**12.5%**	
28%	15.7%	**12.3%**	
30%	18.0%	**12.0%**	
32%	20.5%	**11.5%**	
34%	23.1%	**10.9%**	
36%	25.9%	**10.1%**	
38%	28.9%	**9.1%**	
40%	32.0%	**8.0%**	
42%	35.3%	**6.7%**	
44%	38.7%	**5.3%**	
46%	42.3%	**3.7%**	
48%	46.1%	**1.9%**	
50%	50.0%	**0.0%**	
52%	54.1%	**-2.1%**	
54%	58.3%	**-4.3%**	
56%	62.7%	**-6.7%**	
58%	67.3%	**-9.3%**	
60%	72.0%	**-12.0%**	

Appendix III: Questionnaire

How did you perceive the work task during the experiment on a scale of 1 = very unpleasant to 10 = very pleasant?

Very unpleasant ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very pleasant
1 2 3 4 5 6 7 8 9 10

What influence did the [treatments *E, E+P*: earmarked] tax levied on your willingness to work on a scale from 1 = no influence to 10 = large influence?

No influence ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Large influence
1 2 3 4 5 6 7 8 9 10

How important was the amount of expected gross wage by determining the work duration on a scale from 1 = not important to 10 = very important?

Not important ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very important
1 2 3 4 5 6 7 8 9 10

How important was the amount of expected net wage by determining the work duration on a scale from 1 = not important to 10 = very important?

Not important ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very important
1 2 3 4 5 6 7 8 9 10

Do you think the expected net wage was fair concerning to your work? Please mark the scale from 1 = absolutely unfair to 10 = very fair.

Absolutely unfair ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very fair
1 2 3 4 5 6 7 8 9 10

How do you rate the [treatment *E, E+P*: earmarked] tax rate on a scale from 1 = too low to 1 = too high?

Too low ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ too high
1 2 3 4 5 6 7 8 9 10

How often do you use services provided by the economics library on a scale from 1 = very rarely to 10 = very often?

Very rarely ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very often
1 2 3 4 5 6 7 8 9 10

How often do you use textbooks provided by the economics library on a scale from 1 = very rarely to 10 = very often?

Very rarely ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very often
1 2 3 4 5 6 7 8 9 10

How often do you use e-books provided by the economics library on a scale from 1 = very rarely to 10 = very often?

Very rarely ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very often
1 2 3 4 5 6 7 8 9 10

How important do you consider the further enlargement of the range of online literature of the economics library on a scale from 1 = not important to 10 = very important?

Not important ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very important
1 2 3 4 5 6 7 8 9 10

[Treatment *E, E+P*: Please value your personal advantage of earmarking the tax on a scale from 1 = no advantage to 10 = great advantage.]

No advantage ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Great advantage
1 2 3 4 5 6 7 8 9 10

[*Control* treatment: Imagine that the tax levied in the experiment is earmarked. This means that your tax liability will be fully transferred to the economics library of the Freie Universität Berlin. How much do you value your personal advantage of earmarking the tax on a scale from 1 = no advantage to 10 = great advantage?]

No advantage ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Great advantage

1 2 3 4 5 6 7 8 9 10

[Treatment *E*: Suppose that you can decide on the earmarked tax rate. Please mark on the scale how high should be the tax rate of the earmarked tax.]

7%	9%	11%	13%	15%	17%	19%	21%	23%	25%	27%	29%	31%	33%	35%	37%	39%	41%	43%	45%	47%	49%	51%	53%	55%

[*Control* treatment: Imagine that the tax levied in the experiment is earmarked. This means that your tax liability will be fully transferred to the economics library of the Freie Universität Berlin. Suppose that you can decide on the earmarked tax rate. Please mark on the scale how high should be the tax rate of the earmarked tax.]

7%	9%	11%	13%	15%	17%	19%	21%	23%	25%	27%	29%	31%	33%	35%	37%	39%	41%	43%	45%	47%	49%	51%	53%	55%

[Treatment *E+P*: As how important do you consider your participation in the decision on the rate of the earmarked tax on a scale from 1 = not important to 10 = very important?]

Not important ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very important

1 2 3 4 5 6 7 8 9 10

Please enter the tax rate of the tax that was valid for both you and the other participants in the experiment:

Please enter your course of studies and your current study semester:

For undergraduates: Do you plan an application for a master study at the Freie Universität Berlin?

How old are you?

Are you female or male?

What is your monthly income that becomes available after taxation and social security contributions?

Have you participated in a similar experiment at the Department of Economics at the Freie Universität Berlin?

6 The information content of large book-tax differences – Empirical evidence from Germany

Matthias Sünwoldt

Abstract

The purpose of this paper is to examine the association between large temporary book-tax differences (tBTDs) and earnings persistence for German public companies. Prior literature finds ambiguous results. I extend prior research by precisely controlling for net operating loss carryforwards (NOLCs) that likely distort the tBTD and persistence relation. I observe that earnings persistence is not higher or lower when firms exhibit large positive tBTDs, and I find that large positive tBTDs are not associated with upward earnings management. With respect to temporary tax avoidance as a further potential explanation for large positive tBTDs, I conclude that these tBTDs can be used as a signal for temporary tax avoidance. In contrast, I find relatively low earnings persistence when firms exhibit large negative tBTDs. Hand-collected disclosure information reveals that various accounts affect the large negative level of tBTDs. The lower earnings persistence mainly results from direct and indirect consequences of the 2008 crisis, changing market conditions, and a database misspecification. Because these are somewhat external effects, I suggest that the observed lower earnings persistence is rather spurious.

Keywords

Book-Tax Differences · Earnings Quality · Earnings Persistence · Earnings Management · Tax Avoidance

JEL Classification

H20 · H25 · M41

I thank Sebastian Matenaer, Thomas Hoppe, and several seminar participants at the 2012 ZEW Mannheim Empirical Taxation Research Workshop, and the 2015 Vienna Doctoral Research Seminar for helpful comments and suggestions. I gratefully acknowledge financial support from KPMG.

6.1 Introduction

Tax positions in financial statements (especially deferred tax assets and liabilities and the income tax expense) have been extensively studied in recent years (Graham et al. 2012). This concerns in particular the two, often associated, questions whether these tax positions contain incremental information content and whether and how they are used to manage earnings. The deferred tax expense as a proxy for temporary differences between book income and taxable could provide incremental information content to earnings quality because book income contains temporary accruals that are often excluded from taxable income. Assuming that discretion is higher for accruals than for cash flows and assuming that managers use this discretion to increase book income, earnings quality is reduced. Moreover, even when book and taxable income contain the same kind of accruals (e.g., depreciation), it is likely that managers solely increase the book income to avoid higher tax payments. Nevertheless, temporary tax avoidance as smoothing or lowering the taxable income could also increase the deferred tax expense without affecting book income.

Previous empirical research dealing with the information content of large book-tax differences (BTDs), which include primary temporary and permanent differences between book income and taxable income, shows mixed results. On the one hand, a relatively clear negative relation between large temporary book-tax differences (tBTDs) and earnings persistence as a dimension of earnings quality is revealed (e.g., Hanlon 2005, Blaylock et al. 2012, Tang and Firth 2012, and for Germany, Krummet 2011). On the other hand, Jackson (2015) finds little evidence that large positive tBTDs are associated with upward earnings management. This contradicts the results presented by Blaylock et al. (2012), who find that large positive tBTDs are positively correlated with upward earnings management. Moreover, it has been argued that the low earnings quality is likely influenced by factors other than large tBTDs, including profitability, size, age, and large transitory items (Graham et al. 2012), and that the calculation of tBTDs is distorted by the recognition of net operating loss carryforwards (NOLCs), even when controlling for NOLCs or changes in the valuation allowance (Guenther 2011). Additionally, Lev and Nissim (2004) fail to find a relation between BTDs and earnings growth. Furthermore, it is theoretically unclear why firm-years with large negative tBTDs contain lower earnings persistence, as shown by Hanlon (2005) and Blaylock et al. (2012).

The purpose of this paper is to examine the association between large positive and large negative tBTDs and earnings persistence for German public companies. As described, prior literature does find ambiguous results, and therefore, an additional study seems to be of value. Beside the rather ambiguous prior literature, I expect a significant negative relation between large positive tBTDs and earnings persistence, as shown by Hanlon (2005). Moreover, and in line with Blaylock et al. (2012), I expect that the negative relation is primarily influenced by firm-years categorized as upward earnings management observations. Furthermore, I examine whether tBTDs are correlated with a broader measure of tax avoidance, the cash effective tax rate (CashETR). I expect no differences in earnings persistence for firm-years with large positive tBTDs when these differences are primarily caused by tax avoidance activities. Furthermore, I suppose that the lower earnings persistence for firms with large *negative* tBTDs is mainly caused by underlying real economic effects. Therefore, I analyze the causes

of large negative tBTDs and the link between these differences and low earnings persistence based on hand-collected disclosure information for a small subsample of observations.[33]

I extend prior research by precisely controlling for the deferred tax expense that is caused by NOLCs. Under IFRS (International Financial Reporting Standards), unused NOLCs are recognized as a deferred tax asset when it is probable that the NOLCs can be utilized against future taxable income (IAS 12.36). This recognition generally affects the deferred tax expense, which is, grossed by the statutory tax rate, the common measure for tBTDs. The NOLCs-portion of deferred tax expense does not reflect income measurement differences between book and taxable income, which are the basis for this information content study. The NOLCs-portion solely reflects that today's negative taxable income can be utilized against future positive taxable income. Thus, a deferred tax expense could even arise when there are no income measurement differences (e.g., book income equals taxable income and both are negative). Hanlon (2005) excludes observations that exhibit a net operating loss (Compustat item #52). Following this approach, Guenther (2011) shows that still more than 30% of the influential observations regarding the negative relation between large positive BTDs and earnings persistence exhibit deferred tax assets related to NOLCs. Therefore, precisely controlling for NOLCs seems necessary. Because the NOLCs-portion of deferred tax expense is not separately provided by Datastream or Compustat, I use hand-collected disclosure information for the full sample of 987 firm-year observations.

Surprisingly, I do not observe lower earnings persistence when firms exhibit large positive tBTDs, and I do not find lower earnings persistence for firms that are identified as (presumed) upward earnings managing firms according to several approaches of the Jones model. My results differ from the results presented by Hanlon (2005) and Blaylock et al. (2012) and confirm rather critical prior literature, e.g., Guenther (2011). With respect to tax avoidance as a potential explanation for large positive tBTDs, I find that 27.9% of the firm-years that exhibit large positive tBTDs are categorized as tax avoiders compared to 16.8% (18.6%) of firm-years that exhibit large negative tBTDs (small tBTDs). This indicates that large positive tBTDs can be used as a signal of tax avoidance. As predicted, I do not observe higher or lower earnings persistence when large positive tBTDs likely result from tax avoidance.

In contrast, I find relatively low earnings persistence when firms exhibit large negative tBTDs. Although this is in line with prior literature, it is unclear what causes this relation. Upward earnings management, mainly through discretionary accruals, cannot be used as an explanation in this case. A negative relation between income decreasing earnings management and earnings persistence is a priori unconvincing. I use hand-collected disclosure information for influential observations and find that various accounts affect the level of tBTDs. Thus, for the majority of cases, I do not observe a plain or obvious relation between these high negative tBTDs and earnings persistence. The low earnings persistence mainly results from direct and indirect consequences of the 2008 crisis, changing market conditions, and a database misspecification. Because these are rather external effects, I conclude that the

[33] Guenther (2011) also uses hand-collected disclosure information to determine causes for large positive or large negative tBTDs, but the relation between large tBTDs and low earnings persistence is not examined at the firm-year level.

observed negative relation can be explained by internal management activities only for a small part of the observations.

I contribute to the literature in the following ways. First, I provide additional international evidence on the relation among large tBTDs, earnings quality measured as earnings persistence, upward earnings management, and tax avoidance. Second, I precisely control for distortion due to NOLCs by using hand-collected disclosure data. Third, I investigate the supposed underlying real effects of lower earnings persistence when firms have large negative tBTDs.

6.2 Related literature and hypotheses

6.2.1 Earnings persistence as a proxy for earnings quality

A variety of measures are used in the literature to determine the quality of earnings (see Dechow et al. 2010 for a review). These measures include, for example, earnings persistence, total accruals, discretionary accruals, and income smoothing. In the present work, I use an earnings persistence approach to measure the quality of earnings. The earnings persistence instrument was previously used by Hanlon (2005) and Blaylock et al. (2012) because earnings persistence is, at least theoretically, related to the concept of value relevance (Barth and Hutton 2004), and both Hanlon (2005) and Blaylock et al. (2010) examine investors' assessment of earnings persistence. Earnings persistence describes the sustainability and reproducibility of today's earnings (Dechow et al. 2010). Thus, higher earnings persistence is related to higher decision usefulness in the context of equity valuation. I use the earnings persistence approach in line with Hanlon (2005) and Blaylock et al. (2012). In other words, by using earnings persistence as a proxy for earnings quality, I assume that earnings persistence exhibits decision usefulness with respect to equity valuation. The earnings persistence model can be extended by separating today's earnings into a cash component and an accruals component. Sloan (1996) shows that the cash component exhibits higher earnings persistence than the accruals component. He argues that the differences arise at least partly because the accruals component comes along with a higher degree of subjectivity, or discretion, compared to the cash component. Consequently, earnings management should especially influence the accruals component. Richardson et al. (2005) confirm the results of Sloan (1996). Moreover, they extend the study and find that less reliable components of accruals, such as the change in receivables and inventory or the change in PPE and intangibles, lead to lower earnings persistence.

The studies mentioned above are related to firms that report under US-GAAP (United States Generally Accepted Accounting Principles). Dechow et al. (2010) note that the degree of earnings persistence depends on the accounting system. Atwood et al. (2011) find that the persistence of earnings is not higher or lower for companies reporting under IFRS compared to reporting under US-GAAP or other local accounting principles. Kaserer and Klinger (2008) show that no significant differences arise between US-GAAP and IFRS reporting companies with respect to the different persistence of the cash and accruals component. Thus, I assume that earnings persistence as a proxy for earnings quality is generally applicable to IFRS.

6.2.2 Earnings persistence and BTDs

Large positive tBTDs could include information on earnings quality (Revsine et al. 2005, p. 712-713): upward earnings management may causes a decrease in deferred tax assets or an increase in deferred tax liabilities, both leading to larger positive tBTDs, as well as lower earnings quality. Lev and Nissim (2004) examine the influence of the "tax-to-book income ratio" on earnings growth over a period of five years. They demonstrate that this ratio, which captures temporary and permanent BTDs, is positively correlated with the growth rate of future earnings. Moreover, they show that the temporary component of BTDs exhibits a rather modest ability to predict earnings or stock returns. Hanlon (2005) first analyzes the relation between earnings persistence and large tBTDs. She finds that earnings are less persistent when both large negative and large positive tBTDs occur. The results of Hanlon (2005) and Lev and Nissim (2004) apparently contradict each other with respect to the effect of tBTDs on future earnings. In this context, Jackson (2015) finds that tBTDs are negatively correlated with future pre-tax book income supporting the findings presented by Hanlon (2005).[34] Moreover, he observes a negative correlation between large permanent BTDs and future net book income. Thus, he also confirms the results presented by Lev and Nissim (2004).[35] Based on these results, it seems clear that the analysis of earnings persistence or earnings growth should be based on tBTDs. In this context, Chi et al. (2014) extend the work of Lev and Nissim (2004) and find, in addition to evidence that investors misprice information contained in BTDs, that the ratio of after-tax tBTDs to after-tax book income is negatively related to earnings growth.

In a replication of the Hanlon (2005) study, Guenther (2011) analyzes in detail observations that exhibit large positive tBTDs and low earnings persistence. When he controls for age, large transitory items, large accruals, and high levels of pre-tax ROA, the relation between large tBTDs and low earnings persistence vanishes. Moreover, he finds that a large proportion of the deferred tax expense arises from NOLCs, even when controlling for NOLCs or changes in valuation allowance. Drake (2012) links the relation between tBTDs and earnings persistence with the firm life cycle and finds that after controlling for the firm life cycle stage, the association between large positive tBTDs and lower earnings persistence no longer exists. In line with Guenther (2011), she argues that firm specific characteristics, in this case the firm life cycle, determine the relation between large positive tBTDs and earnings persistence.

Some studies already extent the BTDs and persistence relation to other countries. For instance, Tang and Firth (2012) find that Chinese firms with large positive or large negative BTDs exhibit less earnings persistence. Additionally, they observe that earnings persistence is significantly lower for firms with large abnormal BTDs (differences that occur rather due to earnings management and tax avoiding activities) compared to firms with large normal BTDs (differences that occur rather due to regulatory differences between accounting and tax rules).

---

[34] Jackson (2015) uses separate proxies for temporary and permanent BTDs as independent regression variables whereas Lev and Nissim (2004) use proxies for total (capturing temporary and permanent BTDs) and temporary BTDs, which makes the interpretation of the results more complicated.

[35] Lev and Nissim (2004) observe a positive relationship between the "tax-to-book income ratio" and future net earnings. Since the "tax-to-book income ratio" and the instrument used by Jackson (2015) act oppositely, the economic statement of both studies is identical.

With respect to tBTDs, Martinez and de Souza (2015) confirm the findings presented by Hanlon (2005) for Brazilian and Krummet (2011) for German public companies.

The majority of the aforementioned studies use financial statement information from firms that report under US-GAAP. Both IFRS and US-GAAP determine the recognition and measurement of deferred taxes in accordance with the temporary differences approach.[36] Thus, I assume that the prior findings are c.p. transferable to firms reporting under IFRS. Because Hanlon (2005) and Blaylock et al. (2012) find that large positive and large negative tBTDs are related to lower earnings persistence, I predict that this relation holds for German IFRS financial statements:

H1a: The persistence of earnings is lower for firm-years with large positive or large negative tBTDs compared to firm-years with small tBTDs.

H1b: The persistence of pre-tax accruals for future earnings is lower for firm-years with large positive or large negative tBTDs compared to firm-years with small tBTDs.

6.2.3 Earnings persistence, BTDs, and earnings management

From a theoretical perspective, upward earnings management could be a potential source for large positive tBTDs if the corresponding tax base is not determined simultaneously. tBTDs would simply not occur if the taxable income were determined in accordance with the book income (full book-tax conformity). On the one hand, a firm could report a high book income and a rather low taxable income to keep the tax burden at a low level. On the other hand, the firm could increase the taxable income to avoid reporting large tBTDs, which could be interpreted as a "red flag" by both shareholders and the tax authority (Hanlon and Heitzman 2010, p. 130-131).

Mills and Newberry (2001) show that public firms generally exhibit higher financial reporting costs resulting in larger BTDs. Moreover, opportunistic earnings management to reach bonus thresholds leads to larger BTDs. Phillips et al. (2003) find that the deferred tax expense is informative about earnings management beyond total accruals and abnormal accruals derived from different approaches of the Jones model if the management tries to prevent reporting a loss or an earnings decrease. Blaylock et al. (2012) use a modified Jones model to identify earnings management in connection with large positive tBTDs. They find that earnings persistence is especially low when large positive tBTDs likely arise from upward earnings management. Accordingly, they confirm the findings of Phillips et al. (2003).[37] Jackson (2015) observes a negative association between tBTDs and future changes in pretax earnings. Interestingly, he finds only weak evidence that earnings management contributes to the association between tBTDs and future earnings changes. For instance, the interaction of tBTDs and earnings management, derived from a modified Jones model, has a significant effect on 3-year changes in pretax income but not on 1-and 5-year changes. Moreover, as soon as earnings management is defined as avoiding a loss or avoiding an earnings decline, he does not observe a significant interaction effect between these earnings management measures and

[36] Strictly speaking, minor differences arise between IFRS and US-GAAP, but fundamental differences do not occur (KPMG 2015, p. 55-58).

[37] The calculation of tBTDs by Blaylock et al. (2012) is based on the deferred tax expense.

tBTDs. Hanlon et al. (2012) notice that BTDs are associated with higher audit fees and they conclude that large BTDs are a good proxy for earnings management. Moreover, they argue that tBTDs are especially informative on pre-tax earnings quality. By using a large survey, Heltzer and Shelton (2011) observe that accountants perceive BTDs to be caused by both, upward earnings management and downward management of the taxable income.

Again, the aforementioned studies use financial statement accounts derived under US-GAAP. Jeanjean and Stołowy (2008) investigate to what extend the change from local-GAAP to IFRS changes the earnings management in three different countries (Australia, France, and the U.K.). The authors conclude that the implementation of IFRS does not lead to a decline in earnings management. Liu et al. (2014) find no significant differences between German US-GAAP and German IFRS firms with respect to earnings management measured as earnings smoothing through accruals. However, the results seem to depend on whether the firm's adoption of IFRS is mandatory or voluntary. Barth et al. (2008) and Christensen et al. (2015) find a decrease in earnings management defined as smoothing for German firms voluntarily adopting IFRS. In contrast, they find support for at least a weak increase in smoothing for firms that adopt IFRS mandatorily. Because my sample solely contains firm-years after IFRS became mandatory in 2005, I argue that IFRS firms and US-GAAP firms do not differ essentially with respect to earnings management.

Although Jackson (2015) finds only weak evidence for the association among tBTDs, earnings management, and future earnings, I follow Blaylock et al. (2012) and predict that upward earnings management induces large positive tBTDs, that the upward earnings management leads to lower earnings persistence, and that especially the accrual component of today's earnings exhibits lower earnings persistence in this context. I argue that precisely controlling for deferred taxes resulting from NOLCs should result in lower distortion with respect to the association among large tBTDs, earnings management, and earnings persistence:

H2a: The persistence of earnings is lower for firm-years with large positive tBTDs if the large positive tBTDs are caused primarily by upward earnings management compared to firm-years with large positive tBTDs and less earnings management.

H2b: The persistence of pre-tax accruals for future earnings is lower for firm-years with large positive tBTDs if the large positive tBTDs are caused primarily by upward earnings management compared to firm-years with large positive tBTDs and less earnings management.

6.2.4 Earnings persistence, BTDs, and tax avoidance

The aforementioned studies on BTDs use, at least indirectly, the taxable income as a given reference point to derive evidence whether large BTDs are associated with earnings management. Nevertheless, the taxable income itself could also be managed. In this context, temporary tax avoidance seems to be a likely source for large positive tBTDs. Again, two scenarios are possible. On the one hand, a firm could report low taxable income through tax avoidance and a rather high book income. In this case, large positive tBTDs need to be reported if the tax avoidance strategy does not lead to permanent differences. On the other

hand, the firm could lower the book income to avoid the reporting of tBTDs. This could be the case if the expected costs of recognizing a deferred tax expense would exceed the expected costs of reporting a lower book income, for example, if investors or tax authorities would interpret a large deferred tax expense as a "red flag" (Hanlon and Heitzman 2010, p. 130-131). Thus, tBTDs could only contain information on non-conforming tax avoidance. Moreover, tBTDs capture solely temporary avoiding strategies but they do not capture strategies that lead to permanent tax savings. In other words, tBTDs could be used as an tax avoiding signal only for a specific portion of possible avoiding strategies.

Mills (1998) shows that adjustments by the Internal Revenue Service (IRS) are positively correlated with large positive BTDs. This holds for total BTDs based on company's tax return as well as for tBTDs based on the deferred tax expense. However, it remains unclear whether the large positive BTDs are interpreted as a "red flag" in the selection of companies for tax audits by the IRS or whether large positive BTDs actually are an indicator for tax avoidance strategies. In this regard, Wilson (2009) observes that large positive total, permanent, and temporary BTDs are useful indicators for tax avoidance strategies ("tax shelters"). He concludes that the analyzed tax shelter activities result in a mix of permanent and temporary BTDs. As a result, both studies reveal that tax avoiding activities are often of permanent and temporary nature and that these activities appear simultaneously.

Wilson (2009) also shows that earnings management and tax avoidance are often exercised simultaneously. Accordingly, Frank et al. (2009) find an association between upwards earnings management and tax aggressiveness. In line with that, Tang and Firth (2011) conclude that their abnormal BTDs measure captures both book income and taxable income manipulations. In contrast to these findings, Blaylock et al. (2012) do not observe a high correlation between earnings management and tax avoidance.

When firms report large positive tBTDs because of tax avoidance, lower earnings persistence should theoretically not be observable. In this context, Blaylock et al. (2012) find that large positive tBTD firms exhibit slightly higher earnings persistence when the large positive tBTDs primarily occur due to tax avoidance compared to large positive tBTD firms with less tax avoidance. They argue that the relatively high persistence occurs because large positive tBTD firms engaging in less earnings management. However, this contradicts the observations of Wilson (2009) and Frank et al. (2009). I predict that tax avoidance does not lead to lower earnings persistence when earnings management is low:

H3: *The persistence of earnings is not higher or lower for firm-years with large positive tBTDs if the large positive tBTDs are caused primarily by tax avoidance activities compared to firm-years with large positive tBTDs that occur for other reasons than earnings management and tax avoidance*.

6.3 Research design

The research design is largely in line with Hanlon (2005) and Blaylock et al. (2012). I use the net of NOLC deferred tax expense (the sum of domestic and foreign deferred tax expense minus the deferred tax expense resulting from NOLCs) grossed by the statutory tax rate (38%

until 2006 and 30% since 2007) and divided by average total assets as a proxy for tBTDs.[38] One main extension to prior studies is the precise elimination of the deferred tax expense resulting from NOLCs.[39]

In 2008, the German corporate statutory tax rate decreased from approximately 38% to approximately 30%. Although this decrease does not influence grossing the deferred tax expense as a simple linear transformation, it influences the amount of deferred tax expense itself, especially in 2007, through firms revaluating deferred tax assets and deferred tax liabilities. Depending on the current level of deferred tax assets and liabilities, this leads to a relatively higher or lower deferred tax expense. Because this external effect could distort the results, I control for year fixed effects and run further robustness checks.

In line with Hanlon (2005) and Blaylock et al. (2012), I rank the observations on tBTDs and categorize the observations in three different groups. Group LPBTD (large positive tBTDs) contains observations from the top quintile of yearly tBTDs. Group LNBTD (large negative tBTDs) contains observations from the bottom quintile of yearly tBTDs. All other observations are in group SBTD (small tBTDs).

To test whether large tBTDs contain information about earnings persistence, I estimate the following equation by pooled OLS:

$$PTBI_{it+1} = \gamma_0 + \gamma_1 PTBI_{it} + \gamma_2 LPBTD_{it} + \gamma_3 LNBTD_{it} + \gamma_4 LPBTD_{it} * PTBI_{it} + \gamma_5 LNBTD_{it} * PTBI_{it} + \gamma_n \sum control + \epsilon_{it+1} \quad (5)$$

with $PTBI_{it+1}(PTBI_{it})$ as pre-tax book income of firm i in yea $t + 1$r (t) scaled by average total assets, $LPBTD_{it}$ $(LNBTD_{it})$ as a dummy variable set equal to one for firm-years with scaled tBTDs in the highest (lowest) quintile in year t and $\sum control$ as a vector of control variables, namely, year and industry dummy variables. I expect negative signs for coefficients γ_4 and γ_5.

I portion the pre-tax book income into pre-tax cash flow and pre-tax accruals to analyze earnings persistence in more detail. Prior literature reports that today's accruals exhibit less earnings persistence compared to today's cash flow (Sloan 1996, Richardson et al. 2005, Hanlon 2005). The extended equation reads as follows:

$$PTBI_{it+1} = \gamma_0 + \gamma_1 PTAC_{it} + \gamma_2 PTCF_{it} + \gamma_3 LPBTD_{it} + \gamma_4 LNBTD_{it} + \gamma_5 LPBTD_{it} * PTAC_{it} + \gamma_6 LNBTD_{it} * PTAC_{it} + \gamma_7 LPBTD_{it} * PTCF_{it} + \gamma_8 LNBTD_{it} * PTCF_{it} + \gamma_n \sum control + \epsilon_{it+1} \quad (6)$$

with $PTAC_{it}(PTCF_{it})$ defined as pre-tax accruals (pre-tax cash flow) of firm i in year t scaled by average total assets. I expect negative signs for coefficients γ_5, γ_6, γ_7,and γ_8.

In accordance to Blaylock et al. (2012), I subsequently examine whether large positive tBTDs are the result of upward earnings management or of tax avoidance. As hypothesized, I expect

[38] As of 1 January 2008 the German corporation statutory tax rate is approximately 30%. Because firms valuate their deferred taxes based on tax rates that have been enacted or substantively enacted by the end of the reporting period (IAS 12.47), I grossed up the deferred tax expense in 2007 by 30%. The tax rate is approximated because the actual effective tax rate is especially influenced by the varying German local business tax rate and the amount of foreign income.

[39] See Section 6.4 for details.

a lower earnings quality measured as earnings persistence when large positive tBTDs are the result of upward earnings management and no higher or lower earnings persistence when tBTDs are primarily the result of tax avoidance. I use a cross-sectional Jones model to divide observations into upward earnings management firm-years and baseline firm-years (Jones 1991, Dechow et al. 1995).[40] In detail, I first run the following OLS regression separately for each year: [41]

$$TAC_{it} = \gamma_0 + \gamma_1 1/TotalAssets_{it-1} + \gamma_2 \Delta Revenues_{it} + \gamma_3 PPE_{it} + \epsilon_{it} \quad (7)$$

with TAC_{it} as pre-tax total accruals of firm i in year t scaled by $TotalAssets_{it-1}$[42];
$TotalAssets_{it-1}$ as total assets of firm i in year $t-1$ (lagged total assets);
$\Delta Revenues_{it}$ as change in revenues or net sales of firm i in year t scaled by $TotalAssets_{it-1}$[43];
PPE_{it} as gross property plant and equipment of firm i in year t scaled by $TotalAssets_{it-1}$.

Second, I use the residuals from cross-sectional regressions as firm's i yearly discretionary accruals. In contrast to Blaylock et al. (2012), I do not consider firms' performance as return-on-assets (neither as additional independent variable nor via performance matching). Although, Kothari et al. (2005) conclude that performance-matching is preferable in many of the analyzed cases, my sample consists of too few observations to conduct a well-fitting performance-matching. Moreover, Ayers et al. (2006) and Dechow et al. (2012) argue that performance-matching leads to noise in the discretionary accruals measure, and Keung and Shih (2014) show that performance-matching systematically underestimate discretionary accruals. Nevertheless, Dechow et al. (1995) claim that extreme performance could affect the discretionary accruals estimates. Since a clear recommendation for one of the different approaches does not exists, I use different approaches of the Jones model in the robustness section (see Section 6.5.4).

I categorize the observations into two different groups depending on the calculated discretionary accruals. Group EMpos (large positive discretionary accruals as a signal for upward earnings management) contains firm-years from the top quintile of yearly discretionary accruals. All other observations are categorized in the group EMlow (discretionary accruals around zero or negative).

---

[40] The standard Jones model is a time-series model and not a cross-sectional model. Due to sample restrictions, I run a cross-sectional model. Bartov et al. (2000) find that the cross-sectional Jones model performs even better in detecting earnings management compared to the time-series model.

[41] A preferable strategy would separately regress each year *and* industry. Nevertheless, I regress solely for each year because my sample only contains 987 observations in total and there has to be at least 10 observations in each two-digit industry and year group to run an OLS regression. This condition would hold only for a few industry and year combinations. Moreover, Kothari et al. (2005) find that relaxing the within-industry restriction on Jones model estimations does not essentially change the derived estimations.

[42] I decided not to use the balance-sheet approach used by Kothari et.al. (2005) because Hribar and Collins (2002) show that the error in the balance-sheet approach is correlated with economic characteristics of the firm. Thus, I decided to use the cash-flow approach and to calculate total accruals as residuals. This might mitigate error incurred, for instance, by mergers & acquisitions (Hribar and Collins 2002, Dechow et al. 2010).

[43] I use the change in revenues as an independent variable without any modification because Kothari et al. (2005) demonstrate that the net receivables modification does not significantly increase the model's power. Moreover, I am not convinced that a change in net receivables is fully caused by earnings management.

To test whether the expected lower earnings persistence for LPBTD firm-years results from upward earnings management, I estimate the following equation by pooled OLS within the LPBTD group:

$$PTBI_{it+1} = \gamma_0 + \gamma_1 PTBI_{it} + \gamma_2 EMpos_{it} + \gamma_3 EMpos_{it} * PTBI_{it} + \gamma_n \sum control + \epsilon_{it+1} \quad (8)$$

with $EMpos_{it}$ as a dummy variable set equal to one for firm-years with discretionary accruals in the top quintile and $\sum control$ as a vector of control variables, namely, year and industry variables. I expect a negative sign for coefficient γ_3.

I divide LPBTD firm-years into two subgroups depending on CashETR (cash taxes paid divided by pre-tax book income) to test the alternative explanation that large positive tBTDs are rather the result of tax avoidance. Although tBTDs could be interpreted as a measure for tax avoidance per se, I use the CashETR to identify observations as tax avoider because CashETR captures temporary deferral strategies as well as strategies that are permanent. Thus, the CashETR is a broader measure of tax avoidance. Nevertheless, CashETR as well as tBTDs do not identify conforming tax avoidance. This inherent shortcoming relates to all effective measure of tax avoidance (Hanlon and Heitzman 2010, p. 141). Sine my sample solely include public companies, I suppose that this potential shortcoming is negligible.

The TaxAvoid group (presumable tax-avoiding firms) contains observations from the bottom quintile of CashETR on a yearly basis. All other observations are in the group NoTaxAvoid. I then run the following regression:

$$PTBI_{it+1} = \gamma_0 + \gamma_1 PTBI_{it} + \gamma_2 EMpos_{it} + \gamma_3 TaxAvoid_{it} + \gamma_4 EMpos_{it} * PTBI_{it} + \gamma_5 TaxAvoid_{it} * PTBI_{it} + \gamma_n \sum control + \epsilon_{it+1} \quad (9)$$

with $TaxAvoid_{it}$ as a dummy variable set equal to one for firm-years with CashETR in the lowest quintile. I expect that the coefficient γ_5 is not significantly different from zero.

I use pooled OLS with clustered standard errors at the firm level for all of the earnings persistence equations because of potential heteroscedasticity. The small p-value (0.001) of a Breusch-Pagan test based on Equation (5) indicates that heteroscedasticity might be present.

6.4 Sample selection

I use Thomson Reuters Datastream with Worldscope Fundamentals and hand-collected disclosure information. I examine consolidated financial statements under IFRS for CDAX® companies in Germany from 2004 to 2013. CDAX® includes shares of all domestic companies listed in Prime Standard and General Standard and, thus, represents the German equity market in its entirety. My sample effectively starts in 2005 because I solely use the reported total assets from 2004 to calculate the average total assets for 2005. I effectively start with 2005 because IFRS became mandatory for public German companies in 2005 (for consolidated financial statement, which I use). Moreover, the number of observations before 2005 is rather small. Classification into large tBTDs, into upward earnings management, and into tax avoidance is prepared on an annual basis and thus requires a minimum amount of observations each year.

In February 2015, Datastream included 438 different CDAX® companies.[44] I exclude 54 companies from the financial sector (especially commercial banks and insurance companies but also real estate firms; Worldscope industry group ID 4300) because of different regulation and accounting rules. Additional adjustments are required on a firm-year basis. First, I eliminate all observations with missing or local (US-GAAP or German-GAAP) accounting standards (736 firm-years). Second, I exclude observations with negative pre-tax income (731 firm-years) because loss-years are, on average, less persistent than profitable years, mainly because of curtailments (Hayn 1995, Lawrence et al. 2015) or conservative accounting (Basu 1997). In addition, loss-generating firms likely behave differently in terms of earnings management and tax avoidance. Third, I eliminate observations with required but missing variables (1,386 firm-years).[45] In sum, my sample includes 987 firm-year observations.[46]

The portion of deferred tax expense related to changes in recognized NOLCs as deferred tax assets does not qualify as "true" BTD and should therefore be eliminated. The rational behind that is that the NOLCs-portion of deferred tax expense does not reflect income measurement differences between book and taxable income, which are the basis for this information content study. The NOLCs-portion solely reflects that today's negative taxable income can be utilized against future positive taxable income. Since I focus on income differences, the NOLCs-portion should be excluded.

My final sample does not include observations with negative pre-tax book income in the previous year or in the next year. Thus, distortional effects due to NOLCs should automatically be diminished. Nevertheless, Guenther (2011) shows that NOLCs still affect the deferred tax expense even when loss-years are excluded. Therefore, precisely controlling for NOLCs seems to be required. I use hand-collected disclosure information because Datastream does not provide information on the deferred tax expense resulting from NOLCs. I follow two different approaches to generate this information. As a first approach, I try to directly identify the deferred tax expense that is related to NOLCs from the footnote disclosures. This approach precisely identifies the NOLCs-portion, but the information is not directly mandatory under IFRS. IAS 12.79 demands that firms report the major components of the tax expense separately. This could be, but clearly does not have to be, the NOLCs portion. Thus, I observe this first step information only for 359 firm-years (36.4% of the sample). Consequently, an alternative approach is necessary to derive the required information for the remaining 628 firm-years. For this approach, I use the change in deferred tax assets resulting from NOLCs recognized in the balance sheet as a proxy for the NOLC deferred tax expense. Under IAS 12.81, firms have to report the amount of deferred tax assets resulting from NOLCs, which is recognized in the balance sheet. Thus, if firms fully follow this provision, I should be able to determine the change in deferred tax assets resulting from NOLCs for my full sample. I find that all sample firms follow IAS 12.81 but to various extents. The vast

[44] The actual number of different CDAX® shares is higher than 438 because of preferred stocks. To avoid double entities, I eliminate these shares.

[45] Datastream presents domestic and foreign tax expenses separately. I assume that foreign tax expense equals zero if the actual foreign tax expense value is missing. Otherwise, I would lose another 498 observations (more than half of my final sample).

[46] Calculation: (438 - 54) companies * 10 years = 3,840 firm-years; 3,840 - 736 - 731 - 1,386 = 987 firm-years.

majority of firms directly disclose the amount of deferred tax assets resulting from NOLCs, and the calculation is straightforward. A minority of firms disclose the NOLC deferred tax assets before any write-downs on these deferred tax assets and report these write-downs separately. If this is the case, I use the change in deferred tax assets resulting from NOLCs minus the change in write-downs on deferred tax assets resulting from NOLCs as a proxy for the NOLC deferred tax expense. For 33 firm-years, I observe that the corresponding firms report their NOLC deferred tax assets before any write-downs on deferred tax assets from NOLCs and the aggregated deferred tax assets write-downs separately. These aggregated write-downs likely contain write-downs related to NOLCs and write-downs related to other temporary differences recognized as deferred tax assets. For these 33 observations, I assume that the aggregated write-downs arise solely through NOLC deferred tax assets.[47]

The deferred tax expense from NOLCs derived using the first approach is preferable compared to the proxies from the second approach because the second approach values could be distorted by currency translations and/or consolidation effects. Nevertheless, I am able to derive the first and second approach values for 359 firm-years and find identical values for 183 observations (51.0%). Moreover, I observe for 271 observations (75.5%) less than 25% upper or lower deviation if I compare the first with the second approach. In addition, an overall highly significant correlation between the first and the second approach is observable ($n = 359$; Pearson's p-value < 0.001; Spearman's p-value < 0.001). Thus, a probable distortion by currency translations and/or consolidation effects presumably does not have a strong effect on the calculation of the second approach proxies.

6.5 Results

6.5.1 Descriptive statistics

I start with a brief discussion of the statistics on the effects of the deferred tax expense from NOLCs and then present the main descriptive statistics. I find that 889 firm-years report deferred tax expenses from NOLCs different from zero (90.1% of the sample). This finding is, at first glance, surprising because firm-years with previous, current, or next-year negative pre-tax book income are not included in the full sample. The observed deferred tax expenses from NOLCs are (1) caused by single subsidiaries that generated losses, although the group achieved a profit; (2) caused by changes in the tax group regime along with revaluating the recognition of unused NOLCs as deferred tax assets; (3) caused by recognized NOLCs after acquisitions.

To determine the magnitude of the deferred tax expense from NOLCs in relation to the total deferred tax expense, I calculate a NOLCs-ratio as the absolute value of deferred tax expense from NOLCs to the absolute value of total deferred tax expense (NOLCs portion and the temporary portion). By following this approach, I derive a NOLCs-ratio with values between zero (no deferred tax expense from NOLCs; no distortion) and one (the total amount of

[47] For five firm-years, the aggregated write-downs exceed the deferred tax assets from NOLCs at least for one year. For these observations, I set the deferred tax assets from NOLCs equal to zero. As a robustness test, I assume that the aggregated write-downs arise solely for reasons other than NOLCs. The results remain the same.

deferred tax expense results from NOLCs; extensive distortion). Moreover, I avoid a potential offsetting effect. Consider, for instance, a firm that reports EUR 100 deferred tax expense from temporary differences and EUR 100 deferred tax income (or "negative" expense) from NOLCs. The deferred tax expense would be zero, although the deferred tax income from NOLCs is responsible for 50% of the total deferred tax position. The NOLCs-ratio captures the 50%. The mean (median) NOLCs-ratio is 0.40 (0.39) with a standard deviation equal to 0.30. Thus, on average, 40% of the absolute sum of deferred tax expense from NOLCs and from temporary differences is due to deferred tax expense from NOLCs. Consequently, deferred taxes from NOLCs seem to play a very important role in determining the deferred tax expense in my sample.

In any case, the main results on earnings persistence remain unchanged as long as the elimination of the deferred tax expense from NOLCs does not change the categorization into the three different groups of LPBTD (large positive tBTDs), LNBTD (large negative tBTDs) and SBTD (small tBTDs). I carry out the categorization based on tBTDs with and without controlling for the deferred tax expense from NOLCs to illustrate the influence of this deferred tax proportion on the group categorization. Table 25 presents the number of observations for each group and each categorization strategy as well as the number of observations for the group and strategy interaction. If considering the deferred tax expense from NOLCs would not change the categorization, one would expect zero values besides the main diagonal. It seems that the correction partly influences the categorization. For instance, only 108 out of 195 observations are categorized as large positive tBTD firm-years independently of a correction of deferred tax expense from NOLCs (55.4%). Of the 195 observations, 80 observations are categorized as large positive tBTD firm-years after the NOLC correction is performed, but these observations are categorized as small tBTD firm-years before the correction is carried out. Seven out of 195 observations are categorized as large positive tBTD firm-years after the NOLCs correction, but they are actually categorized as large negative tBTD firm-years before the correction. Summing up, the deferred tax expense from NOLCs is responsible for a large portion of the total deferred tax expense and deferred tax expense from NOLCs considerably influences the categorization based on tBTDs. Thus, a correction of the NOLC portion of deferred tax expense seems necessary. In the following, tBTDs are adjusted for deferred tax expense from NOLCs unless otherwise stated.

Table 18: Group categorization based on tBTDs and NOLCs correction

		Categorization of observations based on tBTDs after NOLCs correction (numbers)			
		LPBTD	LNBTD	SBTD	Total
Categorization of observations based on tBTDs before NOLCs correction (numbers)	LPBTD	108	9	78	195
	LNBTD	7	137	58	202
	SBTD	80	56	454	590
	Total	195	202	590	987

Table 26 presents descriptive statistics separately for firm-years in the LPBTD group (Panel A), the LNBTD group (Panel B) and the SBTD group (Panel C). The categorization based on tBTDs seems to be effective because the LPBTD (LNBTD) group solely contains positive (negative) tBTD values. The minimum value of -4.552 in Panel B seems to be unlikely at first glance. The corresponding firm reports a rather moderate deferred tax income (TEUR 36), but the hand-collected information reveals that the deferred tax expense from NOLCs amounts to TEUR 35,950. Thus, the deferred tax income from differences other than NOLCs, which is the basis for calculating the tBTDs, amounts to TEUR 35,986, corresponding 125% of total assets.[48]

LPBTD firms are more profitable in comparison to SBTD firms ($PTBI_{it}$ interpreted as performance measure; p-value < 0.05) and LNBTD firms (p-value < 0.10).[49] Firms in the LNBTD group are smaller compared to firms in the SBTD group (mean total assets; p-values < 0.10). Pre-tax accruals are, on average, largest for LPBTD firms and smallest for LNBTD firms, but only the differences between LNBTD and SBTD and LNBTD and LPBTD observations are statistically significant (p-values < 0.001).

Surprisingly, the discretionary fraction of pre-tax accruals, derived from the Jones model, is not significantly higher for LPBTD firms compared to SBTD firms. Thus, it seems at least unlikely that the large positive tBTDs are the result of upward earnings management. In contrast, LNBTD firms exhibit on average negative discretionary accruals, which are significantly smaller compared to the discretionary accruals of SBTD firms (p-value < 0.001). That could indicate that LNBTD firms use tBTDs to manage earnings downwards.

The CashETR is significantly different between the three groups, with highest values for LNBTD firms and lowest values for LPBTD firms (p-values < 0.001, except p-value < 0.05 for CashETR between LPBTD and SBTD). One reason could be a higher tax planning activity for LPBTD firms, as predicted. Another reason could be a rather mechanical, or, in other words, non-discretional relation between tBTDs and the effective tax rate. A third explanation could be that firms engage in tax planning that results in large positive tBTDs and low effective tax rates for a certain time. After a while, these firms report large negative tBTDs and large effective tax rates because the tax planning strategies captured by tBTDs are only temporary (mean reversion). However, this may seem unrealistic because these firms could develop or replace tax planning strategies to avoid this change. Since the true tax avoidance is not observable, an unambiguous interpretation is hard to derive.

Summing up, descriptive statistics indicate that upward earnings management seems not to be the driving force behind large positive tBTDs, which is contrary to the findings especially of Blaylock et al. (2012). However, tBTDs are negatively correlated with effective tax rates, which may indicate tax planning activities.

[48] Disclosure information reveals that the high deferred tax expense from NOLCs results from a tax audit and that the high deferred tax income from differences other than NOLCs results mainly from changes in the impairment of receivables. Interestingly, both effects nearly offset one another.

[49] All p-values refer to a mean-comparison test (two-sample and two-tailed t-test).

Table 19: Descriptive statistics separately for LPBTD, LNBTD and SBTD

Variable definition: $tBTDs_{it}$ = sum of foreign and domestic deferred tax expense grossed up by statutory tax rate (since 2007 = 30%, prior to 2007 = 38%) scaled by average total assets ((WC18188+WC18189)/tax rate/$AvAssets_{it}$); $PTBI_{it}$ = pre-tax book income scaled by average total assets (WC01401/$AvAssets_{it}$); $PTAC_{it}$ = pre-tax accruals as pre-tax book income minus pre-tax cash flow scaled by average total assets (WC01401-(WC04860+WC04150)/$AvAssets_{it}$); $PTCF_{it}$ = pre-tax cash flow scaled by average total assets ((WC04860+WC04150)/$AvAssets_{it}$); $AvAssets_{it}$ = average total assets ((one year lagged WC02999+WC02999)/2); $Assets_{it}$ = total assets in Billion EUR (WC02999/1,000) $DisAcc_{it}$ = discretionary accruals from a cross-sectional Jones model (year based); $CashETR_{it}$ = cash taxes paid divided by pre-tax book income (WC04150/WC01401). All Datastream financial statement variables are winsorized at 1st and 99th percentiles.

Panel A: LPBTD group (large positive tBTDs)

Variable	N	Mean	Std. Dev.	Minimum	Median	Maximum
$tBTDs_{it}$	195	0.031	0.032	0.005	0.022	0.255
$PTBI_{it}$	195	0.110	0.074	0.001	0.092	0.353
$PTAC_{it}$	195	-0.013	0.086	-0.239	-0.021	0.526
$PTCF_{it}$	195	0.123	0.092	-0.426	0.112	0.511
$Assets_{it}$	195	7,195	21,305	21	291	133,217
$DisAcc_{it}$	195	0.010	0.101	-0.414	0.006	0.827
$CashETR_{it}$	195	0.260	0.195	0	0.234	1

Panel B: LNBTD group (large negative tBTDs)

Variable	N	Mean	Std. Dev.	Minimum	Median	Maximum
$tBTDs_{it}$	202	-0.057	0.319	-4.552	-0.028	-0.011
$PTBI_{it}$	202	0.097	0.081	0.001	0.073	0.533
$PTAC_{it}$	202	-0.043	0.081	-0.351	-0.043	0.513
$PTCF_{it}$	202	0.139	0.093	-0.186	0.124	0.518
$Assets_{it}$	202	6,738	17,771	21	680	133,217
$DisAcc_{it}$	202	-0.019	0.087	-0.410	-0.054	0.615
$CashETR_{it}$	202	0.368	0.243	0	0.329	1

Panel C: SBTD group (small tBTDs)

Variable	N	Mean	Std. Dev.	Minimum	Median	Maximum
$tBTDs_{it}$	590	-0.001	0.007	-0.022	-0.001	0.018
$PTBI_{it}$	590	0.097	0.081	0.001	0.075	0.544
$PTAC_{it}$	590	-0.023	0.071	-0.346	-0.029	0.400
$PTCF_{it}$	590	0.120	0.094	-0.294	0.102	0.716
$Assets_{it}$	590	10,544	28,132	8	799	133,217
$DisAcc_{it}$	590	0.003	0.077	-0365	0.001	0.497
$CashETR_{it}$	590	0.299	0.196	0	0.278	1

6.5.2 Tests of Hypotheses H1a and H1b

Table 27 presents the results from pooled OLS regression for Hypothesis H1a. The coefficient on $PTBI_{it}$ from model (1) confirms strong earnings persistence. The magnitude of 0.746 is in line with prior findings (Sloan 1996, Hanlon 2005). In model (2), dummy variables and interaction terms for large positive and large negative tBTDs are included. Interestingly, firm-years with large positive tBTDs do not have less persistence earnings compared to firm-years with small tBTDs. The coefficient on $LPBTD_{it}$x$PTBI_{it}$ amounts to -0.047, indicating a lower persistence, but is not significantly different from zero. Thus, I cannot confirm Hypothesis H1a with respect to large positive tBTD firm-years. In contrast, firm-years with large negative tBTDs exhibit lower earnings persistence compared to firm-years with small tBTDs. The coefficient on the interaction $LPBTD_{it}$x$PTBI_{it}$ amounts to -0.158 (p-value = 0.088). The

economic magnitude is rather large, resulting in a decrease in earnings persistence of approximately 20%. With respect to large negative BTD firm-years, the finding is in line with Hypothesis H1a. The results remain qualitatively unchanged when I consider year fixed effects in model (3) and year and industry fixed effects in model (4).

Table 20: OLS estimations for Hypothesis H1a

Variable definition: $PTBI_{it+1}$ ($PTBI_{it}$) = pre-tax income of firm i in year t+1 (t) scaled by average total assets (WC01401/ $AvAssets_{it}$); $LPBTD_{it}$ ($LNBTD_{it}$) = dummy variable equals one if book-tax difference of firm i scaled by average total assets in year t is in the top (bottom) quintile of all tBTDs in year t; $AvAssets_{it}$ = average total assets ((one year lagged WC02999+WC02999)/2); Year Fixed Effects = dummy variable for each year (basis year is 2005); Industry Fixed Effects = dummy variable for each industry group (WC06011 first two digits) if industry group contains at least 30 observations (11 different industry groups). Clustered (firm-level) standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1.

	(1)	(2)	(3)	(4)
Dep. variable:	$PTBI_{it+1}$	$PTBI_{it+1}$	$PTBI_{it+1}$	$PTBI_{it+1}$
$PTBI_{it}$	0.746***	0.789***	0.797***	0.796***
	(0.0371)	(0.0495)	(0.0495)	(0.0493)
$LPBTD_{it}$		0.00265	0.00170	0.00333
		(0.00562)	(0.00551)	(0.00579)
$LNBTD_{it}$		0.0232**	0.0239**	0.0253**
		(0.00962)	(0.00965)	(0.00975)
$LPBTD_{it}$ x $PTBI_{it}$		**-0.0468**	**-0.0392**	**-0.0441**
		(0.0657)	(0.0648)	(0.0663)
$LNBTD_{it}$ x $PTBI_{it}$		**-0.158***	**-0.166***	**-0.170***
		(0.0919)	(0.0931)	(0.0927)
Constant	0.0196***	0.0142***	0.0238***	0.0245***
	(0.00330)	(0.00379)	(0.00587)	(0.00664)
Year Fixed Effects	No	No	Yes	Yes
Industry Fixed Effects	No	No	No	Yes
Observations	987	987	987	987
R-squared	0.616	0.623	0.640	0.644

I present the results for Hypothesis H1b in Table 28. In this setting, today's pre-tax income is separated into pre-tax accruals and pre-tax cash flow. In line with prior literature, I find that the accruals fraction of pre-tax book income is less persistent than the cash flow fraction of pre-tax book income (coefficient on $PTAC_{it}$ ($PTCF_{it}$) in model (1) = 0.669 (0.775); F-test reveals a significant difference at p-value < 0.001). In model (2), neither the interaction $LPBTD_{it}xPTAC_{it}$ nor the interaction $LPBTD_{it}xPTCF_{it}$ provide significant coefficients. Thus, I have to reject H1b for large positive tBTD firm-years. The results for large negative tBTD firm-years are again different. I find a significant coefficient on $LNBTD_{it}xPTAC_{it}$ (coefficient = -0.247; p-value = 0.033) but no significant coefficient on $LNBTD_{it}xPTCF_{it}$ (coefficient = 0.103; p-value = 0.214). It seems that pre-tax accruals are strongly responsible for the lower earnings persistence, whereas pre-tax cash flows explain far less. The results remain qualitatively unchanged when I add year fixed effects in model (3) and year and industry fixed effects in model (4).

Table 21: OLS estimations for Hypothesis H1b

Variable definition: $PTBI_{it+1}$ = pre-tax income of firm i in year t+1 scaled by average total assets (WC01401/ $AvAssets_{it}$); $PTAC_{it}$ = pre-tax accruals of firm i in year t as pre-tax book income minus pre-tax cash flow scaled by average total assets (WC01401-(WC04860+WC04150)/$AvAssets_{it}$); $PTCF_{it}$ = pre-tax cash flow of firm i in year t scaled by average total assets ((WC04860+WC04150)/$AvAssets_{it}$); $LPBTD_{it}$ ($LNBTD_{it}$) = dummy variable equals one if book-tax difference of firm i scaled by average total assets in year t is in the top (bottom) quintile of all tBTDs in year t; $AvAssets_{it}$ = average total assets ((one year lagged WC02999+WC02999)/2); Year Fixed Effects = dummy variable for each year (basis year 2005); Industry Fixed Effects = dummy variable for each industry group (WC06011 first two digits) if industry group contains at least 30 observations (11 different groups). Clustered (firm-level) standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1.

	(1)	(2)	(3)	(4)
Dep. variable:	$PTBI_{it+1}$	$PTBI_{it+1}$	$PTBI_{it+1}$	$PTBI_{it+1}$
$PTAC_{it}$	0.669***	0.744***	0.753***	0.755***
	(0.0371)	(0.0632)	(0.0628)	(0.0625)
$PTCF_{it}$	0.775***	0.801***	0.807***	0.806***
	(0.0371)	(0.0493)	(0.0492)	(0.0488)
$LPBTD_{it}$		-0.000603	-0.000201	0.00131
		(0.00665)	(0.00647)	(0.00667)
$LNBTD_{it}$		0.0106	0.0122	0.0135
		(0.0101)	(0.0101)	(0.0103)
$LPBTD_{it}$ x $PTAC_{it}$		**-0.0610**	**-0.0428**	**-0.0465**
		(0.0710)	(0.0697)	(0.0712)
$LNBTD_{it}$ x $PTAC_{it}$		**-0.247****	**-0.250****	**-0.255****
		(0.115)	(0.111)	(0.112)
$LPBTD_{it}$ x $PTCF_{it}$		**-0.0186**	**-0.0210**	**-0.0258**
		(0.0728)	(0.0707)	(0.0722)
$LNBTD_{it}$ x $PTCF_{it}$		**-0.103**	**-0.115**	**-0.119**
		(0.0824)	(0.0846)	(0.0842)
Constant	0.0141***	0.0118***	0.0221***	0.0230***
	(0.00356)	(0.00442)	(0.00645)	(0.00715)
Year Fixed Effects	No	No	Yes	Yes
Industry Fixed Effects	No	No	No	Yes
Observations	987	987	987	987
R-squared	0.627	0.635	0.650	0.653

Because the above-stated results are partly surprising, I run several robustness checks. First, I successfully replicate the results reported by Hanlon (2005) for US firms and Krummet (2011) for German firms to test whether my research approach is in line with prior studies (untabulated).

Second, I run the same OLS regressions but without adjusting the overall tBTDs for the NOLC portion of deferred tax expense. This approach is similar to the prior work of Hanlon (2005) and Blaylock et al. (2012). As described above, I observe that NOLCs are responsible for a large portion of total tBTDs and that controlling for these NOLCs slightly changes the categorization results. Nevertheless, the OLS results remain qualitatively the same for LPBTD firms. I do not observe lower or higher earnings persistence when firms exhibit large positive tBTDs, even when I do not adjust for NOLCs (untabulated results). In contrast, I find much lower earnings persistence for firms with high negative tBTDs (coefficients on $LNBTD_{it}$x$PTBI_{it}$ between -0.219 and -0.230; p-values between 0.024 and 0.032, depending on model specifications). This also holds for the accruals fraction of pre-tax book income

(coefficients on $LNBTD_{it}xPTAC_{it}$ between -0.305 and -0.316; p-values between 0.003 and 0.006, depending on model specifications) and the cash flow fraction of pre-tax book income (coefficients on $LNBTD_{it}xPTCF_{it}$ between -0.170 and -0.182; p-values between 0.073 and 0.090, depending on model specifications). It turns out that the deferred tax expense from NOLCs does not influence the OLS results with respect to LPBTD firms but strengthens the lower earnings persistence for LNBTD firms.

Third, I exclude 2007 and 2008 observations because of changes in the German corporate tax system due to the 2008 tax reform (reduced tax rate but generally broadened tax base). This leads to a relatively higher or lower deferred tax expense through revaluating existing deferred tax positions, especially in 2007. This external effect could distort the results. I still do not observe lower earnings persistence for LPBTD firm-years, and I still find significantly lower earnings persistence for the accrual component of LNBTD firm-years (p-values < 0.10). In contrast, the coefficient on $LNBTD_{it}xPTBI_{it}$ is no longer significantly different from zero.[50]

Fourth, I include the total amount of accruals scaled by average total assets as additional independent variable and as interaction with $PTBI_{it}$ or $PTAC_{it}$ and $PTCF_{it}$ because earnings persistence is decreasing in the absolute size of accruals (Dechow and Ge 2006, Blaylock et al. 2012). The results reveal no lower earnings persistence for LPBTD firm-years but a significantly lower earnings persistence for LNBTD firm-years with respect to $PTBI_{it}$ (coefficients on $LNBTD_{it}xPTBI_{it}$ between -0.149 and -0.165; p-values between 0.041 and 0.059, depending on model specifications) as well as for $PTAC_{it}$ (coefficients on $LNBTD_{it}xPTAC_{it}$ between -0.236 and -0.250; p-values between 0.024 and 0.014, depending on model specifications) and $PTCF_{it}$ (coefficients on $LNBTD_{it}xPTCF_{it}$ between -0.115 and -0.132; p-values between 0.133 and 0.089, depending on model specifications).

Sixth, I use the grossed change in deferred taxes derived from the balance sheet (change in Datastream ID WC03263) as a proxy for tBTDs. An increase in recognized deferred taxes results from an increased gap between book income and taxable income whereby book-income exceeds taxable income. Thus, this increase may be driven by earnings management resulting in lower earnings persistence. One disadvantage is that this proxy does not include recognized deferred tax assets in terms of deductible temporary differences. However, a large advantage is that this proxy should not include other *pseudo* tBTDs, for instance, arising from recognized tax credits. Once more, the results remain qualitatively stable. I find lower earnings persistence for LNBTD firm-years and no difference in earnings persistence for LPBTD firm-years.

Seventh, I use a time and entity fixed effects model to test equations (1) and (2) instead of pooled OLS with clustered standard errors.[51] Again, the results remain the same. I do not observe lower earnings persistence when firms exhibit large positive tBTDs, and I find a lower earnings persistence when firms exhibit large negative tBTDs (p-values between 0.052 and 0.070).

---

[50] Dropping the 2007 and 2008 observations reduces the sample to 748 firm-years. Thus, the reduced sample size may influence the results.

[51] I do not use the fixed effects model as the default model because it is likely biased due to the lagged values of the dependent variable being used as a right-hand-side variable.

Summing up, my findings are quite robust. Notwithstanding, special items may affect the results. Hanlon (2005) controls for special items by substituting operating earnings after depreciation for pre-tax book income. This excludes the effect of special items. Although this is indeed an important robustness check, it is not possible to run this test with IFRS financial statements. IFRS does not follow the concept of special items at all. An alternative approach would be to control for profit or loss from discontinued operations. Datastream delivers pre-tax income (WC01401) and income from discontinued operations (WC01505) separately. Thus, WC01401 does not contain income from discontinued operations. Therefore, my variable $PTBI_{it}$ controls for discontinued operations per se. As an alternative, I generate scatter plots as added variable plots, including $PTAC_{it}$, $PTCF_{it}$, and $PTBI_{it+1}$, to address the special items problem. In detail, I concentrate on outliers as influential observations in the three groups. These outliers may affect coefficients and standard errors. I estimate the following OLS equation separately for each group to generate the plots.

$$PTBI_{it+1} = \gamma_0 + \gamma_1 PTAC_{it} + \gamma_2 PTCF_{it} + \epsilon_{it+1} \quad (10)$$

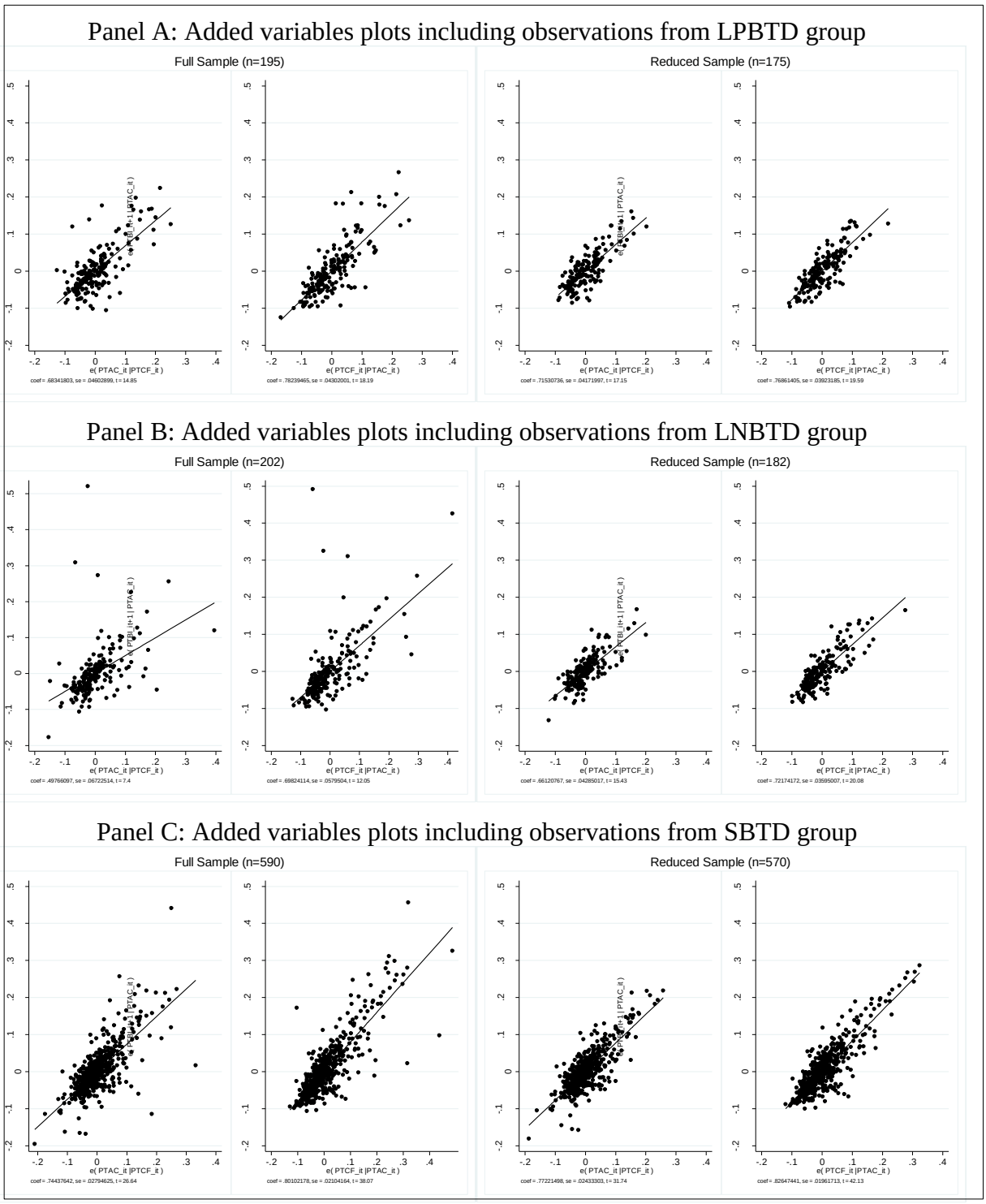
Figure 21 presents added variable plots. There are potential outliers in each of the three groups when I consider the full sample (987 observations), but with different extents. The groups SBTD and LPBTD are rather equal and do not exhibit clear outliers (graphs on the left in Panels A and C). Against that, group LNBTD seems to exhibit some outliers that might influence the above-presented results (plot on the left in Panel B). Thus, I estimate Equation (10) for the full sample and predict Cook's D (Cook 1977).[52] I categorize an observation as influential if Cook's D exceeds the value of 0.00405.[53] In total, 60 firm-years are identified as influential observations. The plots on the right in Figure 21 display added variable plots for each group without these influential observations.

[52] Cook's D combines information about potential outliers (large residuals) and leverage (observations with extreme values for right-hand-side variables). The higher the Cook's D, the more influential the observation.

[53] This typical threshold is computed as 4/number of observations (Bollen and Jackman 1990).

Figure 14: Added variables plots separately for each group

Added variables plots show $PTBI_{it+1}$ by $PTAC_{it}$ after adjusting for $PTCF_{it}$ in the model and vice versa. For instance, computing the residuals of regressing $PTCF_{it}$ against $PTBI_{it+1}$ but omitting $PTAC_{it}$, then computing the residuals from regressing $PTAC_{it}$. against $PTCF_{it}$, and finally, plotting the former residuals against the latter residuals to obtain the plot for $PTAC_{it}$. I run OLS regression Equation (10) separately for each group. The line plotted has the same slope as the coefficient for $PTAC_{it}$ or $PTCF_{it}$. Full sample plots contain all observations. Reduced sample plots contain observations under the condition that Cook's D is at most 4/987. Observations above this threshold are excluded.



I exclude observations with Crook's D above 0.00405 and run OLS regressions as a robustness check. I present the results in Table 29 and Table 30.

Table 22: OLS estimations for Hypothesis H1a after excluding influential observations
Variable definition: See Table 27. Clustered (firm-level) standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1.

	(1)	(2)	(3)	(4)
Dep. variable:	$PTBI_{it+1}$	$PTBI_{it+1}$	$PTBI_{it+1}$	$PTBI_{it+1}$
$PTBI_{it}$	0.783***	0.813***	0.820***	0.816***
	(0.0197)	(0.0209)	(0.0207)	(0.0210)
$LPBTD_{it}$		0.00191	0.00194	0.00152
		(0.00455)	(0.00455)	(0.00458)
$LNBTD_{it}$		0.0135***	0.0152***	0.0157***
		(0.00379)	(0.00373)	(0.00378)
$LPBTD_{it}$ x $PTBI_{it}$		**-0.0647**	**-0.0665**	**-0.0592**
		(0.0517)	(0.0519)	(0.0522)
$LNBTD_{it}$ x $PTBI_{it}$		**-0.106****	**-0.127*****	**-0.128*****
		(0.0436)	(0.0445)	(0.0449)
Constant	0.0141***	0.0113***	0.0236***	0.0246***
	(0.00153)	(0.00177)	(0.00417)	(0.00461)
Year Fixed Effects	No	No	Yes	Yes
Industry Fixed Effects	No	No	No	Yes
Observations	927	927	927	927
R-squared	0.737	0.741	0.755	0.757

Again, I do not find significantly lower earnings persistence for LPBTD firm-years with respect to current pre-tax book income in Table 29 (coefficients on $LPBTD_{it}xPTBI_{it}$) or for pre-tax accruals and pre-tax cash flow as components of pre-tax book income in Table 30 (coefficients on $LPBTD_{it}xPTAC_{it}$ and $LPBTD_{it}xPTCF_{it}$). The coefficients for LNBTD firm-years reveal significant differences. LNBTD firm-years exhibit significantly lower earnings persistence with respect to pre-tax book income in Table 29, with coefficients on $LNBTD_{it}xPTBI_{it}$ between -0.106 and -0.128 (p-values between 0.005 and 0.016). The results in Table 30 reveal that both the coefficient on $LNBTD_{it}xPTCF_{it}$ and the coefficients on $LNBTD_{it}xPTAC_{it}$, are significant at p-values between 0.006 and 0.054. Again, pre-tax accruals and pre-tax cash flow seem to exhibit reduced earnings persistence when large negative tBTDs occur. Nevertheless, the results should be treated with caution because of two potential caveats. First, the full sample already addresses the outlier-effect by winsorizing financial statement data at the 1st and 99th percentiles. Second, declaring single observations as outliers is a very subjective strategy, especially if the underlying effects are not fully discovered. To address the latter point, I examine hand-collected disclosure information for a sub-sample of firm-years with large negative tBTDs and rather strong effects on earnings persistence. Recall that Figure 21 reveals clear outliers in group LNBTD.

Table 23: OLS estimations for Hypothesis H1b after excluding influential observations

Variable definition: See Table 28. Clustered (firm-level) standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1.

	(1)	(2)	(3)	(4)
Dep. variable:	$PTBI_{it+1}$	$PTBI_{it+1}$	$PTBI_{it+1}$	$PTBI_{it+1}$
$PTAC_{it}$	0.740***	0.772***	0.781***	0.775***
	(0.0236)	(0.0307)	(0.0298)	(0.0308)
$PTCF_{it}$	0.799***	0.826***	0.832***	0.828***
	(0.0201)	(0.0217)	(0.0214)	(0.0212)
$LPBTD_{it}$		0.00164	0.00216	0.00178
		(0.00482)	(0.00478)	(0.00478)
$LNBTD_{it}$		0.0119***	0.0134***	0.0145***
		(0.00438)	(0.00440)	(0.00451)
$LPBTD_{it}$ x $PTAC_{it}$		**-0.0569**	**-0.0536**	**-0.0430**
		(0.0597)	(0.0597)	(0.0601)
$LNBTD_{it}$ x $PTAC_{it}$		**-0.111***	**-0.135****	**-0.131****
		(0.0573)	(0.0541)	(0.0537)
$LPBTD_{it}$ x $PTCF_{it}$		**-0.0579**	**-0.0636**	**-0.0568**
		(0.0535)	(0.0533)	(0.0533)
$LNBTD_{it}$ x $PTCF_{it}$		**-0.105****	**-0.125*****	**-0.129*****
		(0.0445)	(0.0456)	(0.0462)
Constant	0.0110***	0.00881***	0.0214***	0.0222***
	(0.00174)	(0.00214)	(0.00429)	(0.00476)
Year Fixed Effects	No	No	Yes	Yes
Industry Fixed Effects	No	No	No	Yes
Observations	927	927	927	927
R-squared	0.742	0.745	0.758	0.760

6.5.3 Influential observations with large negative tBTDs

Finding the likely sources for why firm-years with large negative tBTDs contain lower earnings persistence is an exploratory approach without a certain theoretical method. I focus on hand-collected disclosure information for a small sub-sample of firm-years with large negative tBTDs and low earnings persistence. I consider observations with Cook's D exceeding 4/987 and a decrease in pre-tax book income scaled by average total assets ≥ 20%.[54] I analyze current and next-year financial statements for each firm-year to find causes for the decrease in income. In sum, I examine 16 financial statements and eight different firm-years.[55]

Table 31 lists the main causes for the large negative tBTDs and for the decrease in pre-tax book-income for each influential firm-year. It was not possible to identify a prime cause for large negative tBTDs. A variety of accounts, for example, accounts receivable, inventories, intangible assets, or provisions, are reflected in deferred tax income. Moreover, differences in

[54] I considered potential outliers for the full sample earlier above. Since I am interested in analyzing the lower earnings persistence for observations in the LNBTD group, I now consider observations with negative change in income.

[55] This subsample may at first appear rather small. I estimate Equations (5) and (6) after eliminating these eight firm-years. I do not observe lower earnings persistence for firm-years in the LNBTD group. The coefficients on $LNBTD_{it}xPTBI_{it}$, $LNBTD_{it}xPTAC_{it}$,and $LNBTD_{it}xPTCF_{it}$ are smaller and are no longer statistically different from zero (p-values between 0.630 and 0.973). Thus, these observations strongly influence the full sample results and should therefore be examined in detail.

revenue recognition due to the percentage-of-completion method and further accounting differences (shareholder loans and consolidation measures) lead to deferred tax incoome and, thus, to large negative tBTDs. One firm reported a large negative tBTD due to the 2008 German tax reform.

It is also important to determine the causes for the decline in profits in addition to identifying potential causes for large negative tBTDs. It should be examined whether there is a systematic link between large negative tBTDs and decreasing profits. I find that two firms report a decrease in pre-tax book income because of direct or indirect consequences of the financial crisis that started in 2007/2008. The decrease in income is connected with declining sales and higher expenses for restructuring measures (e.g., personnel restructuring). For these observations, large negative tBTDs arise, at least to some extent, because of the 2008 German tax reform. Both the tax reform and the financial crisis are rather exogenous and arise independently from each other. However, the tax reform, specifically the tax rate drop from approximately 38% to 30%, led to a revaluation of deferred tax assets and deferred tax liabilities, especially in 2007. The (economic) crisis started in 2008. Thus, the link between large negative tBTDs and low earnings persistence is rather accidental.

I observe that two firm-years are attributable to companies in the solar or photovoltaic industry. This industry is, to some extent, interesting because of strong growth rates in the past and currently high competitive pressure and reduced public subsidies. For these observations, the recognition of work in progress based on the percentage-of-completion method and provisions and other liabilities lead to large negative tBTDs. Thus, I do not observe a clear link between large negative tBTDs and decreasing income for these firms.

Moreover, I find a database misspecification that affects one firm-year. Pre-tax income (WC01401) should capture income from continued operations and WC01505 should separately capture income from discontinued operations. I find one case in which WC01401 contains pre-tax income from continued *and* discontinued operations in 2006 and only income from continued operations in 2007. This different handling leads to a measurement error with respect to earnings persistence.

Summing up, I do not find a clear link between large negative tBTDs and lower earnings persistence, besides the reform and crisis relation. As described above, excluding 2007 and 2008 observations weakens the OLS-results. The coefficient on $LNBTD_{it}xPTAC_{it}$ is, then, only significant at the 10% level.

Table 24: List of influential observations with large negative tBTDs and decreasing pre-tax book income

ISIN	Year	Causes of large negative tBTDs	Causes of decreased income
DE0007757007	2007	Main cause is a decrease in deferred tax liabilities (inventories, receivables and other current assets, and consolidation measures)	**Extraordinary item**: Increase in total sales in 2008 but strong decrease in book income. Main cause is the recognition of a provision for the likely payment of an antitrust fine (TEUR 22,000).
DE000A0BVU93	2008	Main cause is a decrease in deferred tax liabilities. In detail, accounting for work in progress on the basis of the percentage-of-completion method leads to TEUR 3,186 less deferred tax liabilities compared to 2007.	**Industry effect:** Slight increase in total sales in 2009 but strong decrease in net income (60% less compared to 2008). Causes: foreign markets sales drop, restructuring expenses, and higher material costs and personnel costs. Business activity is manufacturing photovoltaic systems.
DE000A0KFUJ5	2008	Main cause is decrease in deferred tax liabilities. In detail, TEUR 192 due to elimination of different accounting rules for finance leases and TEUR 35 due to tax rate change (2008 German tax reform).	**Crisis effect:** Slight decrease in sales in 2009 and strong decrease in net income. Provision for bad debts and personnel restructuring as consequences of the financial crisis.
DE0007856023	2007	Main cause is decrease in deferred tax liabilities. In detail, TEUR 5,500 due to tax rate change (2008 German tax reform).	**Crisis effect**: Strong decrease in net income (50% less compared to 2006). Causes: higher material costs, increased research and development expenditures, and reduced orders at the end of 2008 (automotive supplier).

Table 24 (continued)

DE000A0WMPJ6	2010	Main cause is increase in deferred tax assets. In detail, TEUR 5,576 due to temporary differences resulting from restructuring intangibles assets (but no extraordinary depreciation), TEUR 4,483 due to differences in valuation of accounts receivable, and TEUR 1,005 due to differences in valuation of inventories (likely production costs).	**Extraordinarily** high net income in 2010 (TEUR 192,496). Slight decrease in sales in 2011 (TEUR 610,960) and decrease in net income (TEUR 79,536). Two main causes: Decline in demand and reduction of public subsidies.
DE0005156004	2006	Decrease in deferred tax liabilities resulting from a repayment of a shareholder loan (financing of previous acquisition).	**Database error**: Worldscope ID WC01401 contains pre-tax income from continued *and* discontinued operations in 2006. In 2007, only income from continued operations is considered. Thus, pre-tax income is in fact higher in 2007 compared to 2006 and not lower, as incorrectly displayed.
DE0007846867	2005	Main cause is a decrease in deferred tax liabilities. In detail, TEUR 600 due to differences in valuation of provisions and TEUR 433 due to consolidation measures, which influences book income.	**Scaling effect**: Slight decrease in book income before dividing by total assets (higher expenditures on research and development and on general management). Increase in total assets of 108% due to capital increase (IPO) leads to a sharp decrease in income scaled by total assets.
DE000A0DJ6J9	2011	Main cause is an increase in deferred tax assets especially related to provisions and other liabilities. No further explanations in footnotes.	**Industry effect:** Strongly growing company until 2010. Decline in sales in 2011 and 2012 and strong decline in net income. Business activity is manufacturing photovoltaic systems. This is a market with high competitive pressure in the last few years. In particular, Asian competitors are producing modules at lower cost compared to German companies.

6.5.4 Tests of Hypotheses H2a, H2b, and H3

The descriptive statistics and results for Hypotheses H1a and H1b reveal that upward earnings management seems not to be responsible for the large positive tBTDs in group LPBTD. Descriptive statistics show no significant difference between firm-years in the LPBTD group compared to the LNBTD or the SBTD groups with respect to discretionary accruals. OLS regressions in Section 6.5.2 reveal no lower earnings persistence when firms exhibit large positive tBTDs. Nevertheless, it is possible that a subsample of firms in the LPBTD group still engage in short term upward earnings management, resulting in large positive tBTDs and resulting in lower earnings persistence, as stated in Hypotheses H2a and H2b.

However, Table 32 reveals no significantly lower earnings persistence for firm-years in the LPBTD group presumably engaging in upward earnings management. The coefficient on $EMpos_{it}xPTBI_{it}$ is not statistically significantly different from zero. Thus, I do not find any evidence to confirm Hypothesis H2a. The same holds for the components pre-tax accruals and pre-tax cash flow. The corresponding results are reported in Table 33. Accordingly, I do not find any evidence to confirm Hypothesis H2b.

Table 25: OLS estimations regarding Hypothesis H2a

Variable definition: $PTBI_{it+1}$ ($PTBI_{it}$) = pre-tax book income of firm i in year t+1 (t) scaled by average total assets(WC01401/$AvAssets_{it}$); $EMpos_{it}$ = dummy variable equals one if discretionary accruals of firm i in year t are in the top quintile of all discretionary accruals in year t. Discretionary accruals are derived from a cross-sectional Jones model (year based); $AvAssets_{it}$ = average total assets ((one year lagged WC02999+WC02999)/2); Year Fixed Effects = dummy variable for each year (basis year 2005); Industry Fixed Effects = dummy variable for each industry group (WC06011 first two digits) if industry group contains at least 30 observations (11 different groups). Clustered (firm-level) standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1.

	(1)	(2)	(3)
Dep. variable:	$PTBI_{it+1}$	$PTBI_{it+1}$	$PTBI_{it+1}$
$PTBI_{it}$	0.759***	0.774***	0.746***
	(0.0772)	(0.0758)	(0.0798)
$EMpos_{it}$	-0.0131	-0.00554	-0.00794
	(0.0110)	(0.0119)	(0.0130)
$EMpos_{it}$ x $PTBI_{it}$	**0.0162**	**-0.00859**	**0.0125**
	(0.112)	(0.114)	(0.124)
Constant	0.0177***	0.0263**	0.0256**
	(0.00629)	(0.0108)	(0.0115)
Year Fixed Effects	No	Yes	Yes
Industry Fixed Effects	No	No	Yes
Observations	195	195	195
R-squared	0.633	0.671	0.681

Table 26: OLS estimations regarding Hypothesis H2b

Variable definition: $PTBI_{it+1}$ = pre-tax book income of firm i in year t+1 scaled by average total assets(WC01401/$AvAssets_{it}$); $PTAC_{it}$ = pre-tax accruals of firm i in year t as pre-tax book income minus pre-tax cash flow scaled by average total assets(WC01401-(WC04860+WC04150)/$AvAssets_{it}$); $PTCF_{it}$ = pre-tax cash flow of firm i in year t scaled by average total assets((WC04860+WC04150)/$AvAssets_{it}$); $EMpos_{it}$ = dummy variable equals one if discretionary accruals of firm i in year t are in the top quintile of all discretionary accruals in year t. Discretionary accruals are derived from a cross-sectional Jones model (year based); $AvAssets_{it}$ = average total assets ((one year lagged WC02999+WC02999)/2); Year Fixed Effects = dummy variable for each year (basis year 2005); Industry Fixed Effects = dummy variable for each industry group (WC06011 first two digits) if single group contains at least 30 observations (11 different groups). Clustered (firm-level) standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1.

	(1)	(2)	(3)
Dep. variable:	$PTBI_{it+1}$	$PTBI_{it+1}$	$PTBI_{it+1}$
$PTAC_{it}$	0.612***	0.669***	0.653***
	(0.127)	(0.111)	(0.120)
$PTCF_{it}$	0.768***	0.778***	0.752***
	(0.0867)	(0.0829)	(0.0863)
$EMpos_{it}$	-0.00347	0.00110	-0.00231
	(0.0120)	(0.0130)	(0.0140)
$EMpos_{it}$ x $PTAC_{it}$	**0.123**	**0.0616**	**0.0872**
	(0.153)	(0.148)	(0.167)
$EMpos_{it}$ x $PTCF_{it}$	**0.0261**	**0.00147**	**0.0162**
	(0.117)	(0.116)	(0.128)
Constant	0.00989	0.0219*	0.0220*
	(0.00786)	(0.0114)	(0.0120)
Year Fixed Effects	No	Yes	Yes
Industry Fixed Effects	No	No	Yes
Observations	195	195	195
R-squared	0.645	0.677	0.685

I run several robustness checks to confirm the results under different assumptions and specifications. I use different approaches of the Jones model. In detail, I include current, and alternatively, lagged return on assets as a right-hand-side variable in Equation (7) and I conducted performance-matching based on the nearest neighbor observation's current, and alternatively, lagged return on asset because extreme performance likely distort the calculation of discretionary accruals (see Section 6.3 and Dechow et al. 1995, Kothari et al. 2005). Additionally, I run modifications for all different approaches in which I use the change in sales minus change in net receivables as independent variable instead of change in sales (Dechow et al. 1995). The results remain similar. Moreover, I use total accruals as non-cash working capital and depreciation derived from the balance sheet instead of the difference between pre-tax book income and pre-tax cash flow. Again, the results remain similar. This holds for all analyzed versions of the cross-sectional Jones model. Furthermore, I examine the distribution of EMpos firms over the LPBTD, LNBTD, and SBTD groups for the default Jones model approach. I theoretically expect and also observe a higher fraction of EMpos firms in the LPBTD group (24.6%) and a lower fraction in the LNBTD group (14.9%), each compared to the fraction of SBTD firms (19.8%). The corresponding Chi²-tests reveal a significant difference only between the LPBTD and the LNBTD group (p-value = 0.014).

Summing up, based on the results, it seems rather unlikely that large positive tBTDs of German public companies are associated with upward earnings management derived from different approaches of the Jones model. Thus, in this context it seems plausible that firms in my sample do not exhibit lower earnings persistence when large positive tBTDs occur. I additionally test whether upward earnings management is associated with an, on average, lower earnings persistence for the full sample besides large positive tBTDs. I find some evidence that earnings management leads to lower earnings persistence. The default approach of the Jones model as well as the modification that controls for performance as an further right-hand-side variable reveal a lower earnings persistence for the pre-tax book income when large positive abnormal accruals occur (untabulated p-values between 0.023 and 0.039). This also holds when I use the change in sales minus change in net receivables as independent variable instead of change in sales (untabulated p-values between 0.036 and 0.098). The performance matched (modified) Jones model does not reveal overall lower earnings persistence. Thus, the evidence for overall short-term upward earnings management is mixed. One reason could be the financial crisis of 2008. Cimini (2015), for example, observes a decrease in abnormal accruals during the financial crisis between 2008 and 2012 for the majority of EU countries, including Germany. He argues that conditional conservatism during the financial crisis should raise earnings quality and that the closer monitoring activity of the auditor during the crisis is related to an increase in the quality of financial reporting.

Large positive tBTDs could alternatively occur due to tax avoidance. Table 34 presents the results from OLS regression for Hypothesis H3. As predicted, I do not observe lower earnings persistence when large positive tBTDs are primarily the result of presumed tax avoidance activities. The coefficient on $TaxAvoid_{it}xPTBI_{it}$ is not significantly different from zero. Because some firm-years are categorized as upward earnings managers *and* tax avoiders, I exclusively categorize these observations as earnings managers as robustness check. The results remain similar.

Next, I test whether tax avoidance activities result in large positive tBTDs. I find that 29.7% of the firm-years in the LPBTD group are categorized as tax avoiders (CashETR in the bottom quintile). This fraction is significantly different compared to the fraction of tax avoiders in the LNBTD (16.8%) or SBTD (18.6%) groups (Chi^2-test; p-values < 0.01). This indicates that large positive tBTDs may be usable as a signal of tax avoidance (Blaylock et al. 2012, Wilson 2009).

Table 27: OLS estimations regarding Hypothesis H3

Variable definition: $PTBI_{it+1}$ ($PTBI_{it}$) = pre-tax book income of firm i in year t+1 (t) scaled by average total assets (WC01401/$AvAssets_{it}$); $EMpos_{it}$ = dummy variable equals one if discretionary accruals of firm i in year t are in the top quintile of all discretionary accruals in year t. Discretionary accruals are derived from a cross-sectional Jones model (year based). $TaxAvoid_{it}$ = dummy variable equals one if CashETR of firm i in year t are in the bottom quintile of all CashETR each year; $AvAssets_{it}$ = average total assets ((one year lagged WC02999+WC02999)/2); Year Fixed Effects = dummy variable for each year (basis year 2005); Industry Fixed Effects = dummy variable for each industry group (WC06011 first two digits) if single group contains at least 30 observations (11 different groups). Clustered (firm-level) standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1.

	(1)	(2)	(3)
Dep. variable:	$PTBI_{it+1}$	$PTBI_{it+1}$	$PTBI_{it+1}$
$PTBI_{it}$	0.812***	0.817***	0.789***
	(0.0806)	(0.0799)	(0.0898)
$EMpos_{it}$	-0.00547	0.000296	-0.00134
	(0.0120)	(0.0129)	(0.0144)
$TaxAvoid_{it}$	0.0167	0.0134	0.0147
	(0.0129)	(0.0123)	(0.0134)
$EMpos_{it}$ x $PTBI_{it}$	**-0.00326**	**-0.0245**	**-0.00669**
	(0.114)	(0.116)	(0.130)
$TaxAvoid_{it}$ x $PTBI_{it}$	**-0.131**	**-0.106**	**-0.107**
	(0.0967)	(0.0924)	(0.105)
Constant	0.0121*	0.0232*	0.0231*
	(0.00663)	(0.0128)	(0.0132)
Year Fixed Effects	No	Yes	Yes
Industry Fixed Effects	No	No	Yes
Observations	195	195	195
R-squared	0.642	0.677	0.686

In untabulated tests, I substitute CurrentETR for CashETR. The results remain similar. Moreover, I substitute 3Year-CashETR for CashETR and 3Year-CurrentETR for CurrentETR. The variable 3Year-CashETR (3Year-CurrentETR) is measured as the sum of cash taxes paid (current tax expense) over the previous, current, and next year divided by the sum of pre-tax book income over the previous, current, and next year.[56] Again, the results remain similar. Furthermore, I treat observations that are identified as upward earnings managers *and* tax avoiders solely as earnings management observations (in line with Blaylock et al. 2012). Because earnings management seems not to be the driving force behind large positive tBTDs, I neither expect nor find any qualitatively difference. The fraction of firm-years identified as tax avoider in the large positive BTD group is still significantly different (larger) compared to the fraction of tax avoiders in the LNBTD or SBTD groups, but with slightly larger p-values (Chi^2-test; p-values < 0.05).

[56] In contrast to Blaylock et al. (2012) and Dyreng et al. (2008), I do not use the previous three years to calculate 3Year-CashETR or 3Year-CurrentETR because this approach would lead to a halving of the sample. Nevertheless, the results remain qualitatively unchanged when I use the previous three years with 469 usable observations.

6.6 Conclusion

This study provides evidence regarding the association between large tBTDs and earnings persistence as a proxy for earnings quality by using a sample of 987 firm-year observations from German CDAX®-listed IFRS companies. Since Guenther (2011) shows that tBTDs are likely distorted by the deferred tax expense resulting from NOLCs, I precisely control for this potentially distortional effect by using hand-collected disclosure information. I find that 90% of the firm-years in my sample report a deferred tax expense from NOLCs and that, on average, 40% of the total deferred tax expense occur because of deferred tax expense from NOLCs. However, these deferred taxes from NOLCs seem not to be responsible for the association between large positive tBTDs and earnings persistence. In contrast to prior literature, I do not observe a lower earnings persistence when these firms exhibit large positive tBTDs (Hanlon 2005, Blaylock et al. 2012, Tang and Firth 2012), even when I precisely control for a potential distortional effect through deferred taxes from NOLCs. Moreover, and in contrast to Blaylock et al. (2012) but in line with the findings of Jackson (2015), I do not observe lower earnings persistence for large positive BTD firms that are identified as upward earnings managing firms.

In line with Blaylock et al. (2012), I find some evidence that the large positive tBTDs arise because of tax-avoiding activities. By using CashETR, CurrentETR, 3YearCashETR, and 3YearCurrentETR as proxies for tax avoidance, I observe that a statistically significantly larger fraction of the firm-years in the LPBTD group are categorized as tax avoiders compared to the fractions in the SBTD group and the LNBTD group. Thus, large positive tBTDs seem to be more associated with tax-avoiding activities and less associated with upward earnings management. One reason could be a higher earnings quality due to the financial crisis of 2008, as reported by Cimini (2015). Nevertheless, my proxies for tax avoidance are rather broad and, thus, likely capture both "true" and "spurious" tax avoidance due to special tax law regulations that are less manageable (e.g., shorter useful life of assets for tax purposes, tax credits for certain industries).

I observe relatively low earnings persistence when firms exhibit large negative tBTDs. Although this is in line with prior literature (Hanlon 2005; Blaylock et al. 2012), it is unclear what causes this relation. I find that controlling for deferred tax expense from NOLCs reduces the negative association between large negative tBTDs and earnings persistence to some extent, but not completely. Thus, I use hand-collected disclosure information for influential observations and find that various accounts affect the level of negative tBTDs. Moreover, I do not observe a plain or obvious relation between high negative tBTDs and earnings persistence. The low earnings persistence mainly results from direct and indirect consequences of the 2008 crisis, changing market conditions, and a database misspecification. Because these are rather external effects, I suggest that the remaining negative relation between large negative tBTDs and earnings persistence is rather spurious. Thus, my study contributes to the somewhat critical prior research stating that the association between large tBTDs and earnings persistence is likely related to firm-specific characteristics (Guenther 2011; Drake 2012).

References

Atwood, T. J., M. S. Drake, J. N. Myers, and L. A. Myers. 2011. Do earnings reported under IFRS tell us more about future earnings and cash flows? *Journal of Accounting and Public Policy* 30(2): 103-121.

Ayers, B. C., J. Jiang, and P. E. Yeung. 2006. Discretionary Accruals and Earnings Management: An Analysis of Pseudo Earnings Targets. *The Accounting Review* 81(3): 617-652.

Barth, M. E., and A. P. Hutton. 2004. Analyst earnings forecast revisions and the pricing of accruals. *Review of Accounting Studies* 9(1): 59-96.

Barth, M. E., W. R. Landsman, and M. H. Lang. 2008. International accounting standards and accounting quality. *Journal of Accounting Research* 46(2): 467-498.

Bartov, E., F. A. Gul, and J. S. L. Tsui. 2000. Discretionary-accruals models and audit qualifications. *Journal of Accounting and Economics* 30(3): 421-452.

Basu, S. 1997. The conservatism principle and the asymmetric timeliness of earnings. *Journal of Accounting and Economics* 24(1): 3-37.

Blaylock, B., T. Shevlin, and W. Wilson. 2012. Tax avoidance, large positive book-tax differences, and earnings persistence. *The Accounting Review* 87(1): 91-120.

Bollen, K. A., and R. W. Jackman. 1990. Regression diagnostics: An expository treatment of outliers and influential cases, in Fox, J., and Long, J. S. (eds.). Modern Methods of Data Analysis (257-291). Newbury Park, CA: Sage.

Chi, S. S., M. Pincus, and S. H. Teoh. 2014. Mispricing of book-tax differences and the trading behavior of short sellers and insiders. *The Accounting Review* 89(2): 511-543.

Cimini, R. 2015. How has the financial crisis affected earnings management? A European study. *Applied Economics* 47(3): 302-317.

Cook, R. D. 1977. Detection of influential observations in linear regression. *Technometrics (American Statistical Association)* 19(1): 15-18.

Christensen, H. B., E. Lee, M. Walker, and C. Zeng. 2015. Incentives or standards: What determines accounting quality changes around IFRS adoption? *European Accounting Review* 24(1): 31-61.

Dechow, P. M., and W. Ge. 2006. The persistence of earnings and cash flows and the role of special items: Implications for the accrual anomaly. *Review of Accounting Studies* 11(2-3): 253-296.

Dechow, P. M., W. Ge, and C. Schrand. 2010. Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of Accounting and Economics* 50(2-3): 344-401.

Dechow, P. M., A. P. Hutton, J. H. Kim, and R. G. Sloan. 2012. Detecting Earnings Management: A New Approach. *Journal of Accounting Research* 50(2): 275-334.

Dechow, P. M., R. G. Sloan, and A. P. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70(2): 193-225.

Drake, K. D. 2012. Does firm life cycle explain the relation between book-tax differences and earnings persistence? Dissertation Arizona State University. http://hdl.handle.net/2286/R.I.15021, 29. Dec. 2015.

Dyreng, S. D., M. Hanlon, and E. L. Maydew. 2008. Long-run corporate tax avoidance. *The Accounting Review* 83(1): 61-82.

Frank, M. M., L. J. Lynch, and S. O. Rego. 2009. Tax reporting aggressiveness and its relation to aggressive financial reporting. *The Accounting Review* 84(2): 467-496.

Graham, J. R., J. S. Raedy, and D. A. Shackelford. 2012. Research in accounting for income taxes. *Journal of Accounting and Economics* 53(1-2): 412-434.

Guenther, D. A. 2011. What do we learn from large book-tax differences? Working Paper Lundquist College of Business, University of Oregon. http://gsm.ucdavis.edu/sites/main/files/file-attachments/guenther_btds_03_28_11.pdf, 15. May 2015.

Hayn, C. 1995. The information content of losses. *Journal of Accounting and Economics* 20(2): 125-153.

Hanlon, M. 2005. The persistence and pricing of earnings, accruals, and cash flows when firms have large book-tax differences. *The Accounting Review* 80(1): 137-166.

Hanlon, M., and S. Heitzman. 2010. A review of tax research. *Journal of Accounting & Economics* 50(2-3): 127-178.

Hanlon, M., G. V. Krishnan, and L. F. Mills. 2012. Audit Fees and Book-Tax Differences. *Journal of the American Taxation Association* 34(1): 55-86.

Heltzer, W., and S. W. Shelton. 2011. The book–tax divide: Perceptions from the field. *Research in Accounting Regulation* 23(2): 130-137.

Hribar, P., and D. W. Collins. 2002. Errors in estimating accruals: Implications for empirical research. *Journal of Accounting Research* 40(1): 105-134.

Jackson, M. 2015. Book-tax differences and future earnings changes. *Journal of the American Taxation Association* 37(2): 49-73.

Jeanjean, T., and H. Stolowy. 2008. Do accounting standards matter? An exploratory analysis of earnings management before and after IFRS adoption. *Journal of Accounting and Public Policy* 27(6): 480-494.

Jones, J. J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29(2): 193-228.

Keung, E., and M. S. H. Shih. 2014. Measuring discretionary accruals: are ROA-matched models better than the original Jones-type models? *Review of Accounting Studies* 19(2): 736-768.

Kaserer, C., and C. Klingler. 2008. The accrual anomaly under different accounting standards - Lessons learned from the German experiment. *Journal of Business Finance & Accounting* 35(7-8): 837-859.

Kothari, S. P., A. J. Leone, and C. E. Wasley. 2005. Performance matched discretionary accrual measures. *Journal of Accounting and Economics* 39(1): 163-197.

KPMG. 2015. IFRS compared to US GAAP: An Overview. https://assets.kpmg.com/content/dam/kpmg/pdf/2015/12/US-GAAP-comparison-2015-overview.pdf, 29. Dec. 2015.

Krummet, F. 2011. Der Informationsgehalt von latenten Steuern nach IAS 12: Eine empirische Untersuchung zur Extrapolierbarkeit des IFRS-Ergebnisses in deutschen Konzernabschlüssen. 1st Edition, Lohmar-Köln.

Lawrence, A., R. G. Sloan and Y. Sun. 2015. Why are losses less persistent than profits? Curtailments versus conservatism. Boston U. School of Management Research Paper No. 2410590. http://ssrn.com/abstract=2410590, 29. Dec. 2015.

Lev, B., and D. Nissim. 2004. Taxable income, future earnings, and equity values. *The Accounting Review* 79(4): 1039-1074.

Liu, C., C. Y. Yuen, L. J. Yao (posthumously), and S. H. Chan. 2014. Differences in earnings management between firms using US GAAP and IAS/IFRS. *Review of Accounting and Finance* 13(2): 134-155.

Martinez, A. L., and T. B. T. de Souza. 2015. Book-tax differences, earnings persistence and tax planning before and after the adoption of IFRS in Brazil. Working Paper related to the IX Congresso anpcont. http://congressos.anpcont.org.br/ix/anais/files/2015-05/cue301.pdf, 29. Dec. 2015.

Mills, L. F. 1998. Book-tax differences and internal revenue service adjustments. *Journal of Accounting Research* 36(2): 343-356.

Mills, L. R., and K. J. Newberry. 2001. The influence of tax and nontax costs on book-tax reporting differences: Public and private firms. *Journal of the American Taxation Association* 23(1): 1-19.

Phillips, J., M. Pincus, and S. O. Rego. 2003. Earnings management: New evidence based on deferred tax expense. *The Accounting Review* 78(2): 491-521.

Revsine, L., D. W. Collins, and W. B. Johnson. 2005. Financial reporting and analysis. 3rd Edition, New Jersey.

Richardson, S., R. Sloan, M. Soliman, and I. Tuna. 2005. Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics* 39(3): 437-485.

Sloan, R. G. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71(3): 289-315.

Tang, T. and M. Firth. 2011. Can book–tax differences capture earnings management and tax management? Empirical evidence from China. *The International Journal of Accounting* 46(2): 175-204.

Tang, T. and M. Firth. 2012. Earnings persistence and stock market reactions to the different information in book-tax differences: Evidence from China. *The International Journal of Accounting* 47(3): 369-397.

Wilson, R. J. 2009. An examination of corporate tax shelter participants. *The Accounting Review* 84(3): 969-999.

Appendix

English summary

This thesis examines, on the one hand, whether and how tax system characteristics affect individuals' behavior in an experimental environment, and, on the other hand, whether large positive or large negative temporary book-tax differences exhibit information content regarding the persistence of earnings.

The first study empirically test whether declaring tax minimization as legal or illegal affects individuals' tax minimization behavior. In line with our theoretical prediction, we find that labeling a tax minimization opportunity as unambiguously illegal results in significantly less tax minimization compared to labeling tax minimization as unambiguously legal. As soon as we consider detection risk, negative detection consequences (penalties in the case of evasion and interest charges in the case of avoidance) and implicit monitoring, we observe no difference between legal and illegal tax minimization. Moreover, we observe that affective priming reinforces a legality effect. Taken together, our study shows that legality can have strong effects on individuals' behavior. In line with the expressive law approach, defining the borderline between legality and illegality can be used to affect moral costs.

The second study examines whether moral evaluations of tax evasion are egoistically biased. In a real-effort experiment, we demonstrate that individuals' tax morale is subject to a self-serving bias. Individuals with the opportunity to evade taxes consider tax evasion less unethical compared to those without this opportunity. We also find no spillover effects of evasion opportunities on other moral evaluations (e.g., bribery or lying).

The third study shows that, due to mental accounting, the timing of taxation (immediate versus deferred taxation) is important even if there are no wealth differences between alternatives. Subjects in the deferred tax treatment perceive their wage as significantly more fair and this perception (indirectly) increases work effort. Moreover, subjects in the deferred tax treatment make less risky investments. Thus, the presumed neutrality regarding the timing of taxation does not hold.

The fourth study addresses the question of whether two important tax system characteristics – earmarking the revenues of a tax and voting rights in setting the tax rate – have an impact on tax avoidance and real labor supply decisions. We find that earmarking significantly lowers tax avoidance. Democratic participation, as voting on the rate of the earmarked tax, reduces tax avoidance even further. We also show that the individual tax rate vote affects the subsequent avoidance decision. Those who voted for a high tax rate avoid significantly fewer taxes. In contrast, we do not observe any significant impact of earmarking and participation rights on labor supply. This result suggests that a feature of a tax could influence different forms of taxpayers' behavioral responses in different ways. Taxpayers' avoidance responses seem to be much more elastic than their real responses.

The fifth study analyzes whether large positive or negative temporary book-tax differences (tBTDs) exhibit information content regarding the persistence of earnings. I do not observe lower earnings persistence when firms exhibit large positive tBTDs, and I do not find lower

earnings persistence for firms that are identified as (presumed) upward earnings managing firms according to several approaches of the Jones model. With respect to tax avoidance as a potential explanation for large positive tBTDs, I find that large positive tBTDs can be used as a signal of tax avoidance. As predicted, I do not observe higher or lower earnings persistence when large positive tBTDs likely result from tax avoidance. I find relatively low earnings persistence when firms exhibit large negative tBTDs. Based on hand-collected disclosure information for influential observations, I show that various accounts affect the level of tBTDs. The low earnings persistence mainly results from direct and indirect consequences of the 2008 crisis, changing market conditions, and a database misspecification. Because these are rather external effects, I conclude that the observed negative relation can be explained by internal management activities only for a small part of the observations.

German summary

Die vorliegende Dissertation untersucht zum einen, ob und wie verschiedene Eigenschaften eines Steuersystem das Verhalten von Steuerpflichtigen in einem experimentellen Umfeld beeinflussen und zum anderen, inwieweit hohe positive oder hohe negative temporäre Differenzen zwischen handelsrechtlichen und steuerlichen Gewinnen (temporary book-tax differences) einen Informationsgehalt bezüglich der Ergebnisbeständigkeit aufweisen.

Die erste Studie geht der Frage nach, ob die Legalität einer Steuervermeidungsmöglichkeit – „illegale Steuerhinterziehung“ vs. „legale Steuer-gestaltung“ – c. p. zu unterschiedlichen Steuervermeidungsentscheidungen führt. Wir finden heraus, dass dies der Fall ist, wenn keine finanziellen Sanktionen drohen. Die Ausgestaltung der Steuervermeidungsmöglichkeit als illegale Steuerhinterziehung führt zu einer deutlichen Reduktion in der Höhe der vermiedenen Steuer im Vergleich zu einer Ausgestaltung als legale Nutzung eines Steuerschlupflochs. Sobald wir allerdings identische finanzielle Sanktionen unter Unsicherheit implementieren, ist ein Unterschied zwischen legaler und illegaler Steuervermeidung nicht mehr sichtbar. Moralisches Priming führt wiederum zu einem sichtbaren Unterschied. Folglich spielt die Legalität von steuerlichen Vermeidungsalternativen insbesondere eine Rolle, wenn negative finanzielle Konsequenzen nicht vorhanden sind bzw. für den Steuerpflichtigen nicht oder nur bedingt erkennbar sind.

Die zweite Studie untersucht, inwieweit die Möglichkeit der Steuerhinterziehung auch die moralische Bewertung von Steuerhinterziehungen beeinflusst. Wir finden heraus, dass Probanden mit der Möglichkeit der Steuerhinterziehung diese im Durchschnitt als weniger unethisch bewerten als Probanden ohne diese Möglichkeit der Hinterziehung. Ein Übertragungseffekt auf andere moralische Vergehen lässt sich indes nicht nachweisen.

Die dritte Studie zeigt, dass der der Zeitpunkt der Besteuerung (sofortige vs. nachgelagerte Besteuerung) die Arbeits- und Investitionsentscheidungen von Probanden trotz rationaler Indifferenz beeinflusst. Konkret finden wir heraus, dass der Zeitpunkt der Besteuerung die Fairnesswahrnehmung bezüglich des Arbeitslohns stark beeinflusst, der Lohn wird bei nachgelagerter Besteuerung als fairer empfunden, und dass diese Fairnesswahrnehmung als Mediator das individuelle Arbeitsangebot beeinflusst. Wir beobachten folglich einen

indirekten Effekt des Zeitpunkts der Besteuerung auf das Arbeitsangebot. Darüber hinaus zeigen wir, dass die nachgelagerte Besteuerung zu einem Anstieg der Risikobereitschaft führt. Wir argumentieren, dass Mental Accounting eine Erklärung für das beobachtete Verhalten ist.

Die vierte Studie konzentriert sich ebenfalls auf das individuelle Arbeitsangebot und geht der Frage nach, inwieweit die Eigenschaften eines Steuersystems „Zweckbindung der Steuer“ und „demokratische Mitbestimmung“ die individuelle Arbeitsangebotsentscheidung und die legale Steuervermeidungsentscheidung von Probanden beeinflussen. Wir finden heraus, dass das Steuervermeidungsverhalten geringer ist, wenn die Steuer einer Zweckbindung unterliegt. Demokratische Mitbestimmung in Form einer Mitsprache bei der Festlegung des Steuersatzes führt zu einer weiteren, allerdings etwas schwächeren, Reduktion der Steuervermeidung. Das individuelle Arbeitsangebot reagiert dagegen nicht signifikant auf die untersuchten Steuersystemeigenschaften.

Die fünfte Studie analysiert inwieweit hohe positive oder hohe negative temporäre Differenzen zwischen handelsrechtlichen und steuerlichen Gewinnen einen Informationsgehalt bezüglich der Ergebnisbeständigkeit aufweisen. Es zeigt sich, dass dies für deutsche börsennotierte Unternehmen zwischen 2005 und 2013 mit hohen positiven Differenzen nicht der Fall ist. Das Ergebnis ist robust gegenüber einer Vielzahl an alternativen Tests und Erweiterungen (unter anderem Einteilung der Beobachtungen anhand der potentiellen Quellen für hohe positive temporäre Differenzen *Bilanzpolitik* und *Steuervermeidung*). Darüber hinaus wird ersichtlich, dass die vermeintlichen Differenzen zwischen handelsrechtlichen und steuerlichen Gewinnen verzerrt werden durch latente Steuern auf aktivierte Verlustvorträge. Eine im Durchschnitt geringere Ergebnisbeständigkeit bei Vorliegen von hohen negativen Differenzen lässt sich auf verschiedene, eher externe, Effekte zurückführen (Wirtschaftskrise, veränderte Marktbedingungen, Datenbankfehler). Insgesamt zeigt sich, dass hohe temporäre Differenzen kein sicherer Proxy für eine geringere Ergebnisbeständigkeit sind.

General References

Allingham, M. G., Sandmo, A., 1972. Income Tax Evasion: A Theoretical Analysis. *Journal of Public Economics* 1(3/4), 323–338.

Alm, J., Torgler, B., 2006. Culture differences and tax morale in the United States and in Europe. *Journal of Economic Psychology* 27, 224–246.

Alm, J., McClelland, G. H., Schulze, W. D., 1992. Why do people pay taxes? *Journal of Public Economics* 48(1), 21–38.

Atwood, T. J., Drake, M. S., Myers, J. N., Myers, L. A., 2012. Home Country Tax System Characteristics and Corporate Tax Avoidance International Evidence. *The Accounting Review* 87(6), 1831–1860.

Baldry, J. C., 1986. Tax evasion is not a gamble: A report on two experiments. *Economics Letters* 22(4), 333–335.

Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., 2015. Does front-loading taxation increase savings? Evidence from Roth 401 (k) introductions. *Journal of Public Economics* (forthcoming).

Blaufus, K., Bob, J., Hundsdoerfer, J., Kiesewetter, D., Weimann, J., 2013. Decision heuristics and tax perception – An analysis of a tax-cut-cum-base-broadening policy. *Journal of Economic Psychology* 35, 1–16.

Blaylock, B., Shevlin, T., Wilson, W., 2012. Tax avoidance, large positive book-tax differences, and earnings persistence. *The Accounting Review* 87(1), 91–120.

Blumkin, T., Ruffle, B. J., Ganun, Y., 2012. Are income and consumption taxes ever really equivalent? Evidence from a real-effort experiment with real goods. *European Economic Review* 56(6), 1200–1219.

Bobek, D. D., Hatfield, R. C., 2003. An investigation of the theory of planned behavior and the role of moral obligation in tax compliance. *Behavioral Research in Accounting* 15(1), 13–38.

Brosig-Koch, J., Helbach, C., Ockenfels, A., Weimann, J., 2011. Still different after all these years: Solidarity behavior in East and West Germany. *Journal of Public Economics* 95(11), 1373–1376.

Chetty, R., 2015. Behavioral Economics and Public Policy: A Pragmatic Perspective. *American Economic Review* 105(5), 1–33.

Chetty, R., Friedman, J. N., Saez, E., 2013. Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings. *American Economic Review* 103(7), 2683–2721.

Chetty, R., Looney, A., Kroft, K., 2009. Salience and taxation: theory and evidence. *American Economic Review* 99(4), 1145–1177.

Cooter, R., 1998. Expressive law and economics. *Journal of Legal Studies* 27(2), 585–608.

Cooter, R., 2000. Do good laws make good citizens: An economic analysis of internalized norms. *Virginia Law Review* 86(8), 1577–1601.

Cummings, R. G., Martinez-Vazquez, J., McKee, M., Torgler, B., 2009. Tax morale affects tax compliance: Evidence from surveys and an artefactual field experiment. *Journal of Economic Behavior & Organization* 70(3), 447–457.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., 2010. Are risk aversion and impatience related to cognitive ability? *American Economic Review* 100(3), 1238–1260.

Durham, Y., Manly, T. S., Ritsema, C., 2014. The effects of income source, context, and income level on tax compliance decisions in a dynamic experiment. *Journal of Economic Psychology* 40, 220–233.

Dwenger, N., Kleven, H., Rasul, I., Rincke, J., 2015. Extrinsic and intrinsic motivations for tax compliance: Evidence from a field experiment in Germany. *American Economic Journal: Economic Policy* (forthcoming).

Falk, A., Kosfeld, M., 2006. The hidden costs of control. Am*erican Economic Review* 96(5), 1611–1630.

Falsetta, D., Rupert, T. J., Wright, A. M., 2013. The Effect of the Timing and Direction of Capital Gain Tax Changes on Investment in Risky Assets. *The Accounting Review* 88 (2), 499–520.

Fehr, E., Falk, A., 2002. Psychological foundations of incentives. *European Economic Review* 46(4–5), 687–724.

Fochmann, M., Weimann, J., 2013. The effects of tax salience and tax experience on individual work efforts in a framed field experiment. *FinanzArchiv: Public Finance Analysis* 69(4), 511–542.

Fochmann, M., Weimann, J., Blaufus, K., Hundsdoerfer, J., Kiesewetter, D., 2013. Net Wage Illusion in a Real-Effort Experiment. Sca*ndinavian Journal of Economics* 115(2), 476–484.

Gneezy, U., Rustichini, A., (2000) A fine is a price. *Journal of Legal Studies* 29(1), 1–17.

Graham, J. R., 2006. A Review of Taxes and Corporate Finance. *Foundations and Trends in Finance* 1(7), 573–691.

Graham, J. R., Raedy, J. S., Shackelford, D. A., 2012. Research in accounting for income taxes. *Journal of Accounting and Economics* 53(1–2), 412–434.

Guenther, D. A., 2011. What do we learn from large book-tax differences? Working Paper Lundquist College of Business, University of Oregon. http://gsm.ucdavis.edu/sites/main/files/file-attachments/guenther_btds_03_28_11.pdf, 15. April 2016.

Hanlon, M., 2005. The persistence and pricing of earnings, accruals, and cash flows when firms have large book-tax differences. *The Accounting Review* 80(1), 137–166.

Hanlon, M., Heitzman, S., 2010. A review of tax research. *Journal of Accounting and Economics* 50(2–3), 127–178.

Hundsdoerfer, J., Kiesewetter, D., Sureth C., 2008. Forschungsergebnisse in der Betriebswirtschaftlichen Steuerlehre – eine Bestandsaufnahme. *Zeitschrift für Betriebswirtschaft* 78(1), 61–139.

Jackson, M., 2015. Book-tax differences and future earnings changes. *Journal of the American Taxation Association* 37(2), 49–73.

Kastlunger, B., Lozza, E., Kirchler, E., Schabmann A., 2013. Powerful authorities and trusting citizens: The Slippery Slope Framework and tax compliance in Italy. *Journal of Economic Psychology* 34, 36–45.

Kirchler, E., Hoelzl, E., Wahl, I., 2008. Enforced versus voluntary tax compliance: The "slippery slope" framework. *Journal of Economic Psychology* 29, 210–225.

Kirchler, E., Maciejovsky, B., Schneider, F., 2003. Everyday representations of tax avoidance, tax evasion, and tax flight: Do legal differences matter? *Journal of Economic Psychology* 24, 535–553.

Kogler, C., Batrancea, L., Nichita, A., Pantya, J., Belianin, A., Kirchler, E., 2013. Trust and power as determinants of tax compliance: Testing the assumptions of the slippery slope framework in Austria, Hungary, Romania and Russia. *Journal of Economic Psychology* 34, 169–180.

Kogler, C., Muehlbacher, S., Kirchler, E., 2015. Testing the "slippery slope framework" among self-employed taxpayers. *Economics of Governance* 16(2), 125–142.

Kredentser, M. S., Fabrigar, L. R., Smith, S. M., Fulton, K., 2012. Following what people think we should do versus what people actually do. *Social Psychological and Personality Science* 3(3), 341–347.

Krummet, F., 2011. Der Informationsgehalt von latenten Steuern nach IAS 12: Eine empirische Untersuchung zur Extrapolierbarkeit des IFRS-Ergebnisses in deutschen Konzernabschlüssen. 1st Edition, Lohmar-Köln.

Lev, B., Nissim, D., 2004. Taxable income, future earnings, and equity values. *The Accounting* Review 79(4), 1039–1074.

Maydew, E. L., 2001. Empirical tax research in accounting: A discussion. *Journal of Accounting and Economics* 31(1), 389–403.

Mazar, N., Amir, O., Ariely, D., 2008. The Dishonesty of Honest People. A Theory of Self-Concept Maintenance. *Journal of Marketing Research*. 45(6), 633–644.

Prinz, A., Muehlbacher, S., Kirchler, E., 2014. The slippery slope framework on tax compliance: An attempt to formalization. *Journal of Economic Psychology* 40, 20–34.

Richardson, G., Taylor, G., Lanis, R., 2013. The impact of board of director oversight characteristics on corporate tax aggressiveness: An empirical analysis. *Journal of Accounting and Public Policy* 32(3), 68–88.

Russo, F. F., 2013. Tax morale and tax evasion reports. *Economics Letters* 121(1), 110–114.

Sausgruber, R., Tyran, J. R., 2005. Testing the Mill hypothesis of fiscal illusion. *Public Choice* 122(1–2), 39–68.

Sillamaa, M. A., 1999. How work effort responds to wage taxation: An experimental test of a zero top marginal tax rate. *Journal of Public Economics* 73(1), 12–134.

Srinivasan, T. N., 1973. Tax evasion: A model. *Journal of Public Economics* 2(4), 339–346.

Sunstein, C., 1996. Social norms and social roles. *Columbia Law Review* 96(4), 903–968.

Swenson, C., 1988. Taxpayer behavior in response to taxation: An experimental analysis. *Journal of Accounting and Public Policy* 7(1), 1–28.

Tang, T., Firth, M., 2011. Can book–tax differences capture earnings management and tax management? Empirical evidence from China. *The International Journal of Accounting* 46(2), 175–204.

Thaler, R. H., 1985. Mental accounting and consumer choice. *Marketing Science* 4(3), 199–214.

Thaler, R. H., 1990. Anomalies: Saving, fungibility, and mental accounts. *Journal of Economic Perspectives* 4(1), 193–205.

Thaler, R. H., 1999: Mental Accounting Matters. *Journal of Behavioral Decision Making* 12(3), 183-206.

Declaration of honor

Hiermit erkläre ich, dass ich mich noch keinem Promotionsverfahren unterzogen oder um Zulassung zu einem solchen beworben habe und die Dissertation in der gleichen oder einer anderen Fassung bzw. Überarbeitung einer anderen Fakultät, einem Prüfungsausschuss oder einem Fachvertreter an einer anderen Hochschule nicht bereits zur Überprüfung vorgelegen hat.

Hiermit erkläre ich, dass ich für die Dissertation folgende Hilfsmittel und Hilfen verwendet habe: Die englischsprachigen Texte wurden von einem Language Editor auf sprachliche Richtigkeit geprüft. Es wurden die Programme Stata und z-Tree verwendet. An den jeweiligen Studien haben die auf den Seiten III–IV angegebenen Koautoren mitgewirkt. Die einzelnen Studien wurden auf Konferenzen vorgestellt. Entsprechend wurden hilfreiche Kommentare von Konferenzteilnehmern eingearbeitet. Die jeweiligen Konferenzen und Personen sind zu Beginn der Studien angegeben. Bei bereits veröffentlichten Studien wurden Kommentare von anonymen Gutachtern aufgenommen.

Auf dieser Grundlage habe ich die Arbeit selbstständig verfasst.

Berlin, 21.04.2016 Matthias Sünwoldt

Curriculum vitae

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.