Arbeitskreis Quantitative Steuerlehre
Quantitative Research in Taxation – Discussion Papers

Martin Fochmann, Nadja Müller, Michael Overesch

Less Cheating? The Effects of Prefilled Forms on Compliance Behavior

arqus Discussion Paper No. 227

July 2018

www.arqus.info
ISSN 1861-8944

Less Cheating? The Effects of Prefilled Forms on Compliance Behavior[a]

Martin Fochmann[b], Nadja Müller[c], Michael Overesch[d]

University of Cologne

July 2, 2018

Abstract

As a consequence of the digital transformation, individuals are often confronted with prefilled forms or prefilled data entry masks. In situations where cheating and lying are of concern, prefilling and defaults might reduce dishonest behavior. In a controlled experiment, we investigate how correctly and incorrectly prefilled forms influence compliance behavior. We frame our experiment as filing the annual income tax return. We show that correct prefilling enhances compliance. However, in cases of incorrect prefilling, we observe asymmetric effects. If prefilled income is lower than true income, we find no positive compliance effect, and compliance is on the same level as with blank forms. If prefilled income is higher than true income, prefilling still has a positive effect on compliance. In that case, compliance is on the same level as with correctly prefilled forms and higher than with blank forms. Our study contributes to the literature on cheating and lying by showing that prefilled forms and defaults affect compliance by changing the moral costs of dishonest behavior.

Keywords

Dishonesty, Defaults, Prefilled Forms, Tax Compliance, Behavioral Economics

JEL-Classification

C91, D14, H26

---

[a] We thank Matthias Sutter, Kay Blaufus, Abdolkarim Sadrieh, Christoph Engel, Lisa Bruttel, Pete Lunn, Frank Hechtner, Peter Mohr, Jochen Hundsdoerfer, Rainer Niemann, Manuel Schwaninger, Florian Stolley, Christine Straberg and participants of the 2018 ESA Conference, 2018 EAA Conference, 2018 VHB Conference, 2018 Nudging and Information Conference, 2017 GfeW Conference, 2017 arqus Conference, and research seminars at the University of Potsdam and Max Planck Institute for Research on Collective Goods in Bonn for helpful comments and suggestions.

[b] Martin Fochmann, University of Cologne, Business School, Albertus-Magnus-Place, D-50923 Köln, Germany, fochmann@wiso.uni-koeln.de.

[c] Nadja Müller, University of Cologne, Business School, Albertus-Magnus-Place, D-50923 Köln, Germany, nadja.mueller@wiso.uni-koeln.de.

[d] Michael Overesch, University of Cologne, Business School, Albertus-Magnus-Place, D-50923 Köln, Germany, overesch@wiso.uni-koeln.de.

1 Introduction

As a consequence of the digital transformation, massive amounts of data are collected, processed and analyzed in almost all industries and life situations. One benefit of this process is that individuals (e.g., managers, employees) automatically receive relevant information from an information system that supports their decision making. This unprecedented availability of and access to information has not only changed workflows and decision making in organizations and institutions, it has also changed how individuals complete forms and data entry masks. If cheating and lying are of concern in such situations, prefilling and defaults might reduce dishonest behavior. However, the effect of prefilled forms or prefilled data entry masks on compliance behavior is largely unexplored, even though prefilling has become more and more common in various situations, such as the claiming of travel expenses by employees, the reporting of insured events by policy holders or the collection of financial information from subdivision managers. In a controlled experiment, we investigate how prefilled forms affect compliance behavior, focusing specifically on how incorrectly prefilled values influence compliance.

In our investigation, we consider the situation of filing the annual income tax return. Today, taxpayers often start their tax declarations with prefilled forms. Customers who use tax software or file their tax returns online (with e-filing services) find that electronic tax return programs usually carry over the previous year's values to the subsequent year as an orientation aid. Consequently, a current tax return is initially prefilled with last year's numbers (e.g., salary, business income, expenses, deductions, tax credits). However, prefilled data is often incorrect, as data carried forward from the previous year does not fully reflect the conditions of the current year. Furthermore, prefilling plays a significant role in the work of revenue bodies in many countries, including the UK, Canada, Australia, Germany, France, Italy, Spain and Sweden (OECD, 2017). Automatic data exchanges between the tax authority and employers, social insurance agencies and banks enable systems to create tax returns that are prefilled before they are sent to taxpayers. Consequently, instead of completing blank forms manually, taxpayers receive tax returns that are already prefilled with data. Although tax returns prefilled by tax authorities should be highly trustworthy, errors in data input and data transmission can occur. For example, in the UK, "experts estimate that one in ten returns prefilled using HMRC's data will contain errors, which could lead to people paying too much or too little tax" (Telegraph, 2017). Moreover, innovative tools are using new techniques to produce prefilled tax returns

based merely on scanned bills and receipts. Nevertheless, techniques such as optical character recognition are still associated with errors.

While prefilling aims to help the taxpayer file a legally accurate tax return (OECD, 2017), mis-specified or incorrect numbers that are already included in a prefilled tax return might be associated with additional tax evasion. There is some evidence that the implementation of prefilled tax returns reduces tax compliance costs at the taxpayer's level (Goolsbee, 2006, Klun, 2009, Evans and Tran-Nam, 2010). However, there is almost no evidence regarding how prefilled – and particularly incorrectly prefilled – tax returns affect the compliance behavior of taxpayers. Slemrod (2016), for example, states in his recent literature review on tax compliance that "the direct effect of pre-population on compliance is unclear" (p. 64). Kleven et al. (2011) provide initial empirical results that prefilling can have compliance consequences. They find that tax compliance is higher for income that is subject to third-party reporting (already prefilled) than for self-reported income (no prefilling). In contrast, Kotakorpi and Laamanen (2016) and Gillitzer and Skov (2016) observe that partly prefilled tax returns have either no effect or even a slightly negative effect on reported taxable income.

In a controlled experiment, we investigate how prefilled tax returns affect compliance behavior. We are able to clearly distinguish between correctly and incorrectly prefilled tax returns and can isolate their respective influences on tax compliance. We argue that prefilled tax returns influence the moral costs of tax evasion. More precisely, we expect that a deliberate adjustment of correctly prefilled values increases moral costs, and higher moral costs decrease the level of tax evasion. Thus, in such cases, prefilling enhances tax compliance. However, in the case of incorrectly prefilled tax returns, we argue that moral costs decrease, as the incorrect values are "somebody else's mistake", and the taxpayer can therefore evade taxes while maintaining a positive self-view. Our set-up additionally allows us to study how taxpayers adjust prefilled tax returns and how adjustment behavior varies between different prefilling scenarios. Our main contribution is that we extend the literature by separately investigating how correctly and incorrectly prefilled tax returns affect tax compliance. Furthermore, we contribute to the literature by controlling for several other explanations of why prefilling might influence compliance behavior.

In our lab experiment, we implement and vary different prefilling scenarios simultaneously. We only vary how tax returns are prefilled, and we ensure that all monetary aspects such as tax rate, audit probability, and penalties are kept constant. Thus, monetary costs are not influenced by the prefilling of the tax returns, and we can exclude that prefilling changes tax compliance behavior through a change in the audit probability or penalty. This experimental

design allows us to provide a clean test of how prefilling affects tax compliance behavior and provides a high level of internal validity. Consequently, we feel confident that conducting a laboratory experiment is an appropriate method of answering our research question.[1]

Our experiment consists of two parts. First, participants are given a real effort task to earn their pre-tax income. In the second part, participants have to file a tax return with six income fields, one for each round of the real effort task. Participants have the opportunity to evade taxes: if they report a lower income than actually earned, they reduce their tax liability. However, there is a certain probability that participants will be audited. If a participant is caught evading, she is punished with a fine.

In our study, we use four treatments. In the first treatment, participants have to file a blank tax return (i.e., no income field is prefilled). In the other three treatments, participants are confronted with a prefilled tax return. Each participant can easily adjust the prefilled values to enter their own values. In the second treatment, each income field of the tax return is correctly prefilled. In the third and fourth treatments, we prefill randomly chosen income fields with an incorrect income. In the third treatment, the prefilled income is lower than the true income; in the fourth treatment, the prefilled income is higher than the true income. The remaining income fields are still correctly prefilled. This feature allows us to study our research questions not only in a between-subject design but also in a within-subject design.

Our main results are fourfold. First, we find that subjects are more tax compliant when the tax return is correctly prefilled compared to when the tax return is blank. This effect is meaningful and economically significant. In fact, we observe an increase of 15.5 percentage points in the average compliance level. Thus, correct prefilling enhances tax compliance. Second, if the income items are incorrectly prefilled with income below the true income, tax compliance is significantly lower than for correctly prefilled tax returns. The compliance level observed is the same as that for blank tax returns.

Third, the tax compliance level is higher for prefilled tax returns where the reported income is incorrectly high. The tax compliance level does not significantly differ from that of correctly prefilled tax returns. This finding implies an asymmetric effect of incorrect prefilling: If taxpayers would benefit from incorrect prefilling, prefilling has no effect on tax compliance level when compared with blank tax returns. In contrast, if taxpayers would suffer from

[1] Alm et al. (2015) report evidence that laboratory experiments in the field of tax compliance exhibit external validity. They show that behavioral patterns of subjects in the laboratory correspond to decision-making in naturally occurring settings. For a detailed discussion of the use of laboratory experiments for tax research issues, see also Alm (2010).

incorrect prefilling, prefilling still has a positive effect on compliance compared to the case of blank tax returns, and compliance does not differ from compliance in the case of correctly prefilled tax returns.

Fourth, we find that individuals are aware of the incorrectly prefilled values and actively adjust them in their tax returns. Thus, we can rule out that a lack of awareness explains our findings. Moreover, as more than ninety percent of the incorrect values are adjusted by participants, we provide clear evidence that subjects do not stick to the default option and that the observed compliance behavior is not the result of a passive choice. Consequently, the default effect can be excluded as an explanation. Interestingly, we observe different adjustment behaviors depending on the type of prefilled values. Prefilled income higher than true income is almost always adjusted downward. For prefilled income below true income, however, adjustments depend on the deviation from true income. If the deviation is small, subjects do not adjust prefilled values, but if the deviation is large, subjects adjust lower prefilled values upward.

Our results have several implications. Most importantly, we show that prefilled values and defaults have a meaningful influence on compliance behavior. This finding contributes to the general literature on cheating and lying behavior, where prefilling effects have not yet been studied (Gneezy, 2005, Ariely, 2012, Erat and Gneezy, 2012, Battigalli et al., 2013). However, as a consequence of the digital transformation, individuals and decision makers in organizations and institutions are often confronted with prefilled data and prefilled forms in almost all industries. In situations where dishonesty is of concern, prefilling and defaults might reduce dishonest behavior. For example, financial accounting and reporting software (e.g., SAP ERP) automatically generates several financial indicators on the performance of a manager's division, and those indicators are used for budget and investment decisions. Our results suggest that the presence of those defaults and the quality of those defaults both influence the willingness to misreport such indicators. Furthermore, our results suggest that prefilling changes compliance behavior by changing the moral costs of cheating, lying and evading. So far, this effect is neither identified in the general lying and cheating literature nor identified as a potential explanation for how prefilling influences decision making in the tax compliance literature.

Our results suggest that prefilling forms or data entry masks is preferable over non-prefilling. Providing correctly prefilled forms enhances compliance compared to the use of blank forms. Although the positive compliance effect induced by correctly prefilled forms vanishes in the case of beneficially incorrectly prefilled forms, compliance is still at the same level as we observed with blank forms. Furthermore, our results show that prefilled values are

not revised if the prefilled values deviate only slightly from true values and the individual benefits from the incorrect prefilling. Consequently, due to errors that can occur by prefilling, we cannot expect to achieve the high compliance levels associated with correctly prefilled forms.

Furthermore, as prefilling tax returns is a service that the tax administration provides to taxpayers, our study is related to the service paradigm literature. For example, Alm et al. (2010) show that services from the tax administration (e.g., agency-provided information) have a positive and significant impact on compliance behavior. Our paper also contributes to the slippery slope literature (Kirchler et al., 2008) and to the trust paradigm literature (Alm and Torgler, 2011). Both bodies of literature show that tax compliance is influenced not only by the power of the tax authority to conduct audits, collect taxes and punish tax evasion (enforcement paradigm) but also by the trust in authority. While an increase in the power of authority leads to higher enforced tax compliance, an increase in the trust in authority leads to higher voluntary tax compliance (Wahl et al., 2010, Kastlunger et al., 2013, Kogler et al., 2013). If tax returns are correctly prefilled by the tax authority, the trust in authority might increase, and therefore tax compliance might increase as well. However, if tax returns are incorrectly prefilled, the trust in authority might be lower, leading to a decrease in tax compliance.

The remainder of this paper is structured as follows: In Section 2, we discuss the related literature and develop hypotheses. Section 3 describes the experimental design. We analyze compliance behavior in section 4 and adjustment behavior in section 5. Section 6 concludes.

2 Related Literature and Hypotheses

2.1 Prefilling and Tax Compliance

Our research questions of how prefilled – and particularly how incorrectly prefilled – tax returns influence tax compliance behavior are largely unexplored. Although some papers focus on prefilling (especially third-party reporting), the effect of incorrectly prefilled tax returns has not yet been studied. Kleven et al. (2011) analyze data from a tax enforcement experiment in Denmark and find that tax evasion is very low for income subject to third-party reporting and thus already prefilled in tax returns; however, they find that tax evasion is substantial for self-reported (i.e., not prefilled) income.

Kotakorpi and Laamanen (2016) use data from a natural experiment in Finland and examine tax reporting behavior when taxpayers receive prefilled tax returns. The authors observe that prefilling increases the number of deductions claimed but not the number of

income items reported. Rather, the authors find a significant reduction in the number of reported items that were not prefilled. More importantly, on an aggregated level, they do not find that prefilled tax returns influence total taxable income or taxes paid.

Gillitzer and Skov (2016) use data from the Danish tax authority and examine the case of prefilled deductions. Contrary to their expectations, they find that the number of tax deductions claimed doubles and that the total value of deductions increases if tax-deductible charitable contributions are already prefilled in the tax return. The authors suggest that taxpayers neglect to claim their tax-deductible charitable contributions if they are not already prefilled.

Our study substantially differs from previous studies. First, we use a laboratory experiment that enables us to focus on the influence of prefilled tax returns on compliance behavior in a controlled environment. Second, whereas previous studies have had to make the simplifying assumption that items are correctly prefilled, we are able to clearly distinguish between correctly and incorrectly prefilled tax returns and are able to isolate their respective influences on tax compliance. Third, previous studies have only analyzed the number of items claimed but not the actual compliance level. We design an experiment that enables us to analyze the level of tax compliance in more detail.

Fourth, our design allows us to control for the several potential explanations discussed by the studies mentioned above. Kleven et al. (2011), Gillitzer and Skov (2016) and Kotakorpi and Laamanen (2016) suggest that compliance is much higher for third-party reported (i.e., prefilled) items because the possibility of evading taxes is limited. We exclude this explanation with our experimental design, as our treatments offer the same opportunities for tax evasion in the cases of both prefilled and blank tax forms. Kotakorpi and Laamanen (2016) further discuss complexity effects as a possible explanation for changes in reporting behavior. We control for complexity by keeping the compliance decision in our experiment very simple. Participants have full information, there is no computation necessary, and complexity does not differ between our treatments.

2.2 Hypotheses

Initiated by the seminal papers of Becker (1968), Allingham and Sandmo (1972) and Yitzhaki (1974), a variety of papers have modeled tax evasion.[2] Monetary factors determining tax compliance level are the tax rate, audit probability and penalty in case of a detected tax evasion. More recently, theoretical papers have also been incorporating non-pecuniary factors

[2] See Alm et al. (1995), Andreoni et al. (1998), Torgler (2002), Hofmann et al. (2008), Alm (2012), and Slemrod (2016) for excellent literature reviews.

such as the moral costs of tax evasion (e.g., Fortin et al., 2007, Gordon, 1989, Kim, 2003, Myles and Naylor, 1996, Traxler, 2010). Tax morale is modeled as a social norm of tax compliance. Violating this norm creates moral costs. These studies provide robust evidence that tax compliance increases if moral costs increase.

In our study, the only variation is the prefilling in our treatments, and all monetary aspects are kept constant.[3] Prefilling has no influence on tax rate, audit probability, or penalty. Consequently, our experimental design ensures that monetary costs are not affected by different prefilling scenarios. Particularly, we can exclude that prefilling changes tax compliance behavior through a change in the audit probability or penalty.

However, we argue that prefilled tax returns influence the moral costs of tax evasion. In particular, if a tax return is prefilled with *correct* values, tax evasion requires a deliberate adjustment of those prefilled values. We expect that the act of replacing correct values already included in the tax return with incorrect numbers increases the moral costs associated with tax evasion. The tax compliance literature shows that higher moral costs increase the level of tax compliance. We therefore expect tax compliance to be higher in the case of correctly prefilled tax returns (i.e., prefilled income equals true income) compared to blank tax returns (i.e., tax returns are not prefilled). We define a correctly prefilled tax return as a prefilled tax return containing truthful values for all fields, thus resulting in an accurate tax base and tax liability for the taxpayer.

An additional explanation for there being different compliance behavior in the case of prefilled tax returns is the default effect. Several studies find evidence for this effect, which describes the preference of individuals to stay with a preset default option (in our case, prefilled income) rather than actively adjusting the default (Mazar and Hawkings, 2015). If individuals stick to the default option and thus do not adjust correctly prefilled values, they will submit a legally accurate tax return and consequently behave in a tax compliant way. Again, compared to the case without prefilled income (blank tax returns), we thus expect higher tax compliance in the correctly prefilled scenario.

Furthermore, tax compliance behavior might also be influenced by the anchoring effect if tax returns are already prefilled. According to Tversky and Kahneman (1974), the anchoring

[3] In accordance with the experimental tax compliance literature, we expect the tax compliance level to be greater than zero but below full compliance. We set the fiscal parameters (tax rate, audit probability and fine multiplier) in such a way that a purely payoff-maximizing (i.e., without moral concerns) and risk-neutral subject will always declare zero income (full evasion). Depending on the actual risk attitude, a purely payoff-maximizing and risk-averse subject will choose full evasion, full compliance or something in between.

effect describes the influence of an initially presented value on decision making.[4] This influence can be caused by an insufficient adjustment of this value, which serves as a starting point, so that the final decision is assimilated toward this starting point (Tversky and Kahneman, 1974, Epley and Gilovich, 2001). Another suggestion is that the influence of an initially presented value is caused by selective accessibility, i.e., an unconscious activation of knowledge that is consistent with the presented anchor (Strack and Mussweiler, 1997, Chapman and Johnson, 1999). Regardless of the underlying mechanisms that account for the anchoring effect, one could expect that taxpayers completing returns are biased toward the starting values of the prefilled tax return. Thus, compared to the blank tax return scenario (without any starting value), one could expect tax compliance to be higher in the correctly prefilled tax return scenario.

As all explanations tend in the same direction, we formulate the following hypothesis:

Hypothesis 1: *The tax compliance level is higher when taxpayers are given correctly prefilled tax returns compared to blank tax returns.*

Second, we analyze how incorrectly prefilled tax returns that would result in tax savings for the taxpayers (i.e., prefilled income is lower than true income) affect tax compliance. If prefilled values are not adjusted, a taxpayer declares less income than is actually earned, resulting in a lower tax liability and thus resulting in tax evasion. Compared to correctly prefilled tax returns, we expect that *incorrect prefilling* has an opposite effect on moral costs. In the case of incorrectly prefilled tax returns, license to cheat and responsibility shifting may reduce the moral costs of tax evasion. As income is already incorrectly prefilled (e.g., similar to an incorrect data transmission by the tax authority), taxpayers might feel they have the license to behave in a tax-evasive way, or they may feel less responsible for submitting incorrect values. As the lower income value already entered in the return is "somebody else's mistake", an individual can evade taxes while maintaining a positive self-view (Mazar et al., 2008). Consequently, the moral costs of tax evasion are lower compared to the moral costs of tax evasion with correctly prefilled tax returns. Thus, we would expect a lower compliance level if prefilled tax returns are incorrect, particularly if the prefilled income is below the true income.

If tax compliance behavior were influenced by anchoring effects, we would expect taxpayers to be assimilated toward the starting values of the prefilled tax returns. The prefilled lower income operates as an anchor toward lower income reporting. Consequently, if tax returns

[4] For a detailed literature overview see Furnham and Boo (2011).

are incorrectly prefilled with income values beneficial to the taxpayer, we expect tax compliance to be lower compared to compliance when tax returns are correctly prefilled. The same hypothesis will result if we assume that taxpayers are influenced by the default effect. If individuals have a preference to stay with a preset default and therefore do not adjust the prefilled values, tax compliance will be lower for lower prefilled than for correctly prefilled tax returns.

Overall, we formulate the following hypothesis:

> ***Hypothesis 2:*** *Compared to the tax compliance level in the case of correctly prefilled tax returns, the tax compliance level is lower if taxpayers are given incorrectly prefilled tax returns that contain income below the true income.*

Third, we analyze how incorrectly prefilled tax returns that would result in tax disadvantages for the taxpayers (i.e., prefilled income is higher than true income) affect tax compliance. In particular, if prefilled values are not adjusted by a participant, she declares more income than she actually earned, and her corresponding tax liability is higher.

In the case of higher prefilled tax returns, a downward adjustment from a higher prefilled value to the correct value is not associated with monetary costs because no taxes are evaded. Consequently, downward adjustments – at least to the correct value – are to be expected. An adjustment below the correct value is associated with costs. However, the associated monetary costs are identical to those in the correctly prefilled scenario, while the moral costs might be lower due to the incorrect prefilling. Consequently, from this perspective, the tax compliance level might be even lower for higher prefilled tax returns than for correctly prefilled tax returns.

> ***Hypothesis 3a:*** *Compared to the tax compliance level in the case of correctly prefilled tax returns, the tax compliance level does not differ or is even lower if taxpayers are confronted with incorrectly higher prefilled tax returns.*

If we assume that behavior is affected by anchoring effects, we would expect taxpayers to be assimilated toward the prefilled starting values. In particular, the higher prefilled income works as an anchor toward higher income reporting. Consequently, the anchoring effect should lead to a higher tax compliance level for higher prefilled than for correctly prefilled tax returns. If taxpayers are assumed to have a preference to stay with a preset default (default effect) and

therefore do not adjust the prefilled values, the same hypothesis will result. We formulate the following competing hypothesis:

> ***Hypothesis 3b:*** *Compared to the tax compliance level in the case of correctly prefilled tax returns, the tax compliance level is higher if taxpayers are confronted with incorrectly higher prefilled tax returns.*

Table 1 provides an overview of the potential behavioral responses to different prefillings of tax returns and the expected effects on tax compliance.

[Table 1]

3 Experimental Design and Treatments

3.1 Experimental Design

We conduct a laboratory experiment consisting of two parts. The instructions are provided to the participants at the beginning of the experiment (see online appendix A1). In the first part, participants work on a real effort task to earn their pre-tax income. In the second part, participants have to file a tax return to determine their tax liability.

We use a between-subject design with four treatments in which we vary whether and how participants' tax returns are prefilled (see section 3.2 for details). Subjects are randomly distributed to one of the four treatments. After participants have submitted their tax return, they are audited with a certain probability and have to pay a fine in case of a detected tax evasion. After the experiment, participants receive a tax-free show-up fee of 4 € as well as their earned pre-tax income minus their tax liability and minus a potential fine. Before the actual experiment is executed, we measure subjects' willingness to take risk with the Holt and Laury (2002) task (in €). The amount earned in the lottery task is also paid out to each participant at the end of the experiment.

In the first part of the experiment, we use the math puzzle task of Mazar et al. (2008). Participants see matrices with twelve numbers (each with two decimal places) on their screen and have to select the two numbers that add up to ten (e.g., 6.61 + 3.39 = 10). The math puzzle is a search task in which participants have to put in some effort to correctly solve the puzzles to earn money. In each matrix, there are only two numbers that add up to ten. Participants play six rounds of the math puzzle task, each lasting three minutes, with a one-minute break between the rounds. In each round, they can solve a maximum of 20 puzzles. For every correctly solved math puzzle, a participant earns a pre-tax income of 0.42 € (0 € otherwise). After each round,

the participant's number of correctly solved math puzzles and the resulting earned income in that round are displayed. In order to be able to complete the tax return in the second part of the experiment, participants are requested to record their earned income after each round on a piece of paper at their workstation.

After finishing the real effort task, participants file a tax return in the second part of the experiment. A tax of 25% is levied on the declared income. Participants have the opportunity to evade taxes when declaring less income than actually earned. In the instructions, participants are explicitly asked to declare their earned income from part 1. Thus, when declaring less than they earned, participants engage in tax evasion. Unintentional tax evasion by the taxpayer is excluded by design, as our setting is very simple and participants are fully aware of their true income in each round. They do not have to perform any calculations; their only task is to declare the income in the six income fields of the tax return. For this purpose, they are asked to use the records they made on the piece of paper in the first part of the experiment. Nevertheless, if participants desire, they can press a button on the screen to have their actual earned income for each round displayed when they file their tax return.

There is a 30% probability that participants will be audited after they have submitted their tax returns. If a participant is audited and her declared income is lower than her true income, she has to pay a fine that is twice the amount of the evaded taxes. This implies that in case of a detected tax evasion, the subject has to repay the evaded taxes plus additional penalty costs of 100% of the evaded taxes.[5] After subjects are informed about the audit outcome, the experiment is finished and the participants are asked to complete an ex-post questionnaire before they privately receive their payout in cash.

During the experiment, participants receive a "bank account", which is displayed to them after each part of the experiment. First, their total earned income in part 1 is credited to their account. Second, after participants have submitted their tax return, the resulting tax liability is debited to their account. If there is an audit and a tax evasion is detected, the resulting fine is debited to the account. As a final step, the show-up fee of 4 € and the payout from one randomly chosen lottery of the Holt and Laury (2002) task are added. The resulting balance is the participant's total payoff from the experiment.

Before the experiment starts, subjects have to successfully complete a comprehension test. They are asked several questions regarding the puzzle task, pre-tax income determination, tax

[5] In cases where an audit reveals that the declared income is higher than the actual earned income, the subject gets the overpaid taxes back, but no additional costs occur.

liability determination, audit probability and payoff determination. The full set of questions can be found in our online appendix A1.4.

3.2 Treatments

In the second part of the experiment, participants are given a tax return on which their income must be declared. The displayed tax return consists of six income fields, one for the income of each round of the real effort task. Below these fields, the resulting total declared income, as well as the resulting tax liability, is displayed to the participants. In our study, we use the following four treatments (between-subject design) in which we manipulate the tax return that participants have to file. Figure 1 presents our experimental design and highlights the treatment differences.

- *Treatment B (**b**lank)*: income fields are not prefilled, *blank* tax return.
- *Treatment CP (**c**orrectly **p**refilled)*: all income fields of the tax return are *correctly* prefilled (i.e., prefilled income equals true income).
- *Treatment LP (**l**ower **p**refilled)*: 3 income fields of the tax return are *incorrectly lower* prefilled to the advantage of the taxpayer (i.e., prefilled income is *lower* than true income); the other 3 income fields are correctly prefilled.
- *Treatment HP (**h**igher **p**refilled)*: 3 income fields of the tax return are *incorrectly higher* prefilled to the disadvantage of the taxpayer (i.e., prefilled income is *higher* than true income); the other 3 income fields are correctly prefilled.

In treatment B, participants have to file a blank tax return, meaning that no income field is prefilled. In the other three treatments, participants receive a tax return where all six income fields are already prefilled. All prefilled values can be adjusted by participants. In treatment CP, each income field of the tax return is correctly prefilled (i.e., contains participants' true income from each round of the real effort task). In both treatments LP and HP, we manipulate three randomly chosen income fields by prefilling an incorrect income. In treatment LP (HP), the incorrect income is lower (higher) than the corresponding actually earned income. The other three unmanipulated income fields are correctly prefilled. This within-subject variation enables us to analyze how the same subject responds to both an incorrect and correct prefilling.

[Figure 1]

For each incorrectly prefilled value, we randomly vary the level of deviation from the correct amount. Deviation is equally distributed and lies between 0% and 100% of the actually earned income. For treatment LP, this implies that the manipulated prefilled income lies

between 0% (when deviation is 100%) and 100% (when deviation is 0%) of the true income. For treatment HP, this implies that the manipulated prefilled income lies between 100% (when deviation is 0%) and 200% (when deviation is 100%) of the true income. For example, suppose that the actually earned income is 5 €. Consequently, the incorrectly prefilled income lies between 0 and 5 € in treatment LP and between 5 and 10 € in treatment HP. As deviation is determined randomly for each manipulated income field and thus naturally varies across these fields, this setting allows us to examine how different levels of deviation influence taxpayers' compliance behavior.

In the instructions, we ask participants to check whether the prefilled tax return is correct. Each prefilled value of the tax return can be changed as desired by the participant, with a minimum of 0 € and a maximum of 8.40 € (20 correctly solved puzzles à 0.42 €). The same applies to the values the participants are allowed to enter in the blank tax return in treatment B. Empty fields are not allowed. Moreover, participants can always use a computation button on the screen that displays the resulting total declared income and the resulting tax liability based on the inserted income values. There is no computation restriction, so participants can use the button as often as they want to.

We control for all other factors that may influence tax compliance, such as audit probability, fine, and tax rate, by keeping them constant between the different treatments. In our setting, the only variable manipulated is the prefilling of the tax returns. If there are differences in compliance, they can only result from our prefilling manipulation.

3.3 Sample and Data

The experiment was conducted at the computerized experimental laboratory of the University of Cologne (CLER) from August to October 2017. The experimental software was programmed and used with the software z-Tree (Fischbacher, 2007). Participants were recruited with ORSEE (Greiner, 2004). In total, 213 subjects (mainly undergraduate students, 122 females and 91 males) participated and earned, on average, 19.04 € in approximately 97 minutes (approximately 11.78 € per hour). A total of 42 subjects were randomly assigned to treatment B, 43 to treatment CP, 64 to treatment LP and 64 to treatment HP. Over all treatments, we have 1,278 observations (213 subjects with 6 decisions per subject). Table 2 provides an overview on the main characteristics of our participants.

[Table 2]

4 Results: Declared Income

4.1 Descriptive Statistics and Nonparametric Statistics

Our measure for tax compliance is the declared share of true income. If a taxpayer declares the income she actually earned, she complies with the tax law. If she declares an income lower (higher) than the income actually earned, she acts in a tax evasive (over-compliant) way. To account for earned income heterogeneity, we use a relative measure and calculate the ratio between declared income and actually earned income for each single income field. As one subject declares income in six income fields, we have six observations per subject. For the following descriptive and nonparametric analyses, we calculated an average tax compliance level over these observations for each subject. For treatments LP and HP, this was performed separately for manipulated and unmanipulated income fields. Figure 2 shows the mean values of declared share of income for each treatment.[6]

[Figure 2]

The mean declared share of true income lies between 71.1% and 91.9% in our treatments. In line with the experimental tax compliance literature, average tax compliance is therefore higher than 0% but below 100% in all four of our treatments.

The mean declared share of true income in treatment blank (B) is 72.2%, and it is 87.7% in the correctly prefilled treatment (CP). In line with our first hypothesis, tax compliance significantly increases when tax returns are correctly prefilled compared to when blank tax returns are used (Mann-Whitney U-test, $p = 0.048$, two-tailed). Thus, hypothesis 1 is supported and we can formulate the following result:

> ***Result 1:*** *Compared to blank tax returns, the tax compliance level is significantly higher when tax returns are already correctly prefilled.*

The mean declared share of income in treatment LP is 71.1% for the incorrectly prefilled income fields. This share is significantly lower than the mean declared share of 87.7% in treatment CP (Mann-Whitney U-test, $p = 0.011$, two-tailed). Additionally, we can compare the declared share of income in a within-subject comparison, as three random income fields were

[6] In our experiment, the largest deviation possible is 100% of the actually earned income. This implies that the manipulated prefilled values can range from 0% (in treatment LP) to 200% (in treatment HP) of the actually earned income. There is no reason for a declared income higher than 200% of the actually earned income. Therefore, we excluded observations (not subjects) with a declared income greater than 200% of the actually earned income. Please note that all stated results will stay the same if we include all observations in our analysis.

incorrectly prefilled (manipulated) and the other three income fields were correctly prefilled (unmanipulated). For treatment LP, we find that the average declared share of income is significantly lower for the incorrectly prefilled income fields (71.1%) compared to correctly prefilled income fields (79.1%) (Wilcoxon signed-rank test, $p = 0.007$, two-tailed). Thus, hypothesis 2 is supported by both a between-subject (treatment CP vs. LP) and a within-subject (treatment LP) comparison:

> ***Result 2:*** *Compared to correctly prefilled tax returns, the tax compliance level is significantly lower if tax returns are incorrectly prefilled with income values that are below the true income.*

In treatment HP, we observe a mean declared share of income of 91.9% for the incorrectly prefilled (manipulated) income fields. The difference between the declared share of income for the incorrectly prefilled income fields in treatment HP and correctly prefilled income fields in treatment CP is not statistically significant (Mann-Whitney U-test, $p = 0.738$, two-tailed). In the within-subject comparison we also find that the declared share of income for the incorrectly prefilled income fields (91.9%) does not statistically differ from the declared income for the correctly prefilled income fields (89.8%; Wilcoxon signed-rank test, $p = 0.795$, two-tailed). Thus, we do not observe a significant difference between incorrectly prefilled tax returns and correctly prefilled tax returns – neither in a between-subject comparison nor in a within-subject comparison. This result supports hypothesis 3a, whereas we cannot confirm hypothesis 3b.

> ***Result 3:*** *The tax compliance level does not significantly differ when tax returns are incorrectly prefilled with a higher income compared to correctly prefilled tax returns.*

4.2 Regression Analyses

To corroborate our descriptive and nonparametric results, we also present a linear regression analysis. In contrast to the nonparametric tests, we now consider each observation of the six income fields per subject. The dependent variable is the declared share of true income of each subject in each income field. To account for dependency of the six decisions within one subject, we consider linear regression models with subject-specific random effects.[7] Moreover, we consider a vector of subject characteristics such as gender, age, and risk attitude.

[7] We use linear regression models with random effects where the income field number is the time variable and the subject's identity number is the cross-sectional variable.

Our variables of interest are, however, indicator variables for the treatment variations. Treatment CP serves as the default. Thus, each coefficient of our treatment dummies measures the difference between the respective treatment and treatment CP.

Table 3 reports the regression results. In both models, all observations from treatments B and CP are considered. However, in specification (1), we consider only the incorrectly (manipulated) income fields, while in specification (2), we consider the correctly (unmanipulated) income fields of treatments LP and HP. Model 1 enables us to study whether prefilling manipulation has an effect on tax compliance level. Model 2 analyzes whether differences exist between treatments when income fields are not manipulated in the treatments LP and HP.

In both models, we observe a significant negative coefficient for treatment B. We find a lower tax compliance level in the treatment with blank tax returns compared to the treatment with correctly prefilled tax returns. Thus, we are able to confirm our result 1. Moreover, in model 1, we observe a significant negative coefficient for treatment LP. Hence, we find a lower tax compliance level compared to the treatment with correctly prefilled tax returns if prefilled income values are below true income. In contrast, we do not find a significant coefficient for treatment HP. Consequently, the tax compliance level does not statistically differ from that of correctly prefilled tax returns if the tax returns include prefilled income that is higher than true income. These findings confirm our results 2 and 3 in a between-subject comparison. Statistical significance between treatments B and LP, treatments B and HP, and treatments LP and HP was checked by Wald tests, and the resulting p-values are reported at the bottom of the table.

As a control group, we consider in specification (2) only the correctly prefilled income fields of treatments LP and HP. We find no compliance differences between subjects if income fields are correctly prefilled.

[Table 3]

Table 4 reports the outcome for our within-subject comparison. In model 3 (4) we analyze the differences within treatment LP (HP). In both models, the dummy variable “manipulated” takes the value of one if the declared share of income refers to a manipulated income field (0 otherwise). In model 3 (treatment LP), we observe a significant negative coefficient for the manipulation dummy on the declared share of income. Thus, the declared share of income is significantly lower for the lower prefilled income fields than for the correctly prefilled income fields. In contrast, we do not find a significant coefficient in model 4 (treatment HP). The declared share of income for the higher prefilled income fields does not differ from the correctly

prefilled income fields. As a result, we are also able to confirm our stated results 2 and 3 in a within-subject comparison by the regression analysis.

[Table 4]

In all models, we include individual specific variables such as gender, age, and risk attitude. In total, we incorporate the 15 individual variables reported in table 2 in our regressions. The individual variables are not displayed, but the complete set of all regression results is presented in our online appendix A2 (tables A2.1 and A2.2). In line with the literature on tax compliance, we observe that women are significantly more tax compliant than men and that subjects are less compliant the more risk-seeking they are. Furthermore, we find evidence that subjects with higher tax morale are more tax compliant, and we find that the more complex the subjects perceive the tax-related decisions in our experiment to be, the more tax compliant they are.

We can summarize that our explorative and nonparametric findings are supported by our random effects linear regressions. In particular, we are able to confirm our finding that tax compliance is higher in the case of correctly prefilled tax returns compared to blank tax returns (result 1). Furthermore, we find support for the finding that tax compliance is lower in the case of incorrectly prefilled income that is below true income (result 2). Moreover, we find that tax compliance does not differ when tax returns are incorrectly prefilled with income values that are higher than true income (result 3).

4.3 Treatment Blank vs. Different Prefilled Variations

We further analyze the differences between the treatment with blank tax returns and each of the treatments with prefilled tax returns. Table 5 reports the results of linear regression models with random effects for a between-subject comparison, as used in section 4.2. In this case, the non-prefilled treatment B serves as the default. This analysis allows us to compare each prefilled scenario with the case of blank tax forms. Again, we consider observations from either the manipulated (model 5) or unmanipulated (model 6) income fields of treatments LP and HP.

The results again suggest that correctly prefilled tax returns significantly increase compliance compared to blank tax returns. When comparing blank tax returns with incorrectly prefilled tax returns, we find no significant difference in compliance if the prefilled income is below the true income. Thus, the positive compliance effect induced by correctly prefilled tax returns vanishes if the prefilled tax return contains errors that lead to tax savings for the taxpayer.

If the tax return is prefilled with income values that are higher than the true income, tax compliance is significantly higher than in the case of non-prefilled tax returns. As the difference between treatment CP and treatment HP is statistically insignificant (see table 3 and section 4.2), errors in the tax return that lead to tax disadvantages for the taxpayer do not further increase compliance compared with correctly prefilled tax returns.

[Table 5]

4.4 Robustness Tests

Perceived Audit Probability

There is ample evidence that an increased audit probability increases tax compliance (see Torgler, 2002, for an overview). Although our prefilling variation has no influence on the objective audit probability, prefilling might increase the subjective probability of an audit. If the tax return is prefilled by the tax authority, a taxpayer might believe that changing the prefilled values will increase the probability that the tax authority will conduct an audit. To exclude this as an explanation, we first checked whether participants were aware of the audit probability of 30%. For this purpose, participants had to state the audit probability before (in our comprehension test) and after the experiment (in our ex-post questionnaire). There was only one person out of our 213 participants who gave an incorrect answer on the ex-post questionnaire.[8] Second, and more importantly, we asked our participants in our ex-post questionnaire: "How did you perceive the audit probability in the experiment?" (10-point Likert scale from "very low" to "very high"). Over all treatments, the mean answer was 3.6. Differences across treatments are small and statistically insignificant (Kruskal-Wallis test, $p = 0.291$, two-tailed).[9] Consequently, we are able to exclude that compliance differences across treatments are driven by different levels of perceived audit probability.

Individual Variables across Treatments

We also checked whether the distribution of individual characteristics (such as gender, risk attitude, tax morale, and age) differs across treatments. For each of our 15 individual variables reported in table 2, we either applied the chi-squared test (for binominal variables) or Kruskal-Wallis test (for ordinal and interval variables). We do not observe that our treatments differ significantly with respect to individual variables (all p-values above 0.1, two-tailed). The only

[8] That incorrect answer was 10% and was given in treatment HP.

[9] Mean values across treatments are 3.43 in treatment B, 3.26 in CP, 3.95 in LP and 3.61 in HP.

exception occurs with respect to fairness (Kruskal-Wallis test, $p = 0.036$, two-tailed). In particular, in our ex-post questionnaire, we asked: "How fair did you perceive the tax and control system that was applied in the experiment to be?" (11-point Likert scale from "very unfair" to "very fair"). Mean values across treatments are 6.38 in treatment B, 6.67 in CP, 6.02 in LP and 5.41 in HP. Pairwise comparison reveals significant differences between treatments B and HP as well as between treatments CP and HP (Mann-Whitney U-test, both p-values below 0.05, two-tailed). Consequently, the tax regime is perceived as significantly less fair in treatment HP, where the tax returns are incorrectly prefilled to the disadvantage of the subjects.[10]

Influence of Tax Morale on Treatment Effects

In a robustness test, we analyze how our treatment effects are influenced by the individual tax morale of our subjects. In an ex-post questionnaire, we use an adapted question from the World Values Survey that is widely applied in the tax compliance literature (e.g., Slemrod 2003, Alm and Torgler, 2006). We asked: "How do you evaluate the following statement?: Cheating on tax if you have the chance..." Answers were given on a 10-point Likert scale from "…is always justifiable" = 0 to "…is never justifiable" = 9. The observed median (mean) tax morale level in our sample is 8 (7.15).

We use a median split to categorize our subjects into two subsamples. Subjects with a tax morale level of 8 and higher are categorized in the "high tax morale" subsample. Subjects with a level below the median are categorized in the "low tax morale" subsample. Figure 5 shows the mean declared share of income in all treatments for both subsamples. For the subsample with high tax morale (Panel A), we observe the same results (including statistical significance) as reported for the whole sample. For the subsample with low tax morale (Panel B), however, we do not observe any significant differences between our treatments – neither for the between-subject nor for the within-subject comparisons. Therefore, we can conclude that prefilling has a strong impact on compliance behavior for individuals with high tax morale but not for individuals with low tax morale.

[Figure 3]

[10] The ex-post questionnaire was completed after all decisions were made and after the information was provided regarding whether the subject's tax declaration was audited. As the number of audits did not differ significantly across treatments (chi squared test, $p = 0.459$, two-tailed), the lower perceived fairness observed in treatment HP cannot be attributed to a higher audit frequency in this treatment.

Influence of Earned Income on Tax Compliance

In the tax compliance literature, there is much evidence that tax compliance decreases with income (see for example Grundmann and Lambsdorff, 2017, for a recent literature review and recent results). In line with this finding, we observe a significantly negative correlation between earned income (in €) and (relative) declared share of true income. Over all treatments, Spearman's rank correlation coefficient equals -0.23 and is highly significant ($p < 0.01$, two-tailed).[11] We reran our regression analyses and added earned income as a control variable (not reported). All our stated results are robust to this variation. In all regressions, the coefficient of earned income is significantly negative (all p-values below 0.01, two-tailed).

Consistent Tax Compliance Behavior

As common in tax compliance experiments, we observe that participants consistently choose either full compliance or full evasion. Over all treatments, 48.8% of our subjects declared the actual income truthfully in each of the six income fields (always full compliance). In contrast, 12.7% of our participants chose to always declare zero income (always full evasion). Consequently, 61.5% of all subjects revealed consistent tax compliance behavior over the six compliance decisions. See table 6 for treatment details. Although distributions do not differ significantly across treatments (two-tailed, chi-squared test, all p-values above 0.3), we reran our regression analyses and excluded subjects who revealed such a consistent behavior. Regression results are reported in table A2.3 and A2.4 in our online appendix A2. As expected, treatment effect sizes increase. Nevertheless, all our stated results are robust to this variation.

[Table 6]

5 Results: Adjustments of Prefilled Values

In this section, we analyze whether and, if so, to what extent prefilled tax returns are adjusted. Adjustments are defined by the difference between declared and prefilled income. To account for income heterogeneity, we again use a relative measure and normalize to actually earned income.[12] Participants' upward (downward) adjustment of prefilled income of +10% (-

[11] Spearman's rank correlation coefficient lies between -0.19 and -0.26 in each treatment and is highly significant (all p-values below 0.01, two-tailed).

[12] We normalize to actually earned income because this allows a clear comparison between the adjustments across treatments. An alternative would be to normalize to the prefilled value (i.e., difference between declared and prefilled values divided by prefilled value). However, this makes a direct comparison across treatments difficult. For example, suppose that earned income is 10 and prefilled values are 6 in treatments LP and 14 in

10%) indicates that the declared share of income is increased (reduced) by 10% of the actually earned income. If, for example, the earned income is 10 and the prefilled value 6, the declared share of income is 7 (5) in case of a 10% upward (downward) adjustment. Table 7 shows the distribution and mean values of the adjustments made in our treatments for all observations, as well as for the correctly and incorrectly prefilled income fields separately.

[Table 7]

Most importantly, we observe that almost all incorrectly prefilled income fields are adjusted (approx. 93%). Only 6.3% (6.9%) of incorrectly prefilled values are accepted by the subjects as prefilled in treatment LP (HP). In contrast, 64% to 70% of all correctly prefilled values are accepted and not adjusted. This gives clear evidence that subjects are aware of the incorrectly prefilled values and that they do not remain passive or stick to the prefilled values.

Result 4: *Subjects are aware of incorrectly prefilled income fields and adjust them.*

With respect to adjustment differences across our scenarios, we observe a similar decision pattern in each treatment for correctly prefilled income fields. In all three prefilled treatments, 64% to 70% of these income fields are unadjusted, 24% to 34% are adjusted downward, and only a negligible minority (3% to 6%) are adjusted upward. The mean adjustments are negative in all treatments and lie between -0.12 and -0.21.

In contrast, crucial differences are observed for the incorrectly prefilled income fields. Whereas almost all incorrectly prefilled income fields (92%) are adjusted downward in treatment HP, only 27% are adjusted downward in treatment LP. In the latter, the majority (67%) is adjusted upward. This leads to a positive mean adjustment level of +0.23 in treatment LP and a negative level of -0.60 in HP.[13]

HP (i.e., prefilled values deviate by -40% and +40%, respectively). If an adjustment is made by 1 in both treatments, the alternative measure will be +/-16.67% for treatment LP and +/-7.14% for HP. As prefilled values are (by design) always lower in LP than in HP, the measure will always be higher in LP, although adjustments are identical in absolute terms. With our measure, this issue is avoided and reveals adjustments of +/-10% for both treatments.

[13] Interestingly, in treatment LP, the positive mean adjustment level of +0.23 in the case of incorrectly prefilled values is approximately as high as the negative adjustment level of -0.21 in the case of correctly prefilled values. Thus, adjustments cancel each other out on an aggregated level (-0.01). In contrast, subjects on average adjust prefilled income downward in cases of correctly (-0.12) and incorrectly (-0.60) prefilled values in treatment HP. Consequently, the mean adjustment level is negative (-0.36) on an aggregated level.

Result 5: *Whereas in treatment HP, almost all incorrectly prefilled income fields are adjusted downward, in treatment LP, the majority of income fields are adjusted upward.*

Figure 4 shows the mean adjustments for the incorrectly prefilled income fields in treatments LP and HP. This figure also exhibits the difference between prefilled and earned income and between declared and earned income (each normalized to actually earned income).

[Figure 4]

As prefilled values deviate from correct income randomly between 0% and 100% of actually earned income (equal distribution), the normalized difference between prefilled and earned income is, on average, approximately -0.5 in treatment LP and +0.5 in treatment HP. In treatment LP, we find that the observed mean adjustments for the incorrectly prefilled income fields (+0.23) do not compensate the prefilling error (-0.52). Thus, a negative (normalized) difference between declared and true income results (-0.29). In treatment HP, the observed mean adjustments (-0.60) overcompensate the prefilling error (+0.52). Accordingly, we observe a negative (normalized) difference between declared and true income (-0.08). However, overcompensation is not large enough to achieve the low tax compliance level observed in treatment LP (-0.29 vs. -0.08). Consequently, declared share of income is ultimately lower in treatment LP than in HP.

Result 6: *In treatment LP, adjustments do not compensate lower prefilling. In treatment HP, adjustments overcompensate higher prefilling.*

In a final step, we analyze how the level of deviation of prefilled income from true income influences adjustment behavior. Figure 5 shows the adjustments made by subjects depending on the deviation of prefilled income from actually earned income. In treatment HP, we observe an almost linear relationship between adjustments and deviations. Consequently, prefilled values are adjusted in accordance with the deviation level: the larger the prefilled income deviates from true income, the larger the downward adjustment. On average, the difference between declared and prefilled income equals -0.35 for small deviation levels (deviations below or equal to 50%) and -0.80 for large deviation levels (deviations above 50%). In both cases, adjustments differ significantly from zero (two-tailed, one sample median test, $p < 0.001$).

In treatment LP, we observe no linear relationship. For small deviation levels, we find almost no adjustments. If adjustments are made, they are statistically insignificant (two-tailed,

one sample median test, $p = 0.545$). In contrast, for large deviation levels, we find nearly constant adjustments in a range of +0.4 to +0.6. The mean difference between declared and prefilled income equals +0.51 and differs significantly from zero ($p < 0.001$). Consequently, if the prefilled income deviates only slightly from the true income, subjects keep the lower prefilled value without adjusting it. Above a deviation of 50%, subjects adjust the prefilled income upwards.

Result 7: *Subjects make no adjustments and keep lower prefilled income values if the prefilled values do not differ too greatly from the true values. For large deviations from true income, however, subjects adjust lower prefilled income values upward. In contrast, higher prefilled income values are almost always adjusted downward.*

[Figure 5]

As a robustness test, we reran our regression analyses separately for small (below or equal to 50%) and large (above 50%) deviations. The regression results are reported in tables A2.5 and A2.6 in our online appendix A2. For both subsamples, we observe the same pattern as before, and all our stated results are robust to this variation. Consequently, whereas adjustments depend on the deviation level, the relative declared share of income does not.

6 Conclusion

As a consequence of the digital transformation, the process of completing forms has fundamentally changed. In this study, we conduct a controlled experiment with 213 participants to analyze how prefilled forms influence compliance behavior. We frame our experiment as filing the annual income tax return.

We observe that tax compliance is significantly higher when tax returns are correctly prefilled compared to when tax returns are not prefilled (i.e., blank tax returns). Consequently, correct prefilling enhances tax compliance. More importantly, we study the influence of incorrectly prefilled tax returns and find asymmetric results. On the one hand, if prefilled income is lower than true income, we find no positive effect of prefilling: compared to the setting with blank tax returns, tax compliance is on the same level. However, compared to the setting with correctly prefilled tax returns, compliance is significantly lower when prefilled income is lower than true income. On the other hand, if prefilled income is higher than true

income, prefilling still has a positive effect: compliance is on the same level as with correctly prefilled tax returns and is therefore significantly higher than with blank tax returns.

One merit of our experimental design is that it enables us to examine how individuals adjust prefilled income. We find clear evidence that individuals actively make adjustments. In particular, more than 93% of the incorrectly prefilled values are changed by participants. This result enables us to exclude alternative explanations for differing compliance behavior between our treatments. First, a lack of awareness that prefilled values are incorrect can be excluded as an explanation. Second, the adjustment behavior clearly shows that individuals do not have a tendency to remain passive, and compliance behavior is therefore not the result of a passive choice. Participants do not stick to the default option (default effect) but rather make use of their opportunity to change the incorrectly prefilled values.

When analyzing the adjustment behavior for incorrect income values, we observe different adjustment behaviors depending on the direction of error. We find that higher prefilled income is nearly always adjusted downward. In contrast, adjustments of lower prefilled income depend on how much the prefilled income deviates from the true income. If prefilled values deviate only slightly from true income, individuals make no adjustments and keep the lower prefilled values. For larger deviations, however, subjects adjust lower prefilled values upward. We also observe that, on average, these adjustments do not compensate the initial lower prefilling.

Our study suggests that prefilling can help to increase compliance but that the quality of prefilling matters. In cases where the tax authority is able to prefill tax returns with data from employers, social insurance agencies or banks, the likelihood of an incorrect prefilling might be rather small. However, prefilled values will generally not match with actual values if an electronic tax return program carries forward the values entered in the tax forms of previous years. Our results suggest that particularly in cases of small deviations, prefilled values are not revised if taxpayers benefit from the incorrect prefilling. Additionally, a recent trend is the use of software applications that automatically generate prefilled tax returns after scanning pay slips, bills and receipts (e.g., taxbutler in Germany). However, incorrectly prefilled tax returns caused by technical errors of optical character recognition could lead to serious tax compliance issues, as reported in our study.

Our results also provide an answer to the following question: Should we make use of prefilled forms or prefilled entry masks? Indeed, prefilled forms reduce compliance costs. However, what about compliance behavior on an aggregated level? We observe that correctly and higher prefilled tax returns significantly increase tax compliance compared to blank tax returns. This finding supports the utilization of prefilled forms. By contrast, the positive

compliance effect induced by correctly prefilled tax returns vanishes if the prefilled tax return contains errors that lead to tax savings for the taxpayer. However, tax compliance for lower prefilled tax returns is, on average, at the same level as for blank tax returns. From this perspective, our results suggest that prefilling does not always outperform non-prefilling in every situation but that prefilling is preferable over non-prefilling on an aggregated level. Furthermore, our results indicate that observed prefilling effects are mainly driven by individuals with high tax morale. In this case, prefilling has a strong impact on compliance behavior. For individuals with relatively low tax morale, however, the positive effects of prefilling compared to non-prefilling should not be expected. This finding also implies that the influence of prefilling on compliance behavior might differ across countries.

Our results contribute to the general literature on cheating and lying behavior (Gneezy, 2005, Ariely, 2012). We provide evidence that prefilled forms and defaults change compliance by changing the moral costs of dishonest behavior. So far, this effect has not been studied in either the general lying and cheating literature or in the tax compliance literature.

Our study does have some limitations. First, we assume complete information about true income, while in reality, taxpayers might sometimes have incomplete information about their true income. Uncertainty might amplify the tendency for taxpayers to rely on prefilled values. In this case, the prefilled value implies an informative signal. Relying more on prefilled values increases tax compliance in the case of correctly and higher prefilled tax returns but decreases tax compliance in the case of lower prefilled tax returns. Consequently, the compliance gaps between lower and correctly prefilled tax returns, as well as the gaps between lower and higher prefilled tax returns, should increase further. However, the resulting difference between the correctly and higher prefilled tax return scenarios is not obvious.

Another issue is how information is distributed. In our paper, we assume symmetric information between the taxpayer and the tax authority. In reality, information about the taxpayer's income is often asymmetrically distributed, and the taxpayer can exploit this asymmetry strategically. If prefilled values reflect the information set of the tax authority, we might therefore observe that tax compliance further decreases in the lower prefilled treatment. However, we might not observe significant changes in the higher prefilled treatment because a taxpayer can submit evidence (e.g., receipts, invoices) to clarify why she has not accepted the prefilled values. Further research is also required here.

One limitation of our study is that we use students as subjects. Although this has several strong advantages (e.g., homogenous sample, high cognitive capability, low opportunity costs to ensure incentive compatibility), our results have to be treated with caution regarding external

validity. However, as we are not interested in how subjects solve complex case studies where expertise is crucial or where a special context is important, we decided to use students. There is much evidence that student decision-making does not differ significantly from that of professionals and non-students – especially if the complexity of the applied experimental task is low like it is in our experiment (Alm et al., 2015, Depositario et al., 2009, Remus, 1996, Ashton and Kramer, 1980, Elliot et al., 2007). Therefore, we feel confident that using students as subjects is appropriate in our setting.

References

Allingham, M. G. and Sandmo, A. (1972). Income tax evasion: a theoretical analysis. *Journal of Public Economics*, 1, 323–338.

Alm, J. (2010). Testing behavioral public economics theories in the laboratory. *National Tax Journal*, 63(4), 635–658.

Alm, J. (2012). Measuring, explaining, and controlling tax evasion: lessons from theory, experiments, and field studies. *International Tax and Public Finance*, 19(1), 54–77.

Alm, J., Bloomquist, K. M. and McKee, M. (2015). On the external validity of laboratory tax compliance experiments. *Economic Inquiry*, 53(2), 1170–1186.

Alm, J., Cherry, T., Jones, M. and McKee, M. (2010). Taxpayer information assistance services and tax compliance behavior. *Journal of Economic Psychology*, 31(4), 577–586.

Alm, J., Sanchez, I. and De Juan, A. (1995). Economic and noneconomic factors in tax compliance. *Kyklos*, 48, 3–18.

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

Alm, J. and Torgler, B. (2011). Do ethics matter? Tax compliance and morality. *Journal of Business Ethics*, 101(4), 635–651.

Andreoni, J., Erard, B. and Feinstein, J. (1998). Tax compliance. *Journal of economic literature*, 36, 818–860.

Ariely, D. (2012). The honest truth about dishonesty: How we lie to everyone – especially ourselves. Harper, New York.

Ashton, R. H. and Kramer, S. S. (1980). Students as surrogates in behavioral accounting research: Some Evidence. *Journal of Accounting Research*, 18, 1–13.

Battigalli, P., Charness, G. and Dufwenberg, M. (2013). Deception: The role of guilt. *Journal of Economic Behavior & Organization*, 93, 227–232.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76, 169–217.

Chapman, G. B. and Johnson, E. J. (1999). Anchoring, activation, and the construction of values. *Organizational Behavior and Human Decision Processes*, 79, 115–153.

Depositario, D., Nayga, R., Wu, X. and Laude, T. P. (2009). Should students be used as subjects in experimental auctions? *Economics Letters*, 102, 122–124.

Elliot, W., Hodge, F., Kennedy, J. and Pronk, M. (2007). Are MBA students a good proxy for nonprofessional Investors? *The Accounting Review*, 82, 139–168.

Epley, N. and Gilovich, T. (2001). Putting adjustment back in the anchoring and adjustment heuristic: Differential processing of self-generated and experimenter-provided anchors. *Psychological science*, 12(5), 391–396.

Erat, S. and Gneezy, U. (2012). White lies. *Management Science*, 58, 723–733.

Evans, C. and Tran-Nam, B. (2010). Managing tax system complexity: building bridges through prefilled tax returns. *Australian Tax Forum*, 25, 245−274.

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

Fortin, B., Lacroix, G. and Villeval, M. C. (2007). Tax evasion and social interactions. *Journal of Public Economics*, 91(11-12), 2089–2112.

Furnham, A. and Boo, H. C. (2011). A literature review of the anchoring effect. *Journal of Socio-Economics*, 40, 35–42.

Gillitzer, C. and Skov, P. E. (2016). The use of third-party information reporting for tax deductions: Evidence and implications from charitable deductions in Denmark. EPRU Working Paper Series.

Gneezy, U. (2005). Deception: The role of consequences. *American Economic Review*, 9, 384–394.

Goolsbee, A. (2006). The simple return: Reducing America's tax burden through return-free filing. Brookings Institution Hamilton Project Discussion Paper 2006-04.

Gordon, J. P. (1989). Individual morality and reputation costs as deterrents to tax evasion. *European economic review*, 33(4), 797–805.

Greiner, B. (2004). An online recruitment system for economic experiments. *Forschung und wissenschaftliches Rechnen GWDG Bericht 63*. edited by K. Kremer and V. Macho. Göttingen, Gesellschaft für Wissenschaftliche Datenverarbeitung, 79–93.

Grundmann, S. and Lambsdorff, J. (2017). How income and tax rates provoke cheating – An experimental investigation of tax morale. *Journal of Economic Psychology*, 63, 27–942.

Hofmann, E., Hoelzl, E. and Kirchler, E. (2008). Preconditions of voluntary tax compliance: Knowledge and evaluation of taxation, norms, fairness, and motivation to cooperate. *Journal of Psychology*, 216, 209–217.

Holt, C. A. and Laury, S. (2002). Risk aversion and incentive effects. *American Economic Review*, 92(5), 1644–1655.

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

Kim, Y. (2003). Income distribution and equilibrium multiplicity in a stigma-based model of tax evasion. *Journal of Public Economics*, 87(7-8), 1591–1616.

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

Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S. and Saez, E. (2011). Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica*, 79(3), 651–692.

Klun, M. (2009). Prefilled income tax returns: Reducing compliance costs for personal income taxpayers in Slovenia. *Financial Theory and Practice*, 33(2), 219–233.

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

Kotakorpi, K. and Laamanen, J. P. (2016). Prefilled income tax returns and tax compliance: Evidence from a natural experiment. Tampere Economic Working Paper 104.

Mazar, N., Amir, O. and Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, 45, 633–644.

Mazar, N. and Hawkins, S. A. (2015). Choice architecture in conflicts of interest: Defaults as physical and psychological barriers to (dis)honesty. *Journal of Experimental Social Psychology*, 59, 113–117.

Myles, G. D. and Naylor, R. A. (1996). A model of tax evasion with group conformity and social customs. *European Journal of Political Economy*, 12(1), 49–66.

OECD (2015). The use of online services in tax administration. In: Tax Administration 2015: Comparative Information on OECD and Other Advanced and Emerging Economies. OECD Publishing.

OECD (2017). Third-party data management – the journey from post-assessment crosschecking to pre-filling and no-return approaches. In: Tax Administration 2017: Comparative Information on OECD and Other Advanced and Emerging Economies. OECD Publishing.

Remus, W. (1996). Will behavioral research on managerial decision making generalize to managers? *Managerial and Decision Economics*, 17, 93–101.

Slemrod, J. (2003). Trust in public finance. *Public Finance and Public Policy in the New Century*, ed. S. Cnossen and H.-W. Sinn, 49–88. Cambridge, USA: The MIT Press.

Slemrod, J. (2016). Tax compliance and enforcement: New research and its policy implications. Ross School of Business Working Paper No. 1302.

Strack, F. and Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of personality and social psychology*, 73(3), 437–446.

Telegraph (2017). Taxpayers face chaos as HMRC replaces traditional tax returns with pre-filled forms. 26th September 2017, available online: http://www.telegraph.co.uk/news/2017/09/26/taxpayers-face-chaos-hmrc-replaces-traditional-tax-returns-pre/.

Torgler, B. (2002). Speaking to theorists and searching for facts: Tax morale and tax compliance in experiments. *Journal of Economic Surveys*, 16, 657–683.

Traxler, C. (2010). Social norms and conditional cooperative taxpayers. *European Journal of Political Economy*, 26(1), 89–103.

Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science*, 185, 1124–1131.

Wahl, I., Kastlunger, B. and Kirchler, E. (2010). Trust in authorities and power to enforce tax compliance. An empirical analysis of the "Slippery Slope Framework". *Law & Policy*, 32, 383–406.

Yitzhaki, S. (1974). Income tax evasion: A theoretical analysis. *Journal of Public Economics*, 3(2), 201–202.

Table 1: Development of hypotheses

	Hypothesis 1 (blank vs. correctly prefilled)		**Hypothesis 2** (correctly vs. lower prefilled)		**Hypothesis 3** (correctly vs. higher prefilled)	
	effect	**result**	**effect**	**result**	**effect**	**result**
moral costs	increase in moral costs	compliance increases	decrease in moral costs	compliance decreases	decrease in moral costs	compliance decreases
default effect	accept correctly prefilled value	compliance increases	accept lower prefilled value	compliance decreases	accept higher prefilled value	compliance increases
anchoring effect	decision is assimilated toward correctly prefilled value	compliance increases	decision is assimilated toward lower prefilled value	compliance decreases	decision is assimilated toward higher prefilled value	compliance increases
Total effect		compliance increases		compliance decreases		total effect unclear

Note: This table shows the influence of each discussed effect on tax compliance and the resulting total effect. The table should be read as in the following example: The first entry "increase in moral costs" of the column "Hypothesis 1 (blank vs. correctly prefilled)" denotes that moral costs are increased with correctly prefilled tax returns compared to blank tax returns. The second entry "compliance increases" denotes that tax compliance is increased with correctly prefilled tax returns compared to blank tax returns.

Table 2: Main characteristics of participants

Variable	**description**	**mean**
female	female = 1; male = 0	57.28%
risk attitude	Holt&Laury (2002) risk measure	3.96 / 10
age	in years (19 to 69)	25
economics	study with more than one lecture in economics = 1 (0 otherwise)	43.66%
bachelor	education with a bachelor's degree = 1 (0 otherwise)	54.46%
tax experience	experience with tax returns = 1 (0 otherwise)	40.85%
tax knowledge	tax knowledge = 1 (0 otherwise)	12.68%
income	in Euro (monthly income after fixed costs)	370.35
tax morale	0 to 9; low tax morale = 0; high tax morale = 9	7.15
fairness	0 to 10; low perceived fairness of tax and control system in experiment = 0; high perceived fairness of tax and control system in experiment = 10	6.04
decision complexity	0 to 10; low perceived decision complexity in experiment = 0; high perceived decision complexity in experiment = 10	1.93
joy	0 to 10; felt no joy during experiment = 0; felt high joy during experiment = 10	4.40
anger	0 to 10; felt no anger during experiment = 0; felt high anger during experiment = 10	4.56
fear	0 to 10; felt no fear during experiment = 0; felt high fear during experiment = 10	1.92
guilt	0 to 10; felt no guilt during experiment = 0; felt high guilt during experiment = 10	1.48

Note: This table provides an overview of the individual characteristics of the 213 participants in our experiment.

Table 3: Random effects linear regressions, *between*-subject comparison (dependent variable: declared share of true income)

	model 1	model 2
	manipulated	unmanipulated
blank (B)	-0.16**	-0.16**
	(0.08)	(0.08)
lower prefilled (LP)	-0.15**	-0.07
	(0.07)	(0.07)
higher prefilled (HP)	0.03	0.01
	(0.07)	(0.08)
individual controls	yes	yes
constant	0.72***	0.73***
	(0.21)	(0.21)
no. of observations	858	853
no. of subjects	211	210
R-squared		
within	0.0000	0.0000
between	0.3036	0.2640
overall	0.2635	0.2320
Wald test:		
B = LP	p = 0.9173	p = 0.2215
B = HP	p = 0.0089	p = 0.0196
LP = HP	p = 0.0054	p = 0.2129

Note: In this table, the results of random effects linear regressions are presented with the declared share of true income as dependent variable (regression coefficients, standard errors in parentheses). To analyze the influence of the prefilling manipulation in our experiment, we regress on dummy variables for the treatments B, LP and HP while treatment CP serves as the default. In both models, all observations from treatments B and CP are considered. However, in model 1, only the observations of the manipulated values of treatments LP and HP are included in the regression, while in model 2, only the observations of the unmanipulated values of treatments LP and HP are included. *** $p \leq 0.01$, ** $p \leq 0.05$, * $p \leq 0.1$.

Table 4: Random effects linear regressions, *within*-subject comparison (dependent variable: declared share of true income)

	model 3	model 4
	only treatment LP	only treatment HP
manipulated	-0.08***	0.02
	(0.02)	(0.02)
individual controls	yes	yes
constant	0.97**	0.67**
	(0.46)	(0.30)
no. of observations	380	373
no. of subjects	64	64
R-squared		
within	0.0564	0.0023
between	0.3845	0.4129
overall	0.3301	0.3318

Note: In this table, the results of random effects linear regressions are presented with the declared share of true income as dependent variable (regression coefficients, standard errors in parentheses). To analyze the differences within treatments, we use the dummy variable "manipulated", which takes the value 1 if the declared income refers to a manipulated income field (0 otherwise). *** $p \leq 0.01$, ** $p \leq 0.05$, * $p \leq 0.1$.

Table 5: Random effects linear regressions, *between*-subject comparison (dependent variable: declared share of true income)

	model 5	model 6
	manipulated	unmanipulated
correctly prefilled (CP)	0.16**	0.16**
	(0.08)	(0.08)
lower prefilled (LP)	0.01	0.09
	(0.07)	(0.07)
higher prefilled (HP)	0.19***	0.17**
	(0.07)	(0.07)
individual controls	yes	yes
constant	0.57***	0.57***
	(0.20)	(0.20)
no. of observations	858	853
no. of subjects	211	210
R-squared		
within	0.0000	0.0000
between	0.3036	0.2640
overall	0.2635	0.2320

Note: In this table, the results of random effects linear regressions are presented with the declared share of true income as dependent variable (regression coefficients, standard errors in parentheses). To analyze the influence of the prefilling manipulation in our experiment, we regress on dummy variables for the treatments CP, LP and HP, while treatment B serves as the default. In both models, all observations from treatments B and CP are considered. However, in model 1, only the observations of the manipulated values of treatments LP and HP are included in the regression, while in model 2, only the observations of the unmanipulated values of treatments LP and HP are included. *** $p \leq 0.01$, ** $p \leq 0.05$, * $p \leq 0.1$.

Table 6: Consistent tax compliance behavior

	all treatments	treatment B	treatment CP	treatment LP	treatment HP
always full compliance	48.8%	45.2%	58.1%	42.2%	51.6%
always full evasion	12.7%	16.7%	11.6%	15.6%	7.8%
always full compliance or full evasion	61.5%	61.9%	69.8%	57.8%	59.4%

Note: This table presents an overview of the share of subjects that reveal consistent tax compliance behavior in our experiment. Consistent means that the subjects chose either full compliance or full evasion for all six income fields in their tax return.

Table 7: Difference between declared and prefilled income (adjustments)

		correctly prefilled			incorrectly prefilled		all observations		
adjustment		CP	LP	HP	LP	HP	CP	LP	HP
downward	%	23.5%	33.7%	25.5%	26.8%	92.1%	23.5%	30.3%	59.3%
	mean	-0.80	-0.69	-0.59	-0.43	-0.66	-0.80	-0.57	-0.64
no adjustment	%	**70.4%**	**63.7%**	**69.6%**	**6.3%**	**6.9%**	70.4%	35.0%	37.8%
	mean	0	0	0	0	0	0	0	0
upward	%	6.1%	2.6%	4.9%	66.8%	1.1%	6.1%	34.7%	2.9%
	mean	0.61	0.68	0.66	0.52	0.24	0.61	0.52	0.58
total	%	100%	100%	100%	100%	100%	100%	100%	100%
	mean	-0.15	-0.21	-0.12	0.23	-0.60	-0.15	0.01	-0.36

Note: This table shows the distribution and mean values of the adjustments of prefilled values for all observations as well as for correctly and incorrectly prefilled income fields separately. Adjustments are defined as the difference between declared and prefilled income (normalized).

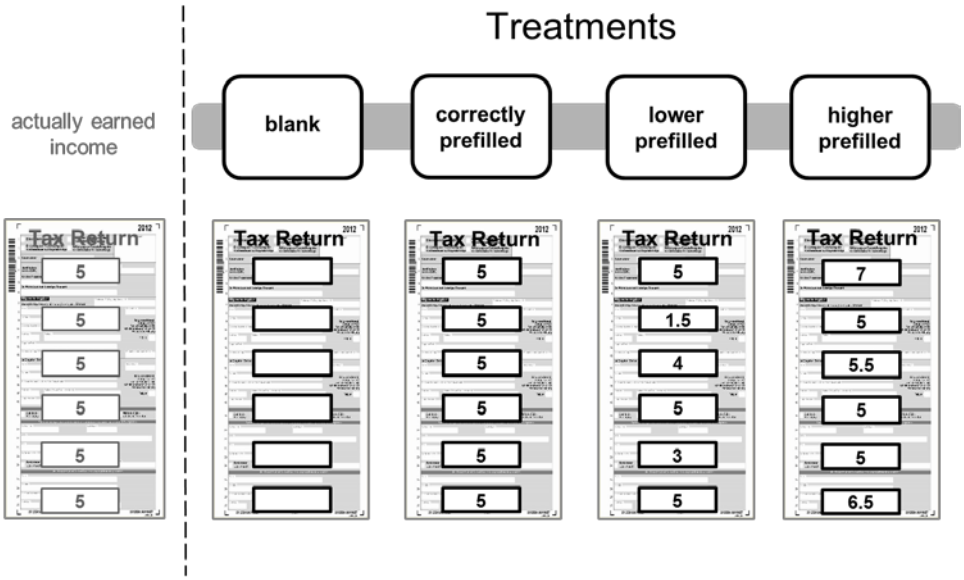


Figure 1: Experimental design

Note: This figure highlights the differences among the treatments. Subjects are randomly distributed to one of the four treatments.

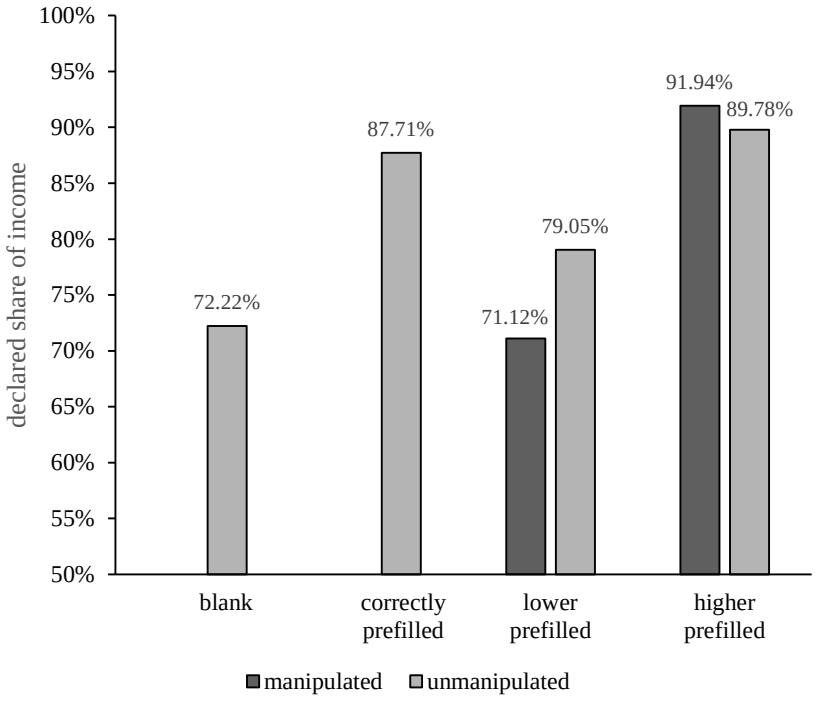


Figure 2: Declared share of income in our treatments

Note: This figure shows the relative declared income for each treatment on average. For treatments LP and HP, the average declared income is shown separately for correctly and incorrectly prefilled income fields. As we have several observations per subject, we calculated an average tax compliance level for each subject.

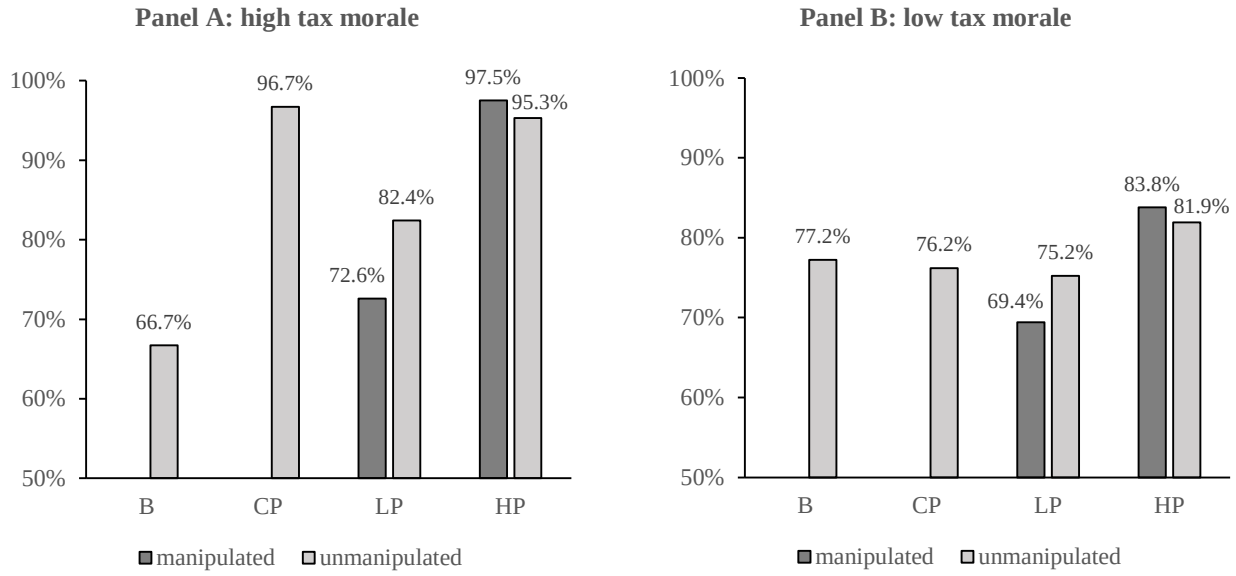


Figure 3: Declared share of income for low and high tax morale subsamples

Note: This figure shows the relative declared income for each treatment, on average, separately for subjects with high tax morale (Panel A) and subjects with low tax morale (Panel B). For treatments LP and HP, the average declared income is shown separately for correctly and incorrectly prefilled income fields. As we have several observations per subject, we calculated an average tax compliance level for each subject.

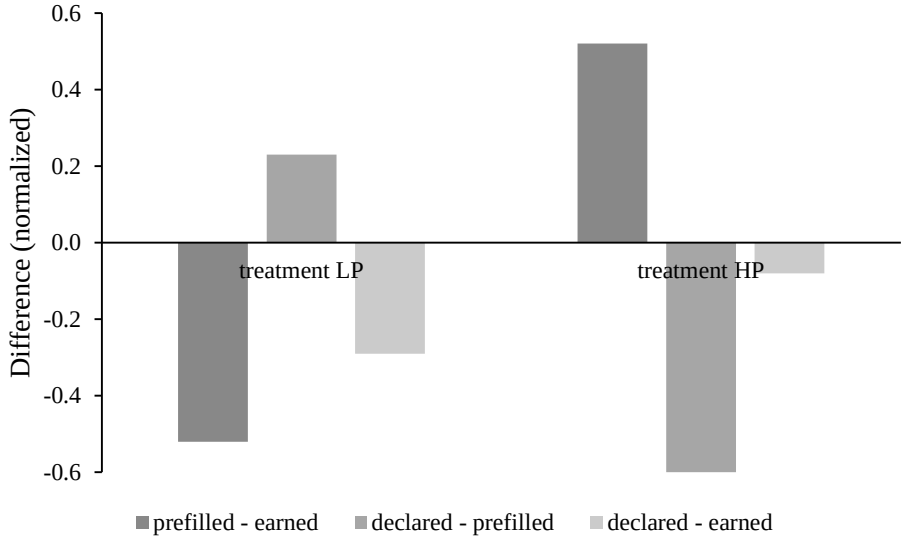


Figure 4: Adjustments in case of incorrectly prefilled tax returns

Note: This figure shows the mean adjustments for the incorrectly prefilled income fields in treatments LP and HP (declared - prefilled). Adjustments are defined as the difference between declared and prefilled income (normalized). The figure also presents the difference between prefilled and earned income and between declared and earned income (each normalized to actually earned income).

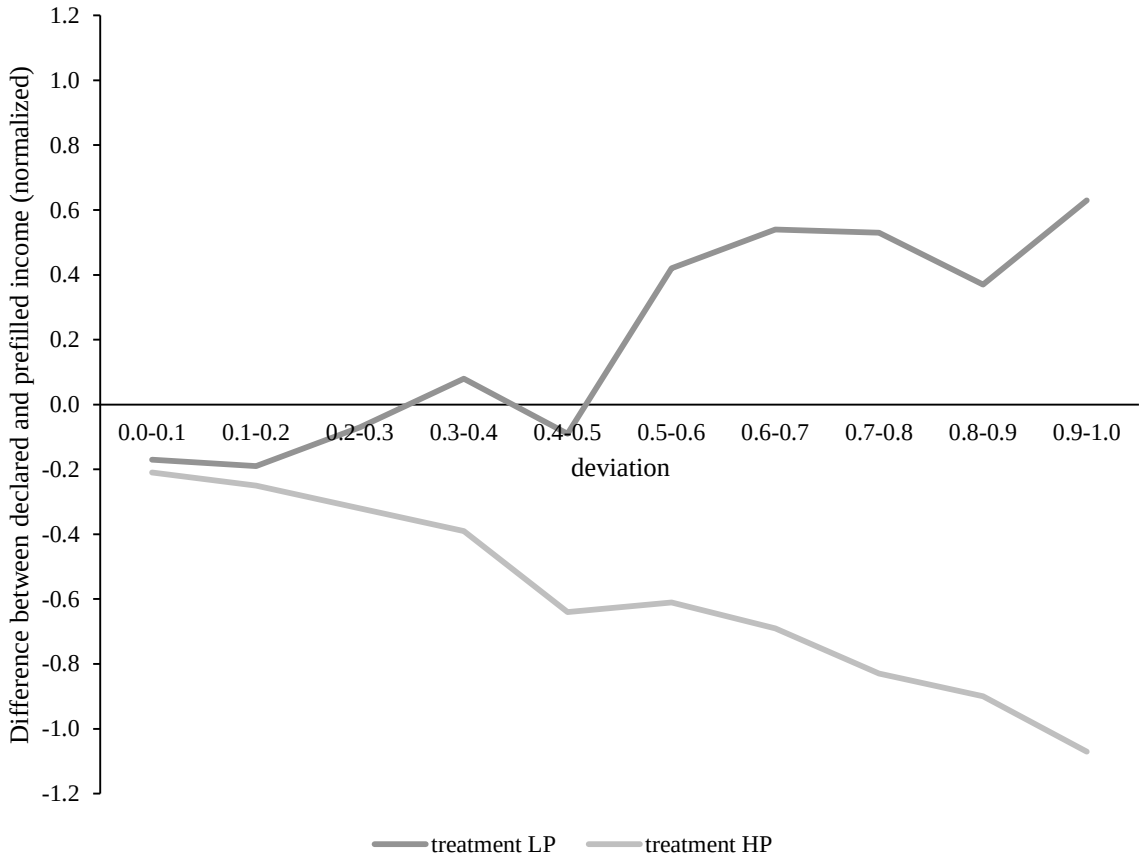


Figure 5: Deviation and adjustments

Note: This figure shows subjects' adjustments dependent on the deviation of prefilled income from true income for treatments LP and HP. Adjustments are defined as the difference between declared and prefilled income (normalized).

Appendix (Online Appendix, not intended for publication)

A1 Instructions

Appendix A1 includes the translated instructions (from German). All participants received the general instructions in print. Before the actual experiment was executed, subjects participated in the Holt and Laury (2002) task. The instructions for this task (first experiment) were displayed on the computer screen. After that, participants received the specific instructions for the actual (second) experiment in print.

A1.1 General Instructions

Thank you for participating in this experimental study. For your participation you receive a participation fee of 4 Euros.

The study consists of two experiments in which you have the possibility to earn money and a questionnaire at the end of the study. How much money you earn depends on your decisions and on chance. The instructions explain how you can influence how much money you earn in this study by your decisions.

It is important that you understand the instructions. Please do not hesitate to ask questions. If you have a question, please raise your hand. We will come to you to answer your question. Please do not ask your question loudly. You can write on the instructions or set markers. Please do not take the instructions home, but give them back to us at the end of the study.

The analysis of the experiment will be anonymous. We will never link your name with the data generated in the experiment. You will not learn the identity of any other participant, neither before nor after the experiment. Also the other participants will not learn your identity. At the end of the experiment, you have to sign a receipt to confirm the payments you received. This receipt will only be used for accounting purposes.

We would like to inform you that you are not allowed to communicate with other participants or leave your seat throughout the whole experiment. Please switch off your mobile phone and put it in your bag.

The calculator, the pen and the sheet of paper (for notes) that are lying on your desk, can be used.

At the end of the study you will receive your payout privately and in cash. Your total payout consists of your payout of the first experiment, plus your payout of the second experiment and the participation fee.

The instructions for the first experiment will be displayed on your computer screen.

A1.2 Instructions for the Holt and Laury (2002) Task

Please choose one of the two lotteries A or B in each of the following 10 decision situations.

You will make a decision for all 10 situations, but your payout from the first experiment is determined only by the one situation that is randomly drawn by the computer after the second experiment.

In each situation, you can either earn 2.00 € or 1.60 € from lottery A and either 3.85 € or 0.10 € from lottery B. The probabilities of winning, however, vary from situation to situation. The further down you move in the table, the higher is the probability of the higher payment and the lower is the probability of the lower payment.

After the first experiment and the second experiment are completed, the computer randomly draws (with the same probability) one of the 10 decision situations. After that, the computer determines your payout from the lottery that you have chosen in this decision situation by a second random draw. For that, the computer uses the probabilities for the higher payment and the lower payment according to the chosen decision situation.

decision	Lottery A	Your decision		Lottery B
		A	B	
1.	2.00 € with 10% or 1.60 € with 90%	o	o	3.85 € with 10% or 0.10 € with 90%
2.	2.00 € with 20% or 1.60 € with 80%	o	o	3.85 € with 20% or 0.10 € with 80%
3.	2.00 € with 30% or 1.60 € with 70%	o	o	3.85 € with 30% or 0.10 € with 70%
4.	2.00 € with 40% or 1.60 € with 60%	o	o	3.85 € with 40% or 0.10 € with 60%
5.	2.00 € with 50% or 1.60 € with 50%	o	o	3.85 € with 50% or 0.10 € with 50%
6.	2.00 € with 60% or 1.60 € with 40%	o	o	3.85 € with 60% or 0.10 € with 40%
7.	2.00 € with 70% or 1.60 € with 30%	o	o	3.85 € with 70% or 0.10 € with 30%
8.	2.00 € with 80% or 1.60 € with 20%	o	o	3.85 € with 80% or 0.10 € with 20%
9.	2.00 € with 90% or 1.60 € with 10%	o	o	3.85 € with 90% or 0.10 € with 10%
10.	2.00 € with 100% or 1.60 € with 0%	o	o	3.85 € with 100% or 0.10 € with 0%

A1.3 Instructions for Main Experiment

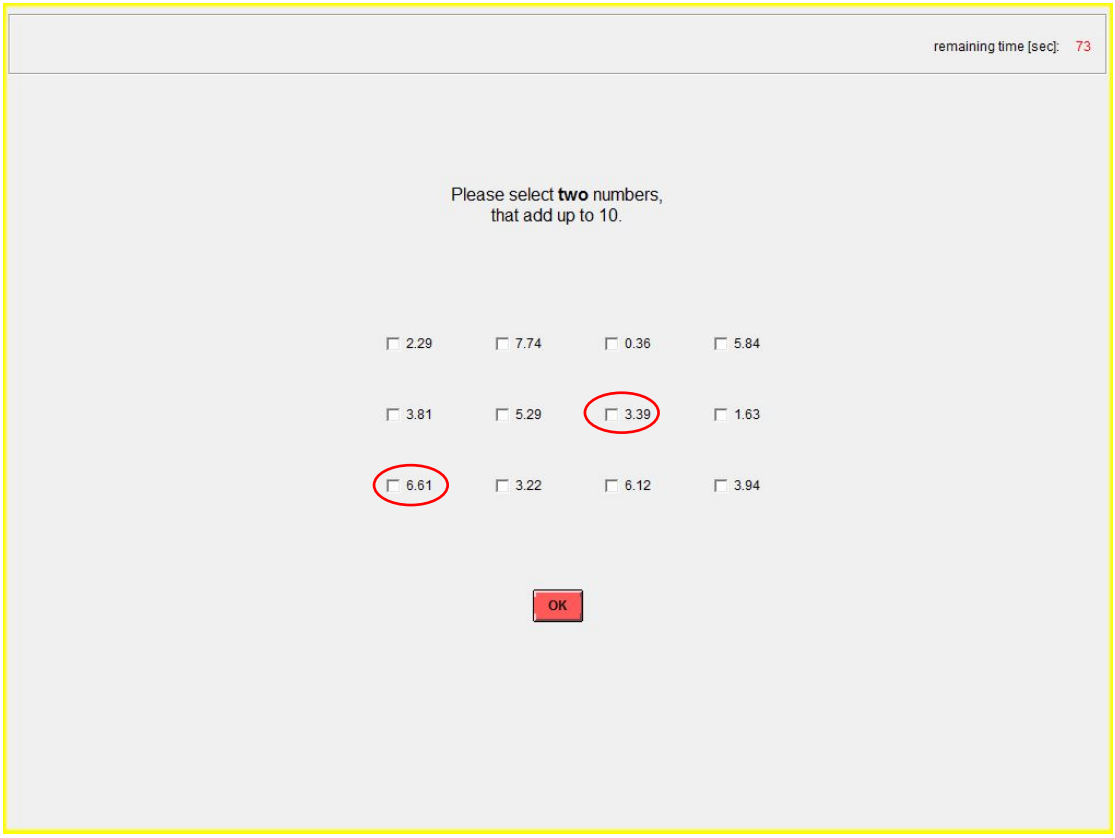
The second experiment consists of two parts. In part 1 of the experiment you have the possibility to earn money. In part 2 of the experiment you complete a tax return over your income.

Part 1

Puzzle Task

The first part of the experiment consists of a math puzzle task. The screen of your computer will display a puzzle with a matrix of 12 numbers. The numbers have two decimal places. These two numbers, that add up to 10, should be found in the matrix. In each matrix with 12 numbers, there are only 2 numbers that exactly add up to 10.

The screen with the matrix looks as follows:



Part 1 of the experiment consists of 6 rounds and in each round you have 3 minutes time to solve as many math puzzles as you like. A maximum of 20 puzzles can be solved in each round. After each round you have a one-minute break before the next round starts.

Your earned income depends on the number of puzzles you solved correctly. For each correctly solved puzzle, you receive an income of 0.42 €. For each incorrectly solved puzzle, you receive an income of 0 €. If you do not solve any puzzle correctly in one round, you earn 0 € in that round. If you solve every puzzle correctly in one round, you earn 8.40 € in that round. After each round, your number of correctly solved puzzles and the resulting income earned in that round will be displayed to you. Your income in each round therefore is as follows:

Earned income in each round =
number of correctly solved puzzles in the corresponding round x 0.42 €

Your total earned income results when adding up the earned income of all six rounds.

After you have earned your pre-tax income (= total earned income) in part 1 of the experiment, you have to complete a tax return on your income in the second part of the experiment. In order to be able to complete the tax return, please note down your earned income after each round on the provided sheet of paper at your place. The sheet of paper serves to help you completing the tax return. Please do not give the sheet of paper back to us, but take it home or dispose of the paper.

Part 2

Tax return

In this study, you have to complete a fictional tax return on your earned income, in order to determine a fictional tax. This means, in part 2 you shall declare the income you earned in part 1. There will be a tax of 25% on the declared income that was stated by you in the tax return. The tax revenue will be used inter alia to finance future research projects at the University of Cologne.

[Treatment B (blank)

After you have finished part 1 of the experiment, your tax return will be displayed on your screen. The tax return consists of 6 rows, one for the income of each round. At the end of each row you find a field in which you enter the income you want to declare. Your declared income can be lower, equal to, or higher than your actually earned income.

After you have filled out the 6 income fields, you can press the button “Submit tax return”, in order to submit your tax return. After submission, you cannot modify your tax return anymore.

Before submitting your tax return, you can also press the button “Compute tax liability” in order to get the total declared income and the resulting tax liability on the basis of your provided information in the tax return displayed. If you want to modify your data, press the button “Modify tax return”. You can then change your data in the six income fields. You can press “Compute tax liability” and “Modify tax return” until you are ready with your tax return. As soon as you want to submit your tax return, press the button “Submit tax return”.]

[Treatment CP, LP and HP (prefilled)

After you have finished part 1 of the experiment, your tax return will be displayed on your screen. The tax return consists of 6 rows, one for the income of each round. The income of each round is already prefilled by the computer. Additionally, you get the total declared income and the resulting tax liability displayed. The prefilled values can deviate from your actually earned income. Please check the prefilled values in the tax return. Your declared income can be lower, equal to, or higher than your actually earned income.

To submit your tax return, please press the button “Submit tax return”. After submission, you cannot modify your tax return anymore.

If you want to modify the data in your tax return, press the button “Modify tax return”. You can then change the values in the six income fields. If you want to get the new resulting total declared income and the resulting tax liability displayed, press the button “Compute tax liability”. You can press “Compute tax liability” and “Modify tax return” until you are ready with your tax return. As soon as you want to submit your tax return, press the button “Submit tax return”.]

The tax payable amounts to 25% of your total declared income:

Tax = 0.25 x total declared income

Audit of the tax return

With a probability of 30% your tax return is audited. If you are audited and the declared income of a round does not coincide with your actually earned income of the corresponding round, you have to repay the undeclared tax. Additionally, a fine is charged at the same amount.

Tax repayment = sum of undeclared taxes

Fine = sum of undeclared taxes

The undeclared tax for the income of each round is:

Undeclared tax = 0.25 x (earned income – declared income)

If you are audited and you declared a higher income than you actually earned in that round, you get a tax refund of the overpaid tax. There will be no fine in that case.

Your personal payout of the 2. Experiment

During the 2. Experiment you will receive a personal "bank account", on which your payout relevant amounts will be posted during the experiment. Your account balance will be displayed to you after each part of the experiment. After part 1, your total earned income of the six rounds will be posted to your account and displayed to you. After you have submitted your tax return, the resulting tax burden of the tax return is posted to your account. If there is an audit and the declared income for each round does not coincide with the actually earned income in the corresponding round, the tax repayment and the resulting fine is debited to the account. If you receive a tax refund, it will be posted to your account.

Your account balance at the end of the 2. experiment is your personal payout of the 2. experiment. Your payout is calculated as follows:

Payout from the 2. experiment = *total earned income in part 1*
- *tax liability from part 2*
- *possible tax repayment*
- *possible fine*
+ *possible tax refund*

Before the 2. experiment begins, you are asked to answer some questions at your computer. The answering of the questions serves to verify your understanding and is not payoff relevant.

A1.4 Comprehension Test

In our comprehension test, subjects have to correctly answer the following questions:

1. How many experiments does the study include?
   Possible answers:
   - *1*
   - *2*
   - *3*
   - *4*
   - *More than 4*

2. How is the earned income for each round determined?
   Possible answers:
   - *The earned income for each round is determined by multiplying the number of correctly solved puzzles in that round with 0.42 €.*
   - *The earned income for each round is determined by multiplying the total number of correctly solved puzzles in all rounds with 0.42 €.*
   - *The earned income for each round is determined by multiplying the number of correctly solved puzzles in that round with 0.24 €.*

3. When is a puzzle solved correctly?
   Possible answers:
   - *When the number 10 is chosen.*
   - *When so many numbers are chosen, that they sum up to ten.*
   - *When those two numbers are chosen, that add up to ten.*

4. Which of the following statements regarding the computation of the tax liability is correct?
   Possible answers:
   - *The tax amounts to 25% of the total declared income.*
   - *The tax amounts to 25% of the total earned income.*

5. With what probability (in percent) will there be an audit of the tax return?
   Possible answer: between 0 and 100%

6. How is your personal payout for the second experiment determined when there is no audit?
   Possible answers:
   - *Payout = total earned income in part 1*
   - *Payout = total declared income in part 2 – tax liability in part 2*
   - *Payout = fix payout of 4 €*
   - *Payout = total earned income in part 1 – tax liability in part 2*

A2 Regressions

Table A2.1: Random effects linear regressions, between-subject comparison (dependent variable: declared share of income)

	model 1	model 2
	manipulated	unmanipulated
blank (B)	-0.16**	-0.16**
	(0.08)	(0.08)
lower prefilled (LP)	-0.15**	-0.07
	(0.07)	(0.07)
higher prefilled (HP)	0.03	0.01
	(0.07)	(0.08)
risk attitude	-0.05***	-0.06***
	(0.02)	(0.02)
female	0.21***	0.17***
	(0.05)	(0.05)
age	0.01	0.01
	(0.00)	(0.00)
economics	-0.09*	-0.08
	(0.05)	(0.05)
bachelor	0.09	0.11*
	(0.05)	(0.06)
tax experience	-0.08	-0.08
	(0.05)	(0.06)
tax knowledge	-0.10	-0.07
	(0.08)	(0.08)
tax morality	0.02*	0.03*
	(0.01)	(0.01)
fairness	-0.01	-0.00
	(0.01)	(0.01)
decision complexity	0.05***	0.05***
	(0.01)	(0.01)
joy	-0.01	-0.00
	(0.01)	(0.01)
anger	-0.02*	-0.02*
	(0.01)	(0.01)
fear	-0.01	-0.01
	(0.01)	(0.01)
guilt	-0.00	0.00
	(0.01)	(0.01)
income	0.00	0.00
	(0.00)	(0.00)
constant	0.72***	0.73***
	(0.21)	(0.21)
no. of observations	858	853
no. of subjects	211	210
R-squared		
within	0.0000	0.0000
between	0.3036	0.2640
overall	0.2635	0.2320
Wald test:		
B = LP	p = 0.9173	p = 0.2215
B = HP	p = 0.0089	p = 0.0196
LP = HP	p = 0.0054	p = 0.2129

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A2.2: Random effects linear regressions, within-subject comparison (dependent variable: declared share of income)

	model 3	model4
	only treatment LP	only treatment HP
manipulated	-0.08***	0.02
	(0.02)	(0.02)
risk attitude	-0.08**	-0.07**
	(0.04)	(0.03)
female	0.28***	0.17**
	(0.10)	(0.08)
age	-0.01	0.00
	(0.01)	(0.01)
economics	-0.10	0.07
	(0.11)	(0.09)
bachelor	0.13	0.08
	(0.11)	(0.09)
tax experience	-0.04	-0.01
	(0.10)	(0.09)
tax knowledge	0.03	-0.10
	(0.16)	(0.12)
tax morality	0.03	0.04*
	(0.02)	(0.02)
fairness	-0.03	0.00
	(0.02)	(0.02)
decision complexity	0.09***	0.06***
	(0.03)	(0.02)
joy	0.01	-0.01
	(0.02)	(0.02)
anger	-0.01	-0.02
	(0.02)	(0.01)
fear	-0.02	-0.01
	(0.03)	(0.02)
guilt	-0.02	0.02
	(0.03)	(0.02)
income	0.00	0.00
	(0.00)	(0.00)
constant	0.97**	0.67**
	(0.46)	(0.30)
no. of observations	380	373
no. of subjects	64	64
R-squared		
within	0.0564	0.0023
between	0.3845	0.4129
overall	0.3301	0.3318

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A2.3: Random effects linear regressions without subjects revealing a consistent tax compliance behavior, *between*-subject comparison (dependent variable: declared share of income)

	model 1	model 2
	manipulated	unmanipulated
blank (B)	-0.30**	-0.25*
	(0.14)	(0.15)
lower prefilled (LP)	-0.29**	-0.08
	(0.12)	(0.13)
higher prefilled (HP)	-0.13	-0.16
	(0.13)	(0.14)
individual controls	yes	yes
constant	0.73**	0.63*
	(0.33)	(0.37)
no. of observations	305	300
no. of subjects	80	79
R-squared		
within	0.0000	0.0000
between	0.5379	0.3840
overall	0.4163	0.2847
Wald test:		
B = LP	p = 0.9165	p = 0.1461
B = HP	p = 0.1103	p = 0.4177
LP = HP	p = 0.1023	p = 0.4739

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A2.4: Random effects linear regressions without subjects revealing a consistent tax compliance behavior, *within*-subject comparison (dependent variable: declared share of income)

	model 3	model4
	only treatment LP	only treatment HP
manipulated	-0.19***	0.04
	(0.04)	(0.05)
individual controls	yes	yes
constant	-0.22***	2.38***
	(1.34)	(0.81)
no. of observations	158	145
no. of subjects	27	26
R-squared		
within	0.1361	0.0061
between	0.5282	0.8265
overall	0.3290	0.4892

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A2.5: Random effects linear regressions, separated for low and high deviation, between-subject comparison (dependent variable: declared share of income)

	model 1	model 2
	manipulated	manipulated
	deviation <= 0.5	deviation > 0.5
blank (B)	-0.15**	-0.15*
	(0.07)	(0.09)
lower prefilled (LP)	-0.20***	-0.16*
	(0.07)	(0.08)
higher prefilled (HP)	0.04	0.07
	(0.07)	(0.08)
individual controls	yes	yes
constant	0.77***	0.72***
	(0.20)	(0.23)
no. of observations	653	684
no. of subjects	192	197
R-squared		
within	0.0000	0.0000
between	0.3618	0.2642
overall	0.2870	0.2228
Wald test:		
B = LP	p = 0.5329	p = 0.9399
B = HP	p = 0.0059	p = 0.0057
LP = HP	p = 0.0005	p = 0.0025

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A2.6: Random effects linear regressions, separated for low and high deviation, within-subject comparison (dependent variable: declared share of income)

	model 3		model 4	
	only treatment LP	only treatment LP	only treatment HP	only treatment HP
	deviation <= 0.5	deviation > 0.5	deviation <= 0.5	deviation > 0.5
manipulated	-0.09***	-0.08***	0.003	0.03
	(0.02)	(0.02)	(0.02)	(0.02)
individual controls	yes	yes	yes	yes
constant	1.01**	0.96**	0.66**	0.69**
	(0.44)	(0.47)	(0.32)	(0.32)
no. of observations	280	290	268	289
no. of subjects	64	64	64	64
R-squared				
within	0.0596	0.0495	0.0000	0.0063
between	0.4121	0.3646	0.3913	0.4078
overall	0.3631	0.3150	0.3433	0.3108

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Impressum:
Arbeitskreis Quantitative Steuerlehre, arqus, e.V.
Vorstand: Prof. Dr. Ralf Maiterth (Vorsitzender),
Prof. Dr. Kay Blaufus, Prof. Dr. Dr. Andreas Löffler
Sitz des Vereins: Berlin

Herausgeber: Kay Blaufus, Jochen Hundsdoerfer, Martin Jacob, Dirk Kiesewetter, Rolf J. König, Lutz Kruschwitz, Andreas Löffler, Ralf Maiterth, Heiko Müller, Jens Müller, Rainer Niemann, Deborah Schanz, Sebastian Schanz, Caren Sureth-Sloane, Corinna Treisch

Kontaktadresse:
Prof. Dr. Caren Sureth-Sloane, Universität Paderborn,
Fakultät für Wirtschaftswissenschaften,
Warburger Str. 100, 33098 Paderborn,
www.arqus.info, Email: info@arqus.info

ISSN 1861-8944