Enhancing Tax Compliance of Individuals: Effective Audit Systems and Anti-Tax-Evasion Mechanisms

Inauguraldissertation
zur Erlangung des Doktorgrades
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Universität zu Köln

2020

vorgelegt von
Tobias Julian Kölle, M.Sc.
aus München

Referent: Prof. Dr. Michael Overesch, Universität zu Köln
Korreferent: Prof. Dr. Martin Fochmann, Freie Universität Berlin
Tag der Promotion: 18.12.2020

Vorwort

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Seminar für ABWL und Unternehmensbesteuerung der Universität zu Köln. Sie wurde im September 2020 von der Wirtschafts- und Sozialwissenschaftlichen Fakultät der Universität zu Köln als Dissertation angenommen.

Mein besonderer Dank gilt meinem Doktorvater Herrn Prof. Dr. Michael Overesch, der die vorliegende Arbeit angeregt und mich hilfsbereit unterstützt hat. Darüber hinaus danke ich Herrn Prof. Dr. Martin Fochmann für den intensiven fachlichen Austausch in zahlreichen Diskussionen und seine wertvollen Anregungen. Ich danke außerdem Frau Prof. Dr. Bettina Rockenbach für den stets inspirierenden Gedankenaustausch. Für die Zusammenarbeit an meinen Forschungsprojekten danke ich außerdem Prof. Dr. Peter Mohr. Zudem danke ich Herrn Prof. Dr. Martin Fochmann für die Erstellung des Zweitgutachtens und Herrn Prof. Dr. Carsten Homburg für die Übernahme des Vorsitzes der Prüfungskommission.

Meine Forschung wäre ohne das Kölner Laboratorium für experimentelle Wirtschaftsforschung nicht möglich gewesen. Aus diesem Grund danke ich insbesondere dem damaligen Labormanager Herrn Dr. Thomas Lauer und meiner Kollegin Frau Dr. Nadja Müller, die mich bei den ersten Experimenten unterstützt hat. Großer Dank gebührt auch meinen weiteren Kolleginnen und Kollegen an den Lehrstühlen der Area Accounting and Taxation der Universität zu Köln. Neben der guten Zusammenarbeit und den vielfältigen Teamaktivitäten werden mir vor allem die vielen Verrücktheiten mit meinem Büronachbarn Mathias, das gelegentliche Feierabendbier mit Max, die Karnevalsfeiern mit Sabine und Tanja, die vielfältigen Unterstützungen durch Hubertus und Sina, und die regelmäßigen Mensabesuche mit Julia in bester Erinnerung bleiben.

Mein besonderer Dank gilt meiner Lebensgefährtin Diana, meiner Schwester Miriam und meinen Eltern Ursel und Jens, die mich auf meinem Lebensweg stets liebevoll unterstützen.

Köln, im September 2020 *Tobias Kölle*

Content

CHAPTER 1

Introduction

Content

1.1 Motivation and object of research

Tax evasion is a pervasive problem for any modern state in the world. As any modern state in the world relies on taxes, governments have to deal with the question of how to secure their tax revenues. Major tax scandals like the so called "Panama Papers" illustrate the difficulties that arise for states collecting their tax revenues (The Economist, 2016a, 2016b). In fact, current estimates on the tax gap resulting from tax evasion show urgent need for action. For example, Murphy (2019) estimates the tax gap resulting from tax evasion for the European Union. His estimations suggest a tax gap of 825 billion Euro a year which equals roughly 1,650 Euro per capita. In order to combat tax evasion, governments have to find an answer to the fundamental question of tax compliance research: Why do people pay or evade taxes?

At a first glance, governments expect people to pay their taxes because paying taxes is legally obligated and non-compliance gets penalized in case of detection. Based on this thought, Allingham and Sandmo (1972) presented the first formalized model of tax compliance. The model is based on the economics-of-crime approach by Becker (1968) and relies on three variables: audit probability, tax rate and fine. Given these three variables of deterrence, a rational taxpayer decides on her compliance by considering only financial aspects. Consequently, an individual only pays taxes if monetary benefits of tax evasion are bigger than monetary costs of tax evasion. With its simplistic structure, the model became the cornerstone of tax compliance research and explains its importance in tax compliance literature to this day. Furthermore, the simplistic structure allows a straight forward analysis of tax compliance and gives direct policy recommendations for governments. As compliance depends solely on enforcement, governments enhance tax compliance by increasing audit probabilities, fines and/or tax rates (Yitzhaki, 1974).

However, reality shows that the individual decision on tax compliance is far more complex and not solely a financial gamble.[1] On the one hand, the purely economic analysis predicts much higher levels of tax evasion than actually observed, given reasonable input parameters for the three deterrence variables (e.g. Slemrod, 2007 and Kleven et al., 2011). The reality and the prediction of the model are only in agreement if we assume

[1] A fact already noted by Allingham and Sandmo (1972) themselves. They state that the model gives "too little attention to nonpecuniary factors in the taxpayer's decision"(p.326).

abnormally high risk aversion (see for example the criticism of Kirchler, 2007; Andreoni et al., 1998, Graetz and Wilde, 1985 and Torgler, 2007).[2]

On the other, there is increasing empirical evidence that an increase in audit probability, tax rate and/or fine does not necessarily increase tax compliance. For example, evidence suggests that tax compliance increases with rising audit probabilities but decreases after a certain threshold (e.g. Mendoza et al., 2017, supported by Slemrod et al., 2001 and already by Weck-Hannemann and Pommerehne, 1989). This raises doubts about the policy implications of the model mentioned above.

In sum, there is evidence for missing variables in the standard model. Further shortcomings are that the standard model relies on controversial assumptions of the neoclassical economic model like perfect rationality, outcome orientation and egoism of individuals (Alm, 2019). As a result, a more promising framework of tax compliance has to implement behavioral economics, i.e. non-financial aspects that influence tax compliance. All these non-financial aspects are sometimes gathered under the umbrella term of tax morale (Luttmer and Singhal, 2014) or more precisely under the umbrella term of non-monetary costs of tax evasion.[3] The non-monetary costs of tax evasion are less specific and capture a wide range of variables like social norm, trust, fairness, participation rights, complexity and knowledge, patriotism, and moral emotions. All in all, there is evidence for the importance of non-monetary costs of tax evasion in the tax compliance puzzle. Undoubtedly, a reasonable framework of tax compliance has to implement these costs.

With the rough outline of tax compliance research in mind, one can distinguish individual research fields within the research of tax compliance. First, there is research focusing on the three variables of the standard tax compliance model, i.e. audit probability, tax rate and fine. Even though the predictions derived from the standard model are rather poor, there is no doubt that the three variables of deterrence are important drivers of tax compliance (see for an overview Alm, 2019). There is much evidence that tax compliance increases by increasing the audit probability (e.g. Andreoni et al., 1998). However, this effect is non-linear and seems to diminish with higher audit rates (e.g. Mendoza et al., 2017), the type of audit and the perception of audits (e.g. Alm et al., 1993 and Alm et al., 1992b). The tax rate affects tax compliance, however the effect is

[2] Given these insights, Alm et al. (1992b) formulated that the puzzle of tax compliance behavior is why people pay taxes and not why they evade them.
[3] Contrary, the deterrence variables of the standard tax evasion model are the monetary costs of tax evasion.

rather small and evidence is still mixed (increase in tax rates reduces tax compliance, e.g. Clotfelter, 1983. Contrary e.g. Alm et al., 1995). Finally, there is evidence that higher fines lead to more compliance, however, the effect on compliance is rather small (e.g. Alm et al., 1992a; Alm et al., 1992b). All in all, the three variables of deterrence have been extensively studied. However, final evidence, especially on tax rates and fines, is still missing.

Second, there is research regarding mechanisms that enhance tax compliance. For example, researchers investigated the effects of third-party reporting and withholding tax (e.g. Kleven et al., 2011; Gillitzer and Skov, 2018; Carrillo et al., 2017; Slemrod et al., 2017; Adhikari et al., 2016). More generally, there are studies that investigate the effect of prefilled tax returns on tax compliance. On the one hand, prefilled tax returns are due to automatic data exchanges between the tax authority and employers, social insurance agencies and banks (third party reporting). On the other hand, electronic tax declaration programs usually carry over the previous year's values to the subsequent year and therefore prefill the current tax return with last year's numbers. However, literature provides mixed evidence regarding the effects of prefilling on tax compliance (e.g. Fochmann et al., 2018; Kotakorpi and Laamanen, 2016; Duncan and Li, 2018; Fonseca and Grimshaw, 2017; Gillitzer and Skov, 2018).

Third, there is research regarding new variables that affect tax compliance besides the standard model of tax evasion (i.e. audit probability, tax rate and fine). For example, researchers investigated the effect of rewards on tax compliance (e.g. Bazart and Pickhardt, 2011; Dwenger et al., 2016; Koessler et al., 2019; Falkinger and Walther, 1991) which gets already utilized by governments (The Economist, 2019). However, more recently research has expanded to (psychological) factors influencing tax compliance, e.g. social norm, trust, fairness, participation rights, complexity and knowledge, patriotism, moral emotions, and public disclosure (e.g. Luttmer and Singhal, 2014; Kirchler et al., 2008; Hofmann et al., 2008; Torgler, 2002; Dwenger et al., 2016; Bø et al., 2015). Research regarding these factors might be especially interesting for governments combating tax evasion as emphasized by the OECD (OECD, 2019). Arguably, a manipulation of these variables might be more cost-efficient with a better marginal utility and might promote a positive mood (higher trust, more fairness etc.) instead of an antagonistic climate (Kirchler et al., 2008).

Fourth, there is research regarding complex interdependence between all the variables affecting tax compliance. One of the latest developments is research dealing with crowding effects. For example, there is already evidence for crowding out of intrinsic

motivation in settings not related with tax compliance by adding penalties (e.g. Gneezy and Rustichini, 2000a) or bonuses (e.g. Gneezy and Rustichini, 2000b; Titmuss, 1970). However, the existence of the effect is controversial in tax compliance research. It is unclear whether the introduction of audits with a penalty for non-compliance or a bonus for compliance increases or decreases compliance (Boyer et al., 2014; Dwenger et al., 2016).

The three essays of this thesis aim to address research gaps described above. The first essay *"Trust them, threaten them, or lure them? Effective audit systems to promote compliance"* is co-authored by Martin Fochmann, Chair for Accounting and Taxation at the Free University of Berlin, Peter N.C. Mohr, Junior-Professor for Neuroeconomics at the Free University of Berlin, and Bettina Rockenbach, Chair for Experimental and Behavioral Economics at the University of Cologne. In this paper, we investigate experimentally the effects of (1) increasing the detection probability of non-compliance, (2) monetary incentives to promote compliance (bonuses and penalties), and the interaction of (1) and (2). Most importantly, we find that compliance decreases with audit systems that penalize non-compliance with a low detection probability compared to a situation without any audits. Likewise, audit systems that reward compliance with a low detection probability decrease compliance. Consequently, the intended deterrence has a diametrically opposed effect than desired. We interpret this as a crowding out of intrinsic motivation to be honest. Only a penalty system with a high detection probability ensures significantly higher compliance than without audits. Among others, I was responsible for all aspects of data collection (I designed, programmed, organized and executed the laboratory experiments), empirical data analysis and writing the scientific paper. However, our joint work was characterized by strong collaboration and a constant exchange of ideas.

The second essay *"Combating overreporting of deductions in tax returns: Prefilling and restricting the deductibility of expenditures"* is co-authored by Martin Fochmann, Chair for Accounting and Taxation at the Free University of Berlin, Frank Hechtner, Chair for Business Taxation at the Friedrich-Alexander University Erlangen-Nürnberg, and Michael Overesch, Chair of Business Taxation at the University of Cologne. We experimentally analyze three anti-tax-evasion mechanisms: (1) prefilling of deductions in tax returns, 2) restricting tax evasion opportunities by either disallowing or 3) limiting the deductibility of expenditures. In our study we focus on deductions, in particular on how individuals report expenditures in a tax return. Deductions are rather underrepresented in tax compliance literature. In fact, researchers focused primarily on income tax evasion

although overreporting of deductions might be the only way to evade taxes for many people (such as typical wage earners) due to third party income reporting. We find that prefilling increases tax compliance compared to blank forms. Disallowing the deductibility of expenditure items is an ineffective mechanism to combat tax evasion as individuals shift their tax evasion activities from the disallowed item to non-restricted items. In contrast, we find that just limiting the deductibility of expenditures avoids this evasion-shift-effect and finally enhances overall tax compliance. Among others, I was responsible for all aspects of data collection (I designed, programmed, organized and executed the laboratory experiments), empirical data analysis and writing the scientific paper. Throughout the project, my co-authors and me had intense discussions and an inspiring work flow.

The third essay *"Systemization and review of non-monetary costs of tax evasion"* is single-authored and thus my sole responsibility. In this paper, I systemize and review tax compliance literature focusing on non-monetary costs of tax evasion. I conclude that a solution approach for the puzzle of tax compliance has to consider not only monetary costs but also non-monetary costs of tax evasion. Backed up with a theoretical foundation of monetary/non-monetary costs of tax evasion, I review the influence of social norm, trust, fairness, participation rights, complexity and knowledge, patriotism, and moral emotions on tax compliance behavior. Moreover, I discuss interdependencies within these sources of non-monetary costs of tax evasion on the one hand and between non-monetary and monetary costs of tax evasion on the other. All in all, these interdependencies can mutually amplify or cancel out, e.g. perceived complexity can have a negative effect on perception of fairness (Carnes and Cuccia, 1996) or a high audit probability can diminish procedural fairness (Farrar et al., 2019) . Finally, I emphasize ideas for future research.

1.2 Trust them, threaten them, or lure them? Effective audit systems to promote compliance

1.2.1 Research question and design

We investigate the effects of (1) increasing the detection probability of non-compliance, (2) monetary incentives to promote compliance (bonuses and penalties), and the interaction of (1) and (2). In a laboratory experiment, participants are requested to report information on a privately observed die roll. The participants receive a payoff that is directly related to the reported die roll (and not to the actual die roll seen), which creates

a monetary incentive to misreport. In our baseline we study compliance with the request for truthful reporting without any audits. Consequently, we investigate participants' intrinsic motivation to comply with the request for truthful reporting (Dwenger et al., 2016, Wang and Murnighan, 2017). Our treatments implement an audit system and vary two dimensions. In one dimension the probability of detecting non-compliance (i.e., misreporting) is varied. In the other dimension we vary the incentives of the participants to comply. In particular, we vary whether detected truthful reporting is awarded a bonus or detected misreporting is penalized (Torgler, 2002, Nosenzo et al., 2016, Alm, 2019, Fabbri et al., 2019). The interaction of both treatment dimensions allows us to evaluate the effects of combinations of detection probabilities and incentives on compliance with the request for truthful reporting. Moreover, our experimental design enables us to disentangle the effects of intrinsic and extrinsic motivation to report truthfully when introducing audits.

1.2.2 Results and contribution to the literature

We find that truthful reporting decreases with the introduction of audit systems that penalize non-compliance (or reward compliance) with a low detection probability, compared to the baseline scenario of no audits. Therefore, the intended deterrence has a diametrically opposed effect than desired. We interpret this as a crowding out of intrinsic motivation to be honest (Frey and Oberholzer-Gee, 1997, Gneezy and Rustichini, 2000a, 2000b, Gneezy et al., 2011). Increasing the detection probability increases the extrinsic motivation to report truthfully and results in increased compliance, both in penalty and in bonus systems. However, only a penalty system with a high detection probability overcompensates the observed gap between compliance in a no audit system and audits with low detection probabilities. A bonus system with a high detection probability just closes the gap. Moreover, a penalty system with high detection probability seems to be most efficient when taking the potential costs of bonus and penalty systems into account. Finally, our control treatments identify the influence of the explicitly stated request to report truthfully ("compliance request"). Without the compliance request, truthful reporting does not vanish completely, but decreases.

Based on our experimental analyses we stress two important facts. In circumstances in which detection probability would be very low (either because the detection of non-compliance is very difficult or very costly) our findings advice against introducing audits. Rather, one might make use of individuals' intrinsic motivation to follow an explicit

compliance request ("trust them"). Yet, when it is (technically) possible to implement a high detection probability, audit systems which penalize non-compliance outperform no audits ("threaten them"). Furthermore, directly targeting non-compliers through punishment has a stronger effect than rewarding the compliers ("lure them").

Even though we motivate our research with compliance of workers in firms, the results can be interpreted more generally and are also applicable for tax compliance research for at least two reasons. First, we focus on situations when compliance is not in the workers' individual interest, which is in line with taxpayers deciding on their tax compliance behavior. Second, we abstract from any specific context and study the situation more generally in a controlled laboratory experiment.

1.3 Combating overreporting of deductions in tax returns: Prefilling and restricting the deductibility of expenditures

1.3.1 Research question and design

We examine the influence of three anti-tax-evasion mechanisms: 1) prefilling of deductions in tax returns, 2) restricting tax evasion opportunities by either disallowing or 3) limiting the deductibility of expenditures. We address our research questions with a laboratory experiment in a controlled environment and analyze expenditure items that are substantial in real life. In the first part of the experiment, participants face a real effort task to generate their pre-tax income. In this part, participants can optionally buy a tool that simplifies the real effort task. This expenditure takes the form of more general work-related expenditures. Additionally, a fixed percentage of generated income is withheld as a fictional social insurance contribution. In the second part of the experiment, participants can optionally donate part of their income to real life institutions. In the third part of the experiment, participants have to file a tax return by reporting their deduction items. Participants are also asked to claim a commuting allowance. The declared income is already prefilled in the tax return and cannot be manipulated by the participants. Participants can evade taxes by declaring higher deductions than their true expenditures. However, with a given probability a participant's tax return will be audited and participants get punished with a fine in the case of tax evasion.

In our baseline treatment, all 4 expenditure items are deductible and the tax return is blank. In the other treatments we prefill the tax return, vary the deductibility of the 4 items (i.e. disallow the deductibility of expenditures) or limit the deductibility of

expenditures. In the later, participants are only allowed to deduct a limited amount of Eurocent per kilometer for the commuting allowance.

1.3.2 Results and contribution to the literature

Our results suggest that prefilled deductions enhance tax compliance by decreasing the item-specific tax evasion level – in particular for items preferred for tax evasion. The positive effect of prefilling might be primarily driven by higher non-monetary costs of tax evasion as our data suggests that the subjective perception of audit probabilities remains constant across treatments.

Disallowing the deductibility of an expenditure item (i.e., cutting the number of tax evasion opportunities) is an ineffective mechanism to combat tax evasion. In fact, individuals shift their tax evasion activities from the disallowed item to non-restricted items (evasion-shift-effect). However, our results suggest that limiting the deductibility (in contrast to disallowing the deductibility completely) avoids this evasion-shift-effect and finally reduces overall tax evasion. One explanation for this observation might be that disallowing the deductibility of expenditures reduces perceived procedural fairness. This might explain why we observe a shift-effect in the former case, but not in the latter one.

We conclude that policy makers trying to combat tax evasion should make use of prefilled deductions in tax returns. Moreover, policy makers should disallow the deductibility of expenditures with caution, as our observations revealed an evasion-shift-effect. However, only limiting the deductible amount might avoid this effect.

1.4 Systemization and review of non-monetary costs of tax evasion

1.4.1 Research question and design

The third essay “Systemization and review of non-monetary costs of tax evasion” focuses on the non-monetary costs of tax evasion. The pervasive problem with tax evasion of individuals was commonly examined by analyzing the monetary costs of tax evasion (i.e. audit probability, fine and tax rate). However, empirical data emphasizes that only a fraction of compliance can be explained by monetary costs of tax evasion. Over the last years of research, it has become apparent that non-monetary costs of tax evasion explain the other fraction of observed compliance. Consequently, research regarding these non-monetary costs of tax evasion is increasing. However, research referred to non-monetary costs of tax evasion in a rather selective fashion. My literature review aims to provide a more systematic approach on the non-monetary costs of tax evasion and encourages to

develop a common terminology. It selects, organizes and integrates information from roughly 170 papers into a comprehensive framework of tax compliance from the perspective of an economist. In particular, it reviews seven sources of non-monetary costs of tax evasion, social norm, trust, fairness, participation rights, complexity and knowledge, patriotism, and moral emotions. Furthermore, it discusses interdependencies within these sources of non-monetary costs of tax evasion on the one hand and between non-monetary and monetary costs of tax evasion on the other.

1.4.2 Results and contribution to the literature

My literature review contributes to the tax compliance literature by shedding light on the jungle of non-monetary costs of tax evasion. It develops a comprehensive framework of tax compliance from the perspective of an economist. In particular, one can still think of tax evasion as a tradeoff between benefits and costs of tax evasion. However, the costs must be divided into two categories, monetary and non-monetary costs of tax evasion. Furthermore, it provides a systematic review of different sources of non-monetary costs of tax evasion, which enhances comprehension, interpretation and comparison of research. All in all, I show that social norm, trust, fairness, participation rights, complexity and knowledge, patriotism, and moral emotions influence individuals deciding on tax compliance. Accordingly, governments can expand their influence on these sources in order to enhance tax compliance. However, governments have to keep in mind interdependencies within non-monetary costs of tax evasion and between non-monetary and monetary costs of tax evasion. Finally, I contribute to the literature of tax compliance with ideas for future research.

References

Adhikari, B., J. Alm, B. Collins, M. Sebastiani, and E. Wilking. 2016. Taxpayer responses to third-party income reporting: Evidence from a natural experiment in the taxicab industry. Paper read at ASSA Annual Meeting, 2018.

Allingham, M. G., and A. Sandmo. 1972. Income tax evasion: A theoretical analysis. *Journal of Public Economics* 1 (3-4):323-338.

Alm, J. 2019. What Motivates Tax Compliance? *Journal of Economic Surveys* 33 (2):353-388.

Alm, J., M. B. Cronshaw, and M. McKee. 1993. Tax Compliance with Endogenous Audit Selection Rules. *Kyklos* 46 (1):27-45.

Alm, J., B. R. Jackson, and M. McKee. 1992a. Estimating the determinants of taxpayer compliance with experimental data. *National Tax Journal* 45 (1):107-114.

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

Alm, J., I. Sanchez, and A. de Juan. 1995. Economic and Noneconomic Factors in Tax Compliance. *Kyklos* 48 (1):1-18.

Andreoni, J., B. Erard, and J. Feinstein. 1998. Tax compliance. *Journal of Economic Literature* 36 (2):818-860.

Bazart, C., and M. Pickhardt. 2011. Fighting Income Tax Evasion with Positive Rewards. *Public Finance Review* 39 (1):124-149.

Becker, G. S. 1968. Crime and Punishment: An Economic Approach. *Journal of Political Economy* 76 (2):169-217.

Bø, E. E., J. Slemrod, and T. O. Thoresen. 2015. Taxes on the Internet: Deterrence Effects of Public Disclosure. *American Economic Journal: Economic Policy* 7 (1):36-62.

Boyer, P., N. Dwenger, and J. Rincke. 2014. Do Taxes Crowd Out Intrinsic Motivation? Field-Experimental Evidence from Germany. In *Working Paper 2014 – 23*: Working Paper of the Max Planck Institute for Tax Law and Public Finance.

Carnes, G. A., and A. D. Cuccia. 1996. An analysis of the effect of tax complexity and its perceived justification on equity judgments. *The Journal of the American Taxation Association* 18 (2):40-56.

Carrillo, P., D. Pomeranz, and M. Singhal. 2017. Dodging the Taxman. *American Economic Journal: Applied Economics* 9 (2):144-164.

Clotfelter, C. T. 1983. Tax Evasion and Tax Rates: An Analysis of Individual Returns. *The Review of Economics and Statistics* 65 (3):363-373.

Duncan, D., and D. Li. 2018. Liar Liar: Experimental Evidence of the Effect of Confirmation-Reports on Dishonesty. *Southern Economic Journal* 84 (3):742-770.

Dwenger, N., H. J. Kleven, I. Rasul, and J. Rincke. 2016. Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany. *American Economic Journal. Economic Policy* 8 (3):203-232.

Fabbri, M., P. Nicola Barbieri, and M. Bigoni. 2019. Ride Your Luck! A Field Experiment on Lottery-based Incentives for Compliance. *Management Science* 65 (9):4336-4348.

Falkinger, J., and H. Walther. 1991. Rewards Versus Penalties: on a New Policy against Tax Evasion. *Public Finance Quarterly* 19 (1):67-79.

Farrar, J., S. E. Kaplan, and L. Thorne. 2019. The Effect of Interactional Fairness and Detection on Taxpayers' Compliance Intentions. *Journal of Business Ethics* 154 (1):167-180.

Fochmann, M., N. Müller, and M. Overesch. 2018. Less cheating? The effects of prefilled forms on compliance behavior. *SSRN Electronic Journal*.

Fonseca, M. A., and S. B. Grimshaw. 2017. Do behavioral nudges in prepopulated tax forms affect compliance? Experimental evidence with real taxpayers. *Journal of Public Policy & Marketing* 36 (2):213-226.

Frey, B. S., and F. Oberholzer-Gee. 1997. The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding- Out. *The American Economic Review* 87 (4):746-755.

Gillitzer, C., and P. E. Skov. 2018. The use of third-party information reporting for tax deductions: evidence and implications from charitable deductions in Denmark. *Oxford Economic Papers* 70 (3):892-916.

Gneezy, U., S. Meier, and P. Rey-Biel. 2011. When and Why Incentives (Don't) Work to Modify Behavior. *Journal of Economic Perspectives* 25 (4):191-210.

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

———. 2000b. Pay Enough or Don't Pay at All*. *The Quarterly Journal of Economics* 115 (3):791-810.

Graetz, M. J., and L. L. Wilde. 1985. The economics of tax compliance: fact and fantasy. *National Tax Journal* 38 (3):355-363.

Hofmann, E., E. Hoelzl, and E. Kirchler. 2008. Preconditions of Voluntary Tax Compliance: Knowledge and Evaluation of Taxation, Norms, Fairness and Motivation to Cooperate. *Zeitschrift für Psychologie/Journal of Psychology* 216 (4):209-217.

Kirchler, E. 2007. *The economic psychology of tax behaviour*: Cambridge University Press.

Kirchler, E., E. Hoelzl, and I. Wahl. 2008. Enforced versus voluntary tax compliance: The "slippery slope" framework. *Journal of Economic Psychology* 29 (2):210-225.

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez. 2011. Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark. *Econometrica* 79 (3):651-692.

Koessler, A.-K., B. Torgler, L. P. Feld, and B. S. Frey. 2019. Commitment to pay taxes: Results from field and laboratory experiments. *European Economic Review* 115:78-98.

Kotakorpi, K., and J.-P. Laamanen. 2016. Prefilled Income Tax Returns and Tax Compliance: Evidence from a Natural Experiment. In *Tampere Economic Working Papers 104*: University of Tampere, School of Management, Economics.

Luttmer, E. F. P., and M. Singhal. 2014. Tax Morale. *Journal of Economic Perspectives* 28 (4):149-168.

Mendoza, J. P., J. L. Wielhouwer, and E. Kirchler. 2017. The backfiring effect of auditing on tax compliance. *Journal of Economic Psychology* 62:284-294.

Murphy, R. 2019. The European Tax Gap. *A report for the Socialists and Democrats Group in the European Parliament. Global Policy*.

Nosenzo, D., O. Theo, M. Sefton, and A. van der Veen. 2016. Discretionary Sanctions and Rewards in the Repeated Inspection Game. *Management Science* 62 (2):502-517.

OECD. 2019. Tax Morale - What drives people and businesses to pay tax?, edited by O. Publishing. Paris: OECD Publishing, 1-68.

Slemrod, J. 2007. Cheating Ourselves: The Economics of Tax Evasion. *Journal of Economic Perspectives* 21 (1):25-48.

Slemrod, J., M. Blumenthal, and C. Christian. 2001. Taxpayer response to an increased probability of audit: evidence from a controlled experiment in Minnesota. *Journal of Public Economics* 79 (3):455-483.

Slemrod, J., B. Collins, J. L. Hoopes, D. Reck, and M. Sebastiani. 2017. Does credit-card information reporting improve small-business tax compliance? *Journal of Public Economics* 149:1-19.

The Economist. 2016a. Leak of the century; The lesson of the Panama papers. *The Economist*, April 9th 2016.
———. 2016b. The Panama papers; A torrential leak. *The Economist*, April 9th 2016.
———. 2019. Hitting the jackpot. *The Economist*, March 2nd 2019.
Titmuss, R. 1970. *The Gift Relationship*.
Torgler, B. 2002. Speaking to Theorists and Searching for Facts: Tax Morale and Tax Compliance in Experiments. *Journal of Economic Surveys* 16 (5):657-683.
———. 2007. *Tax compliance and tax morale: A theoretical and empirical analysis*: Edward Elgar Publishing.
Wang, L., and J. K. Murnighan. 2017. How Much Does Honesty Cost? Small Bonuses Can Motivate Ethical Behavior. *Management Science* 63 (9):2903-2914.
Weck-Hannemann, H., and W. W. Pommerehne. 1989. Einkommensteuerhinterziehung in der Schweiz: Eine empirische Analyse. *Swiss Journal of Economics and Statistics (SJES)* 125 (IV):515-556.
Yitzhaki, S. 1974. Income tax evasion: A theoretical analysis. *Journal of Public Economics* 3 (2):201-202.

CHAPTER 2

Trust them, threaten them, or lure them? Effective audit systems to promote compliance[1]

Martin Fochmann, Tobias Kölle, Peter N.C. Mohr, Bettina Rockenbach

Abstract

Increasing compliance is a vibrant topic in firms. Recent cases of loss of reputation and customer confidence have underlined the importance of this issue. A critical prerequisite for compliance on the firm level is compliance of the workforce. We address the question of appropriate means for firms to foster compliance of their workforce, in particular, when compliance is not in the monetary interests of workers. We investigate experimentally the effects of (1) increasing the detection probability of non-compliance, (2) monetary incentives to promote compliance (bonuses and penalties), and the interaction of (1) and (2). We find that compared to a situation without any audits, audit systems that penalize non-compliance or reward compliance, but with a low detection probability lead to significantly lower compliance. Increasing the detection probability increases compliance, yet only a penalty system with a high detection probability ensures significantly higher compliance than without audits.

[1] For helpful comments and suggestions thanks are due to Marco Fabbri, Susanna Grundmann, Daniele Nosenzo and Long Wang. Research is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC 2126/1– 390838866.

Content

Tables

Figures

2.1 Introduction

Compliance has become a vital topic for firms. This is not only true for banks who significantly expanded their workforce in this area over the last ten years[2], but also for other branches. Firms' ambitions to comply with environmental protection, socially responsible production and customer data protection, for example, and forgoing corruption, money-laundering and bribery do not (primarily) result from a "moral shift" in companies. Rather, these ambitions are fueled by the insight that reputation, customer confidence, and customer trust are essential success factors that may be easily ruined by transgressions. Facebook faces a serious reputation loss by data privacy braech, in particular in the Cambridge Analytica case. Boeing and Volkswagen lost customer trust and significant market value with scandals that disclosed a lack of responsibility within the company, be it the handling of the "self-certification" system at Boeing or deliberate deception in production at Volkswagen.

A critical prerequisite for compliance on the firm level is compliance of the workforce, which is particularly challenging when compliance is not in the workers' individual interest. Firms have taken different means to foster compliance of their workforce. One is creating awareness. The substantial increase of reporting on compliance issues in firms' communication may serve both the general public and as a means of raising awareness within the firm.[3] But, certainly this is not enough. Other means are compliance officers investigating compliance misconduct within the firm and mechanisms encouraging employees to speak up. It seems that engineers at Boeing had known for more than one year about a flaw in a cockpit warning system of 737 MAX, but the issue had not been reported to Boeing's senior executives, regulators or customers (The Economist, 2019b). Similar reports are given from Volkswagen and Facebook. Thus, increasing the number of compliance officers may increase the detection probability of transgressions, but importantly, this has to be combined with appropriate mechanisms to promote the truthful reporting by employees about firm processes. In particular in situations in which truthful reporting may not be in the individual interest of the employee, because it may risk the own and/or the peers' salary or even their jobs.

[2] At the end of 2018, about 30,000 (or 15%) of the 204,000 employees of Citigroup worked in compliance, risk and other control functions – compared to 4% of employees at the end of 2008. Accordingly, there are 43,000 employees working in compliance at JPMorgan Chase. Also HSBC, which was fined for banking Mexican drug money in the past, has around 5,000 employees in anti-money-laundering compliance (The Economist, 2019a).

[3] For example, banks treat compliance with priority and highlight it much more in their annual reports than they used to (The Economist, 2019a).

In this paper we address the issue of appropriate means for firms to foster compliance, in particular how to promote truthful reporting by employees (e.g., engineers' truthful reporting about critical technical problems). Specifically, we investigate the effects of (introducing) audits with a focus on the interaction between (1) increasing the detection probability (e.g., by hiring more compliance officers) and (2) incentive schemes to promote truthful reporting (penalty and bonus) in particular when it is at odds with individual monetary interest. We abstract from any specific context and study the situation more generally. In a controlled laboratory experiment, participants are requested to report information on a privately observed event. A participant's payoff is directly related to the reported and not to the actual event, which creates a monetary incentive to misreport. In our baseline we study compliance with the request for truthful reporting absent any additional regulations. This allows us to investigate participants' intrinsic motivation to comply with the request for truthful reporting (Dwenger et al., 2016, Wang and Murnighan, 2017). Our ten treatments implement the possibility of auditing and vary two dimensions: in one dimension the probability of detecting non-compliance (i.e., misreporting) is varied, while in the other dimension we vary the incentives of the participants to comply. Specifically, we vary whether detected truthful reporting is awarded a bonus or detected misreporting is penalized (Torgler, 2002, Nosenzo et al., 2016, Alm, 2019, Fabbri et al., 2019). The interaction of both treatment dimensions allows us to evaluate the effects of combinations of detection probabilities and incentives on compliance with the request for truthful reporting. Moreover, with our treatments we are able to disentangle the effects of intrinsic and extrinsic motivation to report truthfully when introducing audits.

We find that compared to the baseline scenario of no audit, the introduction of a penalty or bonus system, yet with a low detection probability, *decreases* truthful reporting. We interpret this as a crowding out of intrinsic motivation to be honest (Frey and Oberholzer-Gee, 1997, Gneezy and Rustichini, 2000a, Gneezy and Rustichini, 2000b, Gneezy et al., 2011). Thus, the intended deterrence has a diametrically opposed effect than desired. Increasing the detection probability increases the extrinsic motivation to report truthfully and results in increased compliance in our data, both in penalty and in bonus systems. Yet, a penalty system with a high detection probability is needed to overcompensate the observed gap between compliance in a no audit system and audits with low detection probabilities. A high detection probability with a bonus system just closes the gap. Also when taking the potential costs of bonus and penalty systems into

account, a penalty system with high detection probability seems to be most efficient. To identify the influence of the explicitly stated request to report truthfully ("compliance request"), we conduct control treatments without the specific compliance request. In these treatments the reporting task was described to subjects as a lottery. We find that without the compliance request, truthful reporting does not vanish completely, but is significantly lower. The synopsis of all findings suggests that when introducing an auditing system to detect non-compliant behavior, a low detection probability may be even detrimental to compliance; only a high detection probability with penalties for detected non-compliance may promote compliance. Explicitly requesting compliant behavior yields a positive effect, an additional "compliance premium".

2.2 Related literature

Compliance with the request for truthful reporting is important for the functioning of firms and societies. The decision of an individual to misreport or to report truthfully depends on both extrinsic and intrinsic motivators. Whereas extrinsic motivation impacts individual behavior from outside by rewarding truthful reporting with a bonus or punishing untruthful reporting with a penalty (Torgler, 2002, Dwenger et al., 2016, Nosenzo et al., 2016), intrinsic motivation comes from inside of an individual (Gneezy et al., 2011, Wang and Murnighan, 2017). Most importantly, introducing extrinsic motivators bears the risk of a crowding out of intrinsic motivation (Frey and Oberholzer-Gee, 1997, Gneezy and Rustichini, 2000b, Fehr and Rockenbach, 2003, Bowles and Polania-Reyes, 2012). Consequently, introducing audits with penalties or bonuses can have detrimental effects on truthful reporting. In the following, we review the related literature.

2.2.1 Extrinsic motivation to report truthfully

Authorities such as firms or legislators often introduce extrinsic motivators such as penalties and/or bonuses to foster compliance with the request for truthful reporting. In the case of penalties, untruthful reporting usually entails the tradeoff between the benefit of successful misreporting and a risky prospect of detection and punishment. Following the economics of crime approach – the standard deterrence model – reporting behavior mainly depends on extrinsic factors such as the probability that misreporting is detected and the resulting sanctions (Becker, 1968, Allingham and Sandmo, 1972). A large literature focuses on the effectiveness of these extrinsic factors to enhance reporting

behavior. Starting point of that literature was the study of tax reporting behavior as tax evasion results in massive social welfare losses in almost all countries worldwide. Excellent literature reviews are provided by Andreoni et al. (1998), Torgler (2002), Alm (2012), Slemrod (2016), and Alm (2019). This literature provides robust evidence that penalty mechanisms perform well in fostering truthful reporting. In particular, higher fines and higher probabilities of getting caught evading enhance tax reporting behavior in general.

More recently, a growing body of literature focuses on truthful reporting in a non-tax context. For example, in a laboratory experiment, Laske et al. (2018) study the probability and size of penalties in the deception game (Gneezy, 2005). In this game, misreporting increases the earnings of the cheater at the expense of another individual. The authors observe that introducing audits decreases the frequency of misreports. Furthermore, they find that misreporting decreases with the size of the fine. However, the results are mixed for probabilities. In the one-shot treatments where an individual is only confronted with one parameter set (between-subject design), the frequency of misreports is insensitive to probability changes. In the treatments where participants are confronted with changing probabilities (within-subject design), the frequency of misreports decreases with the probability. In all settings, the authors find that increasing fines is more effective than increasing the detection probability. Gamliel and Peer (2013) apply the task introduced by Mazar et al. (2008) in which participants have to find two numbers in a matrix that added up to 10 and finally have to report the number of correctly solved matrices. In a between-subject design, they find that introducing a deterrence mechanism with a small detection probability (1/36) does not influence reporting behavior significantly.

Another potential extrinsic motivator are bonuses. In line with conventional economic models (e.g., Becker, 1968), increasing the material reward for truthful reporting should reduce misreporting. Although bonus mechanisms are less studied in the tax evasion literature, some studies show that bonuses can enhance tax reporting behavior (Torgler, 2003, Bazart and Pickhardt, 2011, Alm et al., 2012, Fatas et al., 2015, Dwenger et al., 2016, Koessler et al., 2019). In a non-tax context, Fabbri et al. (2019) combine a lottery-based bonus for truthful behavior with probabilistic sanctions for untruthful behavior in a field experiment with bus passengers in Italy. They show that the introduction of the bonus increases the purchased on-board bus tickets and consequently foster truthful behavior. Other studies provide mixed results (Gneezy et al., 2011 and Wang and Murnighan, 2017 for overviews) and suggest that bonuses for truthful behavior

have a stronger effect on the intrinsic than on the extrinsic motivation (see next subsection for more details). Moreover, it seems that monetary rewards must be large enough to have a significantly positive impact (Gneezy et al., 2011).

In sum, the tax-related literature shows clear support that both detection probability and size of penalties foster truthful reporting whereas the literature focusing on reporting behavior in a non-tax context reveals mixed results. Especially, the question whether audits indeed have a positive effect on reporting behavior remains an open question. Furthermore, only some studies investigated the effect of bonuses on reporting behavior. Therefore, further research is needed to shed more light on the motives of untruthful reporting in case of introducing audits.

2.2.2 Intrinsic motivation to report truthfully

An important result of the honesty literature is that – although misreporting cannot be detected – individuals refrain from being dishonest to the maximum extent or are even completely honest (Mazar et al., 2008, Fischbacher and Föllmi-Heusi, 2013, Grolleau et al., 2016, Gneezy et al., 2018, Abeler et al., 2019). This finding suggests that individuals seem to have an intrinsic motivation to behave truthfully and to comply with norms that is independent of the risk of getting caught.

A prominent approach in this regard is the theory of self-concept maintenance developed by Mazar et al. (2008). This theory suggests that individuals will behave dishonestly enough to profit from norm violation, but honestly enough to maintain a positive self-concept. This explains, for example, the great number of partial liars in their matrix task and in the dice task introduced by Fischbacher and Föllmi-Heusi (2013). Another theoretical approach proposes that individuals receive a non-monetary benefit for "doing the right thing" when being honest (e.g., emotional reward), similar to the concept of warm-glow of giving (Andreoni, 1990) because an honest person forgoes the opportunity to harm someone else. Recently, Abeler et al. (2019) used data from 90 experimental studies and find that individuals lie surprisingly little. They argue that the two main motivations for truth-telling are a preference for being seen as honest and a preference for being honest. Dufwenberg and Dufwenberg (2018) suggest that cheating aversion is a main driver for honest behavior.

Furthermore, several studies argue that norms encourage individuals to report truthfully (e.g., Torgler, 2002, Alm, 2012). It is argued that the moral costs associated with misreporting reduce the utility of reporting dishonestly and therefore enhance truthful

reporting (Fortin et al., 2007, Traxler, 2010). The literature suggests that negative emotional reactions, such as guilt or shame felt when violating a norm, are important drivers (Erard and Feinstein, 1994, Dulleck et al., 2016). In particular, an intended violation of a norm might create an aversive emotional reaction while making a reporting decision, and consequently, it might reduce the tendency to misreport (Kirchler, 2007).

Several factors have been shown to influence the intrinsic motivation to comply with the request to report truthfully. Mazar et al. (2008), for example, propose the attention to standards as an important part of their self-concept maintenance theory. They suggest that when individuals attend to their own moral standards, any norm violation is more likely to be reflected in their self-concept. As a consequence, individuals will behave more honestly when attention to standards increases. In line with this hypothesis, Mazar et al. (2008) find that individuals report performance more truthfully in their matrix task after recalling the Ten Commandments and after signing a (fictive) honor code of their university. This result is in line with Krupka and Weber (2009) who suggest that the effectiveness of norms increases when an individual's attention is drawn to them. Further support for the role of attention to standards is provided by Wang and Murnighan (2017). The authors use the deception game (Gneezy, 2005) to investigate the role of small monetary bonuses for truthful reporting. In line with their hypothesis, Wang and Murnighan (2017) show that truthful reporting significantly increases after offering small bonuses for honesty. Importantly, the effect cannot only be attributed to the extrinsic motivation of the bonus itself, as the observed effect on reporting behavior disappears in a condition where misreporting leads to the same monetary consequences as in the bonus condition, but does not come in the form of a bonus. A small bonus for truthful reporting might thus have a symbolic value that increases the attention to standards and as a consequence reduces misreporting. Several psychological studies suggest that monetary bonuses also have subconscious effects on intrinsic motivation (e.g., Bijleveld et al., 2012). Festinger and Carlsmith (1959), for example, show that a small bonus of $1 increases intrinsic motivation more than a larger bonus of $20.

However, the reverse effect might also occur. Several studies suggest that small bonuses can crowd out intrinsic motivation (e.g., Frey and Oberholzer-Gee, 1997). Crowding-out effects might result when monetary incentives do not increase the attention to standards but to the extrinsic rewards. In line with this hypothesis, small bonuses have been shown to undermine prosocial behavior and voluntary cooperation (e.g., Gneezy and Rustichini, 2000b). In a classic example, Titmuss (1970) shows that monetary incentives

reduce rather than increase blood donations. Gneezy and Rustichini (2000b) find that offering high school students a small compensation decreases their effort in a fund-raising campaign.

Crowding-out-effects might also occur when immoral behavior is punished. Gneezy and Rustichini (2000a) report the results of a field study in a group of day-care centers in which they introduce a monetary fine for parents arriving late to collect their children. In contrast to their expectation, the number of late-coming parents increased after the intervention. Fehr and Rockenbach (2003) show that sanctions reduce altruistic cooperation. The authors introduce the possibility for the investor in a trust game to punish non-cooperative behavior of trustees. Announcing a fine for non-cooperative behavior, however, reduces back-transfers of the trustees rather than increasing them. Gamliel and Peer (2013) introduce a probability of 1/36 that misreporting will be detected into the matrix task introduced by Mazar et al. (2008). Participants receive no money at all for solving matrices in the task when a misreport is detected. The authors observe that participants in the condition with a probabilistic penalty report even more solved matrices than participants in the condition in which a misreport cannot be detected. Although the difference (9.4 vs. 8.7 matrices) is not significant, the authors claim that intrinsic motivation might have been crowded out, because they find no evidence that introducing a probabilistic penalty has a positive effect on reporting behavior. However, it has to be noted that Dwenger et al. (2016) find no evidence for a crowding out of the intrinsic motivation to pay church taxes in a field experiment after introducing deterrence or rewards.

In sum, there is substantial evidence that reporting behavior is largely affected by intrinsic motivations. These motivations are, however, very fragile, depending on situation and context. Moral reminders and small bonuses for truthful reporting have been shown to increase the attention to standards and as a consequence reduce misreporting. However, both bonuses and penalties might also crowd out intrinsic motivation and increase misreporting.

2.3 Hypotheses

The literature review has shown that the effectiveness of audit systems for compliance with the request for truthful reporting may crucially depend both on intrinsic and extrinsic motivations. In this section we deduce hypotheses on how the different designs

of audit systems may affect intrinsic as well as extrinsic motivations and thus influence compliance of individuals.

Intrinsic motivation. The introduction of audits with a penalty for non-compliance or a bonus for compliance can have several effects on the intrinsic motivation to report truthfully. It might draw attention to moral standards and signals that non-compliance violates these standards (Bijleveld et al., 2012, Festinger and Carlsmith, 1959, Wang and Murnighan, 2017). In this case, the higher attention to standards increases intrinsic motivation to maintain a positive self-view (Mazar et al., 2008). Moreover, individuals might perceive the settings in which non-compliance can be punished and compliance can be rewarded as more fair compared to a setting without penalties and bonuses. Higher perceived fairness has been shown to be associated with higher intrinsic motivation and higher compliance levels (Hofmann et al., 2008, Kirchler et al., 2008). Consequently, from this perspective introducing audits might increase intrinsic motivation to report truthfully.

However, the introduction of audits can also lead to a crowding out of intrinsic motivation if monetary incentives (partially) destroy intrinsic motives (Frey and Oberholzer-Gee, 1997, Fehr and Rockenbach, 2003, Gamliel and Peer, 2013). Moreover, if monitoring and sanctioning is perceived as a breach of trust (in one's own morally correct behavior), the settings with audits are seen as less fair compared to settings without resulting in a lower intrinsic motivation (Kirchler et al., 2008). Furthermore, introducing (probabilistic) audits implements a risk dimension into the decision context. Therefore, individuals also have to consider a risk component in their decision making. The attention to this risk component might lead to a lower attention to standards and consequently to a lower intrinsic motivation to report truthfully (crowding-out effect). In these cases, introducing audits decreases intrinsic motivation. As the effects have contrary consequences, we formulate two opposing hypotheses:

Hypothesis 1a: *Introducing audits increases intrinsic motivation to comply with the request to report truthfully and consequently increases compliance.*

Hypothesis 1b: *Introducing audits decreases intrinsic motivation to comply with the request to report truthfully and consequently decreases compliance.*

Extrinsic motivation. Penalizing non-compliant behavior (rewarding compliant behavior) with a certain probability reduces (increases) the expected value of non-compliance (compliance). Therefore, introducing audits with penalty/bonus lowers the

relative attractiveness of non-compliance compared to compliance in monetary terms and thus increases the extrinsic motivation to report truthfully. Moreover, a higher detection probability increases the expected sanction in case of non-compliance and increases the expected reward in case of compliance. Consequently, increasing the detection probability further increases the relative attractiveness of compliance and thus increases the extrinsic motivation. We therefore expect higher compliance levels for higher detection probabilities (Torgler, 2002, Alm, 2012, Laske et al., 2018, Thielmann and Hilbig, 2018). Our hypothesis 2 therefore reads as follows:

Hypothesis 2: *Increasing detection probabilities increases extrinsic motivation to comply with the request to report truthfully and consequently increases compliance.*

Bonus vs. penalty incentive schemes. There are only few studies comparing the effectiveness of bonus and penalty incentive schemes to promote compliance. Liang et al. (2013) analyze survey data collected from 186 employees working in companies applying penalty and bonus mechanisms to regulate mandatory IT usage. They find strong positive effects on compliance for penalties, but no significant effects for bonuses. Stronger effects for penalties on compliance are also found for the inspection game in Nosenzo et al. (2014), but not in Nosenzo et al. (2016). Additionally, there are several studies focusing on behavior in social dilemmas with the common finding that penalties are more effective than bonuses. For example, Andreoni et al. (2003) show in proposer-responder games with costly punishments and rewards that "rewards are much less effective in moving the proposers away from the minimum possible offer" and they suggest that "one might expect less cooperation in societies where good behavior is rewarded than in those where poor behavior is punished" (Andreoni et al., 2003, p. 894). The results of several public goods games also reveal that penalties are more effective than bonuses in encouraging public good contributions (e.g., Sefton et al., 2007, Sutter et al., 2010). In a meta-analysis study involving 187 effect sizes, Balliet et al. (2011) observe a slightly larger effect size estimate of penalties than bonuses. Hossain and List (2012), who find that penalties are more effective than bonuses in increasing work productivity (although the robustness of this finding is challenged by de Quidt et al., 2017), argue that loss aversion might be one fundamental driver for the outperformance of penalties. If individuals are loss averse (Kahneman and Tversky, 1979), a penalty that is perceived as a loss should have a

stronger impact on decision making than a bonus that is perceived as a gain. Our hypothesis 3 therefore reads as follows:

Hypothesis 3: *Penalties are more effective in promoting compliance than bonuses.*

2.4 Experimental design

2.4.1 Experimental setup

In our laboratory experiment we study the effectiveness of different audit systems to promote compliance. Our special focus is on promoting truthful reporting of private information (e.g., by an employee about particular processes in the firm) in a situation in which the reporting individual faces a tradeoff between complying with the request for truthful reporting and a personal monetary benefit from misreporting. In light of previous research, we focus on the effects of the interaction between (1) different detection probabilities and (2) different incentive schemes (penalty and bonus) for promoting truthful reporting. As a benchmark for our analysis, we use a treatment without any audits (No Audit treatment). To study the effect of varying the detection probability, we introduce treatments with detection probabilities of 0%, 30% and 70%. In all these treatments, participants are informed about the incentive scheme (penalty or bonus, see below) before they learn their individual detection probability. The rationale for the 0% detection probability is to disentangle the effects of audits on intrinsic and extrinsic motivation to report truthfully. Usually these effects appear simultaneously, but with a detection probability of 0% the extrinsic motivation is absent as audits are never performed (i.e., no monetary consequences are to be expected). Therefore, only intrinsic motivation plays a role in this condition. Comparing the No Audit benchmark and the 0% condition shows how the introduction of audits impacts the intrinsic motivation to report truthfully and thus allows us to test hypothesis 1. Comparing the 0%, 30% and 70% conditions enables us to test how an increase of the detection probability alters compliance (hypothesis 2). To test the effectiveness of different incentive schemes, we study penalty and bonus treatments in which detected non-compliance is punished (in penalty) and detected compliance is rewarded (in bonus), respectively. Comparing these treatments allows us to test hypothesis 3. All treatments are described in detail in Section 2.4.3.

2.4.2 Experimental task

Our experimental task consists of a one-shot decision that abstracts from any specific context. We use a variant of the dice task introduced by Fischbacher and Föllmi-Heusi (2013) that was already applied in a similar form by Kocher et al. (2018). In this task, participants see a video of a die roll on their computer screen. Participants are informed that each of the six possible outcomes of die rolls is seen with the same probability. The advantages of showing videos of die rolls is that the experimenter can pre-select a random sequence that shows every outcome with the same frequency and that misreporting is detectable on an individual level. A potential drawback may be that the observability by the experimenter may affect lying behavior. Yet, there is evidence that in practice it does not (see for example the dice task with observability performed by Kocher et al., 2018, a similar task with observability by Gneezy et al., 2018 and experiments with an anonymous observer by Baeker and Mechtel, 2015, Houser et al., 2016, van de Ven and Villeval, 2015). The participants' task is to report the die roll outcome shown in the video on the next screen. Participants can report any number between 1 and 6. Importantly, payoffs depend on the number reported.

We use a slightly different payoff structure compared to the original experiment proposed by Fischbacher and Föllmi-Heusi (2013) and the variant used by Kocher et al. (2018). Die rolls 1-6 yield 1-6 Euros, respectively. We changed the original payoff structure of die rolls 1-5 yielding 1-5 Euros, respectively, and die roll 6 yielding 0 Euro to make it as easy as possible to calculate possible outcomes, also in the light of additional penalties and bonuses (see below). As in the original task we ask participants explicitly to report the die roll that was shown on the screen.[4] However, participants have the possibility to misreport.

2.4.3 Experimental treatments

We use a between-subject design and participants are randomly assigned to each of our eleven treatments (see supplementary material for all instructions). The *No Audit* treatment serves as our baseline. The six *Detection* treatments vary the two dimensions detection probability and incentive scheme. Note that in case of an audit the truthfulness of the report is unambiguously detected such that the audit probability corresponds to the detection probability. In a 3x2 factorial design we vary the three detection

[4] In the instructions it says: "Your task is to remember the number you saw in the video and to enter this number into the computer".

probabilities 0%, 30%, and 70% and the two incentives schemes penalty and bonus (see Table 1 for a treatment overview).

Table 1: Treatment overview

	Number of subjects	Detection probability	Compliance request	Penalty/ Bonus payment
Without any deterrence mechanism				
No Audit	162	---	✓	---
Detection and additional penalty in case of detected non-compliance				
Detection 0 Penalty	102	0%	✓	---
Detection 30 Penalty	102	30%	✓	✓
Detection 70 Penalty	102	70%	✓	✓
Detection and additional bonus in case of detected compliance				
Detection 0 Bonus	102	0%	✓	---
Detection 30 Bonus	102	30%	✓	✓
Detection 70 Bonus	102	70%	✓	✓
NCR – no compliance request, payoff structure corresponds to respective penalty treatments				
NCR 30 Penalty	102	30%	---	✓
NCR 70 Penalty	102	70%	---	✓
NCR – no compliance request, payoff structure corresponds to respective bonus treatments				
NCR 30 Bonus	102	30%	---	✓
NCR 70 Bonus	102	70%	---	✓

- *No Audit treatment:* The die roll reported is not audited and hence misreporting cannot be detected. The number reported directly corresponds to the subject's payment (1 = 1 Euro, 2 = 2 Euro, … , 6 = 6 Euro).
- *Detection 0/30/70 Penalty treatments:* The subject's reported die roll is detected with a probability of 0%, 30%, and 70%, respectively. If a misreport is detected, the subject only earns a payment corresponding to the die roll seen (and not to the die roll reported) and additionally has to pay a penalty of 1 Euro. If a truthful report is detected, the subject receives the reported (= seen) die role in Euros.
- *Detection 0/30/70 Bonus treatments:* The subject's reported die roll is detected with a probability of 0%, 30%, and 70%, respectively. If a truthful report is detected, the subject receives the reported (= seen) die role in Euros plus a bonus of 1 Euro. If a misreport is detected, the participant only earns the amount of

money in correspondence with the die roll seen, but no additional penalty has to be paid.

In our design we always "correct" a misreport when detected. This mirrors the situation that in real life a firm will always have to correct the incorrect behavior once it is detected. The treatment variation captures whether or not it additionally rewards or penalizes its employees for compliant or non-compliant behavior. Further note that we do not vary the size of the penalty or bonus (as for example in Laske et al., 2018), as – given the motivation of our research – in work contexts legal restrictions oftentimes narrow down the variability in penalties or bonuses that can be applied to workers.

All instructions were shown on the participants' computer screens. Subjects are first informed about the penalty (or bonus) scheme and thereafter learn their individual detection probability. In the Detection 0 treatments no auditing and thus no detection takes place. These treatments are intended to mirror situations in which firms announce the introduction of bonuses (or fines) for (non-)compliant behavior, but it is clear that detection is extremely unlikely, e.g., due to the lack of an effective infrastructure. These treatments are valuable benchmarks, because *in monetary terms*, the No Audit treatment and the Detection 0 treatments are identical. They only differ in that participants of the Detection 0 treatments have been informed about the penalty (or bonus) scheme before they learn that the individual detection probability is zero. Importantly, note that the pre-decision comprehension test already made the possibility of misreporting salient in all treatments (even in our baseline treatment No Audit). In this test, subjects had to correctly name the consequences of different hypothetical scenarios of truth-telling and lying (see section 2.4.4 and supplementary material M2). Yet, we cannot exclude that introducing the penalty (or bonus) scheme further increases the salience of the possibility of misreporting and thus changes the psychological framework of the in monetary terms identical situations of No Audit and Detection 0.

Expected payoffs and incentives to comply. Suppose a participant has seen die roll s and faces a detection probability $q \geq 0$. In the penalty treatments, the payoff from truthfully reporting s is $\pi_s^P = s$ and the expected payoff from untruthfully reporting $r \neq s$ is $\pi_r^P = (1-q)r + q(s-1)$. In the bonus treatments, the expected payoff from truthful reporting is $\pi_s^B = (1-q)s + q(s+1) = s + q$ and the expected payoff from untruthfully reporting $r \neq s$ is $\pi_r^B = (1-q)r + qs$. Thus, independent of whether or not the participant reports truthfully, the expected payoff in the bonus treatments is by q higher than the expected payoff in the penalty treatments, i.e., $\pi_s^P + q = \pi_s^B$ and $\pi_r^P + q = \pi_r^B$. This linear shift in

the expected payoff functions implies that the strategic incentive to report truthfully is identical across the respective penalty and bonus treatment. Moreover, both in the penalty and the bonus treatments the expected payoff difference between truthful and untruthful reporting is identical, i.e., $\Pi = \pi_s^P - \pi_r^P = \pi_s^B - \pi_r^B = q + (1-q)(s-r)$. This means that the monetary incentives to misreport are identical across a bonus and a penalty treatment with the same detection probability.

Risk-neutral participants who maximize their expected payoff report the seen die roll s truthfully if $\Pi \geq 0 \Leftrightarrow s \geq r - \frac{q}{1-q}$. Thus, for detection probabilities $q \leq \frac{1}{2}$ an expected payoff maximizer will always report r = 6, independent of the die roll seen. For detection probabilities $\frac{1}{2} < q \leq \frac{5}{6}$ an expected payoff maximizer will report the seen die roll s truthfully for "high" die rolls and report r = 6 for "low" die rolls seen. The cut-off between "high" and "low" is $6 - \frac{q}{1-q}$. If the detection probability is $q > \frac{5}{6}$ an expected payoff maximizer will always report the seen die roll s truthfully. In our experimental parametrization, q = 0% serves as our benchmark to No Audit and provides no incentives for truthful reporting, like in No Audit. Accordingly, q = 30% represents a scenario with a strictly positive detection probability, but both in penalty and bonus it provides no incentives for truthful reporting and an expected payoff maximizer reports r = 6. Finally, q = 70% represents a scenario in which there is no incentive for truthful reporting of the "low" die rolls 1, 2, and 3, but an expected payoff maximizer truthfully reports the "high" die rolls 4 and 5 (cut-off 3.67). We refrained from studying $q > \frac{5}{6}$ as there is no conflict between the interests of an expected payoff maximizing participant and e.g., a firm asking for compliance.

No compliance request (NCR). In our remaining four No Compliance Request (NCR) treatments, we reframe the decision situation and refrain from explicitly asking participants to report the die roll seen. Instead, participants are asked to enter "any" number after the video with the die roll is shown.[5] The NCR treatments implement the same payoff structure as the Detection treatments, but without the explicit request to report truthfully.

- *NCR 30/70 Penalty treatments:* If a participant enters the die roll seen, she/he earns the corresponding amount of money for sure. However, if she/he enters a different number, she/he enters a lottery. With probability 30% (70%), she/he

[5] In the instructions it says: "Your task is to remember the number you saw in the video and to enter this number or any other number into the computer. You are free to type in any number".

earns the amount of money corresponding to the number entered. With the complementary probability 70% (30%), however, she/he only earns the amount of money corresponding to the die roll seen minus 1 Euro. Consequently, payoff-consequences reflect those of the Detection 30/70 Penalty treatments.

- *NCR 30/70 Bonus treatments:* Payoff-consequences reflect those of the Detection 30/70 Bonus treatments. If a participant chooses to enter the seen die roll, she/he earns the corresponding amount of money for sure and will get an additional payment of 1 Euro with a probability of 30% (70%). If a participant chooses to enter a different number, she/he earns with a probability of 30% (70%) only a payment corresponding to the seen die roll and with the complementary probability 70% (30%) the payment corresponding to the entered number.

Note that we did not implement probability 0% in the NCR treatments, because by removing the request to report truthfully, there is no potential interaction between honesty and auditing. In the NCR 0 treatment participants would just receive the entered amount for sure.

2.4.4 Experimental procedure, controls and sample

Our laboratory experiment started with the instructions and an incentivized comprehension test (see Figure 1). Both were provided on-screen and participants could switch between instructions and comprehension test with a button. In the comprehension test (see supplementary material M2), participants could earn 1 Euro if they answered all four questions correctly at the first attempt (which was accomplished by roughly 50% of our participants). After all participants answered the comprehension test correctly, the experiment proceeded with showing the die roll video. After participants reported the die roll, they were asked to state their belief about the behavior of others (see below). Finally, participants answered a questionnaire, received a payoff information and their individual payoff in cash.

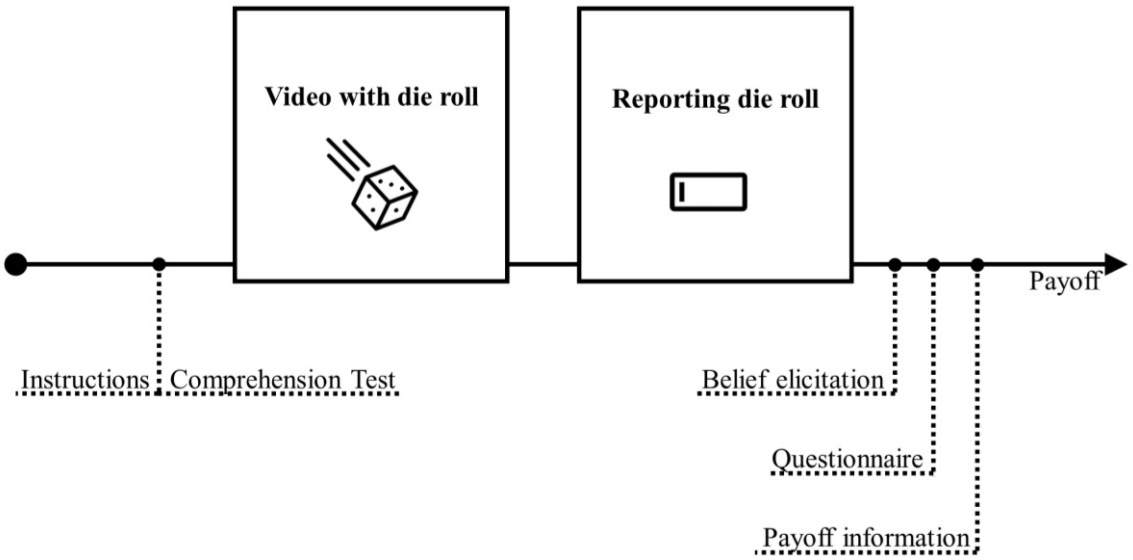


Figure 1: Experimental procedure

Controls. We collected the following controls in our ex-post questionnaire: "age" (in years), "risk attitude" (obtained from the G-SOEP survey; it provides the subject's self-reported general willingness to take a risk, measured on an 11-point scale, where 0 = not willing to take risk and 10 = highly willing to take risk), "female" (1 if participant is female, 0 otherwise), "morale" (we asked about the opinion on the statement: "lying for your own benefit", answers were given on a 10-point Likert scale from "…is always justifiable" = 0 to "…is never justifiable" = 9), "bachelor" (1 if participant studies in a bachelor program, 0 otherwise), "economics" (1 if participant has attended more than one lecture in economics or management, 0 otherwise) "decision complexity" (measures how complex a participant perceived the task in the experiment, 11-point scale from 0 = low perceived decision complexity to 10 = high perceived decision complexity), "monthly income" (monthly income in Euro after fixed costs such as rent) and "political opinion" (participants were asked to state their political opinion on an 8-point scale from 0 = political left to 7 = political right).

Belief elicitation. After reporting the die roll number, we elicited beliefs. We asked the participant to imagine that 60 other participants in this experiment had seen the same die roll as she/he did and asked about the participant's belief about how many of them would have reported a 1, 2, …, or 6. This allows us to study how the participant's own behavior relates to the expected behavior of others.

Sample. The experiment was conducted at the computerized experimental laboratory of the University of Cologne (CLER) from November 2018 to May 2019. The experimental software was programmed and run with z-Tree (Fischbacher, 2007), and the participants

were recruited with ORSEE (Greiner, 2015). In total 1,182 subjects participated and earned, on average, 11.15€ in approximately 45 minutes (approximately an hourly wage of 14.87€). Table 2 provides an overview on the main characteristics of our participants.

Table 2: Main characteristics of our participants

Variable	Description	Mean
Age	in years	25
Risk attitude	self-reported risk attitude 0 to 10	4.97
Female	female = 1; (0 otherwise)	59%
Morale	0 to 9; low morale = 0; high morale = 9	5.21
Bachelor	enrolled in a bachelor program = 1 (0 otherwise)	56%
Economics	participant with more than one lecture in economics = 1 (0 otherwise)	53%
Decision complexity	0 to 10; low perceived decision complexity = 0; high perceived decision complexity = 10	1.92
Monthly income	in Euro (monthly income after fixed costs)	373
Political opinion	0 to 7, political left = 0, political right = 7	2.55

2.5 Results

We start out with a rough overview over our results. Figure 2 displays the average die roll reported and Figure 3 shows the share of misreports, both separated by treatment. In Figure 3, misreport is a binary variable that takes the value 1 if the reported die roll differs from the seen die roll and 0 otherwise (see Appendix A1 for the distribution of data). In all analyses in this section, we only consider the data of participants who observed a die roll lower than 6, as we are only interested in the behavior of participants, who are at least theoretically able to misreport for personal benefits. In fact, all of these subjects reported a die roll of 6 truthfully. The total number of observations used is therefore 985 (= 1,182 – 197).

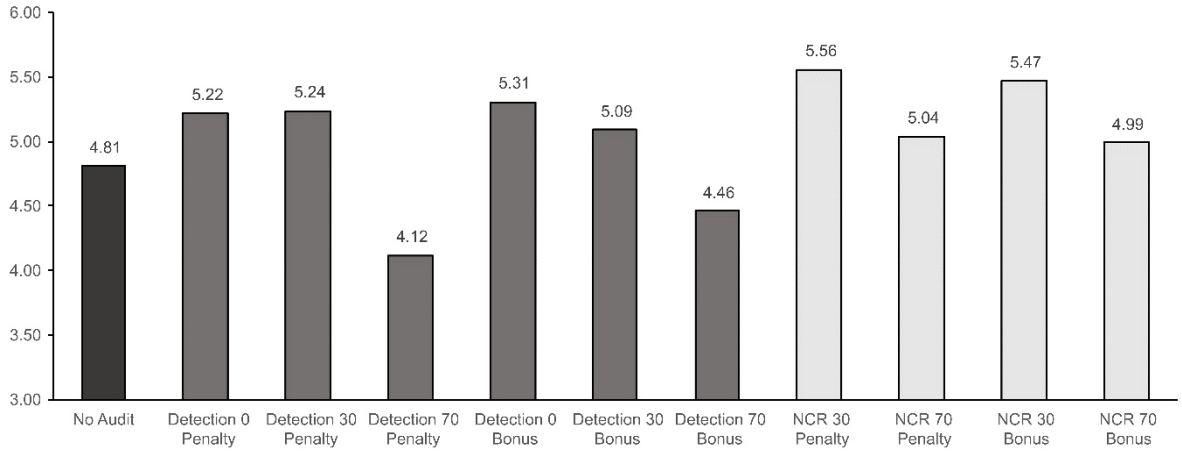


Figure 2: Average reported die roll

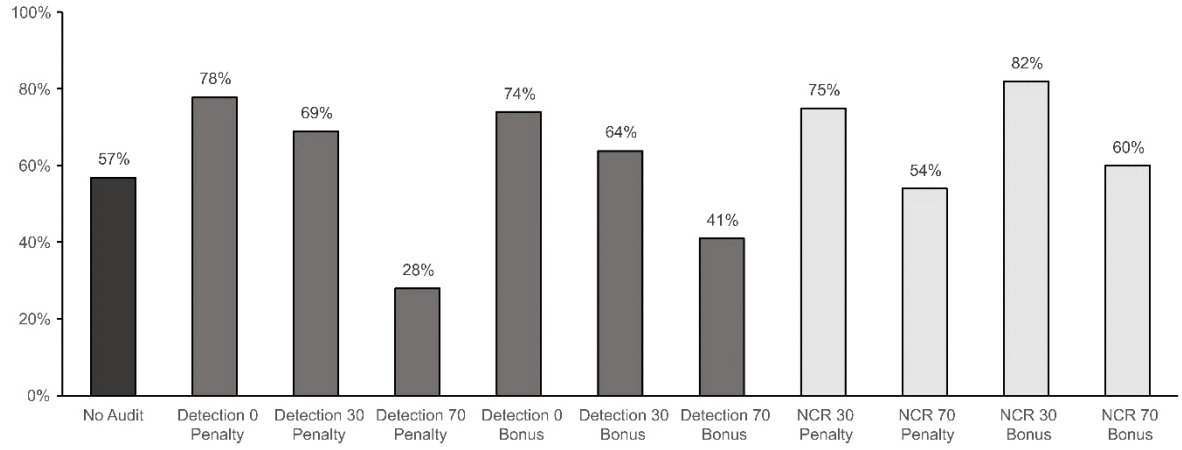


Figure 3: Share of misreports

In the following subsections, we present our main results. The Appendices complement these results with further analyses on the distribution of reported die rolls (Appendix A1), further regression results (Appendix A2), and additional analysis on stated beliefs on the behavior of others (Appendix A3).

2.5.1 Crowding out of intrinsic motivation and crowding in of extrinsic motivation

No Audit vs. Detection 0. In a first step, we compare the three treatments without detection of non-compliant behavior (No Audit, Detection 0 Penalty, and Detection 0 Bonus). In these treatments there is no extrinsic motivation to comply with the request to report truthfully and they therefore allow to study whether and how intrinsic motivation to report truthfully is affected by introducing audits with penalties and bonuses, yet with detection probability zero. In the No Audit treatment the share of misreports (57%) is

significantly lower than in Detection 0 Penalty (78%) and in Detection 0 Bonus (74%), chi^2-test, both p-values ≤ 0.01, two-tailed. Also the average reported die roll increases significantly from 4.81 in No Audit to 5.22 in Detection 0 Penalty and 5.31 in Detection 0 Bonus (Mann-Whitney U-test, both p-values ≤ 0.01, two-tailed). Between Detection 0 Penalty and Detection 0 Bonus we do not observe significant differences in terms of share of misreports (p = 0.591) and average reported die roll (p = 0.604). Remarkably, while we observe some partial liars in No Audit, both in Detection 0 Penalty and Detection 0 Bonus, participants are either honest or lie to the full extent by reporting a 6 (see Figures A1.1 and A1.2 in Appendix A1). Thus, the introduction of an audit system to identify non-compliant behavior, yet with a detection probability of zero seems to crowd out intrinsic motivation to report truthfully, possibly by signaling distrust, and thereby significantly decreases compliance with the request to report truthfully.

Table 3: Regression results

	All treatments (except NCR treatments)	
Dependent variable	**Misreport**	**Reported die roll**
	Model 1	Model 2
Reference group	No Audit	No Audit
Detection 0	1.19***	1.10***
	(0.36)	(0.32)
Detection 30	0.82**	0.79***
	(0.34)	(0.29)
Detection 70	-1.37***	-0.76***
	(0.35)	(0.26)
Detection 0 X Bonus	-0.22	-0.13
	(0.40)	(0.37)
Detection 30 X Bonus	-0.30	-0.25
	(0.38)	(0.32)
Detection 70 X Bonus	0.87**	0.52*
	(0.39)	(0.28)
Controls	Yes	Yes
No. of observations	645	645
Pseudo R-squared	0.2568	0.0980

Note: This table reports regression results (regression coefficients, standard errors in parentheses, reference group: No Audit treatment). In model 1, the dependent variable is our binary variable misreport and we use a logistic regression. In model 2, the dependent variable is the reported die roll and we use an ordered logistic regression. Controls and constants are not reported here, but can be found in Table A2.1 in Appendix A2.

To corroborate these descriptive and nonparametric results, we present regression analyses in Table 3. In model 1, the dependent variable is our binary variable misreport and we use a logistic regression. In model 2, the dependent variable is the reported die roll. As this variable is not metrically scaled, we use an ordered logistic regression. In each model, we use three dummies Detection 0/30/70 that take the value 1 if the decision was

either made in the corresponding Detection 0/30/70 Penalty or Detection 0/30/70 Bonus treatments (0 otherwise) to test the differences between these treatments and the treatment No Audit (which serves as the reference group). To study interaction effects, we use three dummies Detection 0/30/70 X Bonus that take the value 1 if the decision was made in the corresponding Detection 0/30/70 Bonus treatment. Consequently, the coefficient of the dummy Detection 0, for example, measures the difference between Detection 0 Penalty and No Audit. The coefficient of the dummy Detection 0 X Bonus measures the difference between Detection 0 Penalty and Detection 0 Bonus. Moreover, we consider a vector of controls such as die roll seen, gender, age, risk attitude and moral attitude (for details on our controls see section 2.4.4 "controls"). Controls and constants are not reported here, but can be found in Table A2.1 in Appendix A2.

In line with the nonparametric results reported above, we find that participants misreport more frequently and report higher die rolls in our Detection 0 treatments compared to the benchmark No Audit. Differences are significant at the 1%-level. The coefficient of the dummy Detection 0 X Bonus is insignificant. We, thus, find no evidence for a difference between the treatments with penalty and bonus. As a first main result of our study, we therefore conclude that the intrinsic motivation to report truthfully is crowded out by introducing audits with penalties/bonuses, yet with zero detection probability. Consequently, we find support for hypothesis 1b.

Result 1: *Audits with penalties or bonuses, yet with a detection probability of zero, crowd out intrinsic motivation to report truthfully and lead to significantly lower compliance than without any audits.*

Increasing the detection probability. In a second step, we study how misreporting is influenced by an increased detection probability. In Figures 2 and 3, we see that both the share of misreports and the average reported die roll decrease with higher detection probabilities in our Penalty and Bonus treatments. This suggests that a higher detection probability increases the extrinsic motivation to report truthfully (i.e., crowding in of extrinsic motivation), in support of hypothesis 2. However, the difference is rather small (or even non-existent) between detection probabilities of 0% and 30%. A strong effect is only observed with a detection probability of 70%. Only audits with a relatively high detection probability might ensure that misreporting is below the level observed in the No Audit benchmark treatment.

These impressions are supported by the regression analyses reported in Table 3. We find that misreporting and the reported die roll are significantly higher in the treatments with Detection 30 than in the No Audit treatment (Detection 30 dummy). Moreover, the differences between Detection 30 and Detection 0 are not significant (all p-values of Wald tests, conducted after each regression to check whether coefficients differ significantly, are greater than 0.1). Consequently, in case of low probabilities, the extrinsic motivation to report truthfully induced by the monetary incentives does not compensate the initially observed crowding out of intrinsic motivation. This holds for both the Penalty and the Bonus treatments (as the insignificant dummy Detection 30 X Bonus does not provide evidence for a difference in compliance) and provides further support for hypothesis 1b in case of a low detection probability.

Result 2: *Compliance with the request to report truthfully is significantly lower in the treatments with a detection probability of 30% than without any audits. The compliance differences between detection probabilities 0% and 30% are not significant. A detection probability of 30% does not compensate the crowding out of intrinsic motivation induced by audits with penalties or bonuses.*

Yet, with a detection probability of 70% the situation changes. For both incentive schemes, bonus and penalty, the regressions in Table 3 find a significantly lower share of misreports and a significantly lower reported die roll in case of a detection probability of 70% than with 30% and 0% (all p-values of Wald tests are below 0.01). In case of penalty the share of misreports is significantly lower (and the reported die roll is significantly lower) than in the No Audit treatment (coefficients of the dummy Detection 70: model 1: -1.37, model 2: -0.76). Thus, in Detection 70 Penalty the incentives to report truthfully seem strong enough to *overcompensate* the drop in truthful reporting caused by audits. In the bonus setting, the effect is weaker. The interaction dummy Detection 70 X Bonus is significant, but with the opposite sign (model 1: 0.87, model 2: 0.52) and thereby weakens the effect detected for the Detection 70 dummy. Consequently, in Detection 70 Bonus, we find no significant difference compared to the No Audit treatment (p-values of Wald tests are above 0.1).[6] For a high detection probability, we therefore find support for hypothesis

[6] The difference between Detection 70 Bonus and Detection 0 Bonus is significant (p-values of Wald tests are below 0.01).

3 that penalties are more effective in promoting compliance than bonuses. Our results can be summarized as follows:

Result 3: *Audits with a detection probability of 70% yield for both incentives schemes, bonus and penalty, significantly higher compliance with the request to report truthfully than 30% and 0%. In case of penalty with a detection probability of 70% compliance is significantly higher than without any audits and overcompensates the drop in compliance caused by audits. For bonuses, this is not the case.*

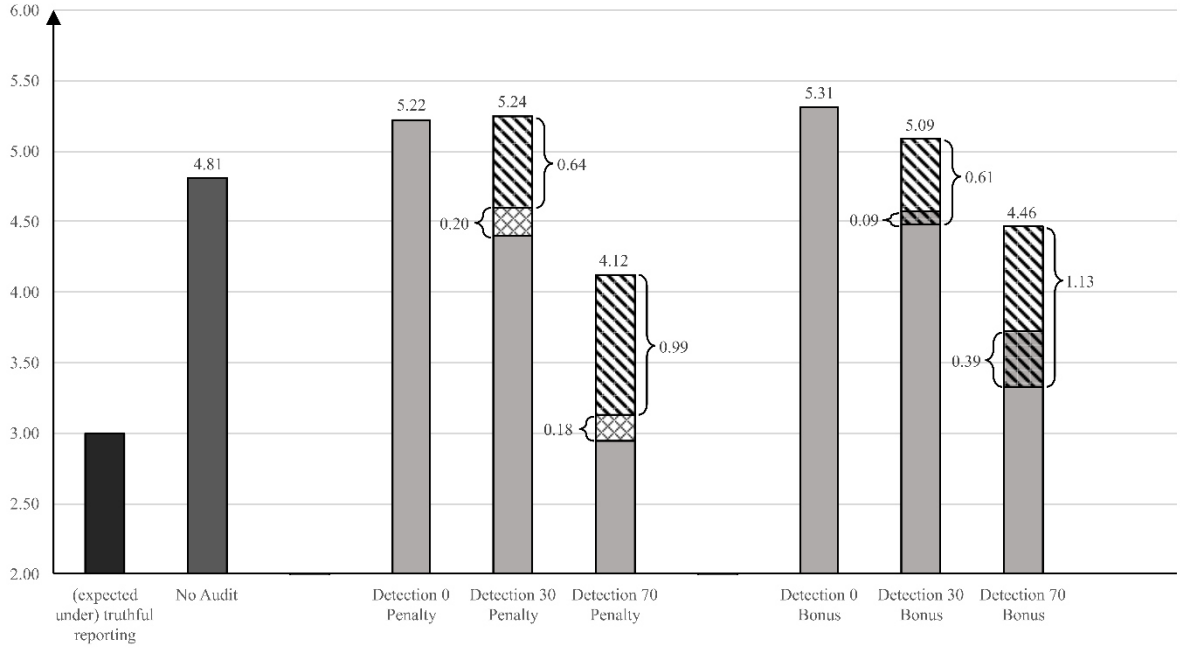
As an important robustness check for our finding that with a detection probability of 70% truthful reporting increases compared to No Audit, we restrict the comparison between No Audit and the Detection 70 treatments to the observed die rolls 1 to 3, when participants' expected payoff is maximized by reporting a 6 (see Table A2.2 in our Appendix A2). This regression corroborates the finding reported above and thus shows that it is not an artifact of high die rolls 4 and 5, for which – in Detection 70 – compliance maximizes expected payoffs.

2.5.2 Profitability of audit systems

So far we have studied the effects of the different systems on fostering compliance. Yet, to assess an audit system's profitability a firm has to weigh the potential benefit from increased compliance against the costs implementing the system. Obviously, both studied audit systems differ in their costs: Whereas bonus payments for compliant behavior lead to additional costs for the firm, penalty payments for non-compliance are in-payments (either direct or indirect by not awarding an otherwise promised payment increase). Additionally, implementing a higher detection probability will most likely incur higher costs, e.g., by hiring more compliance officers.

Participants' payoffs. To assess the profitability of an audit system from a firm's perspective, we first look at participants' payoffs, as they resemble the firms' payments to their workers. Figure 4 shows participants' mean payoffs (full colored bars), mean correction amounts (hatched bars), mean penalty payments (cross-hatched bars) and mean bonus payments (full colored and hatched bars). The full height of a bar shows the hypothetical mean payoff (= mean reported die roll, as shown in Figure 2) in case there were no correction and no penalty/bonus. In the penalty treatments reductions due to corrections and penalty apply. In the bonus treatments, reductions by correction (in case

of a misreport) and additions by bonuses (in case of truthful reporting) apply.[7] Note that the expected payoff under truthful reporting is 3 as we do not consider the data of participants with an observed die roll of 6.



Note: Figure shows the mean payoffs earned by participants (full colored bars), mean correction amount (hatched bars), mean penalty payments (cross-hatched bars) and mean bonus payments (full colored and hatched bars).

Figure 4: Mean payoff per participant

The regression analysis (model 1) in Table 4 shows that, compared to No Audit, participant payoffs are significantly higher in both Detection 0 treatments. Although the level of misreports in Detection 30 is similar to that in Detection 0, audits yield significantly lower payoffs in Detection 30 than in No Audit, yet without a difference between Detection 30 Penalty and Detection 30 Bonus. In the treatments with a detection probability of 70%, the effects of audits are strongest (average payoffs decrease to 2.95 in Detection 70 Penalty and 3.72 in Detection 70 Bonus). Participants' average payoff in Detection 70 Penalty is significantly lower than participants' average payoff in Detection 70 Bonus and close to participants' payoff under truthful reporting.

[7] Suppose for example that a participant observed a die roll of 4 and reported a 6. The implementation of audits with a penalty has two effects: (1) If detected, the misreporting is corrected from 6 to 4 (correction amount = 2) and (2) the participant has to pay a penalty of 1. This results in a participant's payoff of 3. If not detected, the participant has a payoff of 6. In case of audits with a bonus and a detected misreport, the same correction amount of 2 appears, but no penalty. The participant's payoff is therefore 4. If the participant reported truthfully (observed and reported die roll of 4), his/her payoff is 4 under the penalty mechanism as well as under bonus if compliance is not detected. Payoff equals 5 if compliance is detected and the bonus of 1 is paid.

Table 4: Regression results: Participant's payoff and efficiency

	All treatments (except NCR treatments)	
Dependent variable	**Participant's payoff**	**Efficiency (without penalties as direct in-payments)**
	Model 1	Model 2
Reference group	No Audit	No Audit
Detection 0	0.39*	-0.13*
	(0.21)	(0.07)
Detection 30	-0.39*	0.06
	(0.21)	(0.07)
Detection 70	-1.78***	0.52***
	(0.21)	(0.07)
Detection 0 X Bonus	0.11	-0.04
	(0.23)	(0.08)
Detection 30 X Bonus	0.21	-0.00
	(0.23)	(0.08)
Detection 70 X Bonus	0.80***	-0.26***
	(0.23)	(0.08)
Constant	4.44***	0.54***
	(0.43)	(0.15)
Controls	Yes	Yes
No. of observations	645	645
Pseudo R-squared	0.3194	0.3006

Note: This table reports OLS regression results (regression coefficients, standard errors in parentheses, reference group: No Audit treatment). In model 1, the dependent variable is participant's payoff (i.e., after audit, potential penalty and bonus). In model 2, the dependent variable is efficiency. Controls are not reported here, but can be found in Table A2.3 in Appendix A2.

Firm expenditures. Now change perspective and look at participant payoffs as firm expenditures. In No Audit, a firm on average pays 4.81 to its workers, which is by 1.81 higher than expected under truthful reporting (see Figure 4). Compared to No Audit, a firm that considers implementing an audit system with detection probability 30% would expect to reduce its expenditures by 0.21 in Detection 30 Penalty (when penalty payments are not added to the firm's account and by 0.41 when they are added)[8] and by 0.24 in Detection 30 Bonus.[9] When the detection probability is increased to 70%, the expenditure differences to No Audit are even higher: 1.68 for penalty (when penalty payments are not added to the firm's account and 1.86 when they are added) and 1.09 for bonus.[10]

[8] As penalties may not always come as direct in-payments, but rather as forgoing any additional payments, we distinguish between both cases by considering both adding and not adding penalty payments to the firm's account.
[9] Detection 30 Penalty: 0.21 = -0.43 (difference in reported die roll between No Audit (4.81) and Detection 30 Penalty (5.24)) + 0.64 (correction); Detection 30 Bonus: 0.24 = -0.28 (difference in reported die roll between No Audit (4.81) and Detection 30 Bonus (5.09) + 0.61 (correction) – 0.09 (bonus).
[10] Detection 70 Penalty: 1.68 = 0.69 (difference in reported die roll between No Audit (4.81) and Detection 70 Penalty (4.12)) + 0.99 (correction); Detection 70 Bonus: 1.09 = 0.35 (difference in reported die roll between No Audit (4.81) and Detection 70 Bonus (4.46) + 1.13 (correction) – 0.39 (bonus).

So far, we considered workers' payoffs as the only factor in firms' expenditures, yet in real world settings firms will also incur costs by running the audit system. Our results suggest that the advantage of raising the detection probability increases over-proportionally. Moving from a detection probability of 30% to 70% increases the detection probability by factor 2.33, while the firms' expenditure decreases by factor 8 (from 0.21 to 1.68) in case of penalty (when penalty payments are not added to the firm's account and factor 4.54 when they are added) and 4.74 (from 0.24 to 1.09) in case of bonus. Thus, if audit running costs increase linearly in the detection probability, our results provide evidence that it might be beneficial to increase the probability, because expenditures decrease (and consequently the advantage of audits increases) convexly.

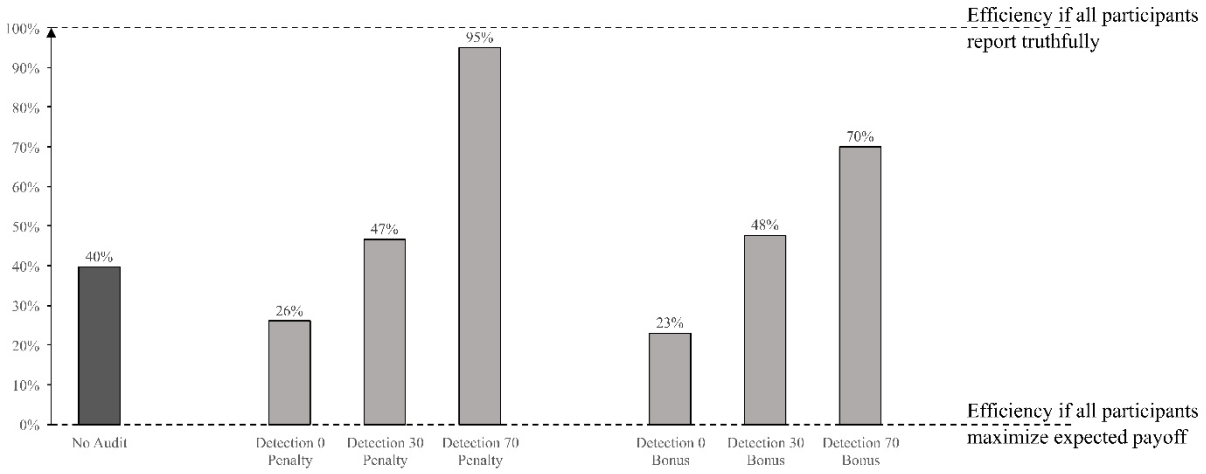
Result 4: *In the treatments with a detection probability of 70% participants' average payoffs are lowest. In Detection 70 Penalty it is significantly lower than in Detection 70 Bonus and close to participants' payoff under truthful reporting. The advantage generated by audits increases convexly in the detection probability.*

Efficiency. A different approach to look at the profitability of audit systems is – instead of comparing them to the No Audit system – to contrast them to two hypothetical benchmarks. One benchmark T is the "ideal case" that all workers report truthfully and the other benchmark M is the "worst case" that all workers report to maximize their individual monetary payoff. In our experimental setting we have $T = 3$ and $M = 6$ in No Audit, Detection 0, and Detection 30 and, $M = 5.4$ in Detection 70.[11] Figure 5 visualizes how close each of the audit systems comes to the benchmark of truthful reporting. The underlying calculation comprises the reported die roll *x*, possible corrections *c*, possible bonus payments *b,* but takes a conservative approach by not adding penalty payments as direct in-payments[12]. Furthermore, it neglects possible audit running costs (see above). The values are normalized such that 0% resembles the case that all participants maximize their expected payoff and 100% resembles the case that all participants report truthfully.

[11] In the treatments No Audit, Detection 0 and 30, the participant's payoff is always maximized by reporting a 6. The same is true for observed die rolls 1, 2 and 3 in the Detection 70 treatments (see section 2.4.3). For die rolls 4 and 5, however, payoff maximization demands truthful reporting. Thus, in 60% of all cases payoff maximization yields a report of 6, in 20% a report of 4 and in 20% a report of 5 resulting in an average report of 5.4 under payoff maximization.

[12] Results are robust when penalties are handled as direct in-payments. Efficiency levels in treatments Detection 30/70 Penalty would be even (slightly) higher.

Thus, our efficiency measure is defined as $\frac{M-(x+b-c)}{M-T}$, which simplifies to $\frac{M-x}{M-T}$ in case of No Audit, Detection 0 Penalty and Detection 0 Bonus.



Note: Figure visualizes for each of the treatments the efficiency by showing how close the firm comes to the benchmark of full compliance. 0% resembles the case if all participants maximize their expected payoff and 100% resembles the case if all participants report truthfully.

Figure 5: Efficiency

Figure 5 shows that No Audit already reaches 40% efficiency. This again demonstrates that there is a considerable level of intrinsic motivation to report truthfully. Efficiency declines to 26% (23%) under Detection 0 in case of penalties (bonuses) and rises to 47% (48%) under Detection 30. Remarkably, Detection 70 Penalty reaches 95% efficiency and clearly differs from Detection 70 Bonus with 70%. Although both Detection 70 treatments increase compliance, the costs of bonus payments make the bonus treatment less efficient. The regression results (model 2) in Table 4 corroborate the impressions from Figure 5.

Result 5: *In case of detection probability 70%, penalties are more efficient than bonuses. Efficiency is highest in Detection 70 Penalty and reaches 95%.*

2.5.3 Analysis of belief elicitation

After a participant reported the die roll, we asked her/him to imagine that 60 other participants in this experiment had seen the same die roll as she/he did and asked for the participant's belief about how many of the others had reported a 1, 2, …, or 6. From this data, we calculated for each participant the belief about the share of misreports by others ("belief share of misreports"). Figure 6 shows the mean belief share of misreports in comparison to the actually observed mean share of misreports for each treatment.

Comparing the belief to the actually observed share of misreports, we find no significant differences in treatment No Audit and in the two treatments Detection 70

Penalty/Bonus (one sample t-test[13], all p-values greater than 0.1, two-tailed). Yet, remarkably, in the Detection 0 and 30 treatments, we observe that the belief share of misreports is always significantly lower than the actually observed share of misreports (all p-values below 0.05). This means that the lower compliance in Detection 0 and 30 compared to No Audit is not anticipated.

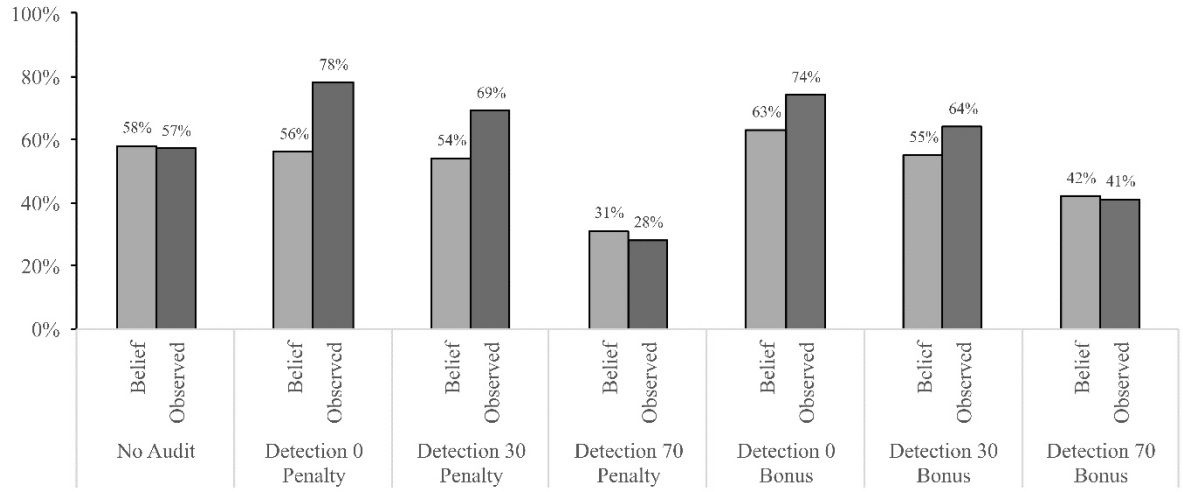


Figure 6: Mean belief share of misreports and actually observed mean share of misreports

Table 5 presents regression analyses with belief share of misreports as dependent variable. We observe no significant belief differences between No Audit and Detection 0 and 30, but a significant decrease in beliefs in Detection 70 (model 1). This supports the finding that compared to No Audit higher compliance in Detection 70 is anticipated, but lower compliance in Detection 0 and 30 is not. Remarkably, we observe significant and positive coefficients of the interaction terms in treatments Detection 0 and Detection 70 indicating that the belief is higher in the Bonus treatments than in the Penalty treatments. This means that participants believe that there is less truthful reporting in the bonus systems than in the respective penalty systems. This is supported by model 2 (with Penalty and Bonus treatments only) in which we incorporated a dummy variable for the Bonus treatments.

[13] Please note that a one sample t-test is applied because on an individual level the belief share of misreports is an interval variable (between 0% and 100%) and misreport a binary variable (0 no misreport, 1 misreport). Therefore, we test for each treatment whether the belief share of misreports equals the observed mean share of misreports (e.g., 0.58 in No Audit treatment).

Result 6: *While beliefs on misreports match actual behavior in No Audit and Detection 70, participants fail to anticipate the lower compliance in Detection 0 and Detection 30 compared to the benchmark No Audit.*

Figure A3.1 in Appendix A3 compares the beliefs of participants who reported truthfully and participants who misreported. In each treatment, we find a much higher belief for the latter group. This indicates that non-compliant individuals expect more non-compliance by others than compliant individuals do. This result is also supported by the regression analysis reported in Table 5 (model 3). The dummy variable misreport (that takes the value 1 in case of a misreport) is positive and highly significant.

Table 5: Regression results: Belief elicitation

	All treatments (except NCR treatments)	**Only Penalty and Bonus treatments**	**All treatments (except NCR treatments)**
Dependent variable	**Belief**	**Belief**	**Belief**
	Model 1	Model 2	Model 3
Reference group	No Audit	Penalty	No Audit
Detection 0	-0.02		
	(0.03)		
Detection 30	-0.03		
	(0.04)		
Detection 70	-0.25***		
	(0.04)		
Detection 0 X Bonus	0.07*		
	(0.04)		
Detection 30 X Bonus	0.01		
	(0.04)		
Detection 70 X Bonus	0.11***		
	(0.04)		
Dummy Bonus treatments		0.07***	
		(0.02)	
Dummy Misreport			0.32***
			(0.02)
Constant	0.84***	0.67***	0.37***
	(0.07)	(0.08)	(0.07)
Controls	Yes	Yes	Yes
No. of observations	645	510	645
Pseudo R-squared	0.2607	0.2063	0.3951

Note: This table reports OLS regression results with belief share of misreports as dependent variable (regression coefficients, standard errors in parentheses, reference group: No Audit treatment in models 1 and 3 and Penalty treatments in model 2). Controls and constants are not reported here, but can be found in Table A2.4 in Appendix A2.

2.5.4 Compliance request premium

In the treatments studied so far, participants were explicitly asked to report the seen die roll, thus to report truthfully ("compliance request"). Potential monetary incentives to comply come on top. To study whether or not this compliance request has any effect on its own, we compare these treatments to the NCR treatments, which yield identical monetary incentives, but do not formulate a compliance request. If the compliance request had no effect on its own, we should not observe any differences between these treatment groups. Figures 7 and 8 show the average reported die roll and share of misreports, respectively, for the treatments with compliance request (CR) and without compliance request (NCR).

Our data suggests that the share of misreports as well as the reported die roll is significantly lower for treatments with compliance request (59%; 4.89) than without (68%; 5.26) (chi^2-test, p = 0.004, two-tailed; Mann-Whitney U-test, p = 0.003, two-tailed). This result remains robust if we exclude treatment No Audit or if we only focus on treatments with a detection probability greater than 0%. This finding is also supported by corresponding regression analyses (see Table 6). In all models, our dummy variable "Compliance request" is negative and significant at the 1%-level. Consequently, participants misreport less frequently and report lower die roll numbers in treatments with a compliance request than without. In particular, this means that although the introduction of audits seems to crowd out intrinsic motivation to report truthfully, it is not crowded out entirely and may be fostered by a non-monetary compliance request. The insignificant coefficient of the interaction dummy "Compliance request X Bonus" does not provide evidence for a compliance request difference between the penalty and bonus system.[14]

Result 7: *For both incentives schemes, bonus and penalty, a compliance request significantly increases compliance.*

[14] All results hold when we restrict the analysis to observed die rolls 1 to 3 (see Table A2.6 in Appendix A2).

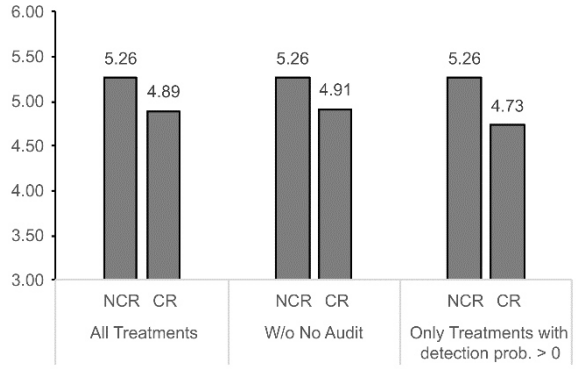


Figure 7: Average reported die roll: Compliance request

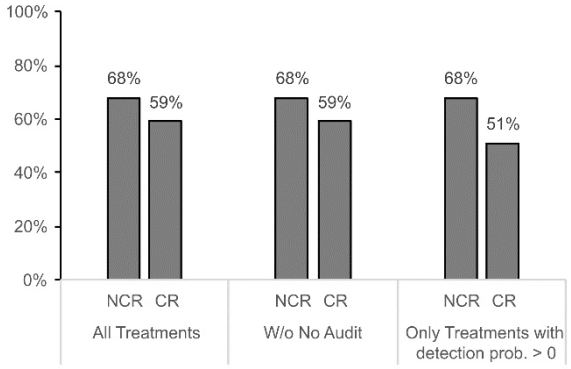


Figure 8: Share of misreports: Compliance request

Table 6: Regression results: Compliance request (reference group: NCR treatments)

	All treatments		**Without No Audit**		**Only TRs with detection probability > 0**	
Dependent variable	**Misreport**	**Reported die roll**	**Misreport**	**Reported die roll**	**Misreport**	**Reported die roll**
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Compliance request	-0.72***	-0.63***	-0.67***	-0.54***	-1.16***	-0.84***
	(0.18)	(0.15)	(0.20)	(0.17)	(0.23)	(0.19)
Compliance request X Bonus	0.20	0.23	0.15	0.13	0.31	0.19
	(0.19)	(0.16)	(0.21)	(0.18)	(0.26)	(0.21)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	982	982	847	847	677	677
Pseudo R-squared	0.1867	0.0645	0.1891	0.0663	0.2432	0.0730

Note: This table reports regression results (regression coefficients, standard errors in parentheses, reference group: NCR treatments). In models 1, 3, and 5, the dependent variable is our binary variable misreport and we use logistic regressions. In models 2, 4, and 6, the dependent variable is the reported die roll and we use ordered logistic regressions. The dummy variable "Compliance request" indicates whether the decision was made in a treatment with compliance request. Controls and constants are not reported here, but can be found in Table A2.5 in Appendix A2.

2.6 Summary and conclusion

Compliance with environmental protection, socially responsible production and customer data protection, for example, but also forgoing corruption, money-laundering and bribery has become a vital topic for firms. Prominent recent examples have shown that transgressions in these areas may seriously harm essential success factors, like the firm's reputation, customer confidence, or customer trust (think e.g. of Boeing, Volkswagen, or Facebook). Against this background, many firms are eager to implement means to foster compliance. Yet, promoting compliance in firm's actions has many facets and potentially requires a change in action, both of the management and the workforce.

This is particularly demanding when the requested change in action is in conflict with individual monetary incentives, because it may risk the own and/or the peers' salary or even their jobs.

What are appropriate means for firms to foster compliance of their workforce? We address this question in a controlled laboratory experiment. In a nutshell, we picture the situation of an individual equipped with valuable information (e.g., a worker knowing details about the production process, relevant for firms' compliance agenda) that is asked to truthfully report this information (e.g., to the management), yet has individual monetary incentives to misreport. In our baseline scenario without any monetary incentives for truth-telling, we find that 43% of participant report truthfully. Hence, in line with previous literature, we observe a high fraction of participants reporting truthfully. Remarkably, in a payoff-equivalent situation with an audit system to detect misreporting, but with a zero detection probability, truthful reporting drops to 22% (in case of a penalty for misreporting) and 26% (in case of a bonus for truthful reporting). Thus, in situations in which the detection of misreporting is extremely low, either because it is technically difficult to find out the truth or because there is no effective detection system, our findings suggest to refrain from implementing an audit system at all and take advantage of the intrinsic motivation to report truthfully. One potential explanation for the observed crowding out of intrinsic motivation might be that introducing audits increases the salience of the possibility of misreporting.

Increasing the detection probability of misreporting increases the extrinsic motivation to report truthfully, which is also detected empirically. Yet, high detection probabilities (70% in our experiment) and a penalty system are needed to make these systems effective. The bonus system fails to increase compliance compared to a no audit system even in case of a high detection probability. The higher efficiency of penalty systems might even amplify in repeated interactions (instead of one-shot decisions as in our study). Then the deterrence effect of penalty systems renders penalties less often necessary, while rewarding the compliers is still and probably increasingly necessary. Our cost-benefit-analysis suggests that systems with high detection probabilities may be cost-efficient, given that the benefit of audits increases convexly, while the costs of increasing the detection probability is most likely concave or linear.

In control experiments we also show that formulating a "compliance request" has a positive effect on truthful reporting. We find significantly higher compliance when requesting to comply, independent of the other measures implemented (detection

probability and penalty or bonus). This underlines the value of rather soft measures like creating awareness and pleas in addition to monetary incentive schemes.[15]

The analysis of the elicited beliefs shows that individuals do not anticipate the crowding out of intrinsic motivation induced by audits. This finding suggests that decision makers in firms might also not expect that the introduction of audits can have detrimental effects on compliance. In contrast, the increase in compliance triggered by audits with a higher detection probability is indeed anticipated. These findings hold for both incentive schemes. Additionally, we observe that non-compliant individuals expect others to be less compliant than compliant individuals do. The interpretation of this "false consensus" effect evidently leads to the question of causality, despite the fact that beliefs are elicited after reporting decisions have been made. This strong correlation between own behavior and believed behavior of others may be caused by beliefs influencing behavior or behavior influencing beliefs. In any case it suggests that guiding individuals' beliefs, e.g. by highlighting compliant behavior of others may have positive effects.

If, based on our experimental analyses, we were to give advice to firms on how to design effective audit systems to promote compliance ("Trust them, threaten them, or lure them?"), we would stress two important facts. In a situation in which the detection of non-compliance is very difficult and/or very costly and therefore the detection probability is very low, our findings speak against introducing audits, but rather making use of individuals' intrinsic motivation to report truthfully and to follow an explicit compliance request ("trust them"). Yet, when it is (technically) possible to implement a high detection probability, audits which penalize non-compliance outperform no audits ("threaten them"). Directly targeting the non-compliers through punishment has a stronger effect than rewarding the compliers ("lure them") and thereby (at best) indirectly addressing non-compliers.

[15] The higher compliance might be the result of an increase of non-monetary (psychological) costs of misreporting (Erard and Feinstein, 1994, Kirchler, 2007, Dulleck et al., 2016).

References

Abeler, J., D. Nosenzo, and C. Raymond. 2019. Preferences for Truth-Telling. *Econometrica* 87 (4):1115-1153.

Allingham, M. G., and A. Sandmo. 1972. Income tax evasion: A theoretical analysis. *Journal of Public Economics* 1 (3-4):323-338.

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.

———. 2019. What Motivates Tax Compliance? *Journal of Economic Surveys* 33 (2):353-388.

Alm, J., T. L. Cherry, M. Jones, and M. McKee. 2012. Social programs as positive inducements for tax participation. *Journal of Economic Behavior & Organization* 84 (1):85-96.

Andreoni, J. 1990. Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. *The Economic Journal* 100 (401):464-477.

Andreoni, J., B. Erard, and J. Feinstein. 1998. Tax compliance. *Journal of Economic Literature* 36 (2):818-860.

Andreoni, J., W. Harbaugh, and L. Vesterlund. 2003. The Carrot or the Stick: Rewards, Punishments, and Cooperation. *The American Economic Review* 93 (3):893-902.

Baeker, A., and M. Mechtel. 2015. Peer settings induce cheating on task performance: IAAEU Discussion Paper Series in Economics.

Balliet, D., L. B. Mulder, and P. A. M. Van Lange. 2011. Reward, punishment, and cooperation: A meta-analysis. *Psychological Bulletin* 137 (4):594-615.

Bazart, C., and M. Pickhardt. 2011. Fighting Income Tax Evasion with Positive Rewards. *Public Finance Review* 39 (1):124-149.

Becker, G. S. 1968. Crime and Punishment: An Economic Approach. *Journal of Political Economy* 76 (2):169-217.

Bijleveld, E., R. Custers, and H. Aarts. 2012. Adaptive reward pursuit: How effort requirements affect unconscious reward responses and conscious reward decisions. *Journal of Experimental Psychology: General* 141 (4):728-742.

Bowles, S., and S. Polania-Reyes. 2012. Economic Incentives and Social Preferences: Substitutes or Complements? *Journal of Economic Literature* 50 (2):368-425.

de Quidt, J., F. Fallucchi, F. Kölle, D. Nosenzo, and S. Quercia. 2017. Bonus versus penalty: How robust are the effects of contract framing? *Journal of the Economic Science Association* 3 (2):174-182.

Dufwenberg, M., and M. A. Dufwenberg. 2018. Lies in disguise – A theoretical analysis of cheating. *Journal of Economic Theory* 175:248-264.

Dulleck, U., J. Fooken, C. Newton, A. Ristl, M. Schaffner, and B. Torgler. 2016. Tax compliance and psychic costs: Behavioral experimental evidence using a physiological marker. *Journal of Public Economics* 134:9-18.

Dwenger, N., H. J. Kleven, I. Rasul, and J. Rincke. 2016. Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany. *American Economic Journal. Economic Policy* 8 (3):203-232.

Erard, B., and J. Feinstein. 1994. The role of moral sentiments and audit perceptions in tax compliance: Carleton University, Department of Economics.

Fabbri, M., P. Nicola Barbieri, and M. Bigoni. 2019. Ride Your Luck! A Field Experiment on Lottery-based Incentives for Compliance. *Management Science* 65 (9):4336-4348.

Fatas, E., D. Nosenzo, M. Sefton, and D. Zizzo. 2015. A Self-Funding Reward Mechanism for Tax Compliance. *SSRN Electronic Journal*.

Fehr, E., and B. Rockenbach. 2003. Detrimental effects of sanctions on human altruism. *Nature* 422 (6928):137-140.

Festinger, L., and J. M. Carlsmith. 1959. Cognitive consequences of forced compliance. *The Journal of Abnormal and Social Psychology* 58 (2):203-210.

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

Fischbacher, U., and F. Föllmi-Heusi. 2013. Lies in Disguise—An Experimental Study on Cheating. *Journal of the European Economic Association* 11 (3):525-547.

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

Frey, B. S., and F. Oberholzer-Gee. 1997. The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding- Out. *The American Economic Review* 87 (4):746-755.

Gamliel, E., and E. Peer. 2013. Explicit risk of getting caught does not affect unethical behavior. *Journal of Applied Social Psychology* 43 (6):1281-1288.

Gneezy, U. 2005. Deception: The role of consequences. *American Economic Review* 95 (1):384-394.

Gneezy, U., A. Kajackaite, and J. Sobel. 2018. Lying Aversion and the Size of the Lie. *American Economic Review* 108 (2):419-453.

Gneezy, U., S. Meier, and P. Rey-Biel. 2011. When and Why Incentives (Don't) Work to Modify Behavior. *Journal of Economic Perspectives* 25 (4):191-210.

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

———. 2000b. Pay Enough or Don't Pay at All*. *The Quarterly Journal of Economics* 115 (3):791-810.

Greiner, B. 2015. Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association* 1 (1):114-125.

Grolleau, G., M. G. Kocher, and A. Sutan. 2016. Cheating and Loss Aversion: Do People Cheat More to Avoid a Loss? *Management Science* 62 (12):3428-3438.

Hofmann, E., E. Hoelzl, and E. Kirchler. 2008. Preconditions of Voluntary Tax Compliance: Knowledge and Evaluation of Taxation, Norms, Fairness and Motivation to Cooperate. *Zeitschrift für Psychologie/Journal of Psychology* 216 (4):209-217.

Hossain, T., and J. A. List. 2012. The Behavioralist Visits the Factory: Increasing Productivity Using Simple Framing Manipulations. *Management Science* 58 (12):2151-2167.

Houser, D., J. A. List, M. Piovesan, A. Samek, and J. Winter. 2016. Dishonesty: From parents to children. *European Economic Review* 82:242-254.

Kahneman, D., and A. Tversky. 1979. Prospect Theory: An analysis of Decision Making Under Risk. *Econometrica* 47 (2):263-291.

Kirchler, E. 2007. *The economic psychology of tax behaviour*: Cambridge University Press.

Kirchler, E., E. Hoelzl, and I. Wahl. 2008. Enforced versus voluntary tax compliance: The "slippery slope" framework. *Journal of Economic Psychology* 29 (2):210-225.

Kocher, M. G., S. Schudy, and L. Spantig. 2018. I Lie? We Lie! Why? Experimental Evidence on a Dishonesty Shift in Groups. *Management Science* 64 (9):3995-4008.

Koessler, A.-K., B. Torgler, L. P. Feld, and B. S. Frey. 2019. Commitment to pay taxes: Results from field and laboratory experiments. *European Economic Review* 115:78-98.

Krupka, E., and R. A. Weber. 2009. The focusing and informational effects of norms on pro-social behavior. *Journal of Economic Psychology* 30 (3):307-320.

Laske, K., S. Saccardo, and U. Gneezy. 2018. Do fines deter unethical behavior? The effect of systematically varying the size and probability of punishment. *SSRN Electronic Journal*.

Liang, H., Y. Xue, and L. Wu. 2013. Ensuring Employees' IT Compliance: Carrot or Stick? *Information Systems Research* 24 (2):279-294.

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

Nosenzo, D., T. Offerman, M. Sefton, and A. van der Veen. 2014. Encouraging Compliance: Bonuses Versus Fines in Inspection Games. *The Journal of Law, Economics, and Organization* 30 (3):623-648.

Nosenzo, D., O. Theo, M. Sefton, and A. van der Veen. 2016. Discretionary Sanctions and Rewards in the Repeated Inspection Game. *Management Science* 62 (2):502-517.

Sefton, M., R. Shupp, and J. M. Walker. 2007. The effect of rewards and sanctions in provision of public goods. *Economic Inquiry* 45 (4):671-690.

Slemrod, J. 2016. Tax Compliance and Enforcement: New Research and Its Policy Implications. In *Ross School of Business Paper No. 1302*.

Sutter, M., S. Haigner, and M. G. Kocher. 2010. Choosing the Carrot or the Stick? Endogenous Institutional Choice in Social Dilemma Situations. *The Review of Economic Studies* 77 (4):1540-1566.

The Economist. 2019a. Big compliance - Rise of the no men. *The Economist*, May 4th 2019.

———. 2019b. Schumpeter - Trust-busting. *The Economist*, June 22nd 2019.

Thielmann, I., and B. E. Hilbig. 2018. Daring dishonesty: On the role of sanctions for (un)ethical behavior. *Journal of Experimental Social Psychology* 79:71-77.

Titmuss, R. 1970. *The Gift Relationship*.

Torgler, B. 2002. Speaking to Theorists and Searching for Facts: Tax Morale and Tax Compliance in Experiments. *Journal of Economic Surveys* 16 (5):657-683.

———. 2003. Beyond punishment: A tax compliance experiment with taxpayers in Costa Rica. *Revista de Análisis Económico* 18 (1):27-56.

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

van de Ven, J., and M. C. Villeval. 2015. Dishonesty under scrutiny. *Journal of the Economic Science Association* 1 (1):86-99.

Wang, L., and J. K. Murnighan. 2017. How Much Does Honesty Cost? Small Bonuses Can Motivate Ethical Behavior. *Management Science* 63 (9):2903-2914.

Appendix

A1 Detailed distribution of reported die roll

Figures A1.1 and A1.2 show the distribution of reported die roll dependent on the seen die roll for each treatment. The presented number in a field denotes the number of participants for the corresponding combination. The color of a field highlights the share of participants who reported the corresponding die roll given the total number of participants with the corresponding die roll seen (i.e., presented number divided by the column sum). The higher this share, the darker the field (see legend). Please note that these figures include participants with a seen die roll of 6 (in contrast to the analyses in section 2.5).

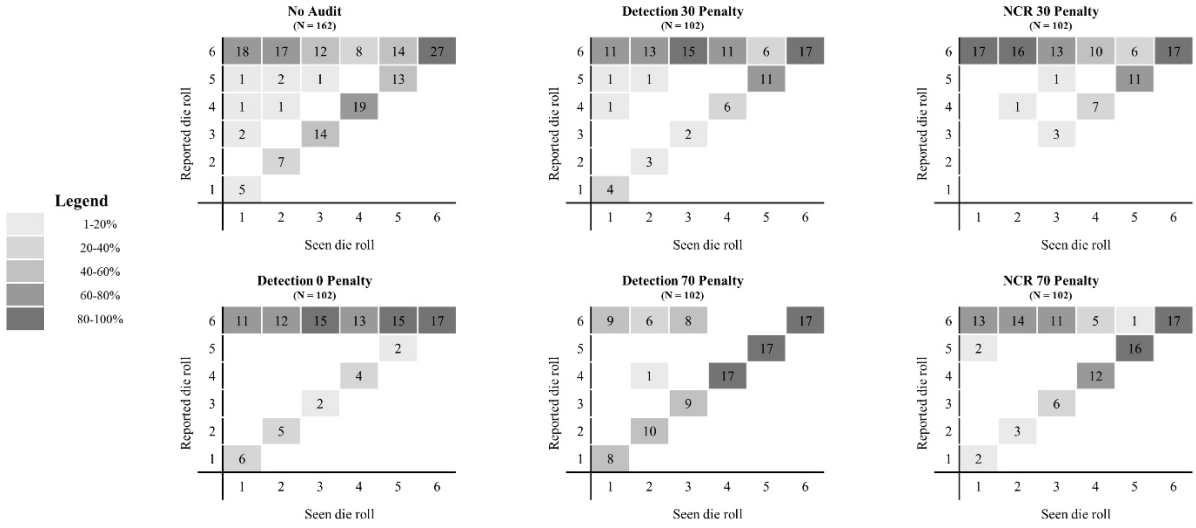


Figure A1.1: Distribution of reported die roll dependent on seen die roll for our benchmark treatment No Audit and all Penalty treatments

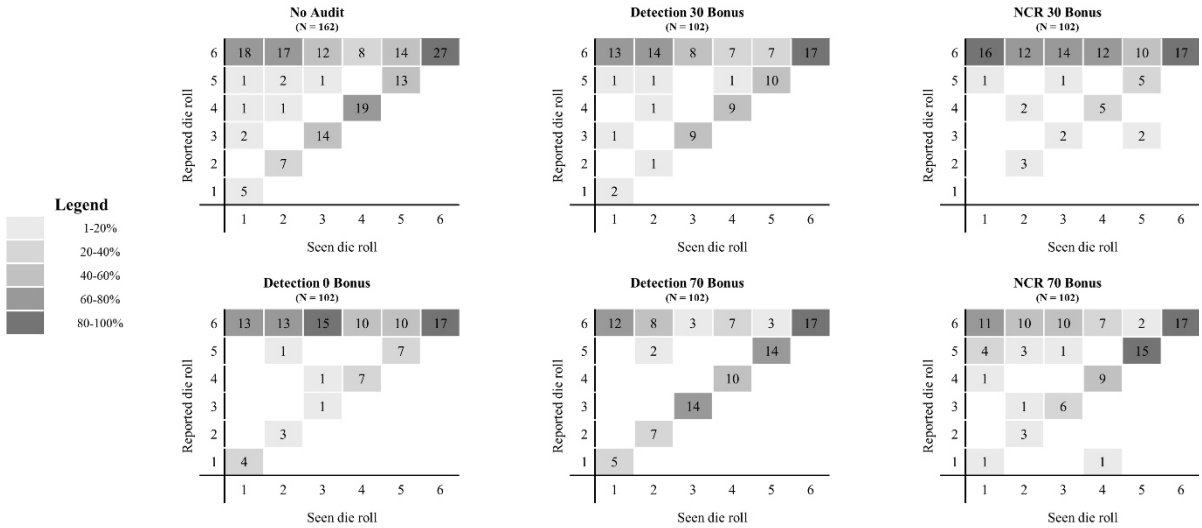


Figure A1.2: Distribution of reported die roll dependent on seen die roll for our benchmark treatment No Audit and all Bonus treatments

Figure A1.3 shows the distribution of reported die roll for each treatment. Please note that participants with a seen die roll of 6 are again not included (following the analyses in section 2.5)

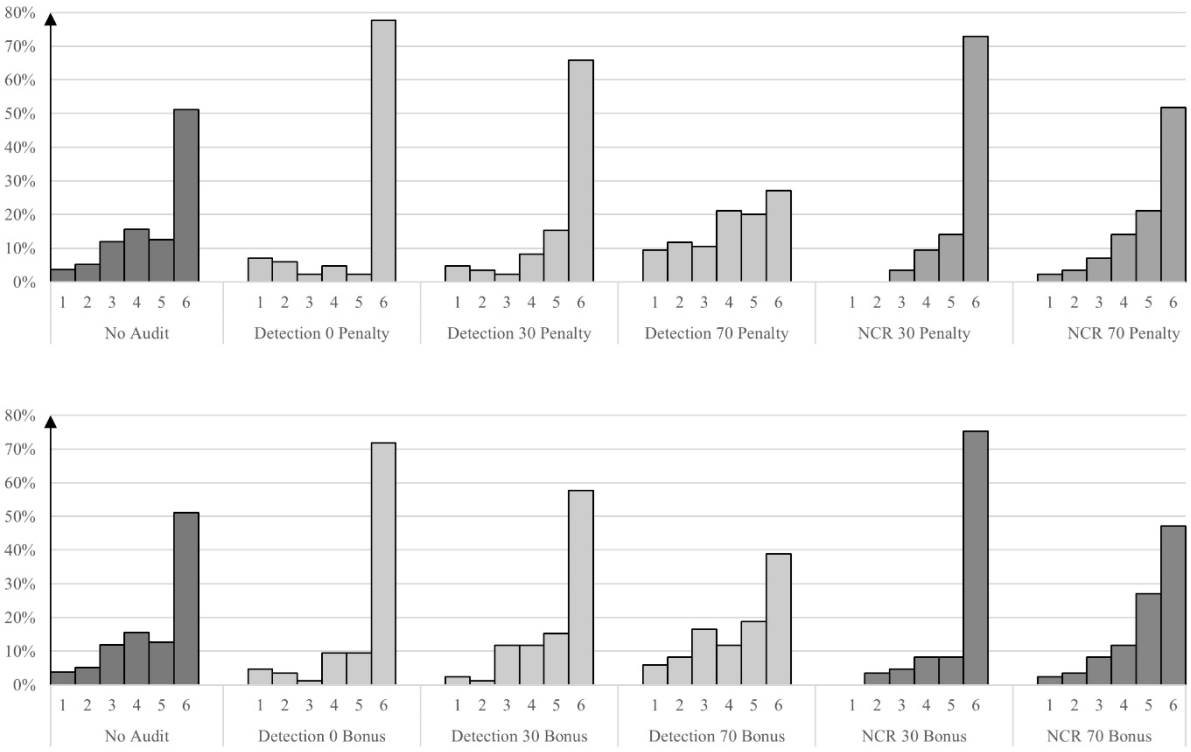


Figure A1.3: Distribution of reported die roll for each treatment

A2 Additional regression results

Table A2.1: Main regression results with constants and controls (corresponds to Table 3)

	All treatments (except NCR treatments)	
Dependent variable	**Misreport**	**Reported die roll**
	Model 1	Model 2
Reference group	No Audit	No Audit
Detection 0	1.19***	1.10***
	(0.36)	(0.32)
Detection 30	0.82**	0.79***
	(0.34)	(0.29)
Detection 70	-1.37***	-0.76***
	(0.35)	(0.26)
Detection 0 X Bonus	-0.22	-0.13
	(0.40)	(0.37)
Detection 30 X Bonus	-0.30	-0.25
	(0.38)	(0.32)
Detection 70 X Bonus	0.87**	0.52*
	(0.39)	(0.28)
Age	-0.05***	-0.04***
	(0.02)	(0.01)
Risk attitude	0.14***	0.10***
	(0.04)	(0.03)
Female	-0.41*	-0.53***
	(0.21)	(0.18)
Morale	-0.22***	-0.17***
	(0.05)	(0.04)
Bachelor	-0.58***	-0.39**
	(0.22)	(0.18)
Economics	0.20	0.21
	(0.21)	(0.17)
Decision complexity	-0.28***	-0.23***
	(0.05)	(0.04)
Monthly income	0.00	0.00
	(0.00)	(0.00)
Die roll seen	-0.56***	0.03
	(0.08)	(0.06)
Political opinion	0.01	-0.01
	(0.08)	(0.07)
Constant cut1		-5.16***
		(0.62)
Constant cut2		-4.29***
		(0.60)
Constant cut3		-3.51***
		(0.59)
Constant cut4		-2.72***
		(0.59)
Constant cut5		-2.00***
		(0.58)
Constant	4.40***	
	(0.73)	
No. of observations	645	645
Pseudo R-squared	0.2568	0.0980

Note: This table shows the same regression results as presented in Table 3, but with all controls and constants displayed (regression coefficients, standard errors in parentheses, reference group: No Audit treatment). In model 1, the dependent variable is our binary variable misreport and we use a logistic regression. In model 2, the dependent variable is the reported die roll and we use an ordered logistic regression.

Table A2.2: Regression results comparing No Audit treatments with Detection 70 treatments for observed die rolls 1 to 3

	Only No Audit and Detection 70 Penalty/Bonus treatments	
Dependent variable	**Misreport**	**Reported die roll**
	Model 1	Model 2
Reference group	No Audit	No Audit
Detection 70	-1.27***	-1.11***
	(0.47)	(0.39)
Detection 70 X Bonus	0.13	0.33
	(0.50)	(0.40)
Controls	Yes	Yes
No. of observations	183	183
Pseudo R-squared	0.2636	0.1189

Note: This table shows regressions comparing the No Audit treatment and the Detection 70 Penalty/Bonus treatments for observed die rolls 1 to 3 (regression coefficients, standard errors in parentheses, controls and constants are not reported, reference group: No Audit treatment). In model 1, the dependent variable is our binary variable misreport and we use a logistic regression. In model 2, the dependent variable is the reported die roll and we use an ordered logistic regression.

Table A2.2 shows no significant interaction effect. For observed die rolls 1 to 3, incentives to comply in both mechanisms seem strong enough to overcompensate the drop in compliance caused by audits. For observed die rolls 1 to 5 (see Table 3) this was only found for the penalty mechanism. However, this disparity does not stand in contrast to our findings formulated in result 3. The difference is mainly driven by the participants with the high observed die rolls 4 and 5. Whereas in Penalty 70 all participants report truthfully (in line with payoff maximization) and consequently compliance increases strongly compared to No Audit (see Figure A1.1), compliance does not change notably in Bonus 70 (see Figure A1.2). This supports our finding that penalties are more effective in promoting compliance than bonuses.

Table A2.3: Regression results: Participant's payoff and efficiency with controls (corresponds to Table 4)

	All treatments (except NCR treatments)	
Dependent variable	**Participant's payoff**	**Efficiency (without penalties as direct in-payments)**
	Model 1	Model 2
Reference group	No Audit	No Audit
Detection 0	0.39*	-0.13*
	(0.21)	(0.07)
Detection 30	-0.39*	0.06
	(0.21)	(0.07)
Detection 70	-1.78***	0.52***
	(0.21)	(0.07)
Detection 0 X Bonus	0.11	-0.04
	(0.23)	(0.08)
Detection 30 X Bonus	0.21	-0.00
	(0.23)	(0.08)
Detection 70 X Bonus	0.80***	-0.26***
	(0.23)	(0.08)
Age	-0.01	0.00
	(0.01)	(0.00)
Risk attitude	0.03	-0.01
	(0.03)	(0.01)
Female	-0.18	0.06
	(0.13)	(0.04)
Morale	-0.07**	0.03***
	(0.03)	(0.01)
Bachelor	-0.06	0.03
	(0.13)	(0.04)
Economics	0.01	-0.00
	(0.13)	(0.04)
Decision complexity	-0.12***	0.04***
	(0.03)	(0.01)
Monthly income	0.00	-0.00
	(0.00)	(0.00)
Die roll seen	0.41***	-0.15***
	(0.04)	(0.01)
Political opinion	-0.02	0.00
	(0.05)	(0.02)
Constant	4.44***	0.54***
	(0.43)	(0.15)
No. of observations	645	645
Pseudo R-squared	0.3194	0.3006

Note: This table shows the same OLS regression results as presented in Table 4, but with all controls displayed (regression coefficients, standard errors in parentheses, reference group: No Audit treatment). In model 1, the dependent variable is participant's payoff (i.e., after audit, potential penalty and bonus). In model 2, the dependent variable is efficiency.

Table A2.4: Regression results on belief elicitation with controls (corresponds to Table 5)

	All treatments (except NCR treatments)	**Only Penalty and Bonus treatments**	**All treatments (except NCR treatments)**
Dependent variable	**Belief**	**Belief**	**Belief**
	Model 1	Model 2	Model 3
Reference group	No Audit	Penalty	No Audit
Detection 0	-0.02		
	(0.03)		
Detection 30	-0.03		
	(0.04)		
Detection 70	-0.25***		
	(0.04)		
Detection 0 X Bonus	0.07*		
	(0.04)		
Detection 30 X Bonus	0.01		
	(0.04)		
Detection 70 X Bonus	0.11***		
	(0.04)		
Dummy Bonus treatments		0.07***	
		(0.02)	
Dummy Misreport			0.32***
			(0.02)
Age	-0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)
Risk attitude	0.01*	0.01**	0.00
	(0.00)	(0.00)	(0.00)
Female	0.01	0.03	0.04**
	(0.02)	(0.02)	(0.02)
Morale	-0.01***	-0.01*	-0.00
	(0.00)	(0.01)	(0.00)
Bachelor	-0.02	0.02	0.01
	(0.02)	(0.03)	(0.02)
Economics	-0.01	0.00	-0.02
	(0.02)	(0.02)	(0.02)
Decision complexity	-0.02***	-0.03***	-0.00
	(0.00)	(0.01)	(0.00)
Monthly income	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)
Die roll seen	-0.07***	-0.08***	-0.04***
	(0.01)	(0.01)	(0.01)
Political opinion	0.01	0.00	0.01
	(0.01)	(0.01)	(0.01)
Constant	0.84***	0.67***	0.37***
	(0.07)	(0.08)	(0.07)
No. of observations	645	510	645
Pseudo R-squared	0.2607	0.2063	0.3951

Note: This table shows the same OLS regression results as presented in Table 5, but with all controls displayed (regression coefficients, standard errors in parentheses, reference group: No Audit treatment in models 1 and 3 and Penalty treatments in model 2).

Table A2.5: Regression results on compliance request with constants and controls (corresponds to Table 6)

	All treatments		**Without No Audit**		**Only TRs with detection probability > 0**	
Dependent variable	**Misreport**	**Reported die roll**	**Misreport**	**Reported die roll**	**Misreport**	**Reported die roll**
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Compliance request	-0.72***	-0.63***	-0.67***	-0.54***	-1.16***	-0.84***
	(0.18)	(0.15)	(0.20)	(0.17)	(0.23)	(0.19)
Compliance request X Bonus	0.20	0.23	0.15	0.13	0.31	0.19
	(0.19)	(0.16)	(0.21)	(0.18)	(0.26)	(0.21)
Age	-0.03**	-0.03***	-0.04**	-0.04***	-0.04**	-0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Risk attitude	0.13***	0.11***	0.13***	0.11***	0.16***	0.14***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)
Female	-0.19	-0.28**	-0.19	-0.27*	-0.28	-0.29*
	(0.16)	(0.14)	(0.18)	(0.15)	(0.20)	(0.17)
Morale	-0.16***	-0.14***	-0.14***	-0.13***	-0.09*	-0.09**
	(0.04)	(0.03)	(0.04)	(0.03)	(0.05)	(0.04)
Bachelor	-0.41**	-0.33**	-0.28	-0.25*	-0.34*	-0.34**
	(0.16)	(0.14)	(0.18)	(0.15)	(0.20)	(0.17)
Economics	0.06	0.12	0.10	0.12	-0.06	-0.01
	(0.16)	(0.14)	(0.17)	(0.15)	(0.20)	(0.17)
Decision complexity	-0.27***	-0.26***	-0.28***	-0.27***	-0.25***	-0.23***
	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)
Monthly income	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Die roll seen	-0.63***	-0.14***	-0.64***	-0.16***	-0.79***	-0.20***
	(0.06)	(0.05)	(0.06)	(0.05)	(0.08)	(0.06)
Political opinion	-0.02	-0.02	-0.04	-0.04	-0.01	0.00
	(0.06)	(0.05)	(0.07)	(0.06)	(0.08)	(0.07)
Constant cut1		-6.03***		-6.17***		-6.09***
		(0.52)		(0.57)		(0.63)
Constant cut2		-5.16***		-5.32***		-5.17***
		(0.51)		(0.55)		(0.61)
Constant cut3		-4.38***		-4.60***		-4.31***
		(0.50)		(0.55)		(0.60)
Constant cut4		-3.61***		-3.85***		-3.53***
		(0.50)		(0.54)		(0.60)
Constant cut5		-2.85***		-3.06***		-2.63***
		(0.49)		(0.53)		(0.59)
Constant	4.75***		4.70***		4.82***	
	(0.59)		(0.63)		(0.73)	
No. of observations	982	982	847	847	677	677
Pseudo R-squared	0.1867	0.0645	0.1891	0.0663	0.2432	0.0730

Note: This table shows the same regression results as presented in Table 6, but with all controls and constants displayed (regression coefficients, standard errors in parentheses, reference group: NCR treatments). In models 1, 3, and 5, the dependent variable is our binary variable misreport and we use logistic regressions. In models 2, 4, and 6, the dependent variable is the reported die roll and we use ordered logistic regressions. The dummy variable "Compliance request" indicates whether the decision was made in a treatment with a compliance request.

Table A2.6: Regression results on compliance request for observed die rolls 1 to 3

	All treatments		**Without No Audit**		**Only TRs with detection probability > 0**	
Dependent variable	**Misreport**	**Reported die roll**	**Misreport**	**Reported die roll**	**Misreport**	**Reported die roll**
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Compliance request	-1.45***	-1.15***	-1.51***	-1.17***	-1.61***	-1.31***
	(0.28)	(0.23)	(0.31)	(0.26)	(0.35)	(0.29)
Compliance request X Bonus	0.11	0.17	0.15	0.17	-0.08	0.09
	(0.26)	(0.23)	(0.28)	(0.25)	(0.35)	(0.30)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	590	590	509	509	407	407
Pseudo R-squared	0.2076	0.1096	0.2110	0.1185	0.2615	0.1335

Note: Regressions shown in Table 6 are rerun for observed die rolls 1 to 3 and this table shows the regression results (regression coefficients, standard errors in parentheses, controls and constants are not reported, reference group: NCR treatments). In models 1, 3, and 5, the dependent variable is our binary variable misreport and we use logistic regressions. In models 2, 4, and 6, the dependent variable is the reported die roll and we use ordered logistic regressions. The dummy variable "Compliance request" indicates whether the decision was made in a treatment with a compliance request.

A3 Additional analysis of belief elicitation

Figure A3.1 compares the beliefs of participants who reported truthfully and participants who misreported.

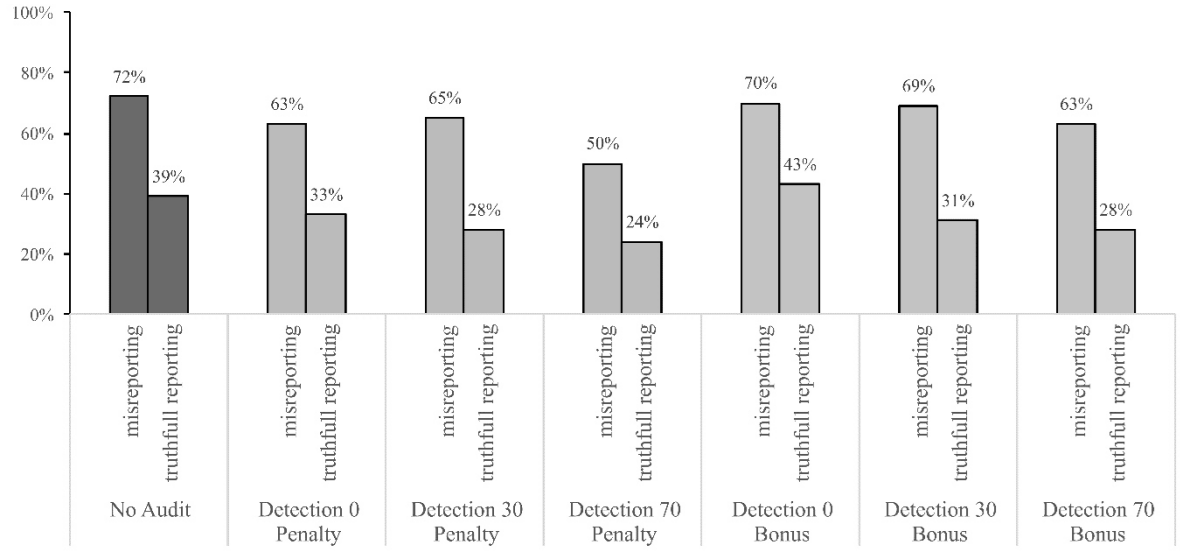
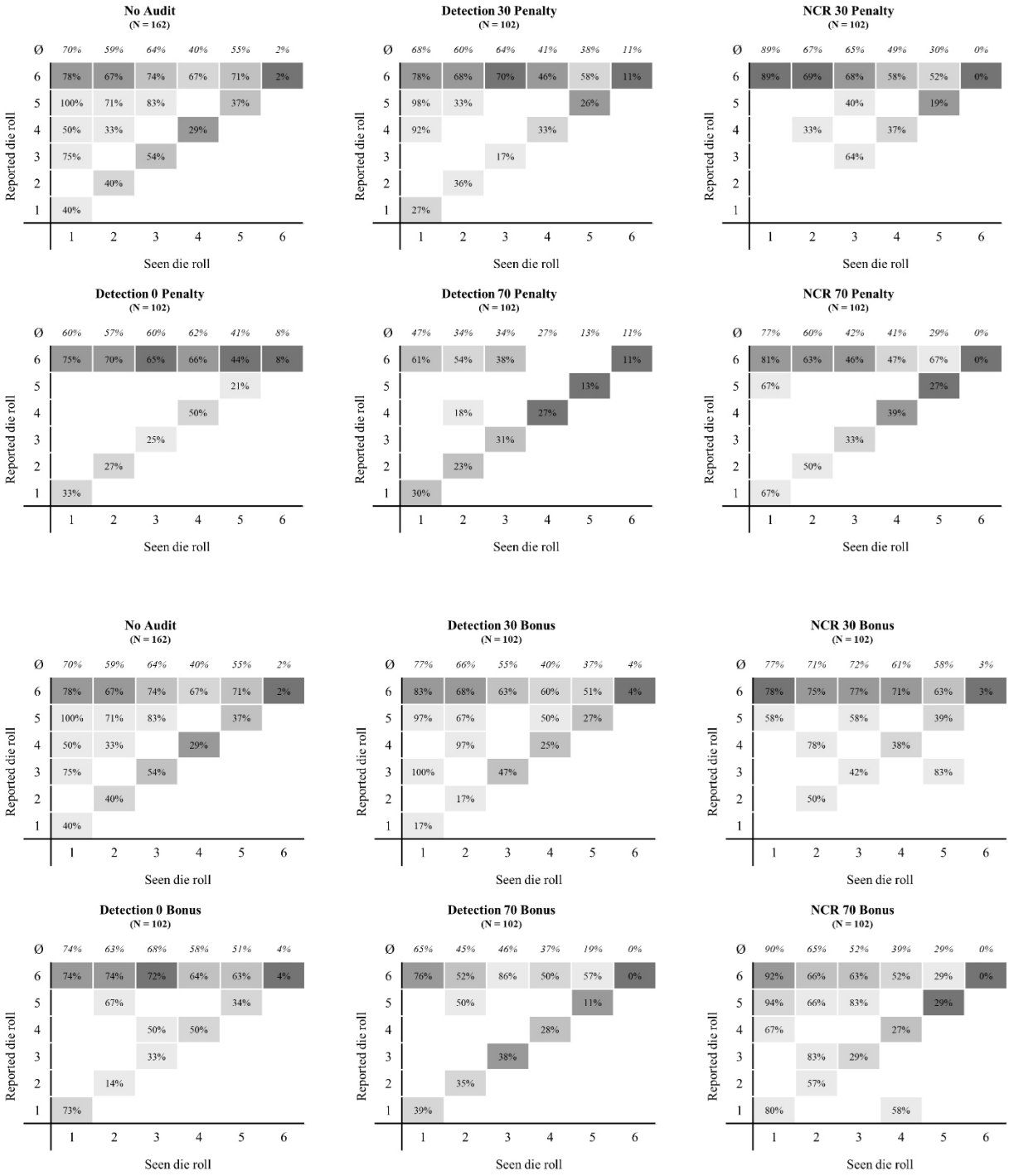


Figure A3.1: Mean belief share of misreports separated by participants who reported truthfully and participants who misreported

Figure A3.2 shows the mean belief of share of misreports by others dependent on seen and reported die roll for each treatment. The presented number in a field denotes the mean belief of share of misreports for the corresponding combination. For example, participants in the No Audit treatment with a seen die roll of 1 and a reported die roll of 6 believe that on average 78% of others – who also see a die roll of 1 – will misreport. Please note that these figures include participants with a seen die roll of 6 (in contrast to the analyses in section 2.5).

On top of each matrix, the mean belief over all participants with the corresponding seen die roll is presented. As expected, we observe the tendency that the belief decreases for higher observed die rolls. This makes sense as participants are asked about compliance behavior of others who see the same die roll as they did. Consequently, the monetary incentive to lie decreases for higher observed die rolls.



Note: The field's color highlights the share of participants who reported the corresponding die roll given the total number of participants with the corresponding die roll seen (see Appendix A1 for more details). Please note that the color reveals no information about the belief, but provides information about the distribution of reported die roll dependent on the seen die roll.

Figure A3.2: Mean belief of share of misreports by others dependent on seen and reported die roll for our benchmark treatment No Audit and all Penalty (upper Panel) and all Bonus (lower Panel) treatments

Supplementary material

M1 Instructions

Supplementary material M1 includes the translated instructions (from German). All participants received the instructions on their computer screen.

M1.1 General instructions

Thank you very much for participating in this experimental study. For your participation, you will receive a lump sum of 4 Euros.

The experimental study consists of one experiment, in which you have the possibility to earn money, and a questionnaire at the end of the study. For answering the questionnaire you will receive a lump sum of 2 Euros. The amount of money you earn in the experiment depends on your decisions in the experiment and on chance. The instructions explain to you how you can influence the amount of money you earn in this study with your decisions.

It is important that you understand the instructions. Hence, please do not hesitate to ask any questions. If you have a question, please raise your hand. We will then come to you to answer your question. Please do not ask your question loudly.

The analysis of the experimental study will be anonymous. We will on no account link your name to the data collected in the experimental study. You will not get to know the identity of any other participant, neither before, nor after the experimental study. Likewise, the other participants will not get to know your identity. At the end of the experimental study, you will have to sign a receipt to confirm the payments you received. This receipt will be used for accounting purposes only.

We would like to point out that you are not allowed to communicate with other participants or leave your seat throughout the whole experimental study. Please make sure to switch off your mobile phone.

At the end of this experimental study, you will receive your payout privately and in cash. Your total payout consists of your payout of the experiment, the lump sum of 2 Euros for answering the questionnaire and the lump sum of 4 Euros for your general participation.

M1.2 Instructions treatment No Audit

At the beginning of this experiment you will see a video with a die roll with one of the following results:

The video with the die roll is chosen randomly from the computer and all 6 videos are drawn with the same probability.

Your task is to remember the number you saw in the video and to enter this number into the computer.

The reported die roll determines your payoff:

Reported die roll	"1"	"2"	"3"	"4"	"5"	"6"
	=	=	=	=	=	=
Euro	1	2	3	4	5	6

Your payoff is therefore:

Payoff = amount of money corresponding to your reported die roll

M1.3 Instructions treatment Detection 0 Penalty

At the beginning of this experiment you will see a video with a die roll with one of the following results:

The video with the die roll is chosen randomly from the computer and all 6 videos are drawn with the same probability.

Your task is to remember the number you saw in the video and to enter this number into the computer.

The reported die roll determines your payoff:

Reported die roll	"1"	"2"	"3"	"4"	"5"	"6"
	=	=	=	=	=	=
Euro	1	2	3	4	5	6

With a certain probability, that will be announced to you shortly, the computer will audit your reported die roll. You cannot influence this process. The computer checks your reported die roll against the seen die roll in the video.

If you get audited and your reported die roll matches with the seen die roll in the video, you will receive the amount of money corresponding to your reported die roll.

If you get audited and your reported die roll does not match with the seen die roll in the video, you will receive the amount of money corresponding to the seen die roll.
Additionally, you have to pay a penalty of 1 Euro.

[Button: Show individual audit probability] (the following text displays)

Your individual audit probability is 0%.

This means, that you will not get audited. If your reported die roll and the seen die roll do not match, you will not experience any negative consequences.

Therefore, your payoff is:

Payoff = amount of money corresponding to your reported die roll

M1.4 Instructions treatment Detection 30/70 Penalty

At the beginning of this experiment you will see a video with a die roll with one of the following results:

The video with the die roll is chosen randomly from the computer and all 6 videos are drawn with the same probability.

Your task is to remember the number you saw in the video and to enter this number into the computer.

The reported die roll determines your payoff:

Reported die roll	"1"	"2"	"3"	"4"	"5"	"6"
	=	=	=	=	=	=
Euro	1	2	3	4	5	6

With a certain probability, that will be announced to you shortly, the computer will audit your reported die roll. You cannot influence this process. The computer checks your reported die roll against the seen die roll in the video.

If you get audited and your reported die roll matches with the seen die roll in the video, you will receive the amount of money corresponding to your reported die roll.

If you get audited and your reported die roll does not match with the seen die roll in the video, you will receive the amount of money corresponding to the seen die roll. Additionally, you have to pay a penalty of 1 Euro.

[Button: Show individual audit probability] (the following text displays)

Your individual audit probability is 30%/70%.

Therefore, your potential payoff is:

"Reported die roll matches seen die roll"

Payoff = amount of money corresponding to your reported die roll

"Reported die roll does not match seen die roll"

Payoff ("no audit", probability of 70%/30%) = amount of money corresponding to your reported die roll

Payoff ("audit", probability of 30%/70%) = amount of money corresponding to the seen die roll – 1 Euro

M1.5 Instructions treatment Detection 0 Bonus

At the beginning of this experiment you will see a video with a die roll with one of the following results:

The video with the die roll is chosen randomly from the computer and all 6 videos are drawn with the same probability.

Your task is to remember the number you saw in the video and to enter this number into the computer.

The reported die roll determines your payoff:

Reported die roll	"1"	"2"	"3"	"4"	"5"	"6"
	=	=	=	=	=	=
Euro	1	2	3	4	5	6

With a certain probability, that will be announced to you shortly, the computer will audit your reported die roll. You cannot influence this process. The computer checks your reported die roll against the seen die roll in the video.

If you get audited and your reported die roll matches with the seen die roll in the video, you will receive the amount of money corresponding to your reported die roll.
Additionally, you get a bonus of 1 Euro.

If you get audited and your reported die roll does not match with the seen die roll in the video, you will receive the amount of money corresponding to the seen die roll.

[Button: Show individual audit probability] (the following text displays)

Your individual audit probability is 0%.

This means, that you will not get audited. If your reported die roll and the seen die roll do not match, you will neither experience negative nor positive consequences.

Therefore, your payoff is:

Payoff = amount of money corresponding to your reported die roll

M1.6 Instructions treatment Detection 30/70 Bonus

At the beginning of this experiment you will see a video with a die roll with one of the following results:

The video with the die roll is chosen randomly from the computer and all 6 videos are drawn with the same probability.

Your task is to remember the number you saw in the video and to enter this number into the computer.

The reported die roll determines your payoff:

Reported die roll	"1"	"2"	"3"	"4"	"5"	"6"
	=	=	=	=	=	=
Euro	1	2	3	4	5	6

With a certain probability, that will be announced to you shortly, the computer will audit your reported die roll. You cannot influence this process. The computer checks your reported die roll against the seen die roll in the video.

If you get audited and your reported die roll matches with the seen die roll in the video, you will receive the amount of money corresponding to your reported die roll.
Additionally, you get a bonus of 1 Euro.

If you get audited and your reported die roll does not match with the seen die roll in the video, you will receive the amount of money corresponding to the seen die roll.

[Button: Show individual audit probability] (the following text displays)

Your individual audit probability is 30%/70%.

Therefore, your potential payoff is:

"Reported die roll matches seen die roll"

Payoff ("no audit", 70%/30% probability) = amount of money corresponding to your reported die roll

Payoff ("audit", 30%/70% probability) = amount of money corresponding to your reported die roll + 1 Euro

"Reported die roll does not match seen die roll"

Payoff ("no audit", probability of 70%/30%) = amount of money corresponding to your reported die roll

Payoff ("audit", probability of 30%/70%) = amount of money corresponding to the seen die roll

M1.7 Instructions treatment NCR 30/70 Penalty

At the beginning of this experiment you will see a video with a die roll with one of the following results:

The video with the die roll is chosen randomly from the computer and all 6 videos are drawn with the same probability.

Your task is to remember the number you saw in the video and to enter this number or any other number into the computer. You are free to type in any number If you decide to report the number you have seen in the video with the die roll, the reported die roll determines your payoff:

Reported die roll	"1"	"2"	"3"	"4"	"5"	"6"
	=	=	=	=	=	=
Euro	1	2	3	4	5	6

If you decide to report a number, that you have not seen in the video with the die roll, you take part of a lottery. Your payoff will be determined by a random process, that you cannot affect.

With a probability of 70%/30% your payoff will be the amount of money corresponding to your reported die roll.

With a probability of 30%/70% your payoff will be the amount of money corresponding to the seen die roll minus 1 Euro.

Therefore, your potential payoff is:

"Reported die roll matches seen die roll"

Payoff = amount of money corresponding to your reported die roll

"Reported die roll does not match seen die roll"

Payoff (probability of 70%/30%) = amount of money corresponding to your reported die roll

Payoff (probability of 30%/70%) = amount of money corresponding to the seen die roll – 1 Euro

M1.8 Instructions treatment NCR 30/70 Bonus

At the beginning of this experiment you will see a video with a die roll with one of the following results:

The video with the die roll is chosen randomly from the computer and all 6 videos are drawn with the same probability.

Your task is to remember the number you saw in the video and to enter this number or any other number into the computer. You are free to type in any number The reported die roll determines your payoff:

Reported die roll	"1"	"2"	"3"	"4"	"5"	"6"
	=	=	=	=	=	=
Euro	1	2	3	4	5	6

If you decide to report the number, that you have seen in the video with the die roll, you take part of a lottery. Your payoff will be determined by a random process, that you cannot affect.

With a probability of 70%/30% your payoff will be the amount of money corresponding to your reported die roll.

With a probability of 30%/70% your payoff will be the amount of money corresponding to your reported die roll plus 1 Euro.

If you decide to report a number, that you have not seen in the video with the die roll, you take again part of a lottery. Again, your payoff will be determined by a random process, that you cannot affect.

With a probability of 70%/30% your payoff will be the amount of money corresponding to your reported die roll.

With a probability of 30%/70% your payoff will be the amount of money corresponding to the seen die roll.

Therefore, your potential payoff is:

"Reported die roll matches seen die roll"

> Payoff (probability of 70%/30%) = amount of money corresponding to your reported die roll
>
> Payoff (probability of 30%/70%) = amount of money corresponding to your reported die roll + 1 Euro

"Reported die roll does not match seen die roll"

Payoff (probability of 70%/30%) = amount of money corresponding to your reported die roll

Payoff (probability of 30%/70%) = amount of money corresponding to the seen die roll

M2 Comprehension tests

M2.1 Comprehension test treatment No Audit/Detection 0 Penalty/Detection 0 Bonus

1. What is your task in this experiment?
   - Reporting the number, that you remembered from the die roll (correct)
   - Reporting any number
   - Reporting a number, that you did not remember from the die roll

2. Imagine you see the following result of the die roll:

And you report a "3".
How many Euros do you get?
[open field] (correct: 3)

3. Imagine you see the following result of the die roll:

And you report a "1".
How many Euros do you get?
[open field] (correct: 1)

4. Imagine you see the following result of the die roll:

And you report a "5".
How many Euros do you get?
[open field] (correct: 5)

M2.2 Comprehension test treatment Detection 30/70 Penalty

1. What is your task in this experiment?
   - Reporting the number, that you remembered from the die roll (correct)
   - Reporting any number
   - Reporting a number, that you did not remember from the die roll

2. Imagine you see the following result of the die roll:

And you report a "3".
How many Euros do you get?
[open field] (correct: 3)

3. Imagine you see the following result of the die roll:

And you report a "1".
How many Euros do you get, if you are audited?
[open field] (correct: 2)
How many Euros do you get, if you are not audited?
[open field] (correct: 1)

4. Imagine you see the following result of the die roll:

And you report a "5".
How many Euros do you get, if you are audited?
[open field] (correct: 2)
How many Euros do you get, if you are not audited?
[open field] (correct: 5)

M2.3 Comprehension test treatment Detection 30/70 Bonus

1. What is your task in this experiment?
   - Reporting the number, that you remembered from the die roll (correct)
   - Reporting any number
   - Reporting a number, that you did not remember from the die roll

2. Imagine you see the following result of the die roll:

And you report a "3".
How many Euros do you get, if you are audited??
[open field] (correct: 4)
How many Euros do you get, if you are not audited?
[open field] (correct: 3)

3. Imagine you see the following result of the die roll:

And you report a "1".
How many Euros do you get, if you are audited??
[open field] (correct: 3)
How many Euros do you get, if you are not audited?
[open field] (correct: 1)

4. Imagine you see the following result of the die roll:

And you report a "5".
How many Euros do you get, if you are audited??
[open field] (correct: 3)
How many Euros do you get, if you are not audited?
[open field] (correct: 5)

M2.4 Comprehension test treatment NCR 30/70 Penalty

1. What is your task in this experiment?
   - Reporting the number, that you remembered from the die roll
   - Reporting any number (correct)
   - Reporting a number, that you did not remember from the die roll

2. Imagine you see the following result of the die roll:

And you report a "3".
How many Euros do you get?
[open field] (correct: 3)

3. Imagine you see the following result of the die roll:

And you report a "1".
How many Euros do you get with a probability of 30%/70%?
[open field] (correct: 2)
How many Euros do you get with a probability of 70%/30%?
[open field] (correct: 1)

4. Imagine you see the following result of the die roll:

And you report a "5".
How many Euros do you get with a probability of 30%/70%?
[open field] (correct: 2)
How many Euros do you get with a probability of 70%/30%?
[open field] (correct: 5)

M2.5 Comprehension test treatment NCR 30/70 Bonus

1. What is your task in this experiment?
   - Reporting the number, that you remembered from the die roll
   - Reporting any number (correct)
   - Reporting a number, that you did not remember from the die roll

2. Imagine you see the following result of the die roll:

And you report a "3".
How many Euros do you get with a probability of 30%/70%?
[open field] (correct: 4)
How many Euros do you get with a probability of 70%/30%?
[open field] (correct: 3)

3. Imagine you see the following result of the die roll:

And you report a "1".
How many Euros do you get with a probability of 30%/70%?
[open field] (correct: 3)
How many Euros do you get with a probability of 70%/30%?
[open field] (correct: 1)

4. Imagine you see the following result of the die roll:

And you report a "5".
How many Euros do you get with a probability of 30%/70%?
[open field] (correct: 3)
How many Euros do you get with a probability of 70%/30%?
[open field] (correct: 5)

CHAPTER 3

Combating overreporting of deductions in tax returns: Prefilling and restricting the deductibility of expenditures[1]

Martin Fochmann, Frank Hechtner, Tobias Kölle, Michael Overesch

Abstract

We experimentally analyze three anti-tax-evasion mechanisms: 1) prefilling of deductions in tax returns, 2) restricting tax evasion opportunities by either disallowing or 3) limiting the deductibility of expenditures. We find that prefilling compared to blank forms reduces tax evasion. Cutting the number of tax evasion opportunities by disallowing the deductibility of expenditure items is an ineffective mechanism to combat tax evasion as individuals shift their tax evasion activities from the disallowed item to other non-restricted items. In contrast, our results suggest that just limiting the deductibility of expenditures avoids this evasion-shift-effect and finally enhances overall tax compliance.

[1] We thank James Alm, Peter N.C. Mohr and Kay Blaufus for helpful comments and suggestions.

Content

Tables

Figures

3.1 Introduction

Tax evasion continues to be a serious problem in society and rising media coverage of evasion scandals heightens the urgency to act. Initiated by the seminal papers of Becker (1968), Allingham and Sandmo (1972) and Yitzhaki (1974), a variety of papers have studied tax evasion and dishonest behavior.[2] Researchers focused primarily on tax evasion of income/earnings although overreporting of deductions might be the only possibility to evade taxes for many people (such as typical wage earners) due to third party income reporting. Consequently, deductions are rather underrepresented in tax compliance literature in view of their importance. In our study, we therefore focus on deductions, in particular on how individuals report expenditures in a tax return. We examine the influence of three mechanisms applied to combat tax evasion behavior: 1) prefilling of deductions in tax returns, 2) restricting tax evasion opportunities by either disallowing the deductibility of expenditures or 3) limiting the deductibility of expenditures. For our purpose, we run an experiment in a controlled environment and analyze expenditure items that are substantial in real life.

Nowadays, taxpayers often start their tax declarations with tax returns in which income and/or deduction items are already prefilled. On the one hand, automatic data exchanges between the tax authority and employers, social insurance agencies and banks allow tax returns to be prefilled (third party reporting). On the other hand, electronic tax declaration programs (e-filing services) usually carry over the previous year's values to the subsequent year (which assists taxpayers at least as an orientation aid) and therefore prefill the current tax return with last year's numbers (e.g., deductions, expenditures, salary, tax credits). A prefilled tax return supports taxpayers to file a legally accurate tax return (Goolsbee, 2006; Klun, 2009; Evans and Tran-Nam, 2010; OECD, 2017). However, literature provides only little and mixed evidence regarding the effects of prefilling on tax compliance (see section 3.2.1). Prefilling of deductions in tax returns is likely to affect the monetary costs associated with tax evasion especially due to a higher (perceived) detection probability of tax fraud in case of third party reporting. As the tax compliance literature already provides robust results that a higher detection probability has a positive impact on tax compliance (Torgler, 2002), we focus on the non-monetary and more psychological consequences of prefilling in our study. We therefore ensure that all

[2] See Andreoni et al. (1998), Torgler (2002), Hofmann et al. (2008), Alm (2012), Slemrod (2016) and Alm (2019) for excellent literature reviews.

monetary aspects such as tax rate, detection probability, and penalties are kept constant. Finally, our data shows that the prefilling of deductions compared to blank forms reduces tax evasion significantly. This highlights the importance of the non-monetary consequences of prefilled tax returns and provides evidence that prefilling of deductions is an effective mechanism to enhance tax compliance.

To combat tax evasion, policy makers frequently restrict tax evasion opportunities – either by disallowing or limiting the deductibility of expenditures. Disallowing the deductibility cuts the number of opportunities to evade taxes. Many countries – like the US, UK, France and Germany – have rather strict rules and consequently disallow the deductibility of, for example, expenditures regarding office space at home, work clothes that are also usable for private purposes, high priced gifts for customers and business clients, and fines. Instead of disallowing the deductibility completely, deductibility of expenditures is sometimes only limited to a certain amount. German tax law for example limits the deductibility of travel expenses, social insurance expenses, food expenses of employees in case of external activities, childcare expenses, and expenses for household-related services. Remarkably, the effects of restricting the deductibility of expenditures on individual tax evasion behavior is unexplored in the literature. However, some recent research in a related context indicates that taxpayers might shift their evasion behavior to compensate for such limitations (Adhikari et al., 2016; Asatryan and Peichl, 2017; Carrillo et al., 2017; Slemrod et al., 2017; Vossler and Gilpatric, 2018).

Our study provides evidence to the literature that cutting the number of tax evasion opportunities by disallowing the deductibility of expenditure items is an ineffective mechanism to combat tax evasion. Our data shows that individuals shift their tax evasion activities from the disallowed item to other non-restricted items (evasion-shift-effect). In contrast, our results suggest that just limiting the deductibility of expenditures avoids this evasion-shift-effect. In this case, tax evasion level of the restricted item and overall tax evasion are reduced. A limited deductibility seems to be an effective mechanism to combat tax evasion.

The remainder of this paper is structured as follows. We briefly discuss the related literature in section 3.2 and develop our hypotheses in section 3.3. In section 3.4 we describe the experimental design and results are presented in section 3.5. Section 3.6 concludes.

3.2 Related literature

3.2.1 Effects of prefilling on tax compliance

Our research question of how prefilling of expenditure items in a tax return influences tax compliance behavior is largely unexplored. Although some papers study prefilling (especially third-party reporting), they mainly focus on the prefilling of income items and show mixed results. A positive effect of third-party reporting is found by Kleven et al. (2011), who analyze data from a tax enforcement experiment in Denmark. Their focus is not on prefilled tax returns directly, but third-party reported data is prefilled by the tax authority in the tax returns. The authors 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. Fochmann et al. (2018) show in a laboratory experiment that a correct prefilling of income items enhances tax compliance compared to a setting without prefilling. In a neutral dice rolling experiment without tax framing and without audit or punishment, Duncan and Li (2018) find that confirmation reports (comparable to correct prefilling) have a positive effect on compliance behavior. However, dishonest behavior cannot be analyzed on an individual level.

In contrast, some studies find no or even a negative effect of prefilling on compliance. 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.

Fonseca and Grimshaw (2017) use an online experiment to study the effects of behavioral nudges on prefilled tax returns. Without nudges, they find that correct prefilling does not increase overall compliance, but that incorrect prefilling reduces compliance. However, this result is mainly driven by the fact that over-compliant participants (i.e., individuals who report a higher taxable income than they actually have and thus pay more taxes), are categorized as non-compliant subjects. In case the tax return is incorrectly prefilled with an income too low (i.e., the prefilled income is lower than the randomly assigned income), they observe that the introduction of a checkbox as a physical barrier to change prefilled fields further decreased compliance, but combining

the checkbox with a fictitious norm message does not influence the overall compliance level.

Bruner et al. (2015) investigate reporting behavior for partly prefilled tax returns and focus on different opportunities for underreporting deductions. In a complex setting, they vary the audit probability, the presence of itemized deductions, and the uncertainty about the correct values. They find that correct as well as incorrect prefilling reduces overall compliance. Gillitzer and Skov (2018) 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 in several dimensions. We use a parsimonious laboratory experiment that enables us to focus on the influence of prefilled expenditure items on compliance behavior in a controlled environment. A laboratory experiment allows us to analyze the level of tax compliance in more detail and excludes that the analysis is biased by undeliberate tax evasion behavior. Our experimental design differs in many ways from that of Bruner et al. (2015), Fonseca and Grimshaw (2017) and Fochmann et al. (2018). We focus on prefilled expenditure items, not mixing it with prefilled income items, in order to clearly focus on the effects for deductions, as previous studies show that reporting behavior may differ for income and deductions (Fochmann and Wolf, 2019). We use a real effort game, so that participants have to earn their income, instead of using a windfall gain. We exclude that uncertainty about the audit probability (as in Fonseca and Grimshaw, 2017) might influence our results.

Our design controls for several potential explanations discussed by the studies mentioned above. Kleven et al. (2011), Gillitzer and Skov (2018) 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 are no computation needs and complexity does not differ between treatments.

3.2.2 Disallowing and limiting the deductibility of expenditures

Disallowing or limiting the deductibility of expenditures (e.g., capping the total amount of expenditures) are frequently discussed topics in the literature. For example, Feldstein (2015) advocates for these mechanisms to tackle rapidly increasing national debt for the United States by restricting the amount of taxes refunded. In 2012, the UK already implemented a single cap on all personal deductions. Expenditures can only be deducted up to an amount of £50,000 or – if greater – 25% of income. This cap stimulated debates on potentially negative effects on for instance charitable donations (Smith, 2012). This is why Schizer (2015) criticizes the idea of a one-size-fits-all cap and suggests to apply different expenditure-specific caps. However, Lowry (2014) estimates that several combinations of deduction limits may shift taxpayers to claim standard deduction instead of itemizing. As a consequence, the expected growth in tax revenues from limiting deductions would be partially offset. All in all, there are multiple dimensions to be considered when limiting expenditures. For example, its effect on income distribution, labor and savings decisions, or planning and administrative costs. However, the effect of limiting expenditures on tax evasion keeps unexplored thus far. It is unclear whether limiting deductions indeed reduces total overdeductions or taxpayers adjust their behavior for example by shifting overdeductions to other (non-restricted) items.

There is some research suggesting that taxpayers might change their behavior to avoid such restrictions. For example, in Chile diesel taxes paid can be fully used as a credit against VAT. However, this is only allowed if diesel is used in industrial activities. Otherwise, if diesel is for example used in freight or public transportation, this rule gets restricted as only a fraction of diesel taxes paid can be claimed as a tax credit for VAT. Agostini and Martínez A (2014) investigated this regulation and show that firms actively manipulate the classification to avoid this restriction. Carrillo et al. (2017) suggest that taxpayers facing third-party reporting of one income item (i.e., tax evasion opportunity gets limited) make offsetting adjustments on other items. In particular, after a policy intervention, Ecuadorian firms increased reported revenue but at the same time also increased reported costs by 96 cents per dollar of revenue adjustment. Such an offsetting-effect is also found by Slemrod et al. (2017) who investigate the response of US sole proprietorships to Form 1099-K that provides the IRS with third-party information about electronic sales. Even though there is, as expected, an increase in reported receipts, taxpayers largely offset this increase “with increased reported expenses, which do not face information reporting, diminishing the impact on reported net taxable income” (p.1).

This finding is supported by Adhikari et al. (2016) who show for the taxicab industry that the increase in receipts due to third-party income reporting of Form 1099-K is offset by an increase in expenses. More evidence is also provided by Asatryan and Peichl (2017) who observe that Armenian firms respond to additional reported income raised by audits with a similar increase in deductions. Vossler and Gilpatric (2018) confirm the offsetting-effect in a controlled laboratory experiment. They show that revealing the item that is targeted in an audit leads individuals to report more truthfully on this item. However, at the same time they evade more on other items.

All in all, these studies do not focus on our research question directly, as the aim of our paper is to study how restricting the deductibility of expenditures impacts tax evasion behavior. However, they do indicate that limiting evasion opportunities might lead taxpayers to adjust their tax evasion behavior to compensate for such restrictions.

3.3 Hypotheses

Tax compliance literature started to focus on how monetary factors such as tax rate, audit probability and fines determine tax compliance behavior (Becker, 1968; Allingham and Sandmo, 1972; Srinivasan, 1973; Yitzhaki, 1974). This literature provides robust evidence that a higher audit/detection probability as well as a higher fine reduces tax evasion (Spicer and Thomas, 1982; Alm et al., 1995; Maciejovsky et al., 2001; Torgler, 2003; Cummings et al., 2009; Fortin et al., 2007; Gërxhani and Schram, 2006). More recently, literature studies how non-monetary and more psychological factors such as social norms, tax morale, fairness concerns, trust and services provided by the tax authority influence tax compliance (Andreoni et al., 1998; Torgler, 2002; Hofmann et al., 2008; Alm, 2012; Alm, 2019).[3] From these findings, new frameworks such as the slippery slope framework (Kirchler et al., 2008) and new paradigms such as the service and trust paradigm (Alm, 2012, Alm, 2019) have evolved. This literature suggests that a higher trust in the tax authority (e.g., by enhancing procedural fairness) and a higher service quality of the tax authority increases tax compliance (Alm, 2019; Hofmann et al., 2008; Kirchler et al., 2008).

[3] Even the theoretical tax compliance literature has already started to consider non-monetary factors to explain/predict tax compliance behavior (e.g., Fortin et al., 2007; Gordon, 1989; Kim, 2003; Myles and Naylor, 1996; Traxler, 2010; Prinz et al., 2014).

3.3.1 Prefilling of deductions

One promising mechanism to enhance tax compliance is prefilling of deductions in tax returns. We argue that prefilling might signal advanced information of the tax authority about the expenditures of an individual. Moreover, individuals might believe that deviating from the prefilled values will increase the probability that the tax authority audits the tax return. Consequently, prefilling of an item should lead to a higher subjective detection probability and might therefore lower tax evasion.

Moreover, prefilling might also increase the non-monetary costs of tax evasion.[4] First, prefilling of tax returns might lead to default effects (Johnson and Goldstein, 2003, Mazar and Hawkins, 2015) or anchoring effects (Tversky and Kahneman, 1974; Epley and Gilovich, 2001; Strack and Mussweiler, 1997; Chapman and Johnson, 1999) that bias individuals toward the prefilled values of the tax return. Correctly prefilled tax returns would then nudge individuals toward more tax compliance. Second, in case of correct prefilling, we argue that the act of replacing correct values with incorrect numbers in order to evade taxes increases the moral costs associated with tax evasion. Third, due to a better service, individuals might perceive a higher procedural fairness when the tax authority prefills tax returns compared to blank forms. Literature provides conclusive evidence that higher procedural fairness is associated with higher tax compliance which can also be operationalized by an increase in the non-monetary costs of tax evasion (Alm, 2019; Hofmann et al., 2008; Kirchler et al., 2008). All three effects might consequently result in a lower tax evasion level.

In line with the results of Fochmann et al. (2018) who analyzed the prefilling of income items and observed a lower tax evasion level with correctly prefilled items than with blank items, we formulate our first hypothesis as follows:

Hypothesis 1: *Prefilling of deductions in tax returns reduces the tax evasion level.*

3.3.2 Restricting the deductibility of expenditures

Disallowing the deductibility of expenditures. Another mechanism to enhance tax compliance might be disallowing the deduction of specific expenditure items. Under the assumption, that taxpayers refrain from shifting tax evasion activities to other non-restricted items, the overall tax evasion level will decrease. We formulate our hypothesis 2a therefore as follows:

[4] See Fochmann et al. (2018) for a more detailed discussion.

Hypothesis 2a: *Disallowing the deductibility of an expenditure item does not affect the tax evasion level of other non-restricted items and reduces the overall tax evasion level.*

However, if taxpayers shift their tax evasion activities to other non-restricted expenditure items, the tax evasion level for those items will increase. This would be in line with the finding that taxpayers increase claimed deductions to offset an increase in reported income due to third-party reporting or audits (Adhikari et al., 2016; Asatryan and Peichl, 2017; Carrillo et al., 2017; Slemrod et al., 2017; Vossler and Gilpatric, 2018). Moreover, it might be that individuals feel unfairly treated when the deductibility of an expenditure item is restricted or even completely disallowed. Consequently, perceived procedural fairness might be reduced (reactance) which results in more tax evasion (Hofmann et al., 2008; Kirchler et al., 2008). As the shift in tax evasion to non-restricted items might undercompensate, compensate, or even overcompensate (due to reactance) the positive effect of a disallowance on tax evasion, we refrain from hypothesizing the influence that this mechanism has on overall tax evasion. We therefore formulate hypothesis 2b as follows:

Hypothesis 2b: *Disallowing the deductibility of an expenditure item increases the tax evasion level of other non-restricted items.*

Limiting the deductibility of expenditures. A third mechanism to combat tax evasion might be to limit the deductibility of expenditures. Thus, the deductibility of expenditures is neither completely allowed nor disallowed, but limited. Again, the effect of such a restriction on the overall tax evasion level depends on whether taxpayers shift tax evasion activities to other non-restricted items or not. Hence, we formulate the following two hypotheses:

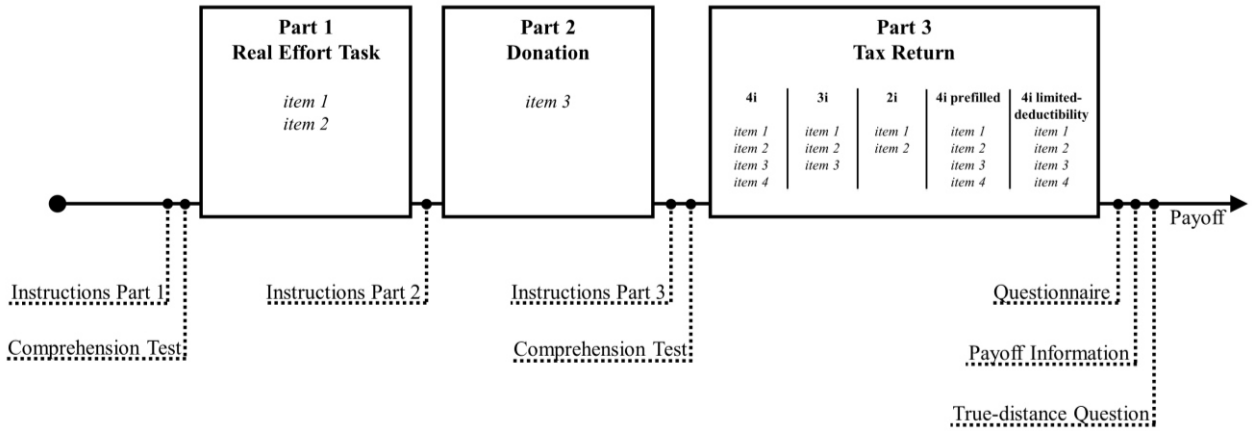
Hypothesis 3a: *Limiting the deductibility of an expenditure item does not affect the tax evasion level of other non-restricted items and reduces the overall tax evasion level.*

Hypothesis 3b: *Limiting the deductibility of an expenditure item increases the tax evasion level of other non-restricted items.*

3.4 Experimental design and treatments

3.4.1 Experimental design of the three main parts

We conduct a laboratory experiment consisting of the following three parts: (1) real effort task, (2) donation and (3) tax return (see Figure 1). The instructions for each part are provided to the participants at the beginning of the corresponding part and are shown in the supplementary material M1.



Note: In this figure, the experimental procedure is shown. Items 1 to 4 refer to our four deduction items used in our experiment (see Table 1 for an overview). The terms 4i, 3i, 2i, 4i prefilled and 4i limited-deductibility refer to our five treatments (see section 3.4.5).

Figure 1: Experimental procedure

In the first part of the experiment (real effort task), participants generate their (pre-tax) income by solving math puzzle tasks introduced by 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 four payoff-relevant rounds of the math puzzle task and in the beginning one testing-round, 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 93 Eurocent (0 Eurocent otherwise). After each round of the real effort task, 22.6% of the generated income is withheld as a fictional social insurance contribution (item 1).[5]

[5] To keep the experiment as simple as possible, participants receive no benefits from social insurance in our experiment.

Before each round, participants can optionally buy a tool (item 2) that simplifies the real effort task.[6] More precisely, the amount of numbers is reduced in all matrices of that round (e.g., from twelve to ten). As the amount of irrelevant numbers is reduced, individuals might solve more puzzles within the given time. The tools cost 37 Eurocent (142 Eurocent, 299 Eurocent) and they reduce the amount of numbers from twelve to eleven (ten, nine), respectively. A simplification tool is valid for one round. Before each round, each participant decides whether she wants to buy one of the three simplification tools.[7]

After each round, the number of correctly solved math puzzles, the amount of withhold social insurance contribution, the expenditures for simplification tools and the resulting earned income in that round are displayed to the participants. To complete the tax return in the third part of the experiment, participants are requested to record the displayed information after each round on a piece of paper at their workstation. Piece of paper and pen are provided to the participants.

In the second part of the experiment (donation), participants can optionally donate part of their generated income to real life institutions (item 3). For this purpose, participants can enter an amount of money which they want to donate. They are asked to select institutions out of a list (e.g., UNICEF and Greenpeace). Again, participants are requested to record the donation amount on a piece of paper at their workstation as a preparation for their tax return.

In the third part (tax return), participants have to file a tax return by reporting their deduction items. Participants are also asked to claim a commuting allowance. They do so by entering the distance between their home and the laboratory in kilometers.[8] For every entered kilometer, participant's taxable income is reduced by 30 Eurocents (commuting allowance, item 4). The most important characteristic of item 4 is the fact that any misreporting is undetectable as the experimenter does not know the true distance. Consequently, this item mirrors expenditures that can only very hardly be verified by the tax authority.[9] Table 1 highlights the item characteristics.

[6] This item mirrors work-related or professional expenditures that might enhance someone's productivity like purchasing a new notebook or attending an advanced training course. Taxable income is usually calculated by subtracting expenditures from earnings (e.g., labor income). This tool and the corresponding expenditures represent a common example for work-related expenditures of employees.

[7] In our experiment, we observed that over all five treatments 57.1% of the participants bought a simplification tool at least during one round.

[8] To enter the distance, participants are allowed to use their smartphones and apps like Google Maps.

[9] In real life, tax authorities can check the plausibility of the entered distance quite easily, but are usually unable to retrace how often the taxpayer has traveled the distance in the taxable period.

The taxable income, that is income minus declared deductions for items 1, 2, 3 and 4, is subject to a tax rate of 40%. The declared income is already prefilled in the tax return and cannot be manipulated by the participants. However, participants have the opportunity to evade taxes if they declare higher deductions than their true expenditures. In the instructions, participants are explicitly asked to declare their true expenditures. Thus, if participants declare higher deductions they engage in tax evasion. Unintentional tax evasion by the taxpayer is virtually excluded by design (our setting is quite simple) and participants are fully aware of their true expenditures.[10]

There is a probability of 30% that a participant will be audited after she has submitted her tax return. If a participant is audited and her declared taxable income is lower than her true taxable 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.[11]

After completion of an ex-post questionnaire (see section 3.4.3), subjects are informed about the audit outcome and their payoff. There is one last question at the very end of the experiment ("true-distance question", details below in section 3.4.4) before participants privately receive their payoff in cash. The payoff consists of a show-up fee of 4 Euro, a reward for correctly answered comprehension tests (see section 3.4.2) and the money earned in the experiment (= pre-tax income minus true expenditures minus tax liability minus potential fine).

[10] To complete the tax return, participants are asked to use the records they made on the piece of paper. Moreover, participants can press a button on the tax return screen to have their actual expenditures displayed to exclude that record errors bias their compliance behavior. In our experiment, we observed that over all five treatments 40.3% of the participants pressed the button at least once.

[11] If an audit reveals that the declared taxable income is higher than the true taxable income, the participant gets back the overpaid taxes and no additional costs occur.

Table 1: Item overview

Item	Description
Item 1 (*social insurance contribution*)	Fixed percentage rate (22.6%) of income. Participants can deduct the social insurance contribution from their tax base in the tax return.
Item 2 (*work-related expenditures*)	Expenditure that occurs when participants buy tools to simplify the income generation. Participants can buy these tools for fixed prices before each of the four rounds of the real effort task. Participants can deduct the total costs of purchased tools in the tax return.
Item 3 (*donation*)	Expenditure that occurs when participants donate part of their generated income to real life institutions (e.g. UNICEF, Greenpeace). In the tax return, participants can deduct their donation.
Item 4 (*commuting allowance*)	Expenditure that captures participant's costs to arrive at the laboratory. Participants are asked to enter the distance from their home to the laboratory. For every entered kilometer, participant's taxable income decreases by 0.30€.

3.4.2 Comprehension tests

Prior to part one and also before part three of the experiment, subjects have to complete a monetary incentivized comprehension test. They are asked several questions regarding the puzzle task, pre-tax income determination, tax liability determination, audit probability and payoff determination. If participants answer the questions correctly on their first (second) try, they receive an additional payment of 1 Euro (0.50 Euro), otherwise 0 Euro. The full set of questions is provided in the supplementary material M2.

3.4.3 Ex-post questionnaire

The tax compliance literature provides evidence that several socio-demographic and attitudinal variables have an influence on tax compliance behavior, such as age (Muehlbacher et al., 2011), gender (Kastlunger et al., 2010), risk attitude (Dulleck et al., 2016; Fochmann and Wolf, 2019), tax morale (Alm, 2019; Kirchler, 2007; Lewis, 1982; Torgler, 2002), income (Grundmann and Graf Lambsdorff, 2017; Gangl and Torgler, 2020) and emotions (Erard and Feinstein, 1994; Bosco and Mittone, 1997; Dulleck et al., 2016; Blaufus et al., 2017; Enachescu et al., 2019). At the end of the experiment (but before participants learn their final payoffs, see Figure 1), participants are therefore asked to answer a questionnaire that collects socio-demographic and attitudinal data. Table 2 provides an overview of relevant variables and the appendix A1 contains additional information on the ex-post questionnaire. The answers to these questions are used as controls in our regression analyses.

3.4.4 True-distance question

After the questionnaire and after participants are informed about the audit outcome and their final payoffs, we display a final question to the participants and ask them to enter the true distance from their home to the laboratory. We explicitly point out to the participants that their answer to this question will not affect their final payoff and that the actual experiment is already completed.[12] This last question enables us to estimate the tax evasion level with item 4 ex-post of the experiment.[13] However, this analysis has to be threatened with caution, because it demands the honesty of the participants. Nevertheless, we feel confident that most participants entered the true value, because of our appeal to be honest and because the participants knew that the actual experiment was over. Furthermore, we observe significant lower and more realistic answers than reported in the tax return of the experiment.[14]

3.4.5 Treatments

Our experiment consists of five treatments (between-subject design). Figure 2 highlights the differences between them. The first treatment allows the deduction of all 4 items in the tax return (base case).

- *Treatment 4i (base case)*: All 4 items are deductible in the tax return.

In the second treatment we prefill each expenditure item with its correct value in the tax return. For item 4 (commuting allowance) we used the median distance reported in the true-distance question in treatment 4i as the prefilled value for item 4 (that is 5km).[15] All monetary aspects such as tax rate, audit probability, and penalties are kept constant by this treatment variation. Thus, we exclude that prefilling changes tax compliance behavior through a change in the audit probability or penalty.

- *Treatment 4i prefilled*: All 4 expenditure items are deductible and the deductions are prefilled in the tax return.

In the next two treatments we disallowed the deduction of specific expenditures:

[12] After the clarification, we literally asked the following question: "In this experiment you were asked to enter the distance from your home to the laboratory and it was up to you to enter a smaller or greater distance. For the analysis of this study we kindly ask you to enter the true distance in kilometers. Again, you are allowed to use your smartphone."

[13] There is technically no other way to obtain this information due to the anonymity of our participants.

[14] Furthermore, considering this "limitation" makes our following results even stronger. If we assume that some participants still report more than the true number of kilometers we underestimate tax evasion in item 4. Consequently, tax evasion in item 4 might be even higher and our already highly significant results even stronger.

[15] Consequently, we conducted treatment 4i prefilled with a certain time-delay after treatment 4i.

- *Treatment 3i (3 items)*: Items 1, 2 and 3 are deductible in the tax return. Item 4 (commuting allowance) is non-deductible.
- *Treatment 2i (2 items):* Items 1 and 2 are deductible in the tax return. Items 3 (donation) and 4 (commuting allowance) are non-deductible.

The remaining treatment matches the first treatment (base case) with one exemption regarding item 4 (commuting allowance). Participants are only allowed to deduct a limited amount of 10 Eurocent per kilometer (instead of 30 Eurocent as in the base case).[16]

- *Treatment 4i limited-deductibility*: All 4 expenditure items are deductible. However, participants are only allowed to deduct a limited amount of 10 Eurocent per kilometer in item 4.

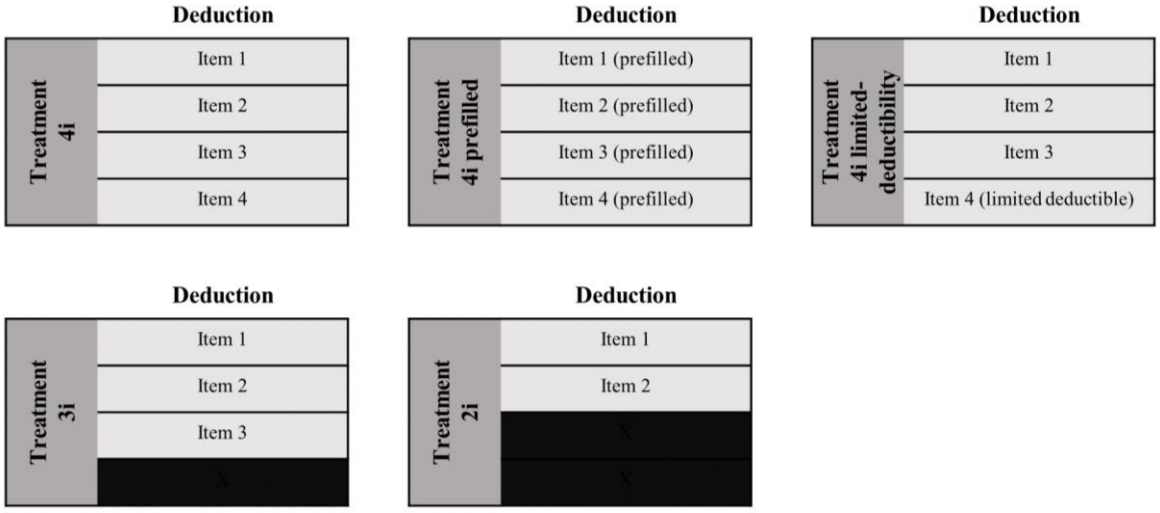


Figure 2: Treatment differences

We control for all other factors that might influence tax compliance, such as fine and tax rate, by keeping them constant between the different treatments. Also the experimental setting of part one and part two remains constant over all treatments to ensure that any observed difference in compliance behavior is due to the treatment manipulations regarding the tax return in part three of the experiment. This implies, for example, that in treatment 2i participants still have the opportunity to donate in part two, but that the donation cannot be deducted in part three.

[16] Participants in this treatment do not know that the commuting allowance is higher in other treatments. In the instructions in this treatment, we only state that the commuting allowance is 10 Eurocent per kilometer (see supplementary material M1.4).

3.4.6 Sample and data

The experiment was conducted at the computerized experimental laboratory of the University of Cologne (CLER) from September to December 2018. The experiment was programmed and executed with the software z-Tree (Fischbacher, 2007). Participants were recruited with ORSEE (Greiner, 2015). In total, 191 subjects (mainly undergraduate students) participated and earned, on average, 22.11 Euro in approximately 90 minutes (approximately 14.74 Euro per hour). A total of 40 subjects were randomly assigned to treatment 4i, 42 to treatment 3i, 35 to treatment 2i, 39 to treatment 4i prefilled and 35 to treatment 4i limited-deductibility. Table 2 provides an overview on the main characteristics of the participants collected in our ex-post questionnaire (see section 3.4.3).

Table 2: Main characteristics of our participants

Variable	Description	Mean
Female	female = 1; male = 0	53.7%
Risk attitude	self-reported risk attitude 0 to 10	4.54
Age	in years	26
Economics	participant with more than one lecture in economics = 1 (0 otherwise)	45.8%
Tax experience	experience with tax returns = 1 (0 otherwise)	42.6%
Monthly income	in Euro (monthly income after fixed costs)	346
Tax morale	0 to 9; low tax morale = 0; high tax morale = 9	7.07
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	5.54
Decision complexity	0 to 10; low perceived decision complexity in experiment = 0; high perceived decision complexity in experiment = 10	2.55
Joy	0 to 10; felt no joy during experiment = 0; felt high joy during experiment = 10	5.44
Anger	0 to 10; felt no anger during experiment = 0; felt high anger during experiment = 10	3.45
Fear	0 to 10; felt no fear during experiment = 0; felt high fear during experiment = 10	2.11
Guilt	0 to 10; felt no guilt during experiment = 0; felt high guilt during experiment = 10	1.88

Note: This table provides an overview of the individual characteristics of the 191 participants in our experiment.

3.5 Results

We use two tax evasion measures to analyze our experimental data. First, we use the interval variable *overdeductions* which measures the absolute level of overdeductions (i.e., declared deductions minus true deductions). Second, we use the dummy variable *evader* which takes the value of 1 if a participant evaded any tax (i.e., was not fully

compliant). Whereas the variable evader measures whether a participant is fully compliant or not, the variable overdeductions also measures the level of non-compliance (i.e., the magnitude of tax evasion). Both variables *overdeductions* and *evader* are calculated for each item separately.[17] Figures 3–7 show the results of the five treatments which we discuss in the following in more detail.

[17] In the rare case of underdeductions, both variables are set to 0.

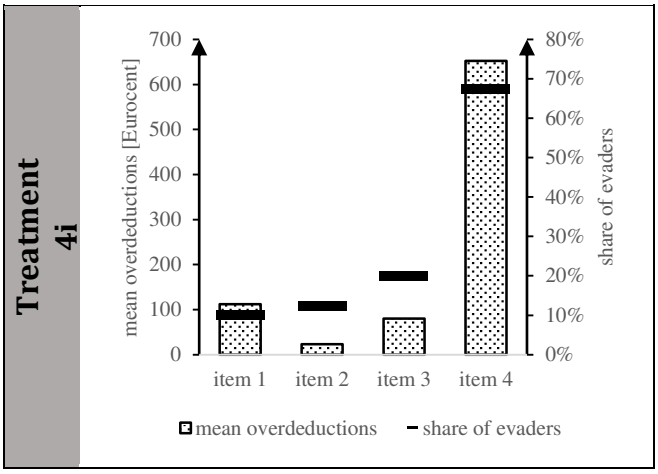


Figure 3: Tax evasion in treatment 4i

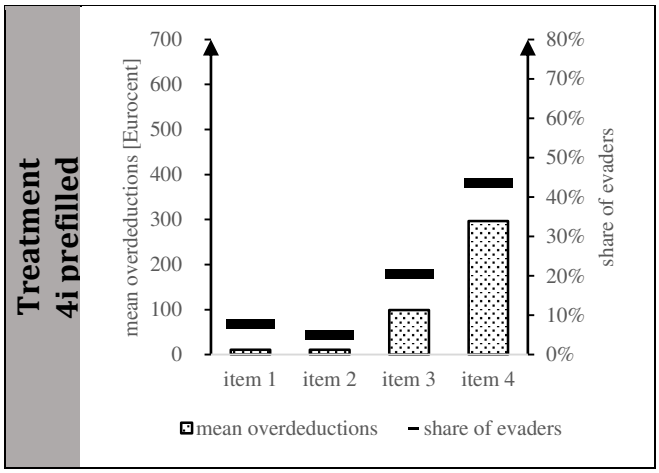


Figure 4: Tax evasion in treatment 4i prefilled

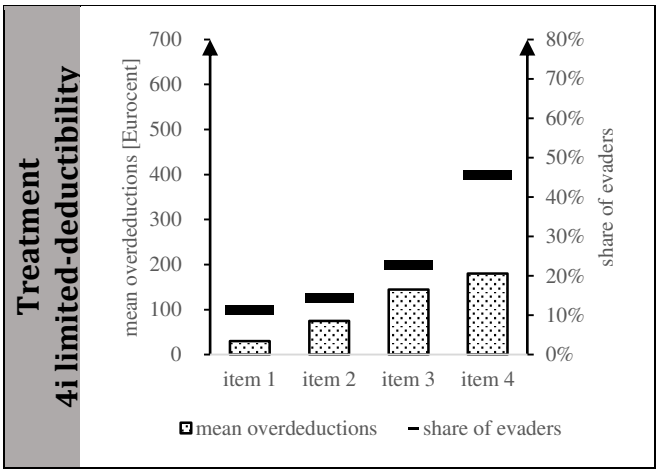


Figure 5: Tax evasion in treatment 4i limited-deductibility

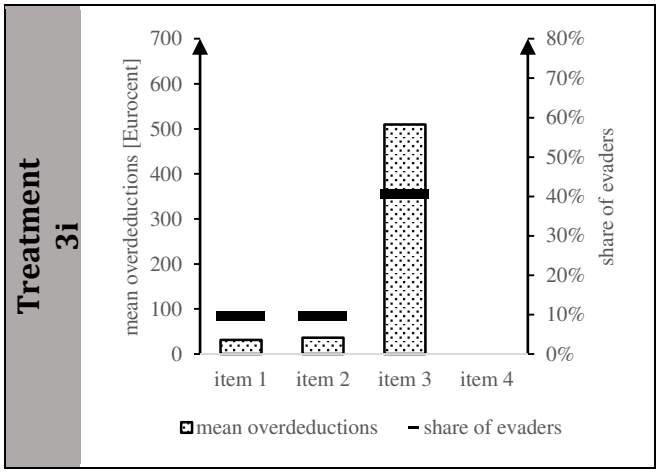


Figure 6: Tax evasion in treatment 3i

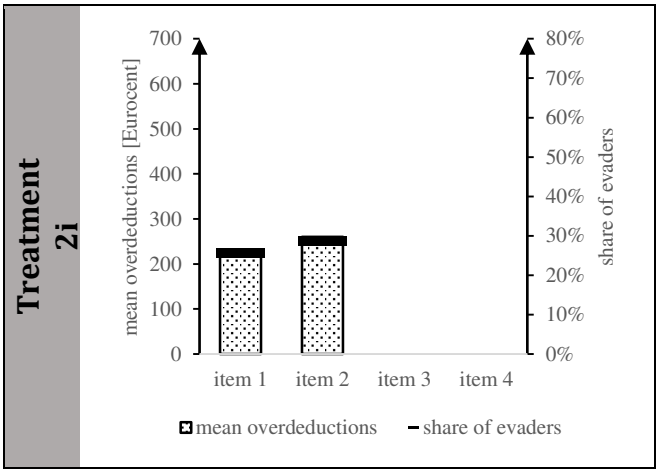


Figure 7: Tax evasion in treatment 2i

3.5.1 Tax evasion within each treatment

We start the presentation of our results by analyzing the item-specific tax evasion levels within each treatment. We test for differences between the items by using the McNemar test for variable evader and the Wilcoxon signed-rank test for variable overdeductions (always two-tailed). We robustly observe that item 4 (commuting allowance) is preferably used for tax evasion – both in terms of share of evaders and overdeductions. As tax evasion cannot be detected with this item in our experiment, this finding is in line with the tax compliance literature. In particular, this literature provides robust evidence that a lower audit/detection probability increases tax evasion (Allingham and Sandmo, 1972; Spicer and Thomas, 1982; Alm et al., 1995; Maciejovsky et al., 2001; Torgler, 2003; Cummings et al., 2009; Fortin et al., 2007; Gërxhani and Schram, 2006).

Treatment 4i (base case). Figure 3 shows that the tax evasion level is higher for item 4 than for the other items in treatment 4i. In particular, 67.5% of the participants evade taxes with item 4 compared to approx. 10 to 20% with items 1, 2 and 3. Mean overdeductions are 112 Eurocent for item 1, 23 for item 2, 80 for item 3 and 652 for item 4. All differences between item 4 and the other items are highly significant (all p-values below 0.001). Between items 1, 2 and 3 differences are insignificant (all p-values above 0.1).

Treatment 3i. Figure 5 shows that 40.5% of participants evade taxes with item 3 compared to 9.5% with items 1 and 2 in treatment 3i. Mean overdeductions yield 510 Eurocent for item 3, 36 for item 2 and 31 for item 1. All differences between item 3 and the other two items are highly significant (all p-values below 0.001). No significant differences are observed between items 1 and 2 (all p-values above 0.1).

Treatment 2i. Figure 6 reveals that 28.6% of participants evade taxes with item 2 compared to 25.7% with item 1 in treatment 2i. Mean overdeductions yield 260 Eurocent for item 2 and 221 for item 1. Differences between item 1 and 2 are insignificant (all p-values above 0.1).

Treatment 4i prefilled. Figure 4 shows that 43.6% of participants evade taxes with item 4 compared to 20.5% with item 3, 5.1% with item 2 and 7.7% with item 1 in treatment 4i prefilled. Mean overdeductions yield 297 Eurocent for item 4, 99 for item 3 and 11 for item 2 and 1. All differences between item 4 and the other three items are significant (all p-values below 0.05). However, our statistical tests also reveal significant differences between items 1 and 3 for overdeductions ($p = 0.0326$) and significant differences

between items 2 and 3 for overdeductions (p = 0.0171) and evaders (p = 0.0313). For all other item combinations, we find no significant differences (all p-values above 0.1).

Treatment 4i limited-deductibility. Figure 7 exhibits that 45.7% of participants evade taxes with item 4 compared to 22.9% with item 3, 14.3% with item 2 and 11.4% with item 1 in treatment 4i limited-deductibility. Mean overdeductions yield 180 Eurocent for item 4, 145 for item 3, 75 for item 2 and 30 for item 1. All differences between item 4 and the other three items are significant (all p-values below 0.05). The only exemption occurs for the comparison between items 4 and 3 where we find no significant difference for the variable overdeductions (p = 0.1994). No significant differences are observed between items 1, 2 and 3 (all p-values above 0.1).

3.5.2 Prefilling

To analyze the effect of prefilled tax returns on tax compliance, we compare treatment 4i with treatment 4i prefilled (see Figures 3 and 4). Whereas item 4 is commonly used by the participants for tax evasion in treatment 4i, we find a strong decrease of tax evasion with this item in treatment 4i prefilled. In particular, overdeductions of item 4 decrease from 652 to 297 and share of evaders from 67.5% to 43.6%. For items 1 and 2, we also observe that prefilling reduces tax evasion, but to a lower extend. Item 3 is unaffected. We explain the small effects for items 1 to 3 by the already low tax evasion levels for these items in treatment 4i.

A low tax evasion benchmark might be less-than-ideal to test the effectiveness of prefilling. For example, we observe a decrease in overdeductions from 112 to 11 Eurocent for item 1. This decrease for item 1 is in absolute terms lower than the decrease for item 4. However, overdeductions are getting close to zero (i.e., indicating no tax evasion) with prefilled tax returns and in relative terms the decrease for item 1 (approx. 90%) is even higher than for item 4 (approx. 54%). The level of total overdeductions (i.e., sum of overdeductions over all available deduction items) decreases from 867 Eurocent in treatment 4i to 418 in treatment 4i prefilled.

The results are also supported by linear regressions (see Table 3) with overdeductions (models 1 to 4) and total overdeductions (sum of overdeductions, model 5) as dependent variables. In all models, we regress on a dummy variable treatment 4i prefilled that equals 1 if the decision was made in this treatment (0 otherwise). Treatment 4i serves as the default. Moreover, we consider a vector of individual characteristics as controls collected in our ex-post questionnaire (see section 3.4.3 and Table 2 for details).

Controls are not reported in Table 3, but the full set of regression results can be found in our appendix A2.

We find significantly lower overdeductions for item 4 (model 4), but not for the other items (models 1 to 3), and a significantly lower level of total overdeductions (model 5) in treatment 4i prefilled than in treatment 4i. The coefficient of the variable treatment 4i prefilled in model 4 points to (on average) lower overdeductions by 339 Eurocent for item 4 in this treatment compared to treatment 4i. Consequently, our findings support hypothesis 1 and provide evidence that prefilling is an effective mechanism to reduce tax evasion. We rerun models 1 to 4 also with the item-specific variable evader as dependent variable (logistic regression). All results are robust to this variation. Regression results are provided in the appendix A2.

Result 1: *Prefilling deductions in the tax return reduces tax evasion.*

Table 3: Regression results: Prefilling of deductions

	Treatment 4i vs 4i prefilled				
Dependent variable	**Over-deductions**	**Over-deductions**	**Over-deductions**	**Over-deductions**	**Total over-deductions**
	Model 1	Model 2	Model 3	Model 4	Model 5
	Item 1	Item 2	Item 3	Item 4	All items
Treatment 4i prefilled	-124.43	-49.67	33.34	-338.69**	-479.47**
	(80.54)	(65.46)	(55.08)	(158.57)	(212.78)
Constant	408.89	342.36	202.07	862.37	1,815.69**
	(335.26)	(272.50)	(229.27)	(660.07)	(885.74)
Controls	Yes	Yes	Yes	Yes	Yes
No. of observations	78	78	78	78	78
R-squared	0.196	0.145	0.182	0.221	0.209

Note: This table presents the results of linear regression models with either overdeductions or total overdeductions as dependent variables (regression coefficients, standard errors in parentheses). Number of observations is determined by the number of subjects participated in the respective treatments. *** p<0.01, ** p<0.05, * p<0.1

3.5.3 Disallowing the deductibility of expenditures: Evasion-shift-effect

We now analyze the effectiveness of deduction-disallowance of specific expenditure items that are commonly used by the participants for tax evasion. For this purpose, we compare the tax evasion levels between treatments 4i, 3i (item 4 not deductible) and 2i (items 3 and 4 not deductible). See Figures 3, 5 and 6 for the respective results.

First, a direct comparison of treatment 4i with treatment 3i reveals that the level of tax evasion in item 3 increases. In particular, the disallowance to deduct item 4 leads to

an increase in overdeductions (share of evaders) in item 3 from 80 Eurocents (20.0%) in treatment 4i to 510 (40.5%) in treatment 3i. Second, a direct comparison of treatment 3i with treatment 2i reveals that the level of tax evasion in items 1 and 2 increases. In particular, the disallowance to deduct item 3 leads to an increase in overdeductions (share of evaders) in item 1 from 31 Eurocents (9.5%) in treatment 3i to 221 (25.7%) in treatment 2i and an increase in overdeductions (share of evaders) in item 2 from 36 Eurocents (9.5%) in treatment 3i to 260 (28.6%) in treatment 2i.

Again, we run linear regressions. The results are presented in Table 4. As dependent variables we again consider overdeductions (models 1 - 5) or total overdeductions (sum of overdeductions, models 6 - 7). The variables of interest are dummy variables for the treatments 3i and 2i. Each treatment variable takes the value of 1 if the decision was made in the respective treatment (0 otherwise). In models 1 to 3 and 6 (4, 5 and 7), we consider the dummy variable for treatment 3i (2i) and set treatment 4i (3i) as the default. Again, we consider individual characteristics as controls.[18]

Our results show that disallowing the deductibility of one item affects the tax evasion level of the remaining items. In particular, tax evasion level of item 3 is significantly higher in treatment 3i – where item 4 is non-deductible – than in treatment 4i (model 3). However, for items 1 and 2 we do not observe significant differences between both treatments (models 1 and 2). Comparing treatments 3i and 2i, we find that the tax evasion levels of item 1 (model 4) and of item 2 (model 5) are significantly higher in treatment 2i.

In both regression models with total overdeductions as dependent variable (models 6 and 7), we fail to find a significant treatment effect. Consequently, disallowing the deductibility of expenditures for one item fails to reduce the overall tax evasion level significantly. In conclusion, we observe an evasion-shift-effect resulting in an increase of tax evasion for at least one of the remaining items. This increase is high enough to achieve a similar level of total overdeductions as before. Consequently, cutting tax evasion opportunities does not reduce total tax evasion. This supports hypothesis 2b and rejects hypothesis 2a.

[18] Please notice that we rerun models 1 to 5 also with the item-specific variable evader as dependent variable (logistic regression). All results are robust to this variation and regression results are provided in the appendix A2.

Result 2: *Disallowing the deductibility of expenditures causes an evasion-shift-effect. Individuals shift overdeductions from the restricted item to other non-restricted items. Overall tax evasion level does not change significantly.*

Table 4: Regression results: Disallowing the deductibility

	Treatment 4i vs 3i			**Treatment 3i vs 2i**		**Treatment 4i vs 3i**	**Treatment 3i vs 2i**
Dependent variable	**Over-deductions**	**Over-deductions**	**Over-deductions**	**Over-deductions**	**Over-deductions**	**Total over-deductions**	**Total over-deductions**
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
	Item 1	Item 2	Item 3	Item 1	Item 2	All items	All items
Treatment 3i	-39.40	-5.91	475.04***			-270.14	
	(80.74)	(68.21)	(152.20)			(225.28)	
Treatment 2i				268.09**	272.67**		8.15
				(104.54)	(104.35)		(227.23)
Constant	248.86	262.40	-858.01	251.99	603.57*	982.43	526.52
	(324.30)	(273.97)	(611.32)	(328.05)	(327.46)	(904.88)	(713.09)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	82	82	82	77	77	82	77
R-squared	0.197	0.138	0.289	0.236	0.234	0.173	0.274

Note: This table presents the results of linear regression models with overdeductions as dependent variable (regression coefficients, standard errors in parentheses). Number of observations is determined by the number of subjects participated in the respective treatments. *** p<0.01, ** p<0.05, * p<0.1

3.5.4 Limiting the deductibility of expenditures: No evasion-shift-effect

Finally, we analyze the effectiveness of limiting the deductibility of an expenditure item. For this purpose, we compare the tax evasion levels between treatment 4i and treatment 4i limited-deductibility (see Figures 3 and 7). For item 4, we observe that overdeductions decrease from 652 to 180 Eurocent. The share of evaders decreases from 67.5% to 45.7%. For the other items, we fail to find any significant differences between the two treatments. Overall tax evasion (i.e., sum of overdeductions) decreases from 867 Eurocent in treatment 4i to 430 in treatment 4i limited-deductibility.

These findings are supported by linear regressions (see Table 5) with overdeductions (models 1 - 4) and total overdeductions (models 5 - 6) as dependent variables. In all models, we regress on a dummy variable treatment limited-deductibility that equals 1 if the decision was made in this treatment (0 otherwise). Treatment 4i serves as the default. Moreover, we consider individual characteristics as controls.[19] We find significantly lower

[19] Please notice that we rerun models 1 to 4 also with the item-specific variable evader as dependent variable (logistic regression). All results are robust to this variation and regression results can be found in the appendix A2.

overdeductions for item 4 (model 4), but not for the other items (models 1 to 3), and a significantly lower level of total overdeductions (models 5 and 6) in treatment 4i limited-deductibility than in treatment 4i. Consequently, we provide evidence that there is no evasion-shift-effect in treatment 4i limited-deductibility. In particular, a limited deductibility for one item does not affect the tax evasion level of the other items and decreases the overall tax evasion level. Consequently, a limited deductibility seems to be an effective mechanism to reduce tax evasion. Our results therefore support hypothesis 3a and reject hypothesis 3b.

Result 3: *Limiting the deductibility of an expenditure causes no evasion-shift-effect. Individuals do not shift overdeductions from the restricted item to other non-restricted items. Overall tax evasion level is reduced significantly.*

In model 6, we additionally include the observations from treatment 4i prefilled to also test for differences between treatment 4i limited-deductibility and 4i prefilled. We observe no significant difference between both treatments in overall tax evasion (checked by Wald test). Previous results are supported.

Table 5: Regression results: Limiting the deductibility

	TR 4i vs 4i limited-deductibility					Treatment 4i vs 4i prefilled vs 4i limited-deductibility
Dependent variable	**Over-deductions**	**Over-deductions**	**Over-deductions**	**Over-deductions**	**Total over-deductions**	**Total over-deductions**
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Item 1	Item 2	Item 3	Item 4	All items	All items
TR 4i limited-deductibility	-50.32	49.61	91.38	-449.78***	-359.11*	-360.95*
	(81.74)	(79.99)	(74.59)	(127.33)	(192.26)	(185.43)
TR 4i prefilled						-443.95**
						(181.84)
Constant	438.71	452.28	473.47	1,847.08***	3,211.54***	1,853.77***
	(372.04)	(364.07)	(339.51)	(579.58)	(875.08)	(675.76)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	75	75	75	75	75	113
R-squared	0.208	0.136	0.171	0.304	0.281	0.228
Wald test						
4i limited-deductibility = 4i prefilled						p = 0.669

Note: This table presents the results of linear regression models with either overdeductions or total overdeductions as dependent variables (regression coefficients, standard errors in parentheses). Number of observations is determined by the number of subjects participated in the respective treatments. *** p<0.01, ** p<0.05, * p<0.1

3.5.5 Robustness tests

Order effects. In treatments 4i and 3i always the last expenditure item (presented in the respective tax return) has the highest level of tax evasion. Although treatment 2i lacks this observation, we conducted two additional treatments with small sample size to provide some evidence that order effects do not bias the observed behavior. In these robustness treatments, we reversed the order of the presented items in the respective tax return. In all other aspects treatment 4i-inverse-order (N=12) follows treatment 4i and treatment 3i-inverse-order (N=19) follows treatment 3i. Again, tax evasion level is highest for item 4 in treatment 4i-inverse-order and highest for item 3 in treatment 3i-inverse-order. Also in line with our previous results, we observe the lowest tax evasion level for items 1 and 2. Therefore, these findings provide some evidence that order effects do not bias the tax evasion decisions of the participants. Please note that the results of the two robustness treatments must be interpreted with caution due to the low number of observations.

Perceived Audit Probability. There is ample evidence that an increased audit probability increases tax compliance (see Torgler, 2002 and Alm, 2019 for an overview). Although

our treatment variation has no influence on the objective audit probability to detect tax fraud, it might be that the subjective perception of the audit probability is affected. For example, Kogler et al. (2016) observe that their experimental manipulations changed the subjective audit probability although the objective audit probability, which was explicitly mentioned to their participants before the experiment started, was unchanged. However, we provide evidence that a change in subjective/perceived audit probability can be excluded as an explanation for our observed treatment differences. 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 our main five treatments, the mean answer was 3.97. Differences across treatments were small and statistically insignificant (Kruskal-Wallis test, $p = 0.679$, two-tailed).

Button to display the actual expenditures. As outlined in section 3.4.1, participants had a button on the tax return screen of the experiment that displayed their actual expenditures (see footnote 9). As a robustness test, we rerun all regressions and included additionally – as a control variable – how often a participant pressed this button. All results are robust to this variation. Moreover, the control variable never shows up significantly in the regressions.

3.6 Discussion and conclusion

We analyzed three anti-tax-evasion mechanisms that focus on deductions: 1) prefilling of deductions in tax returns and 2) restricting tax evasion opportunities by either disallowing or 3) limiting the deductibility of specific expenditures.

Our results suggest that prefilled deductions enhance tax compliance. In particular, item-specific tax evasion level decreases – especially for items preferred for tax evasion – and as a consequence overall tax evasion level is reduced. As we do not observe that the subjective perception of audit probabilities varies significantly across treatments, the positive effect of prefilling might be primarily driven by higher non-monetary costs associated with tax evasion under this mechanism. This finding highlights the importance of non-monetary and psychological factors for the design of tax regulations.

Disallowing the deductibility of one expenditure item (i.e., cutting the number of tax evasion opportunities) is an ineffective mechanism to combat tax evasion. In fact, individuals shift their tax evasion activities from the disallowed item to other non-restricted items (evasion-shift-effect). However, our results suggest that limiting the deductibility (in contrast to disallowing the deductibility completely) avoids this evasion-

shift-effect and finally reduces overall tax evasion. Remarkably, disallowing the deductibility of expenditures completely might lead to a reduction in perceived procedural fairness. This might explain why we find a shift-effect in the former case, but not in the latter one.

We conclude that policy makers trying to combat tax evasion should – if technically feasible – prefill deductions in tax returns. Whereas our results suggest that a disallowance of deductions results in an evasion-shift-effect, policy makers might avoid this effect by only limiting the deductible amount.[20] While we observe similar effects of prefilling tax returns and limiting (not disallowing) the deductibility of expenditures on overall tax evasion level, both mechanisms differ in their approach. Whereas limiting the deductibility constrains the monetary benefit of tax evasion directly, prefilling does not change any tax evasion opportunity or the objective monetary costs or benefits of tax evasion. If technically feasible, prefilling might be easier to implement than changing the tax law to limit the deductibility of expenditures. Whereas the former just requires a change in the administrative process (and is already performed by tax preparation software), the latter needs democratic justification. More importantly, disallowing or limiting the deductibility can be characterized as a lump-sum solution that also affects the tax bill of individuals who would comply anyway. Whereas prefilling comes without these negative consequences for honest taxpayers, it influences the compliance behavior and the effective tax bill of individuals who tend to evade taxes by claiming additional expenditures in their tax statements.

Our study does have limitations. One limitation is that our sample primarily consists of students. 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 complex case studies where special expertise is crucial, we decided to use students. Moreover, 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 in our experiment (Alm et al., 2015; Depositario et al., 2009; Remus, 1996; Ashton and Kramer, 1980; Elliott et al., 2007). Therefore, we feel confident that using students as subjects is appropriate in our setting.

[20] For example, expenditures related to taxable earnings (e.g., labor income) are in many cases (at least somehow) related to the private sphere. Here, the legislator decides whether a full, partial or no deduction of such expenditures is applied.

The statistical power of experiments is an important issue in experimental economics and potentially also in our study. Therefore, we were cautious in interpreting our results – especially when we found no statistically significant differences. In this regard, further research addressing specific findings of our study might be useful. For example, prefilling reduces tax evasion for almost all items, but a statistically significant effect is only observed for item 4 which is the item that is mostly used for tax evasion. Future research might help to identify whether prefilling has a general effect or affects primarily items with already high tax evasion levels.

References

Adhikari, B., J. Alm, B. Collins, M. Sebastiani, and E. Wilking. 2016. Taxpayer responses to third-party income reporting: Evidence from a natural experiment in the taxicab industry. Paper read at ASSA Annual Meeting, 2018.

Agostini, C. A., and C. Martínez A. 2014. Response of Tax Credit Claims to Tax Enforcement: Evidence from a Quasi-Experiment in Chile. *Fiscal Studies* 35 (1):41-65.

Allingham, M. G., and A. Sandmo. 1972. Income tax evasion: A theoretical analysis. *Journal of Public Economics* 1 (3-4):323-338.

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.

———. 2019. What Motivates Tax Compliance? *Journal of Economic Surveys* 33 (2):353-388.

Alm, J., K. M. Bloomquist, and M. McKee. 2015. On the External Validity of Laboratory Tax Compliance Experiments. *Economic Inquiry* 53 (2):1170-1186.

Alm, J., I. Sanchez, and A. de Juan. 1995. Economic and Noneconomic Factors in Tax Compliance. *Kyklos* 48 (1):1-18.

Andreoni, J., B. Erard, and J. Feinstein. 1998. Tax compliance. *Journal of Economic Literature* 36 (2):818-860.

Asatryan, Z., and A. Peichl. 2017. Responses of firms to tax, administrative and accounting rules: Evidence from Armenia. In *CESifo Working Paper No. 6754*.

Ashton, R. H., and S. S. Kramer. 1980. Students As Surrogates in Behavioral Accounting Research: Some Evidence. *Journal of Accounting Research* 18 (1):1-15.

Becker, G. S. 1968. Crime and Punishment: An Economic Approach. *Journal of Political Economy* 76 (2):169-217.

Blaufus, K., J. Bob, P. E. Otto, and N. Wolf. 2017. The Effect of Tax Privacy on Tax Compliance – An Experimental Investigation. *European Accounting Review* 26 (3):561-580.

Bosco, L., and L. Mittone. 1997. Tax Evasion and Moral Constraints: some Experimental Evidence. *Kyklos* 50 (3):297-324.

Bruner, D., M. Jones, M. McKee, and C. Vossler. 2015. Tax Reporting Behavior: Underreporting Opportunities and Prepopulated Tax Returns. In *Working Paper Number 15-11*: Department of Economics, Appalachian State University.

Carrillo, P., D. Pomeranz, and M. Singhal. 2017. Dodging the Taxman. *American Economic Journal: Applied Economics* 9 (2):144-164.

Chapman, G. B., and E. J. Johnson. 1999. Anchoring, Activation, and the Construction of Values. *Organizational Behavior and Human Decision Processes* 79 (2):115-153.

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

Depositario, D. P. T., R. Nayga, X. Wu, and T. P. Laude. 2009. Should students be used as subjects in experimental auctions? *Economics Letters* 102 (2):122-124.

Dulleck, U., J. Fooken, C. Newton, A. Ristl, M. Schaffner, and B. Torgler. 2016. Tax compliance and psychic costs: Behavioral experimental evidence using a physiological marker. *Journal of Public Economics* 134:9-18.

Duncan, D., and D. Li. 2018. Liar Liar: Experimental Evidence of the Effect of Confirmation-Reports on Dishonesty. *Southern Economic Journal* 84 (3):742-770.

Elliott, W. B., F. D. Hodge, J. J. Kennedy, and M. Pronk. 2007. Are M.B.A. Students a Good Proxy for Nonprofessional Investors? *The Accounting Review* 82 (1):139-168.

Enachescu, J., J. Olsen, C. Kogler, M. Zeelenberg, S. M. Breugelmans, and E. Kirchler. 2019. The role of emotions in tax compliance behavior: A mixed-methods approach. *Journal of Economic Psychology* 74:102194.

Epley, N., and T. Gilovich. 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.

Erard, B., and J. Feinstein. 1994. The role of moral sentiments and audit perceptions in tax compliance: Carleton University, Department of Economics.

Evans, C., and B. Tran-Nam. 2010. Managing Tax System Complexity: Building Bridges through Pre-filled Tax Returns. *Australian Tax Forum* (2):245-274.

Feldstein, M. 2015. Raising Revenue by Limiting Tax Expenditures. *Tax Policy and the Economy* 29 (1):1-11.

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

Fochmann, M., N. Müller, and M. Overesch. 2018. Less cheating? The effects of prefilled forms on compliance behavior. *SSRN Electronic Journal*.

Fochmann, M., and N. Wolf. 2019. Framing and salience effects in tax evasion decisions – An experiment on underreporting and overdeducting. *Journal of Economic Psychology* 72:260-277.

Fonseca, M. A., and S. B. Grimshaw. 2017. Do behavioral nudges in prepopulated tax forms affect compliance? Experimental evidence with real taxpayers. *Journal of Public Policy & Marketing* 36 (2):213-226.

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

Gangl, K., and B. Torgler. 2020. How to Achieve Tax Compliance by the Wealthy: A Review of the Literature and Agenda for Policy. *Social Issues and Policy Review* 14 (1):108-151.

Gërxhani, K., and A. Schram. 2006. Tax evasion and income source: A comparative experimental study. *Journal of Economic Psychology* 27 (3):402-422.

Gillitzer, C., and P. E. Skov. 2018. The use of third-party information reporting for tax deductions: evidence and implications from charitable deductions in Denmark. *Oxford Economic Papers* 70 (3):892-916.

Goolsbee, A. 2006. The'Simple Return': Reducing America's Tax Burden Through Return-free Filing. In *Discussion Paper 2006-04*: Brookings Institution.

Gordon, J. P. P. 1989. Individual morality and reputation costs as deterrents to tax evasion. *European Economic Review* 33 (4):797-805.

Greiner, B. 2015. Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association* 1 (1):114-125.

Grundmann, S., and J. Graf Lambsdorff. 2017. How income and tax rates provoke cheating – An experimental investigation of tax morale. *Journal of Economic Psychology* 63:27-42.

Hofmann, E., E. Hoelzl, and E. Kirchler. 2008. Preconditions of Voluntary Tax Compliance: Knowledge and Evaluation of Taxation, Norms, Fairness and Motivation to Cooperate. *Zeitschrift für Psychologie/Journal of Psychology* 216 (4):209-217.

Johnson, E. J., and D. Goldstein. 2003. Do Defaults Save Lives? *Science* 302 (5649):1338-1339.

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

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

Kirchler, E. 2007. *The economic psychology of tax behaviour*: Cambridge University Press.

Kirchler, E., E. Hoelzl, and I. Wahl. 2008. Enforced versus voluntary tax compliance: The "slippery slope" framework. *Journal of Economic Psychology* 29 (2):210-225.

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez. 2011. Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark. *Econometrica* 79 (3):651-692.

Klun, M. 2009. Pre-filled Iincome Tax Returns: Reducing Compliance Ccosts for Personal Income Taxpayers in Slovenia. *Financial Theory and Practice* 33 (2):219-233.

Kogler, C., L. Mittone, and E. Kirchler. 2016. Delayed feedback on tax audits affects compliance and fairness perceptions. *Journal of Economic Behavior & Organization* 124:81-87.

Kotakorpi, K., and J.-P. Laamanen. 2016. Prefilled Income Tax Returns and Tax Compliance: Evidence from a Natural Experiment. In *Tampere Economic Working Papers 104*: University of Tampere, School of Management, Economics.

Lewis, A. 1982. The social psychology of taxation. *British Journal of Social Psychology* 21 (2):151-158.

Lowry, S. 2014. Itemized tax deductions for individuals: Data analysis. Washington, DC: Congressional Research Service.

Maciejovsky, B., E. Kirchler, and H. Schwarzenberger. 2001. Mental accounting and the impact of tax penalty and audit frequency on the declaration of income: An experimental analysis: Humboldt University of Berlin, Interdisciplinary Research Project 373: Quantification and Simulation of Economic Processes.

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

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

Muehlbacher, S., E. Kirchler, and H. Schwarzenberger. 2011. Voluntary versus enforced tax compliance: Empirical evidence for the "slippery slope" framework. *European Journal of Law and Economics* 32:89-97.

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

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.

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

Remus, W. 1996. Will behavioral research on managerial decision making generalize to managers? *Managerial and Decision Economics* 17 (1):93-101.

Schizer, D. M. 2015. Limiting Tax Expenditures. *Tax Law Review* 68 (2):275-354.

Slemrod, J. 2016. Tax Compliance and Enforcement: New Research and Its Policy Implications. In *Ross School of Business Paper No. 1302*.

Slemrod, J., B. Collins, J. L. Hoopes, D. Reck, and M. Sebastiani. 2017. Does credit-card information reporting improve small-business tax compliance? *Journal of Public Economics* 149:1-19.

Smith, S. 2012. Increasing Charitable Giving: What Can We Learn from Economics? *Fiscal Studies* 33 (4):449-466.

Spicer, M. W., and J. E. Thomas. 1982. Audit probabilities and the tax evasion decision: An experimental approach. *Journal of Economic Psychology* 2 (3):241-245.

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

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

Torgler, B. 2002. Speaking to Theorists and Searching for Facts: Tax Morale and Tax Compliance in Experiments. *Journal of Economic Surveys* 16 (5):657-683.

———. 2003. Tax morale and institutions. *SSRN Electronic Journal*.

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

Tversky, A., and D. Kahneman. 1974. Judgment under Uncertainty: Heuristics and Biases. *Science* 185 (4157):1124-1131.

Vossler, C. A., and S. M. Gilpatric. 2018. Endogenous audits, uncertainty, and taxpayer assistance services: Theory and experiments. *Journal of Public Economics* 165:217-229.

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

Appendix

A1 Additional information on ex-post questionnaire

At the end of the experiment (but before participants learn their final payoffs), participants are asked to answer a questionnaire. In particular, the following variables are collected: "age" (in years), "risk attitude" (obtained from the SOEP survey and gives the subject's self-reported general willingness to take a risk, measured on an 11-point scale where 0 = not willing to take risk and 10 = highly willing to take risk), "female" (female = 1, 0 otherwise), "economics" (1 if the participant experienced more than one lecture economics, 0 otherwise), "tax experience" (1 if the participant states that she ever completed a tax declaration in the past, 0 otherwise), "tax morale" (adapted question from the World Values Survey: "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), "fairness" (perceived fairness of tax and control system, measured on an 11-point scale where 0 = low perceived fairness of tax and control system and 10 = high perceived fairness of tax and control system), "decision complexity" (measures how complex a participant perceived the tax-related decisions in the experiment, 11-point scale from 0 = low perceived decision complexity to 10 = high perceived decision complexity), "monthly income" (monthly income in Euro after fixed costs such as rent), and "joy", "anger", "fear" and "guilt" (each variable measures the level of joy/anger/fear/guilt felt by an individual during the experiment, 11-point scale from 0 = felt no joy/anger/fear/guilt during experiment to 10 = felt high joy/anger/fear/guilt during experiment).

A2 Regressions

Table A2.1: Regression results: Prefilling of deductions with controls (corresponds to Table 3)

	Treatment 4i vs 4i prefilled				
Dependent variable	**Over-deductions**	**Over-deductions**	**Over-deductions**	**Over-deductions**	**Total over-deductions**
	Model 1	Model 2	Model 3	Model 4	Model 5
	Item 1	Item 2	Item 3	Item 4	All items
Treatment 4i prefilled	-124.43	-49.67	33.34	-338.69**	-479.47**
	(80.54)	(65.46)	(55.08)	(158.57)	(212.78)
Age	-14.38	-8.59	-4.18	4.62	-22.52
	(8.93)	(7.26)	(6.11)	(17.58)	(23.60)
Risk attitude	8.30	1.50	12.63	10.13	32.56
	(17.74)	(14.42)	(12.13)	(34.92)	(46.86)
Female	-82.53	-19.20	-20.12	37.67	-84.18
	(90.22)	(73.33)	(61.70)	(177.62)	(238.35)
Tax experience	129.30	-14.91	11.64	-40.37	85.66
	(86.91)	(70.64)	(59.43)	(171.11)	(229.61)
Tax morale	7.83	-4.11	-7.28	-7.23	-10.79
	(17.61)	(14.31)	(12.04)	(34.67)	(46.52)
Economics	-38.93	2.90	-28.02	-55.88	-119.93
	(79.96)	(64.99)	(54.68)	(157.43)	(211.26)
Decision complexity	14.58	-3.63	-5.03	-81.99**	-76.06*
	(17.20)	(13.98)	(11.76)	(33.86)	(45.44)
Monthly income	-0.15	0.35*	0.28*	-0.18	0.30
	(0.22)	(0.18)	(0.15)	(0.44)	(0.59)
Fairness	-9.75	-4.13	-7.63	18.07	-3.43
	(18.56)	(15.09)	(12.69)	(36.55)	(49.04)
Joy	14.92	-14.20	-5.94	-9.45	-14.68
	(16.97)	(13.79)	(11.61)	(33.41)	(44.84)
Anger	-13.21	-6.07	-15.52	-11.54	-46.33
	(16.54)	(13.44)	(11.31)	(32.56)	(43.69)
Fear	-18.72	1.99	-3.89	-8.30	-28.92
	(17.62)	(14.32)	(12.05)	(34.68)	(46.54)
Guilt	33.30*	-16.23	28.91**	-11.84	34.14
	(18.38)	(14.94)	(12.57)	(36.19)	(48.56)
Constant	408.89	342.36	202.07	862.37	1,815.69**
	(335.26)	(272.50)	(229.27)	(660.07)	(885.74)
No. of observations	78	78	78	78	78
R-squared	0.196	0.145	0.182	0.221	0.209

Note: This table shows the same regression results as presented in Table 3, but with all controls displayed. The results of linear regression models are presented with either overdeductions or total overdeductions as dependent variables (regression coefficients, standard errors in parentheses). Number of observations is determined by the number of subjects participated in the respective treatments. *** p<0.01, ** p<0.05, * p<0.1

Table A2.2: Regression results: Prefilling of deductions (with variable "evader")

	Treatment 4i vs 4i prefilled			
Dependent variable	**Evader**	**Evader**	**Evader**	**Evader**
	Model 1	Model 2	Model 3	Model 4
	Item 1	Item 2	Item 3	Item 4
Treatment 4i prefilled	-0.78	-166.03	-0.05	-1.00*
	(1.20)	(0.00)	(0.70)	(0.55)
Constant	-4.83	532.99	1.53	1.20
	(6.78)	(0.00)	(3.73)	(2.33)
Controls	Yes	Yes	Yes	Yes
No. of observations	78	78	78	78
R-squared	0.1135	0.1135	0.1135	0.1135

Note: In this table, the results of logistic regression models are presented with evader as dependent variable (regression coefficients, standard errors in parentheses). Number of observations is determined by the number of subjects participated in the respective treatments. *** p<0.01, ** p<0.05, * p<0.1

Table A2.3: Regression results: Disallowing the deductibility (with variable "evader")

	Treatment 4i vs 3i			**Treatment 3i vs 2i**	
Dependent variable	**Evader**	**Evader**	**Evader**	**Evader**	**Evader**
	Model 1	Model 2	Model 3	Model 4	Model 5
	Item 1	Item 2	Item 3	Item 1	Item 2
Treatment 3i	1.80	-0.37	1.33**		
	(3.66)	(1.32)	(0.65)		
Treatment 2i				3.16**	3.18***
				(1.25)	(1.23)
Constant	7.79	-1.03	-3.33	0.92	1.93
	(13.91)	(7.91)	(2.71)	(3.74)	(3.05)
Controls	Yes	Yes	Yes	Yes	Yes
No. of observations	82	82	82	77	77
R-squared	0.3650	0.3650	0.3650	0.3650	0.3650

Note: In this table, the results of logistic regression models are presented with evader as dependent variable (regression coefficients, standard errors in parentheses). Number of observations is determined by the number of subjects participated in the respective treatments. *** p<0.01, ** p<0.05, * p<0.1

Table A2.4: Regression results: Limiting the deductibility (with variable "evader")

	Treatment 4i vs 4i limited-deductibility			
Dependent variable	**Evader**	**Evader**	**Evader**	**Evader**
	Model 1	Model 2	Model 3	Model 4
	Item 1	Item 2	Item 3	Item 4
Treatment 4i limited-deductibility	0.73	0.90	0.33	-0.94
	(1.47)	(1.30)	(0.77)	(0.62)
Constant	7.07	8.22	4.36	4.52
	(12.62)	(7.26)	(4.26)	(2.89)
Controls	Yes	Yes	Yes	Yes
No. of observations	75	75	75	75
R-squared	0.0072	0.0107	0.1028	0.0967

Note: In this table, the results of linear regression models are presented with evader as dependent variables (regression coefficients, standard errors in parentheses). Number of observations is determined by the number of subjects participated in the respective treatments. *** p<0.01, ** p<0.05, * p<0.1

Supplementary material

M1 Instructions

Supplementary material M1 includes the translated instructions (from German). All participants received the general instructions in print (M1.1). Before the experiment was executed, the participants received the instructions for the first part of the experiment (M1.2). The instructions for the second part (decision on donation) were displayed on the computer screen (M1.3). After that, participants received the instructions for the third part of the experiment in print (M1.4).

M1.1 General instructions

Thank you very much for participating in this experimental study. For your participation, you will receive a lump sum of 4 Euros.

The experimental study consists of one experiment, in which you have the possibility to earn money, and a questionnaire at the end of the study. The amount of money you earn depends on your decisions in the experiment and on chance. The instructions explain to you how you can influence the amount of money you earn in this study with your decisions.

It is important that you understand the instructions. Hence, please do not hesitate to ask any questions. If you have a question, please raise your hand. We will then come to you to answer your question. Please do not ask your question loudly. You may write on the instructions and highlight them. Please do not take the instructions home, but return them to us at the end of the study.

The analysis of the experiment will be anonymous. We will on no account link your name to the data collected in the experiment. You will not get to know the identity of any other participant, neither before, nor after the experiment. Likewise, the other participants will not get to know your identity. At the end of the experiment, you will have to sign a receipt to confirm the payments you received. This receipt will only be used for accounting purposes.

We would like to point out that you are not allowed to communicate with other participants or leave your seat throughout the whole experiment. Please make sure to switch off your mobile phone.

The calculator, the pen and the notepad that are lying in front of you may be used.

The currency used in the experiment is ECU, where 100 ECU equal 1 Euro.

At the end of this experimental study, you will receive your payout privately and in cash. Your total payout consists of your payout of the experiment and the lump sum of 4 Euros for your participation.

The instructions for the experiment will be handed out shortly.

M1.2 Instructions for the first part of the experiment

Procedure of the experiment

The experiment consists of three parts:

- In part 1, you solve puzzle tasks and thereby generate income.
- In part 2, you make a decision about a donation.
- In part 3, you fill out a tax return. Based on the tax return, a tax will be determined. The tax amount is among others dependent on your income.

In the beginning of each part, you will receive instructions which describe the respective part of the experiment.

The payout that you will receive at the end of the experiment is dependent on your decisions and statements in the parts 1 to 3.

Credit account

During the experiment, you receive a personal credit account on which your payout relevant amounts will be posted on during the experiment. Your account balance will be displayed to you after each part of the experiment.

Your account balance at the end of the experiment determines your personal payout of the experiment. For this purpose, your account balance will be converted to Euros and added to your lump sum of 4 Euros for your participation and your payouts of the comprehension tests.

General information

Comprehension Test

Prior to part 1 and 3, respectively, you will be asked to do a comprehension test on your computer. The comprehension test consists of different blocks of questions. Each question block consists of up to four questions. For answering those questions correctly, you will receive an additional payout. The following applies:

If you answer all questions of a question block …

- … correctly on the first try, you will receive 1 Euro.
- … correctly on the second try, you will receive 0.50 Euro.
- … correctly on the third or on more tries, you will receive 0 Euro.

Questionnaire

After the experiment, you will be asked to answer a couple of questions on your computer. Please answer the questions thoroughly as they contain important information for us.

General Advice

The used program separates decimal places with a point and not with a comma (e.g. “2.5” instead of “2,5”).

Puzzle Task

The first part of the experiment consists of a mathematical puzzle task. The screen of your computer will display a puzzle with a matrix of 12 numbers. From this matrix, you should select the two numbers that add up to 10. In each matrix with 12 numbers, there are only 2 numbers that exactly add up to 10. All numbers have two decimal places.

The screen with the matrix looks as follows:

Part 1 consists of one non-payout relevant practice round and **4 payout relevant rounds**. In each round, you have 3 minutes time to solve as many mathematical puzzles as you like. A maximum of 20 puzzles can be solved correctly in each round. After each round, you have a break of 1 minute before the next round starts.

Your earned income depends on the number of puzzles you solved correctly. For each puzzle that you solved correctly, you receive an income of 93 ECU. For each puzzle that you solved incorrectly, you receive an income of 0 ECU. Your income in each round is calculated as follows:

Earned income per round = number of correctly solved puzzles in the respective round x 93 ECU

If you solve every puzzle correctly in one round, you will consequently earn 1,860 ECU 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.

Social Insurance Contribution

A fictitious social insurance contribution of 22.6% of your income will be retained. This amount will be deducted automatically and not paid out to you.

Simplifying the puzzle task

In order to simplify the puzzle task, you can reduce the number of selectable numbers so that you can recognize the right combination among the remaining numbers quicker.

Removing numbers involves costs. The costs are:

- "Remove 1 number": 37 ECU per round
- "Remove 2 numbers": 142 ECU per round
- "Remove 3 numbers": 299 ECU per round

Of course you can also choose not to remove any numbers ("Remove no numbers"). Thus, you do not have costs.

At the beginning of each round, you decide how many numbers you want to remove. Your decision applies for all puzzles in that round. The costs incur once per round and are deducted automatically from your income at the end of the round. In the following round you can make a new decision if you want to remove numbers.

On the reverse side of these instructions, you can find examples for all 4 options.

Please note that it is not possible to remove numbers in the practice round.

General overview and notepad

At the end of the first part, your total income of all 4 rounds, the total amount of retained social insurance contributions and your total costs to simplify the puzzle tasks will be displayed to you in a general overview.

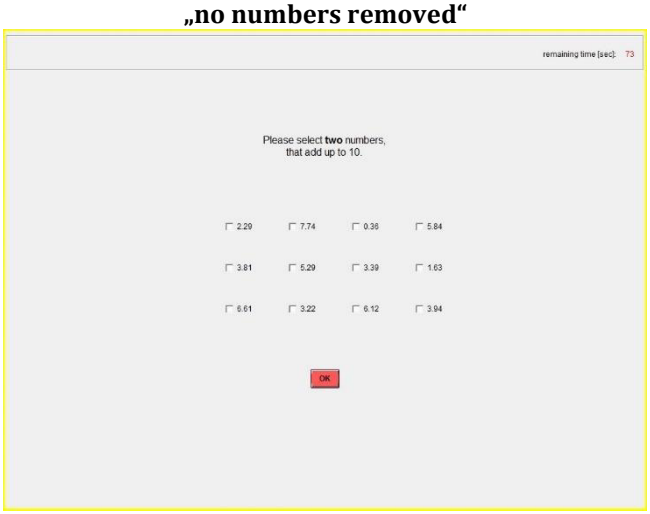
You will need these information when you fill out the tax return **in part 3**. Hence, please write down these information on the notepad at your place.

Please do not return the notepad to us, but take it home or throw it in trash.

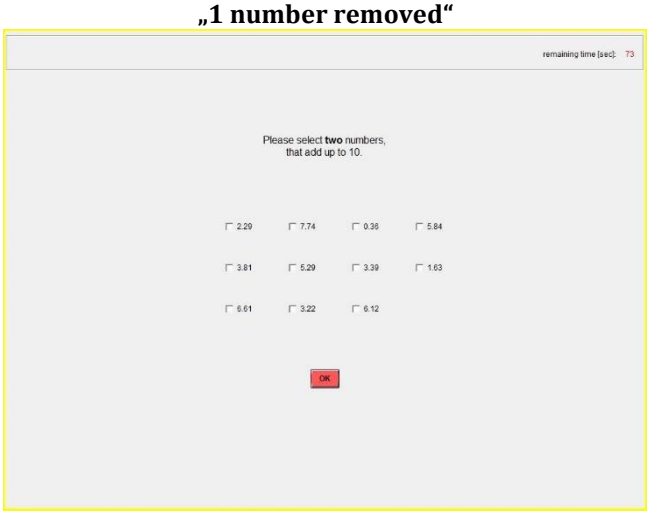
Credit account

After part 1 of the experiment, your total income, which you generated in the 4 rounds of the puzzle task, will be posted on your credit account while the total amount of social insurance contributions and the costs for simplifying the puzzle task will be deducted.

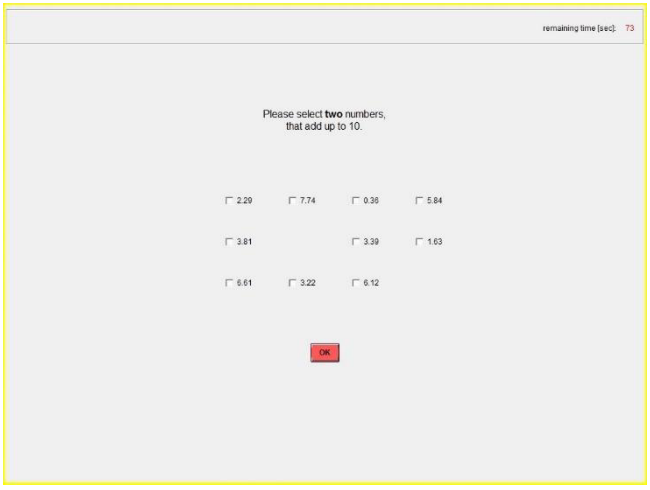
„no numbers removed“



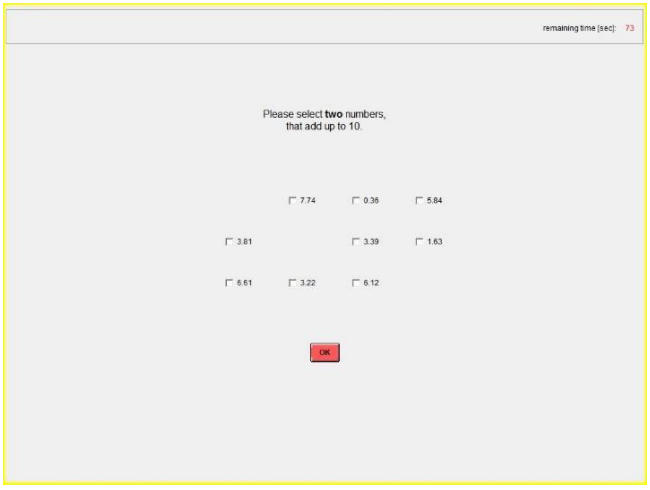
„1 number removed“



„2 numbers removed“



„3 numbers removed“



M1.3 Instructions for the second part of the experiment

In part 2 of the experiment, you have the possibility to donate part of your income.

The computer program will show you various charitable organizations. You may choose one or more organizations. If you choose more than one organization, your donation will be split equally between these organizations. If you do not choose any organization, the computer program will randomly choose one organization.

After the execution of the experimental study, the donations will be transferred from the chair of Behavioral Accounting/Taxation/Finance (Prof. Dr. Martin Fochmann) to the respective organizations. You may later receive proof for these transactions on demand.

Once again, please note the amount that you donated on your notepad for part 3.

Credit account

After part 2, the amount that you donated will be deducted from your credit account.

M1.4 Instructions for the third part of the experiment

Tax return

For this study, you now have to complete a fictional tax return in order that a fictional tax can be determined. This means that in part 3, you shall declare the expenses you had in part 1 and 2.

With your tax return, a declared taxable income will be determined, which depends on your total income of the puzzle task (part 1) and on your declaration of expenses in the following.

The tax rate is 40% on the declared taxable income and will automatically be retained. The tax revenue stays at the University of Cologne.

[Treatments: 4i, 4i prefilled, 4i limited-deductibility]

The declared taxable income is determined according to the following scheme:

Total income of the puzzle task (part 1)
– declared amount of social insurance contributions from part 1
– declared total costs for the simplification of the puzzle tasks from part 1
– declared donation from part 2
– declared commuting allowance
= declared taxable income

[Treatment 3i]

The declared taxable income is determined according to the following scheme:

Total income of the puzzle task (part 1)
– declared amount of social insurance contributions from part 1
– declared total costs for the simplification of the puzzle tasks from part 1
– declared donation from part 2
= declared taxable income

[Treatment 2i]

The declared taxable income is determined according to the following scheme:

Total income of the puzzle task (part 1)
– declared amount of social insurance contributions from part 1
– declared total costs for the simplification of the puzzle tasks from part 1
= declared taxable income

The total income of the puzzle task is prefilled by the computer and cannot be changed.

Your task is to complete the **[Treatments: 4i, 4i prefilled, 4i limited-deductibility:** 4**; Treatment 3i:** 3**; Treatment 2i:** 2**]** expense form fields. Based on your declared expenses, the declared taxable income is determined, which is the basis for the tax calculation. Your declared expenses can be smaller, equal to or larger than your actual expenses.

Please keep in mind to always enter the total expenses in the form fields and not for example the expenses of one particular round of the puzzle task.

[Treatments: 4i, 4i prefilled, 4i limited-deductibility]
One exception to this is the commuting allowance. The commuting allowance compensates the costs for your way from your apartment to the experimental laboratory. For this purpose, please enter the shortest way from your apartment to the experimental laboratory (only one-way). Please bring up to a round figure. The computer will then automatically calculate the deductible amount for the commuting allowance. For every declared kilometer, the commuting allowance is **[Treatment 4i, 4i prefilled:** 30**; Treatment 4i limited-deductibility:** 10**]** ECU. You are allowed to use an app (e.g. Google Maps) on your smartphone to find out the distance in kilometers. The maximum of declarable kilometers is 50.

[Treatment 4i prefilled]
The 4 expense form fields are prefilled by the computer. The prefilled form fields might deviate from your actual expenses. Please check the prefilled expense form fields in your tax return.

After filling out all the **[Treatments: 4i, 4i prefilled, 4i limited-deductibility:** 4**; Treatment 3i:** 3**; Treatment 2i:** 2**]** expense form fields, you can press the button "submit tax return" to submit your tax return. After submitting your tax return, you cannot make any changes.

Before you submit your tax return, you can also press the button "calculate tax". Then, your declared taxable income and the resulting tax will be displayed on your screen. If you want to change your declarations, you can press the button "change tax return". You can press "calculate tax" and "change tax return" as often as you want until you are done with your tax return. When you want to submit your tax return, press the button "submit tax return".

Please keep in mind that your declared expenses must not result in a negative taxable income.

Tax Payment

The payable tax is 40% of your declared taxable income.

$$Tax = 0.40 \times declared\ taxable\ income$$

Audit of the tax return

With a probability of 30%, your tax return will be audited, checking if your declared expenses in the tax return coincide with your actual expenses. If your tax return is audited and the declared expenses do not coincide with your actual expenses, you have to pay a fine. The fine is twice the amount of the not-payed tax.

Not-payed tax = 0.40 x (actual taxable income – declared taxable income)
Fine = 2 x not-payed tax

If the audit comes to the conclusion that you payed too much taxes (because your declared expenses are smaller than your actual expenses), the amount which was payed too much will be refunded. In this case, no fine applies.

Your personal payout from the experiment

After submitting your tax return in part 3, the resulting tax burden of the tax return is deducted from your credit account. If your tax return is audited and the declared taxable income does not coincide with the actually taxable income, the resulting fine is deducted from your credit account as well. If you receive a tax refund, it will be posted on your account.

Your account balance at the end of the experiment is your personal payout from the experiment. Therefore, your payout is determined as follows:

Your payout =
total earned income from the puzzle task (part 1)
– actual total social insurance contributions from part 1
– actual total expenses for the simplification of puzzle tasks from part 1
– actual donation from part 2
– payable tax determined based on the declared taxable income from part 3
– possible fine
+ possible tax refund

Please keep in mind that your declarations in the tax return only affect the payable tax (and possibly the fine or tax refund). The amount of your total income, the amount of social insurance contributions, the amount of expenses to simplify the puzzle tasks or the amount of your donation are not affected.

The payout will be converted to Euros and added to your lump sum of 4 Euro for your participation and your payouts from the comprehension tests. The resulting amount of money will afterwards be paid out to you in cash.

M2 Comprehension Tests

M2.1 Comprehension test before part 1

1. How many payout relevant rounds has the first part of the experiment?

2. Which amount (in ECU) do you earn for every correctly solved puzzle?

3. Which statement is correct?
   - *One round of the puzzle task consists of one puzzle and takes 3 minutes.*
   - *One round of the puzzle task consists of maximal 20 puzzles and takes 3 minutes.*
   - *One round of the puzzle task consists of infinitely many puzzles and takes 3 minutes.*

4. Which statement is correct? In order to simplify the puzzle task, you can delete irrelevant numbers
   - *You can decide on it at the beginning of each single round. Thus, costs can incur in every round.*
   - *You can decide on it at the beginning of each single round. There are no costs.*
   - *You decide on it at the beginning of the experiment for the entire experiment. Thus, costs can incur only once.*
   - *You decide on it at the beginning of the experiment for the entire experiment. There are no costs.*

M2.2 Comprehension test before part 3

[Treatment 4i, 4i prefilled, 4i limited-deductibility]

1. Regarding the commuting allowance, which statement is correct?
   - *The declared kilometers do not affect the declared taxable income.*
   - *For every declared kilometer, the declared taxable income is reduced by 1 ECU.*
   - *For every declared kilometer, the declared taxable income is reduced by* **[Treatment 4i, 4i prefilled:** *30***; Treatment 4i limited-deductibility:** *10***]** *ECU.*

2. Is it possible that your declared expenses in the tax return deviate from your actual expenses (i.e. that they are lower or higher)?
   - *Yes.*
   - *No.*

3. Which of the following statements regarding the calculation of the tax is correct?
   - *The tax amounts to 40% of the actual income.*
   - *The tax amounts to 40% of the declared taxable income.*

4. What is the probability (in percent) of an audit of your tax return?

5. Do your declarations in the tax return affect your actual earned income from the puzzle task?
   - *Yes.*
   - *No.*

[Treatment 4i, 4i prefilled, 4i limited-deductibility, 3i]

6. Do your declarations in the tax return affect the actual amount of your social insurance contribution or the actual amount of your costs to simplify the puzzle task or the actual amount of your donation?

- o *Yes.*
- o *No.*

[Treatment 2i]

6. Do your declarations in the tax return affect the actual amount of your social insurance contribution or the actual amount of your costs to simplify the puzzle task?

- o *Yes.*
- o *No.*

7. Your declarations in the tax return solely affect the amount of tax (and possibly the fine or the tax refund) and not on other matters?

- o *Yes.*
- o *No.*

8. For a person who wants to pay the correct amount of taxes, it applies that:

- o *Declared expenses < actual expenses.*
- o *Declared expenses > actual expenses.*
- o *Declared expenses = actual expenses.*

9. For a person who wants to pay less taxes than the correct amount of taxes, it applies that:

- o *Declared expenses < actual expenses.*
- o *Declared expenses > actual expenses.*
- o *Declared expenses = actual expenses.*

10. For a person who wants to pay more taxes than the correct amount of taxes, it applies that:

- o *Declared expenses < actual expenses.*
- o *Declared expenses > actual expenses.*
- o *Declared expenses = actual expenses.*

11. Do your declarations in the tax return affect your personal payout from this experiment?

- o *Yes, because they affect the amount of taxes to be paid.*
- o *No.*

CHAPTER 4

Systemization and review of non-monetary costs of tax evasion[1]

Tobias Kölle
University of Cologne

September, 2020

Abstract

The pervasive problem with tax evasion of individuals was commonly examined by analyzing the monetary costs of tax evasion (e.g. audit probability, fine and tax rate). Only recently, the analysis has been expanded to non-monetary costs of tax evasion. This paper systemizes this area of research and reviews its findings including latest developments. In particular, the paper reviews seven sources of non-monetary costs of tax evasion, namely social norm, trust, fairness, participation rights, complexity and knowledge, patriotism, and moral emotions. Finally, I discuss interdependencies between these sources of non-monetary costs of tax evasion on the one hand and between non-monetary and monetary costs of tax evasion on the other.

[1] For helpful comments and suggestions thanks are due to Martin Fochmann und Michael Overesch

Content

Tables

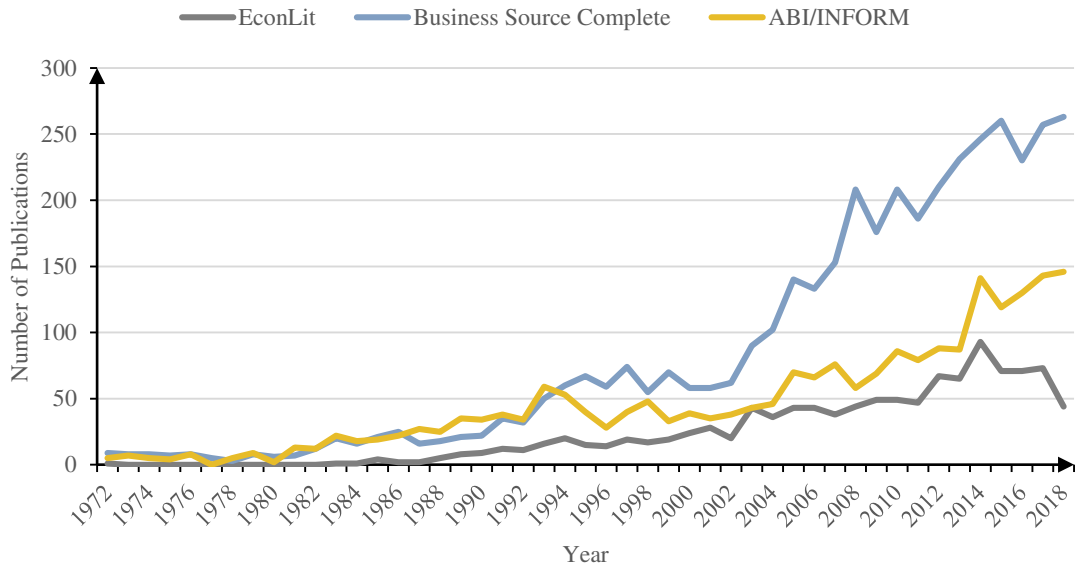
Figures

4.1 Introduction

"Of course not, I'm not an idiot, says Liu Yongli, a chauffeur in Beijing, when asked whether he has ever paid personal income tax" (The Economist, 2018, p.60). And indeed, there is a significant economic damage caused by tax evasion in any modern society, see Slemrod (2016). However, this must not disguise the fact that (fortunately) the vast majority of individuals is tax compliant. At first glance this is not surprising at all. In fact, paying taxes is legally obligated and non-compliance is penalized. However, empirical data emphasizes that only a fraction of widespread compliance can be explained by monetary costs of tax evasion (i.e. deterrence factors like audit probability, fine and tax rate) alone. Over the last years of research, it has become apparent that non-monetary costs of tax evasion explain the other fraction of observed compliance. Accordingly, the decision on paying or evading taxes is not a gamble for individuals maximizing their individual monetary benefit. Rather, the decision on paying or evading taxes is related to sources of non-monetary costs of tax evasion like social norms of compliance, trust in authorities and governments, concern regarding fairness and/or participation rights.

The literature review at hand contributes to the tax compliance literature by shedding light on the jungle of non-monetary costs of tax evasion. It selects, organizes and integrates information from roughly 170 papers into a comprehensive framework of tax compliance from the perspective of an economist. This is of special importance for multiple reasons. First, research regarding tax compliance has developed fast recently (see Figure 1). This is not surprising as tax evasion remains a huge and unsolved problem for any modern state in the world to this day. Institutions like the OECD and the World Bank emphasize the importance of research regarding voluntary tax compliance (namely non-monetary costs of tax evasion) for the future of taxation, see OECD (2019). Furthermore, new technologies like mobile laboratories and new measurement capabilities to measure for example eye movement, blood pressure, heart rate, neuronal activity, or skin resistance offer new possibilities in answering research questions. However, not everything that is possible is necessary and it is important to recap and gather previous research in order to identify and highlight new and promising research tracks. Second, even though the academic literature already contains excellent literature reviews on the monetary costs of tax evasion (mainly deterrence), there is a lack of

literature reviews focusing on the non-monetary costs of tax evasion.[2] Third, research regarding non-monetary costs of tax evasion lacks a uniform technical language. One of the reasons for this phenomenon is that research on non-monetary costs of tax evasion attracts (and requires) researchers from different departments (like Business Administration, Economics, Politics, Law, Psychology and Medicine), who import their own technical language into tax compliance research. This results in confusing labeling and complicates comprehension, interpretation and comparison of research. The problem is made all the more acute by the fact that research has referred to sources of tax evasion in a rather selective fashion. This paper aims to provide a more systematic analysis of sources of non-monetary costs of tax evasion and encourages to develop a common terminology.



Note: Figure shows the number of released (reviewed) publications on tax compliance for the years 1972-2019 as they are listed in the databases EconLit, Business Source Complete and ABI/INFORM for the search term "tax compliance".

Figure 1: Increasing research on tax compliance.

Although the review at hand includes the research of roughly 170 papers, it cannot claim to be comprehensive. Importantly, it focuses on tax compliance related to individuals and small businesses and not on tax compliance of multinational companies. The literature search for this paper was concluded in July 2020.

[2] The literature review at hand integrates in a series of literature reviews regarding tax compliance: For a general review see Alm (2019), for a review of theoretical models see Hashimzade et al. (2013), for a review with a focus on surveys see Slemrod (2007), for reviews with a focus on field experiments see Mascagni (2018) and Hallsworth (2014), for a review with a focus on laboratory experiments see Torgler (2002). For reviews most comparable with the study at hand see Hofmann et al. (2008) and Luttmer and Singhal (2014).

This paper continues as follows: section 4.2 describes the development of costs of tax evasion in theory and introduces non-monetary costs of tax evasion. Section 4.3 outlines the main research methods employed in the context of non-monetary costs of tax evasion and raises important methodological considerations relating to findings as they arise throughout the paper. Section 4.4 reviews seven sources of non-monetary costs of tax evasion: social norm, trust, fairness, participation rights, complexity and knowledge, patriotism, and moral emotions. Section 4.5 covers interdependencies both between different sources of non-monetary costs of tax evasion and between monetary and non-monetary costs of tax evasion. Finally, section 4.6 provides concluding remarks on future research.

4.2 The costs of tax evasion

In the beginning of the tax compliance literature, Allingham and Sandmo (1972) presented the first model of tax compliance, based on the economics-of-crime approach by Becker (1968). It relies on three variables: audit probability, tax rate and fine. Given these three variables of deterrence, a rational taxpayer decides on her compliance. Her decision is therefore only dependent on financial aspects. In other words, evading taxes faces specific monetary costs of tax evasion. Consequently, an individual only pays taxes if monetary benefits of tax evasion are bigger than monetary costs of tax evasion. This simplistic structure allows a straight forward analysis of tax compliance and explains its importance in tax compliance literature to this day. Furthermore, the model implies direct policy recommendations for governments combating tax evasion. As compliance depends solely on enforcement, governments enhance tax compliance by increasing audit probabilities, fines and/or tax rates (Yitzhaki, 1974). However, the substantial drawbacks of the deterrence model cannot be overlooked, even Allingham and Sandmo (1972) themselves indicated its constraints.[3]

There are two major points of criticism against the standard model of deterrence. First, empirical evidence on the three deterrence variables audit probability, fine and tax rate does not support the policy implications of the model mentioned above. Increasing audit probability, fine and/or tax rate does not necessarily result in better tax compliance. For example, recently Mendoza et al. (2017) have observed empirically that tax

[3] Allingham and Sandmo (1972) themselves stated that their model gives "too little attention to nonpecuniary factors in the taxpayer's decision" (p.326).

compliance increases with rising audit probabilities but decreases after a certain threshold (supported by Slemrod et al., 2001 and already by Weck-Hannemann and Pommerehne, 1989).[4]

Second, the prediction of tax compliance based on the deterrence model is rather poor. Given reasonable input parameters for the three deterrence variables, the purely economic analysis predicts much higher levels of tax evasion than actually observed (for example Slemrod, 2007 with archival data and Kleven et al., 2011 with experimental data) or would induce abnormally high risk aversion (see for example the criticism of Kirchler, 2007; Andreoni et al., 1998, Graetz and Wilde, 1985 and Torgler, 2007). Given these insights, Alm et al. (1992) formulated that the puzzle of tax compliance behavior is why people pay taxes and not why they evade them.

The constraints of the deterrence model demonstrate the need for improvement. Several researchers extended the model with factors for example accounting for employer withholding (Kleven et al., 2011 and Alm et al., 2016), rewards on compliance (Falkinger and Walther, 1991) or strategic audit selection rules (see for example recently Kuchumova, 2017). However, these models still capture only the monetary costs of tax evasion. Furthermore, they still rely on controversial assumptions of the neoclassical economic model like perfect rationality, outcome orientation and egoism of individuals.[5]

A more promising way to develop a framework on tax compliance relies on behavioral economics. More precisely, economic models have to accept and consider non-financial aspects that influence tax compliance. All these non-financial aspects are sometimes gathered under the umbrella term of tax morale (Luttmer and Singhal, 2014). However, tax compliance literature failed to develop a common technical language. Researchers use the terms psychic costs, moral costs, intrinsic motivation, social norm, tax ethics, intrinsic tax morale and tax honesty but refer mostly to the same object of investigation.[6] This confusing terminology and the lack of systemization impede comprehension, interpretation and comparison of research. Furthermore, if tax compliance research adopts terminology from other sciences (like psychology) it adopts also already existing

[4] A literature review on the effects of deterrence variables on tax compliance is for example provided by Andreoni et al. (1998).

[5] More constraints are for example bounded rationality and mental accounting (Muehlbacher et al., 2017), fiscal illusion, salience, limited attention and overweighting of probabilities, hyperbolic discounting, reference points, gains versus losses, loss aversion (Rees-Jones, 2014 and Engström et al., 2015), risk-seeking behavior and status quo bias. These and more aspects are discussed in Alm (2019) and Luttmer and Singhal (2014).

[6] For example, Frey (1997b) interprets tax morale as an intrinsic motivation to pay taxes. Contrary, Luttmer and Singhal (2014) integrate intrinsic motivation under the umbrella term of tax morale. And finally both terms are oftentimes used as synonyms (e.g. Alm and Torgler (2006) and Slemrod (2016)).

controversy of these terms. For example, the term intrinsic motivation is based on a concept in psychology and is often used in tax compliance literature nowadays. However, there is a highly spirited debate about this term in psychology itself and its problematic definition.[7] This has reached a point where some researchers in psychology even appeal to abandon this term in general (Rheinberg and Engeser, 2018). From the perspective of an economist, I reject tiptoeing through this minefield. More importantly, from the perspective of an economist, one can still think of tax evasion as a tradeoff between benefits and costs of tax evasion. However, the costs must be divided into two categories, monetary and non-monetary costs of tax evasion. Consequently, the incorporation of these non-financial aspects does not neglect deterrence mechanisms. The individual's utility function has rather a multifaceted nature (Cullis and Lewis, 1997, Alm and Torgler, 2011). Obviously, the deterrence variables of the classic tax evasion model are the monetary costs of tax evasion (also rewards on compliance would be gathered under this term, see Falkinger and Walther, 1991). For example, a higher audit probability increases the monetary costs of tax evasion for an individual. Contrary, the non-monetary costs of tax evasion are less specific and capture variables like social norm, trust, concerns regarding fairness, participation rights and so on. Figure 2 shows this concept graphically.[8]

Conclusion 1: *Solely considering monetary costs of tax evasion fails predicting tax compliance. From the perspective of an economist, a solution approach for the puzzle of tax compliance has to consider not only monetary costs but also non-monetary costs of tax evasion.*

[7] In psychology research Ryan and Deci (2000) define intrinsic motivation as "doing something because it is inherently interesting, satisfying, or enjoyable", contrary to extrinsic motivation that refers to "doing something because it leads to a separable outcome (such as receiving a financial award from a third party)".

[8] In the proposed framework, non-monetary costs of tax evasion equals the term tax morale, which appears frequently in studies for example of Luttmer and Singhal (2014). Loosely spoken, both terms are umbrella terms for any "voluntary" tax compliance (e.g. social norm, trust, fairness and so on).

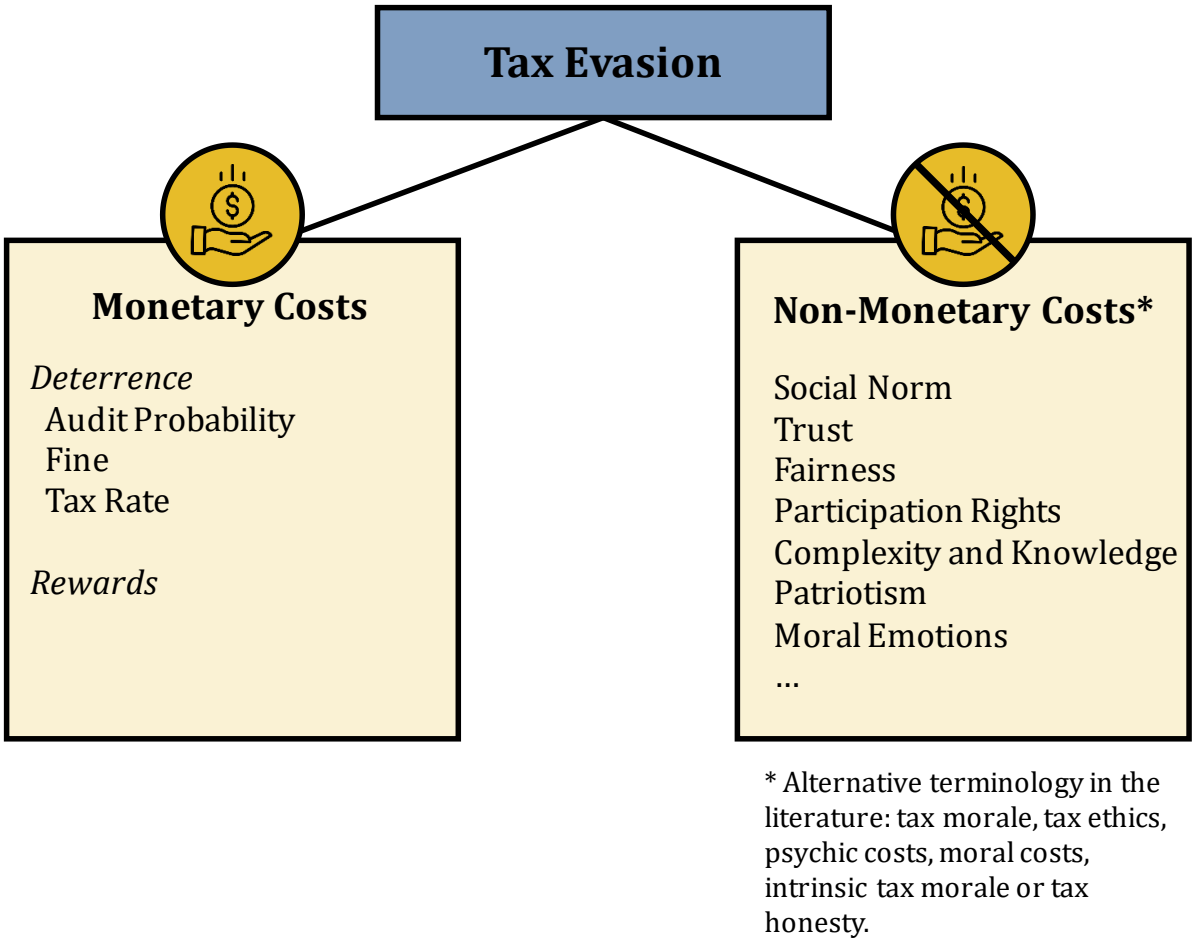


Figure 2: Monetary and non-monetary costs of tax evasion.

The importance and existence of non-monetary costs of tax evasion can be estimated on different ways. Surveys asking for tax attitude like the World Value Survey provide first evidence on its importance. For example, Alm and Torgler (2006) investigated tax attitude in the USA and 15 other European countries with information from the World Value Survey and found that individuals in the USA have the most honest tax attitude, followed by Austria and Switzerland. However, participant's answers in a survey do not necessarily reflect their actual behavior (a common problem with interpreting survey data, see section 4.3.1). Another approach estimating the importance of non-monetary costs of tax evasion is by comparing the gap between predicted level of tax evasion based on the monetary costs of tax evasion and tax evasion observed in reality or experimentally. For example, Dwenger et al. (2016) and, more generally, Fochmann et al. (2020b) observe compliant behavior even in the absence of monetary costs of tax evasion (i.e. without audits).[9] Other researchers measure differences in tax compliance between countries in

[9] Likewise, voluntary cooperation is also observed in standard economic game theory like dictator games (Engel, 2011).

otherwise identical circumstances (i.e. identical monetary costs of tax evasion), which is sometimes labeled as culture effect, see Luttmer and Singhal (2014). On the one hand, researchers conduct identical laboratory experiments in different countries, see for example Alm et al. (1995) comparing Spain and the U.S., Cummings et al. (2009) comparing the U.S., South Africa and Botswana, Gërxhani and Schram (2006) comparing Albania and the Netherlands and Lefebvre et al. (2015) comparing Belgium, France and the Netherlands. Even though there are differences, results remain mixed.[10] On the other hand, researchers measure tax compliance within one country by comparing individuals with migrant background with individuals without migrant background. Halla (2012) observes better tax compliance behavior for American-born individuals in the United States whose countries of ancestry have on average a more honest tax attitude. Kountouris and Remoundou (2013) confirm this finding for first-generation immigrants of European countries. In the U.S., DeBacker et al. (2015) observe that corporations with owners from countries with higher corruption norms evade more taxes. However, cross-culture experiments have to be treated with caution, because experimenter, language, and currency effects might disturb the observed results (see Roth, 1995). Furthermore, there might be already significant regional differences regarding tax attitude within one country. Consequently, participants of such experiments would have to be representative for a whole country. Finally and most importantly, measuring so called culture effects lacks clarity. The rationale behind observed culture differences is notoriously unclear. In all likelihood, factors like social norm and/or trust are the real drivers of observed culture differences.

Recognizing the existence and importance of non-monetary costs of tax evasion as a driver of tax compliance leads to the question of its sources. There is a horrendous number of factors discussed by researchers (even though researchers do not explicitly integrate them under the term of non-monetary costs of tax evasion). However, up to now there is no theoretical framework that examines non-monetary costs of tax evasion

[10] Andrighetto et al. (2016) test for cultural differences of Swedes and Italians in a laboratory setting. They observe that Swedes are more likely to be either completely honest or completely dishonest while Italians are more likely to cheat by a small amount. Zhang et al. (2016) test for cultural differences between participants from United Kingdom and Italy in a laboratory experiment. However, their findings suggest that cross-country differences in tax compliance cannot be explained by a lack of morality amongst southern European taxpayers. Guerra and Harrington (2018) compare results of a laboratory experiment in Denmark and Italy and find out that individual self-reported tax attitude cannot predict actual tax compliance behavior. Although previous research indicates a better tax attitude for citizens of Denmark in comparison with citizens of Italy, the tax compliance observed in the experiment was lower in Denmark.

systematically. Rather, comprehension, interpretation and comparison of research is again disturbed by confusing labeling of the same phenomenon.

One of the most prominent conceptualization of monetary and non-monetary costs of tax evasion (although with other terminology) was developed by Kirchler et al. (2008). They introduced a slippery-slope framework that summarizes variables affecting tax compliance in two dimensions, power and trust. The authors assume that the impact of changes in one dimension depends on the level of the other dimension, what results graphically in a slippery-slope. The power dimension captures enforced compliance (i.e. monetary costs of tax evasion) and the trust dimension captures voluntary compliance (i.e. non-monetary costs of tax evasion). Consequently, tax payments can be increased by increasing trust in authorities and/or power of authorities. The authors argue that other variables like social norm and fairness affect the trust dimension and therefore tax compliance indirectly. Even though the assumption of the slippery-slope framework regarding the interdependency of trust with other variables of non-monetary costs of tax evasion is empirically valid (see section 4.5), it neglects direct effects of these variables on tax compliance and it neglects variables at all, that do not interact with trust.

In line with the slippery-slope framework, the framework proposed in the study at hand is based on two dimensions. However, the focus of the framework proposed in this paper is wider. First, the monetary costs of tax evasion not only capture enforced compliance by financial-deterrence but also compliance by financial rewards. Second, the non-monetary costs of tax evasion not only capture trust but also (the effect of) other variables of non-monetary costs of tax evasion (like social norms, concerns regarding fairness and participation rights) directly.

4.3 Remarks on the Measurement of non-monetary costs of tax evasion

Before I review the different sources of non-monetary costs of tax evasion in more detail, I share some remarks on the methods of data collection for research regarding non-monetary costs of tax evasion. Indeed, the advantages and disadvantages of the different methods might explain differences in results and might complicate its generalization.

Likewise research in the broader field of tax evasion, research of non-monetary costs of tax evasion faces naturally the lack of reliable data. There is a well-known quote of Cowell (1991) summarizing the problem: “Data from official investigations are hardly ever available and data from other sources may be suspect: if you could directly observe and measure a hidden activity, then presumably it could not really have been properly

hidden in the first place" (p.123). Keeping this problem in mind, there are different data sources to find empirical answers to the puzzle of tax compliance and more precisely on the sources of non-monetary costs of tax evasion (testing of theoretical models or uncoupled explorative research). Research distinguishes broadly natural data and data derived from natural data,[11] survey data and experimental data. All of them come with different advantages and drawbacks but the pervasive problem is that of reliable data.[12]

With respect to non-monetary costs of tax evasion, surveys and experimental studies are the most important source of information as natural data regarding non-monetary costs of tax evasion does hardly exist. Experimental studies are further distinguished in laboratory experiments and field experiments. In particular studies based on laboratory experiments are increasing which is also true for other disciplines in social sciences, see Falk and Heckman (2009). In the following I highlight advantages and drawbacks of these three data sources and raise important methodological considerations relating to findings as they arise throughout the paper, the importance of design for credibility and comparability and why results have to be interpreted with some caution.

4.3.1 Surveys

Surveys simply ask individuals on their tax evasion behavior or their hypothetical behavior in a hypothetical situation. Researchers can conduct their own surveys or use data of professional surveys like the World Value Survey. Major advantages of surveys are data availability, large number of observations and relatively low costs. However, the reliability of survey data is questionable. Participants might not remember their true behavior, they might not reveal their true behavior (consciously or unconsciously), and/or the pool of participants might not be representative. Consequently, surveys may not accurately reflect the actual behavior of individuals in reality (Fishbein and Ajzen, 1975). This issue is not limited to the tax compliance literature. In fact, it is also a pervasive problem in other sciences using survey data.

4.3.2 Laboratory experiments

Most laboratory tax compliance experiments follow a similar design. After reading the instructions and answering a (sometimes incentivized) comprehension test, participants earn (or simply receive) their pre-tax income. This is for example done with a real effort

[11] See notably Henderson et al. (2012), who estimate true economic activity with night light observed from outer space; data that can be compared with official accounts in order to estimate shadow economy and tax evasion.
[12] See Alm (2019) for more details on data collection in tax compliance research.

task like the puzzle task of Mazar et al. (2008). Afterwards, participants have to file a tax return. Hypothetical taxes are paid on the reported income and with a certain probability participants are audited. If the audit reveals that a participant evaded taxes (e.g. reported income was lower than true income), she has to pay a fine (often a multiple, e.g. 2, of evaded taxes). This procedure repeats for a certain number of rounds. At the end of the experiment, participants answer a questionnaire that collects additional data (especially demographics) and they receive their payoffs. Given this basic design, researchers manipulate a certain variable (e.g. change in audit probability, tax rate or fine) or integrate new elements (e.g. voting on tax-spending or public good provision). The latter is especially relevant for measuring non-monetary costs. Finally, the treatment-effect can be measured by comparison of the manipulated treatment with the base case.

Laboratory experiments tackle the major drawback of surveys mentioned above, as researchers observe real behavior. The advantages of laboratory experiments are low costs for data (in comparison with field experiments), replicability and high control of environment and parameters. The internal validity is therefore generally high. Indeed, laboratory experiments might be the only possible source of data for a wide range of research questions regarding non-monetary costs of tax evasion that are often times in the intersection of tax evasion and psychology. Problems might arise on the external validity. The interpretation and generalization of the results of laboratory experiments have to be treated with caution. The laboratory setting is artificial and participants know that they are part of a study. Furthermore, the results might be sensitive to the experimental design (Alm, 1991). Laboratory experiments lack a real social context and real consequences so that participants might perceive a "game-environment". Typically, the stack-size is rather small as deciding on compliance with an experimental income of for example 4€ in a laboratory experiment might be fundamentally different than behavior facing an annual tax return. Moreover, penalties like jail cannot be simulated. Especially in the beginnings of laboratory experiments, the total number of participants can be insufficient for statistical power. However, one of the most common criticism on tax compliance experiments, that participants of laboratory experiments are normally students without any tax experience, was rebutted among others by Choo et al. (2016). Importantly, results of laboratory experiments do not predict levels of tax evasion for reality. Rather, they predict a tendency how individuals react to changes in certain variables (that were measured by treatment differences). Accordingly, Alm (2010) concludes that many concerns in the context of laboratory experiments seem largely

unwarranted. The critique on tax compliance experiments is for example summarized and addressed by Alm et al. (2015) and Torgler (2002).

4.3.3 Controlled field experiments

A controlled field experiment is conducted in the natural environment of the participants without their knowledge. Typically, the pool of participants is divided. Some individuals receive a treatment-message (often times an official letter of the tax authority)[13] that indicates for example with respect to non-monetary costs of tax evasion a moral appeal (e.g. Blumenthal et al., 2001, who tried to encourage taxpayers to comply by underlining valuable services financed with taxes), while other individuals receive a neutral message or no message at all. Finally, researchers analyze the effect of the message by comparing the treatment-group with the baseline group.

Field experiments avoid the main disadvantages of laboratory experiments. Field experiments observe real behavior of real taxpayers in their natural environment dealing with real stack sizes. Furthermore, participants are uninformed about being part of a study (avoiding a "game-environment"). Typically, field experiments contain a high number of observations and therefore better statistical power. Consequently, the external validity of field experiments and their relevance for policy makers might be higher. However, a field experiment lacks a high controlled environment like a laboratory setting (for example spillover effects are out of control and likely to exist, see Drago et al., 2020 and Alstadsæter et al., 2019), and the experiments are not directly replicable and usually more expensive. Moreover, treatment-letters might be more a form of nudging instead of learning, as stated out by Mascagni (2018). Furthermore, field experiments are one-time interventions and long-term effects are unknown. Following Mascagni (2018), "variations in compliance […] rely on evaders starting to comply. The fact that evaders may be less affected by moral and social appeals could therefore partly explain the lack of significant results in many studies." Moreover, moral appeals in form of a letter might have a deterrence-effect as well, for example increasing the perceived audit probability (Fellner et al., 2013). Furthermore, a significant time lag between receiving the treatment letter and filing the tax return is likely to exist. Therefore, the effects of such messages might be underestimated. Finally, the design of a controlled field experiment lacks flexibility and

[13] Recently, researchers have experimented with other methods of message transmission, for example visit of a tax official see Doerrenberg and Schmitz (2015) and Ortega and Scartascini (2015), email and SMS see Mascagni et al. (2017).

limits its utility in answering research questions, especially in the context of investigating non-monetary costs of tax evasion. Most parameters of tax compliance cannot be manipulated in field experiments. Consequently, field experiments are restricted by design to specific parameter-manipulations like communication (Slemrod, 2016).

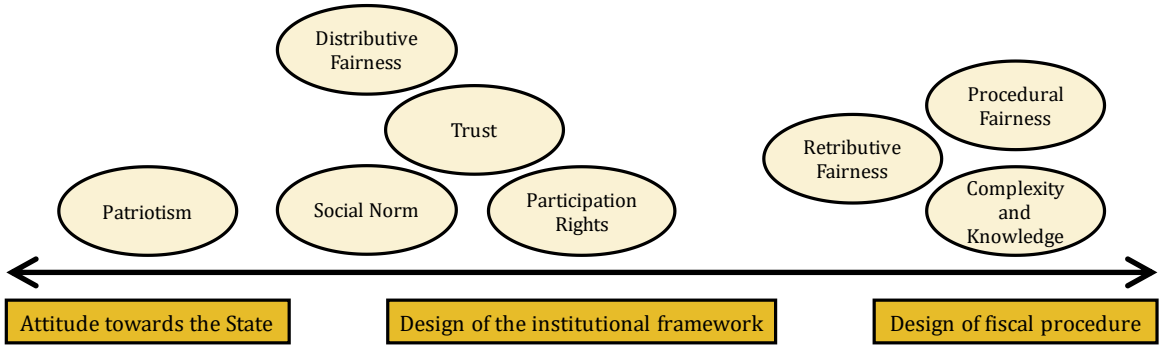
4.4 Sources of non-monetary costs of tax evasion

As described above, non-monetary costs of tax evasion capture for compliance that is not predictable by monetary costs of tax evasion (e.g. deterrence). This "voluntary" compliance has arguably many sources. This adds two questions, (1) what are the sources of non-monetary costs of tax evasion and (2) can the sources of non-monetary costs of tax evasion be manipulated in order to affect tax compliance? Especially the second question is relevant for governments combating tax evasion. Compared with monetary costs of tax evasion, manipulating non-monetary costs of tax evasion may have several advantages. For example, manipulation might be more cost-efficient with a better marginal utility. Furthermore, increasing non-monetary costs of tax evasion (e.g. enhance trust, fairness etc.) might promote a positive mood instead of an antagonistic climate ("cops and robbers attitude", Kirchler et al., 2008).

In the following, I review seven sources of non-monetary costs of tax evasion. I review research regarding social norm, trust, fairness, participation rights, complexity and knowledge, patriotism, and moral emotions (see Figure 2). However, my selection of subtopics is not meant to be exhaustive. Instead it summarizes the most relevant and recent studied factors of non-monetary costs of tax evasion. I distinguish research with respect to the applied method of data collection. As discussed in the previous section each method has certain advantages and drawbacks, which limit the generalizability of findings. Especially evidence based on survey data benefits from support of experimental data.

A systemization of social norm, trust, fairness, participation rights, complexity and knowledge, patriotism, and moral emotions is difficult and challenging. Up to now, there is no such concept in the literature. In the following, I sort the sources of non-monetary costs of tax evasion and allocate them to three categories. The first category is the taxpayer's attitude towards the State in general. Clearly, patriotism can be allocated to this category. The second category deals with the design of the institutional framework. For example, the existence and the degree of participation rights is allocated to this category. The third category deals with the design of fiscal procedure. This category

is mainly determined by tax authorities and captures for example procedural fairness. Figure 3 shows the sources of non-monetary costs of tax evasion and their allocation to these three categories.[14] The only exemption are moral emotions. Moral emotions are not displayed in Figure 3, because they cannot easily be allocated to one of the three categories. Instead, moral emotions would add a new dimension in Figure 3, affecting all categories with respect to individual characteristics.



Note: Figure shows the different sources of non-monetary costs of tax evasion (except moral emotions) allocated to three categories: Attitude towards the State, design of the institutional framework and design of fiscal procedure.

Figure 3: Allocation of different sources of non-monetary costs of tax evasion.

4.4.1 Patriotism

Patriotism is not necessarily the most important determinant of tax compliance. However, it distinguishes itself from other sources of tax compliance as governments have already manipulated it consciously. For example, governments have stimulated patriotism in World War I (Kang and Rockoff, 2015) and World War II (Jones, 1996) in order to enhance tax compliance. However, until recently there was no robust scientific basis for the effectiveness of such manipulations. Furthermore, it was questionable whether effects of patriotism and its manipulation are possible in times of peace.

Based on the survey-findings of Wenzel (2007), that taxpayers who identify with their nation as a whole had the most favorable tax attitude, Konrad and Qari (2012) find a robust positive association between patriotism and tax compliance by analyzing two modules of the International Social Survey Programme (ISSP). Gangl et al. (2016) verify this effect experimentally and investigate whether patriotism can be manipulated by tax authorities. The authors conducted three laboratory experiments (and a survey), in which

[14] Fairness is further distinguished in distributive fairness, retributive fairness and procedural fairness (see section 4.4.4 for details).

they displayed a national flag, national landscapes or national achievements respectively in order to manipulate patriotism. The results suggest that patriotism (on a national and local level) can be manipulated and that patriotism increases tax compliance. Regarding conscious manipulation of patriotism, Gangl et al. (2016) sound a note of caution for undesirable side effects. In their experiment, displaying a national flag increased not only patriotism but also nationalism. Contrary, Meiselman (2018) conducted a controlled field experiment sending messages relating to civic pride to income tax nonfilers. The message relating to civic pride had no effect on response rate. The authors argue that perhaps individuals with whom a message about civic pride would succeed had already filed their tax returns. Similar results and argumentations are provided by Chirico et al. (2019).

Table 1 summarizes the empirical evidence of patriotism affecting non-monetary costs of tax evasion.

Conclusion 2: *Patriotism affects tax compliance positively (a fact that governments already utilized). Accordingly, patriotism seems to increase the non-monetary costs of tax evasion. However, manipulating patriotism may come at the price of nationalism.*

Table 1: Empirical evidence on patriotism affecting non-monetary costs of tax evasion

Patriotism			
Study	**Method**	**Country**	**Main result regarding patriotism**
Wenzel (2007)	Survey	Australia	Respondents who identify with their nation as a whole had the most favorable tax attitude.
Konrad and Qari (2012)	Survey	USA, Austria, Ireland, the Netherlands, Poland, Canada, Portugal, Uruguay	Positive association between patriotism and tax compliance.
Gangl et al. (2016)	Laboratory experiment + survey	Austria	Patriotic materials can impact identification with the community and increase cooperation. Reported and manipulated patriotism indirectly increase tax compliance. Attempts to manipulate patriotism can also lead to nationalism.
Meiselman (2018)	Field experiment	USA	Message relating to civic pride had no effect on response rate.
Chirico et al. (2019)	Field experiment	USA	Message that stressed civic duty did not increase compliance.

4.4.2 Social norm

Following Luttmer and Singhal (2014), social norm is a key driver of tax compliance and therefore of the non-monetary costs of tax evasion. As Kirchler et al. (2008) point out, social norms are related to the behavior of reference groups like friends or vocational

groups. Furthermore, Elster (1989) emphasize that a norm must be shared by other people and partly sustained by their approval and disapproval to be a social norm.[15] More precisely, Cialdini and Trost (1998) characterize social norms as "rules and standards that are understood by members of a group, and that guide and/or constrain social behavior without the force of laws" (p.152). However, the definition of the term "social norm" keeps difficult, lacks uniformity and there is no definition in favor in tax compliance literature.[16]

Surveys support the assumption of the positive influence of social norms on tax compliance. In a survey among adults in Britain, Orviska and Hudson (2002) find a relationship between social norm and self-reported tax evasion. Jimenez and Iyer (2016) confirm this relationship in a survey among U.S. taxpayers. Notably, in a survey in Germany, Doerrenberg and Peichl (2018) asked participants about their attitude towards tax evasion, after some of the participants got additional information on the extent of tax evasion in society (in particular, that 10% evade taxes). They observe that the self-reported attitude towards tax evasion declined, indicating that the relatively high evasion level weakens the social norm of compliance. Therefore, their research indicates that appeals to social norms might have a backfiring effect if one's belief of the behavior of others becomes worse. Accordingly, Torgler and Schneider (2005) observe that a lower perceived compliance leads to a decrease of tax attitude in Austria by analyzing data of the World Values Survey and the European Values Survey. Even more evidence on this effect is provided by Frey and Torgler (2007) and Wenzel (2004). One explanation for this observation is provided by Cullis et al. (2012a) who suggests that individuals code changes as "gains" or "losses" with respect to their reference point. If their reference point is changed, individuals' perception of their entitlement to income changes.

In controlled field experiments, researchers tried to enhance tax compliance by appeals to social norms. Typically, individuals receive a letter with a moral appeal (treatment group) and the observed behavior is compared with individuals receiving no letter (or for example a letter with a simple payment reminder). The results are, however, mixed. Some researchers fail to find an influence of social norm appeals on tax compliance. Blumenthal et al. (2001) conducted a field experiment in Minnesota. Their two different letters with normative appeals sent to large group of taxpayers had no effect. Torgler (2004) and Torgler (2013) find no effect for normative appeals in Switzerland.

[15] In the study of Elster (1989) the term "social norm" is discussed in much more detail with many examples.
[16] For example, Cialdini and Trost (1998) distinguish four types of social norm: personal norms, subjective norms, injunctive norms and descriptive norms. This concept was tested empirically by Bobek et al. (2013), but could not establish itself in tax compliance literature so far.

Fellner et al. (2013) conducted a field study in Austria with potential evaders of TV license fees and fail to increase compliance by moral appeals. Similarly, Castro and Scartascini (2015) conducted a field experiment regarding the property tax in Argentina. However, the two messages on reciprocity and peer effects had on average no effect on compliance. Also Chirico et al. (2019) fail to find an effect of social norm appeals in a field experiment in Philadelphia.

Other researchers find a (small) positive impact of social norm appeals on tax compliance. Del Carpio (2014) conducted a field experiment (along with surveys) on property taxes in Peru, sending a letter with the average rate of compliance. Surprisingly, this appeal on social norm increased compliance by roughly 20%, while a letter with information on enforcement had no additional effect beyond a letter with a simple payment reminder. Bott et al. (2017) support these results with a field experiment in Norway. Their sample included 15,000 taxpayers who were likely to have misreported their foreign income in the previous year. The moral appeal mainly increased the reported income while a deterrence letter mainly increased the share of reporters. Hallsworth et al. (2017) conducted a field experiment in the United Kingdom and including social norm messages in standard payment reminder letter increased payment rates for overdue taxes.

There are different explanations for the mixed results.[17] First, the appeals to social norms might be not powerful enough to change the perceived social norm of individuals (that were possibly formed over many years). Moreover, the effectiveness of field experiments is dependent on other parameters like trust. If an individual has low trust in tax authorities, she might have also low trust in messages that she officially receives from these institutions. Second, even if appeals to social norms affect the social norm of an individual it might not necessarily translate into changes in compliance behavior. Or there might be even a so called backfiring effect of moral appeals. As suggested by Bardach (1989), a moral appeal might signal a rather weak enforcement system that tries to compensate its weakness with rhetoric. Third, there might be a timing effect so that the appeal to social norms only have an effect at the moment of reading the letter (as it is a rather small manipulation) that diminishes over time and is vanished at the moment of filing the tax return.

[17] Noteworthy, by expanding the focus from individuals to corporations, Ariel (2012) even found backfiring effect of moral suasion letters with a field experiment in Israel. However, so far there is only some survey-evidence for backfiring effects of moral suasion letters for individuals (see last paragraph of this section).

However, as shown in laboratory experiments, the manipulation of social norms is possible. For example, Alm et al. (2019) confirm in a laboratory experiment that normative appeals have a small positive impact on tax compliance. Interestingly, Guala and Mittone (2010) demonstrate in a laboratory experiment that the development of social norms does not necessarily take a long period of time. They observe that conventions have a tendency to become social norms. With their experimental design, they show that an equilibrium strategy that emerged in a coordination game can influence behavior in a "social dilemma" game. Finally, the authors conclude that such norms "may be more difficult to disrupt by changing individual incentives than one would assume based on standard rational choice analysis" (p.755).

Table 2 summarizes the empirical evidence of social norms affecting non-monetary costs of tax evasion.

Conclusion 3: *Social norms are a driver of tax compliance, even though its manipulation (e.g. by appeals) is difficult and evidence therefore mixed. Importantly, only a social norm that emphasizes high levels of tax compliance increases tax compliance. Accordingly, a social norm that emphasizes high levels of tax compliance seems to increase the non-monetary costs of tax evasion.*

Table 2: Empirical evidence on social norms affecting non-monetary costs of tax evasion

Social Norm			
Study	**Method**	**Country**	**Main result regarding social norm**
Orviska and Hudson (2002)	Survey	United Kingdom	Sense of civic duty impacts on perceptions of whether tax evasion is right or wrong and hence an individual engaging in tax evasion.
Jimenez and Iyer (2016)	Survey	USA	Social norms influence compliance intentions indirectly through internalization as personal norms. Specifically, as the strength of social norms in favor of tax compliance increases, personal norms of tax compliance also increase, and this leads to a subsequent increase in compliance intentions.
Doerrenberg and Peichl (2018)	Survey	Germany	Manipulating the social norm through information about the general extent of tax evasion has a negative effect on tax attitude, i.e. social norms can backfire if they reveal that a certain behavior is regrettably frequent.
Torgler and Schneider (2005)	Survey	Austria	If people perceive that tax evasion is a common phenomenon, their motivation to contribute to society decreases.
Frey and Torgler (2007)	Survey	30 European Countries	If taxpayers believe tax evasion to be common, tax attitude decreases. Alternatively, if they believe others to be honest, tax attitude increases.
Wenzel (2004)	Survey	Australia	Social norms should be ineffective when identification is weak; social norms might then even backfire when they contrast with one's internalised norms.
Cullis et al. (2012a)	Survey	Italy, United Kingdom	Social norms exert their influence by changing the reference points that individuals use when they code changes as 'gains', or 'losses'.
Blumenthal et al. (2001)	Field experiment	USA	Little or no evidence that either of two normative appeals affects aggregate tax compliance behavior.
Torgler (2004)	Field experiment	Switzerland	Moral suasion has hardly any effect on taxpayers' compliance.
Torgler (2013)	Field experiment	Switzerland	Moral suasion has hardly any effect on taxpayers' compliance.
Fellner et al. (2013)	Field experiment	Austria	Appealing to morals does not enhance compliance on aggregate.
Castro and Scartascini (2015)	Field experiment	Argentina	No average compliance effect for messages regarding reciprocity and peer-effects.
Chirico et al. (2019)	Field experiment	USA	A message that stressed the value of public services, neighbors' compliance, or civic duty did not increase compliance.
Del Carpio (2014)	Field experiment (+ surveys)	Peru	Disclosing information on the level of compliance had a large positive impact on compliance and raised beliefs about compliance.
Bott et al. (2017)	Field experiment	Norway	Moral letter increased the intensive margin, i.e. the amount of reported income, but had on average no effect on the extensive margin, i.e. the share of individuals reporting income.
Hallsworth et al. (2017)	Field experiment	United Kingdom	Including social norm messages in standard reminder letters increases payment rates for overdue tax.
Alm et al. (2019)	Laboratory experiment	USA	Normative appeals generally have a modest and positive impact on tax compliance.

4.4.3 Trust

Trust in government and tax authorities is a key component of tax compliance. There is empirical evidence that trust increases tax compliance and consequently increases the non-monetary costs of tax evasion. On the one hand, evidence is provided by survey data. Following the results of a survey among 26 cantonal tax authorities in Switzerland by Feld

and Frey (2002), the relationship between taxpayers and tax authorities is form of a psychological contract and based on a relationship of trust.[18] Trust as a driver of voluntary tax compliance is also assumed by Fjeldstad (2004). The author argues that the different levels of compliance regarding service charges in South Africa can only be explained by trust in the government. Based on survey data, Bergman (2002) find trust to be a driver of tax compliance by comparing Argentina with Chile. Also Torgler (2003) confirms the positive effect of trust on tax attitude by analyzing data from the World Value Survey and checking its robustness with the Taxpayers Opinion Survey. On the basis of survey data collected from 2,292 tax avoiders in Australia, Murphy (2004) shows that the level of trust in tax authorities among this group was lower than that of the general population (however, mainly driven by procedural fairness perceptions, see section 4.4.4). She concludes that trust in authorities enhances tax compliance. Torgler and Schneider (2005) identify trust as key determinant of tax attitude in Austria by analyzing data of the World Values Survey and the European Values Survey. Also Richardson (2008) confirms the relationship of trust and tax evasion across 47 countries, however on a less sophisticated data basis. More evidence is provided by Hammar et al. (2009), who analyze survey data in Sweden. They observe that especially distrust in politicians fosters tax evasion. Kastlunger et al. (2013) find out in a survey among self-employed Italian taxpayers and entrepreneurs that trust is positively related to voluntary tax compliance.

The results obtained by surveys are supported by the experimental literature. Cummings et al. (2009) conducted field experiments in Botswana and South Africa and find that tax compliance differences can be explained by the percipience of tax administration and individuals assessment of the quality of governance, both variables referring to trust. In laboratory and online experiments, Wahl et al. (2010a) manipulated trust in tax authorities in a hypothetical setting. Participants were asked to imagine being citizen of a fictive country, whose tax authorities are trustworthy or not, depending on the treatment. Even though the results lack consideration of experimental demand effects, the results suggest that trust increases tax attitude and therefore compliance. The experimental setting of Wahl et al. (2010a) was adopted by Kogler et al. (2013). They conducted the laboratory experiments in Austria, Hungary, Romania and Russia and verify the positive association between trust and tax compliance. More support for the mediating effect of trust is provided by laboratory experiments of Wahl et al. (2010b).

[18] A psychological contract is a well-known concept for social psychologists. It provides a clear distinction from a formal contract with agreed explicit and material sanctions.

They conclude that trust in tax authorities is an important precondition for voluntary tax compliance.

Table 3 summarizes the empirical evidence of trust affecting non-monetary costs of tax evasion.

Conclusion 4: *Trust is an important determinant of tax compliance. Higher trust in tax authorities and governments is associated with higher tax compliance. Accordingly, trust seems to increase the non-monetary costs of tax evasion.*

Table 3: Empirical evidence on trust affecting non-monetary costs of tax evasion

Trust			
Study	**Method**	**Country**	**Main result regarding trust**
Bergman (2002)	Survey	Argentina/Chile	Social variables associated with trust and legitimacy have an independent effect on social solidarity.
Feld and Frey (2002)	Survey	Switzerland	The implicit psychological contract between taxpayers and tax authorities in Switzerland is based on a relationship of trust.
Torgler (2003)	Survey	Austria, Belgium, Denmark, Finland, France, Great Britain, Iceland, Ireland, Italy, N. Ireland, Netherlands, Norway, Portugal, Sweden, West and East Germany (World Value Survey); United States (Taxpayer Opinion Survey)	Trust in public officials and the legal system have a positive effect on tax attitude.
Murphy (2004)	Survey	Australia	Distrust leads to active resistance towards regulatory authorities.
Torgler and Schneider (2005)	Survey	Austria	Trust has a positive impact on tax attitude.
Richardson (2008)	Survey	47 Countries	Lower level of trust comes with higher level of tax evasion.
Hammar et al. (2009)	Survey	Sweden	Distrust in politicians fosters tax evasion.
Kastlunger et al. (2013)	Survey	Italy	Trust is positively related to voluntary tax compliance.
Cummings et al. (2009)	Field experiment (+ survey)	Botswana, South Africa	Tax compliance differences can be explained by the percipience of tax administration and individuals assessment of the quality of governance.
Wahl et al. (2010a)	Laboratory experiment	Austria	Trust increases tax compliance.
Kogler et al. (2013)	Laboratory experiment	Austria, Hungary, Romania, Russia	Tax evasion is low in condition of high trust.
Wahl et al. (2010b)	Laboratory experiment	Austria	Trust in tax authorities is an important precondition for voluntary tax compliance.

4.4.4 Fairness

Concerns regarding fairness assumed to be an important part of the reasoning of tax compliance behavior. Already Cowell (1992) and Falkinger (1995) emphasized to

implement fairness into economic models. Again, tax compliance literature uses different terms like fairness, justice and equity, but refers (mostly) to the same object of investigation. The existence of concerns regarding fairness is not surprising and for example revealed in surveys, that observe the widespread perception that the wealthy are avoiding their fair share of taxes (Braithwaite, 2003, Rawlings, 2003 and Kinsey et al., 1991). This highlights that fairness is not only seen on an individual level, but more often on a relative-to-others level (Taylor, 2003). Furthermore, tax evasion might be an act of social protest against a tax system that is perceived to be unfair (Wenzel, 2003).

There are different attempts to conceptualize fairness. For example, Kinsey and Grasmick (1993) distinguish exchange fairness, vertical fairness and horizontal fairness. Exchange fairness refers to the perceived value of tax-funded services relative to one's tax burden. Vertical fairness aims at the burden of taxes among different income brackets, while horizontal fairness aims at the tax burden of taxpayers in the same income bracket. Following social psychology of Tyler and Smith (1998), Wenzel (2003) suggested to distinguish three types of fairness: (1) distributive fairness, (2) procedural fairness and (3) retributive fairness. Distributive fairness refers to resource distribution, procedural fairness refers to the processes of resource distribution and retributive fairness refers to perceived sanctions for breaking social rules and norms. Going into more details, the three dimensions can be analyzed on the individual, group, and/or societal level.

For the distributive fairness, there is evidence that perceived unfairness results in lower tax compliance. In a survey in Scotland, Dean et al. (1980) observed that 26% of taxpayers believed that they paid far too much relative to other taxpayers of the same income bracket. Spicer and Becker (1980) observe experimentally that individuals who suffer under (benefit from) fiscal inequity evade more (less) taxes. Also individuals who perceive low vertical fairness evade more taxes (Kinsey and Grasmick, 1993, Roberts and Hite, 1994, Bazart and Bonein, 2014, Fortin et al., 2007). Besley et al. (2019) observe a massive increase in tax evasion in the United Kingdom with the introduction of the poll tax (which replaced a tax based on property values) in 1990. These results can be traced back on a perception of unfairness because the tax was not related to the ability to pay. More evidence is provided by Hartner-Tiefenthaler et al. (2012) who find a positive relation between tax compliance and the perception of distributive fairness in the European Union. Also in laboratory experiments researchers observed that distributive fairness is a relevant factor for tax compliance (e.g. Trivedi et al., 2003).

The importance of procedural fairness is observed by Wenzel (2002) who conducted a survey in Australia and concludes that taxpayers were more compliant when they perceived the treatment by the Australian Taxation Office as fair and respectful. Also Murphy (2003) demonstrated in a survey in Australia that procedural unfairness undermines regulators' legitimacy and results in resistance against tax authorities. Even more survey-based evidence is provided by Hartner et al. (2008) and Faizal et al. (2017). These findings are especially emphasized in recent research by authors postulating a service-paradigm for tax authorities (e.g. Kirchler, 2007). Also experimentally, Verboon and van Dijke (2011) emphasize that procedural fairness moderates the effect of sanction severity on compliance. In particular, their results suggest that severe sanctions increase compliance only when authorities acted in a fair manner. However, procedural fairness is often times related to other sources of non-monetary costs of tax evasion. Therefore, more evidence for the effects of procedural fairness on tax compliance provides section 4.5.[19]

The retributive fairness is high when individuals agree with the design of punishment for tax evasion, including the process of audits and penalties. For example, there is some evidence for a backfiring effect of high audit probabilities, that might indicate that the retributive fairness decreased (e.g. Mendoza et al., 2017). Cullis et al. (2012b) observe in the laboratory that excessive enforcement decreases individuals' willingness to comply with taxation (which the authors refer to as "spite effect"). There is also research regarding the absence of penalties for non-compliance, namely tax amnesty, and its effects on tax compliance. Tax amnesties allow tax evaders to pay back evaded taxes without being punished. Research suggests that a tax amnesty lowers retributive fairness of honest taxpayers and lowers their tax compliance, following Sausgruber and Winner (2004), Hasseldine (1998) and Alm et al. (1990).

Table 4 summarizes the empirical evidence of fairness affecting non-monetary costs of tax evasion.

Conclusion 5: *Fairness in general is positively associated with tax compliance. Accordingly, higher distributive fairness, procedural fairness and retributive fairness seem to increase the non-monetary costs of tax evasion.*

[19] For example, procedural fairness can be related with trust, social norms and participation rights.

Table 4: Empirical evidence on fairness affecting non-monetary costs of tax evasion

Fairness			
Study	**Method**	**Country**	**Main result regarding fairness**
Dean et al. (1980)	Survey	Scotland	Respondents cited inequity as being in the forefront of reasons for people deciding to evade tax.
Spicer and Becker (1980)	Laboratory experiment	USA	The percentage of taxes evaded was highest among those who were told that their tax rates were higher than average and lowest among those told their tax rates were lower than average.
Kinsey and Grasmick (1993)	Survey	USA	Perceived vertical equity explains decline in the acceptability of tax cheating.
Roberts and Hite (1994)	Survey	USA	In general, respondents' stated preferences for vertical equity approximate the current distribution of the income tax burden, yet there is a relatively high consensus that the income tax is unfair, especially with regard to the ability of wealthy taxpayers to exploit loopholes to avoid paying their fair share, and that respondents regard their own tax burdens as unfair.
Bazart and Bonein (2014)	Laboratory experiment	France	Disadvantageous inequity in tax rates leads to a decrease in the level of reported income while an advantageous inequity decreases the level of evasion. When taxpayers learn the average reported income of their other group members at the end of the period, most of them adjust their current reported income up or down to come closer to the previous reported mean.
Fortin et al. (2007)	Laboratory experiment	France	For a given gross income and a given personal tax rate, the individual will report less when facing a reduction in the mean tax rate of his group. Perceived unfair taxation leads to increased tax evasion.
Besley et al. (2019)	Natural data	United Kingdom	The poll-tax shock is plausibly interpretable as a shock to the intrinsic compliance motive due to the perceived unfairness of the new tax base.
Hartner-Tiefenthaler et al. (2012)	Survey	Austria, Czech Republic, United Kingdom	EU-tax compliance was positively related to distributive justice.
Trivedi et al. (2003)	Laboratory experiment	Canada	Subjects who discovered that they were taxed at a lower level than others (they were beneficiaries of tax inequity) increased their compliance.
Murphy (2003)	Survey	Australia	The feeling of taxpayers to be poorly treated by a tax authority can lead to questioning the legitimacy of the tax authority. This can affect their willingness to comply and may lead to active resistance.
Hartner et al. (2008)	Survey	Australia	When people feel treated in a procedurally fair manner by the tax authority and procedurally fair decision rules are employed, motivational postures of deference increase whereas motivational postures of defiance decrease.
Faizal et al. (2017)	Survey	Malaysia	Procedural justice affect tax compliance. Procedural justice was positively and significantly correlated to trust. However, trust does not mediate the relationship between justice and compliance.
Verboon and van Dijke (2011)	Laboratory experiment + survey	Netherlands	Procedural fairness moderates the effect of sanction severity on compliance with authorities' regulations. Severe sanctions increased compliance with the authority more than mild sanctions, but only when authorities acted in a fair manner.
Mendoza et al. (2017)	Natural data	Many	Backfiring effects of control and deterrence are tightly connected with perceptions of distrust and unfairness. The expected backfiring effect may thus relate to how excessive auditing hinders voluntary compliance.
Cullis et al. (2012b)	Laboratory experiment	Egypt	The behavior of individuals is consistent with the presence of a 'spite effect' when they perceive enforcement as 'excessive'.

4.4.5 Participation rights

It is reasonable to assume that individuals are more likely to comply if they have to follow self-imposed social rules instead of rules that are externally given.[20] That is why the implementation of participation rights in tax compliance theory was already suggested by Weck-Hannemann and Pommerehne (1989). Arguably, participation rights can be seen as some form of procedural fairness (see above) and the structural separation follows reasons of simplicity. In theory, participation rights (e.g. voting-rights on tax spendings or the extent of the tax burden) increase tax compliance. This prediction is supported by empirical evidence.

Early laboratory evidence is provided by Alm et al. (1993). They conducted an experimental study, in which participants voted on how collected taxes were spent. The authors find evidence for a positive compliance effect if individuals vote on tax spending. More recently, Casal et al. (2016) have found experimentally that voting on tax contributions and on tax distribution leads to higher compliance. Lamberton et al. (2018) find evidence in laboratory and online experiments that allowing taxpayers to express non-binding preferences about the way their taxes are used increase compliance by roughly 15%. Also Wahl et al. (2010b) observe better cooperation and tax compliance by voting. Additionally, they show that denying participation-rights on important decisions for taxpayers has even negative effects on tax compliance. These negative effects can be explained by reactance theory of Brehm (1966) that postulates that if behavioral freedoms are reduced (e.g. subjectively important agendas are decided by someone else), individuals engage in reactance.

The impact of voting on tax compliance has been studied extensively in the cantons of Switzerland utilizing cantonal differences in participation rights. Pommerehne and Weck-Hannemann (1996) combined data on declared household income by the Swiss Bureau of Taxation with the corresponding data independently collected from tax statistics in 25 cantons in Switzerland. They observe inter alia that tax compliance is positively correlated with the extent of political participation rights. This results is supported by Torgler (2005) who examined survey data from the International Social Survey Programme (ISSP) and analyzed the impact of direct democracy on tax attitude in Switzerland. He concludes that direct democratic rights influence tax attitude positively. In a survey among the 26 cantonal tax authorities in Switzerland, Feld and Frey (2002)

[20] Consequently, participation rights can be seen as a determinant of procedural fairness as well.

show that tax compliance is higher in cantons with more direct participation rights. Accordingly, Feld and Kirchgässner (2000) reviewed empirical studies on the political culture in Switzerland and conclude that tax evasion is lower in direct than in representative democratic systems.

However, voting is also relevant in the context of deterrence. Alm et al. (1999) conducted an experimental study, in which participants vote on certain parameters of the enforcement regime (tax rate, audit probability or fine multiplier respectively). On the one hand, results suggest that compliance increases if the group selects greater enforcement. On the other, the authors observe a decrease in compliance if greater enforcement is rejected. Feld and Tyran (2002) conducted a laboratory experiment with a more simplistic experimental design (e.g. detection probability always equals 1 and voting takes place only on a general fine). Overall, they confirm the results of Alm et al. (1999) and explain their results by greater legitimacy of the taxation procedure. They argue that voting signals cooperation, increases belief on cooperation and finally increases cooperation itself.[21]

All in all, there is evidence that participation rights increase tax compliance. Arguably, tax evasion should be lower in direct than in representative democratic systems. But is this also true for individuals being overruled in a voting, for example, on tax spendings? Again, the experimental literature provides first evidence. Alm et al. (1993) observe lower levels of compliance for close votes. Also the imposition of unpopular tax spendings lowers compliance. Accordingly, Wahl et al. (2010b) observed a positive effect of voting only if the offered alternatives were relevant to the voters. Finally, Lamberton et al. (2018) find that the positive compliance effect of voting effect vanishes if taxes are allocated across disliked spending categories. In this context, Hunt et al. (2019) examined how election outcomes influence tax attitude; especially relevant at the present time, as numerous studies have documented that political polarization has increased (for example for the United States see Mason (2015)). In particular, Hunt et al. (2019) analyzed partisan reactions to presidential election outcomes in the U.S. and find that election outcomes influence overall positive or negative feelings, trust in government and subsequently tax attitude.

Table 5 summarizes the empirical evidence of participation rights affecting non-monetary costs of tax evasion.

[21] Accordingly, voting is also in other research settings (e.g. public goods experiments) positively related with cooperation, see for example Cinyabuguma et al. (2005).

Conclusion 6: *Participation rights (e.g. voting) increase tax compliance and therefore the non-monetary costs of tax evasion. However, this effect might vanish for overruled voters.*

Table 5: Empirical evidence on participation rights affecting non-monetary costs of tax evasion

Participation Rights			
Study	**Method**	**Country**	**Main result regarding participation rights**
Alm et al. (1993)	Laboratory experiment	USA	Individuals respond positively when tax proceeds are directed toward programs they approve of and when they feel they are active in the decision process.
Casal et al. (2016)	Laboratory experiment	Austria	Having voice on tax contributions and on tax distribution leads to higher compliance.
Lamberton et al. (2018)	Laboratory experiment	USA	Allowing participants to express non-binding preferences over tax spending priorities leads to a 16% increase in compliance. Allowing taxpayers to express their preferences on the distribution of government spending reduces the stated take-up rate of a questionable tax loophole by 15%.
Wahl et al. (2010b)	Laboratory experiment	Austria	Differences in tax payments arose from voting only if the offered alternatives were relevant to the voters. It seems that denying citizens the opportunity to participate in decisions has negative effects rather than the opposite – a positive effect of participation rights on cooperation.
Pommerehne and Weck-Hannemann (1996)	Natural data	Switzerland	Noncompliance is lower when taxpayers have direct control over government budgets.
Torgler (2005)	Survey	Switzerland	Direct democratic rights have a positive effect on tax attitude.
Feld and Frey (2002)	Survey	Switzerland	There is an implicit psychological contract between taxpayers and tax authorities in Switzerland. This holds in particular if voters are directly involved in political decision-making.
Alm et al. (1999)	Laboratory experiment	USA	Individual behavior appears to be affected by the outcome of the vote when the vote is on the enforcement regime. Rejection by the group of greater enforcement decreases compliance drastically. Similarly, compliance increases when the group discusses and selects greater enforcement.
Feld and Tyran (2002)	Laboratory experiment	Switzerland	Tax compliance is higher on average in an endogenous fine treatment in which subjects are allowed to approve or reject the proposal of a fine. The main explanation why people show higher tax compliance if they are allowed to vote on a fine is legitimacy. Subjects who reject the proposal of the fine show a higher compliance rate than subjects in the exogenous fine treatment even if they know that the dominant strategy under the existence of the low fine is non-compliance. Individuals who vote against the fine contribute effectively more if the fine is adopted than individuals voting for the fine contribute in the case the symbolic fine is rejected.
Hunt et al. (2019)	Quasi-Experiment	USA	Election outcomes generate overall positive or negative feelings (i.e., affect balance) among partisans, which influences beliefs about trust in government, and subsequently their tax compliance intentions. Political party moderates the relationship between election outcomes and affect balance in such a way that Democrats experience greater overall positive affect balance when their party wins the election compared to Republicans.

4.4.6 Complexity and knowledge

"The hardest thing in the world to understand is the income tax." is a popular quote from Albert Einstein and reflects the viewpoint of many people who complain about the overwhelming complexity of tax law.[22] There is evidence that complexity affects tax compliance negatively (see for example the studies of Clotfelter, 1983 and Cox and Eger, 2006 who analyze natural data and Niemirowski et al., 2003, Kirchler et al., 2006 and Saad, 2014 who analyze survey data) while higher knowledge affects tax compliance positively (see for example the studies of Groenland and van Veldhoven, 1983, Kirchler and Maciejovsky, 2001 and Kirchler et al., 2006, that are based on survey data and the study of Park and Hyun, 2003 that is based on laboratory data). Basically, there are two approaches to reduce complexity. First, tax rules can be simplified[23] and tax authorities can increase taxpayer services. As shown by Alm et al. (2010) in a laboratory experiment, better administrative services can enhance formerly low tax compliance that was due to high complexity in a tax system.

Second, individuals' knowledge of tax law can be increased in order to reduce perceived complexity. Taxpayers can be educated to obtain a better understanding of the tax rules. In a field experiment in the Netherlands, Nagel et al. (2019) tested the effectiveness of a tax training as a service-oriented approach for entrepreneurs. They find out that the tax training had a positive effect on tax compliance behavior. Also knowledge in form of general education is found to affect tax compliance positively, see Kirchler et al. (2008) and Rodriguez-Justicia and Theilen (2018).

From the perspective of an individual, complexity of tax law can be tackled by hiring tax professionals. Long and Caudill (1987) find that taxpayers mandate professional tax assistances when tax returns are more complex and the marginal tax rate is higher. Furthermore, they observe that taxpayers who mandate professional tax assistance have a lower tax liability than self-preparers.[24] The positive correlation between complexity and paid tax preparers is also supported by Dubin et al. (1992) who highlight that an

[22] The quote is even displayed on the website of the US Internal Revenue Service, see https://www.irs.gov/newsroom/tax-quotes.

[23] However, one has to keep in mind interdependency between complexity and other sources of non-monetary costs of tax evaison. For example, a reduced complexity might have negative effects on distributive fairness or positive effects on procedural fairness (see section 4.5).

[24] Somehow contrary to the results above, studies by Erard (1993) and Erard (1997) indicate that compliance is generally lower if a tax return is prepared by a professional tax advisor. The author interprets this finding as a result of higher aggressiveness of CPAs and lawyers in their reporting practices.

increase in the number of forms increases the demand for professional tax assistance. The results are supported by the analyses of archival tax return data of Christian et al. (1993).

Table 6 summarizes the empirical evidence of complexity and knowledge affecting non-monetary costs of tax evasion.

Conclusion 7: *Complexity and knowledge have oppositional effects on tax compliance. While complexity affects tax compliance negatively, knowledge affects tax compliance positively. Accordingly, higher complexity seems to reduce non-monetary costs of tax evasion and higher knowledge seems to increase non-monetary costs of tax evasion. Both parameters are likely to be interdependent, meaning that for example higher knowledge reduces complexity and vice versa.*

Table 6: Empirical evidence on complexity and knowledge affecting non-monetary costs of tax evasion

Complexity and Knowledge			
Study	**Method**	**Country**	**Main result regarding complexity and knowledge**
Clotfelter (1983)	Natural data	USA	The complexity of tax returns was associated with more underreporting among non-business tax returns.
Niemirowski et al. (2003)	Survey	Australia	The most significant predictor of non-compliance intent (at least for a youth sample) was the belief that the process of taxation is too complex.
Cox and Eger (2006)	Natural data	USA	Tax agency's organizational complexity mediates taxpayer non-compliance.
Kirchler et al. (2006)	Survey	Australia	Taxpayers' intent to report timely and correctly is higher if self-reported tax knowledge is high and tax law is perceived as not too complex.
Saad (2014)	Survey	New Zealand	Participants believed that complexity have partly contributed to non-compliance.
Groenland and van Verldhoven (1983)	Survey	Netherlands	Taxation is positively evaluated by higher educated people with a generalized internal locus of control.
Kirchler and Maciejovsky (2001)	Survey	Austria	Knowledge of the legal principles of Austrian tax law is correlated with tax morality.
Park and Hyun (2003)	Laboratory experiment	Korea	Tax education is an effective tool to induce taxpayers to comply more.
Alm et al. (2010)	Laboratory experiment	USA	Uncertainty due to complexity reduces both the filing and the reporting compliance of an individual. Agency provided information has a positive and significant impact on the tendency of an individual to file a tax return, and also on reporting for individuals who choose to file a return.
Nagel et al. (2019)	Survey + natural data	Netherlands	Training affects specific domains of tax compliant behavior, in particular a positive effect on tax compliant behavior.
Rodriguez-Justicia and Theilen (2018)	Survey	29 European countries	Education has a positive impact on tax attitude for those individuals that are net beneficiaries of the welfare state, and a negative impact for those that are net contributors. The more highly educated individuals exhibit higher levels of tax attitude in countries that have better quality public services, a fairer tax system and higher quality institutions.
Long and Caudill (1987)	Natural data	USA	Professional tax assistance is directly related to the complexity of the tax return. Income tax liability is relatively lower on paid-preparer than self-prepared returns.
Dubin et al. (1992)	Natural data	USA	Increases in the number of forms increase the demand for practitioners and decrease self-preparation.
Christian et al. (1993)	Natural data	USA	The probability of preparer use increases with complexity.

4.4.7 Moral emotions

Moral emotions trace back to evolution and help people in choosing the best strategy in human interactions, following Jacquemet et al. (2019). However, in tax compliance literature the examination of the influence of moral emotions on tax compliance is at its beginnings.

Coricelli et al. (2010) measure emotions in the context of tax compliance by skin conductance responses and self-reports. They find first evidence that misreporting is correlated with higher emotional arousal. Contrary, compliant individuals do not

experience such emotions and they decide more quickly than evaders. Furthermore, the authors remark that emotions experienced after a positive audit are probably associated with private emotions such as guilt, regret, or anger. Accordingly, a public shaming of detected tax evaders (e.g. with pictures) may be associated with pubic emotions such as shame or embarrassment. Somehow contrary, Dulleck et al. (2016) measure psychic stress by heart rate variability and find that higher psychic stress increases tax compliance. They identify three types of taxpayers, (1) high tax attitude and tax compliance but no psychic stress, (2) high tax attitude and tax compliance but high psychic stress and (3) lower tax attitude and tax compliance and a psychic stress level somewhere in between of (1) and (2). Coricelli et al. (2014) correlate shame and guilt as moral emotions with tax evasion and conclude that only shame is correlated with the intensity of evasion. Christian and Alm (2014) run laboratory experiments measuring and promoting the moral emotions sympathy and empathy. They find that these emotions enhance tax compliance. Fochmann et al. (2019) provide evidence that incidental emotions (i.e. emotions not related to actual tax behavior; mood) affect individuals' tax compliance behavior. They find that positive incidental emotions lead to a lower willingness to comply than aversive incidental emotions. Analyzing integral emotions, Lubian and Zarri (2011) find first evidence that individuals pay taxes because they (emotionally) like it. They find strong evidence for positive hedonic effects and conclude that paying taxes is rewarding in itself.

Table 7 summarizes the empirical evidence of moral emotions affecting non-monetary costs of tax evasion.

Conclusion 8: *Moral emotions are likely to affect tax compliance and therefore the non-monetary costs of tax evasion. There are first results about psychic stress influencing tax compliance, that is in line with research regarding negative emotions. However, research in this area of tax compliance is at an early stage.*

Table 7: Empirical evidence on moral emotions affecting non-monetary costs of tax evasion

Moral Emotions			
Study	**Method**	**Country**	**Main result regarding moral emotions**
Coricelli et al. (2010)	Laboratory experiment	France	The intensity of anticipated and anticipatory emotions before reporting income positively correlates with both the decision to cheat and the proportion of evaded income. The experienced emotional arousal after an audit increases with the monetary sanctions and the arousal is even stronger when the evader's picture is publicly displayed. The risk of a public exposure of deception deters evasion whereas the amount of fines encourages evasion. An audit policy that strengthens the emotional dimension of cheating favors compliance.
Dulleck et al. (2016)	Laboratory experiment	Australia	A positive correlation between psychic stress and tax compliance, thus underscoring the importance of moral sentiments for tax compliance.
Coricelli et al. (2014)	Laboratory experiment	France	When cheating is made public and the contravener is not successively reintegrated, the total amount of cheating is significantly increased compared to when cheating is made public but publicity is immediately followed by reintegration. The former condition is associated with more intense negative emotions related to cheating. This suggests that the employment of a social shaming mechanism may be an effective, albeit very sensitive, tool in the hands of policy makers.
Christian and Alm (2014)	Laboratory experiment	USA	The presence of sympathy in most cases encourages more tax compliance. Priming to elicit empathy has a positive impact on tax compliance.
Fochmann et al. (2019)	Laboratory experiment + survey	Germany	Positive incidental emotions lead to a lower willingness to comply than aversive incidental emotions. Participants of a survey show lower tax compliance attitudes on days associated with a positive mood. These findings are supported by the results of a controlled experiment in which incidental emotions are induced by standardized pictures.
Lubian and Zarri (2011)	Survey	Italy	Fiscal honesty generates a higher hedonic payoff than cheating.

4.5 Interdependencies

The previous section examined different sources of non-monetary costs of tax evasion separately and subsequently their effect on tax compliance. Although the influence of some of these sources on tax compliance is rather clear, researchers and policy makers have to interpret these results with caution. Inter alia, because there is some evidence suggesting that the sources of non-monetary costs of tax evasion interact with each other. Consequently, the effects of these sources on tax compliance described above can mutually amplify or cancel out. Moreover, research suggests interdependencies between non-monetary costs of tax evasion and monetary costs of tax evasion. Both interdependencies (within non-monetary costs of tax evasion and between non-monetary and monetary costs of tax evasion) are described in more detail in the following.

However, there is to mention that any correlation between certain sources of non-monetary costs of tax evasion cannot necessarily be interpreted as interdependency. For example, it might be that in a given study the variables for trust and fairness measure the same phenomenon. Moreover, research provides mixed results regarding the direction of correlation and causality between certain sources (e.g. Jimenez and Iyer (2016) for the fairness-trust relation). Subsequently, this underlines once again the importance of the applied research method and in particular its design (see section 4.3).[25]

4.5.1 Interdependencies within non-monetary costs of tax evasion

Following the illustration of Figure 3, the clustering of sources of non-monetary costs of tax evasion gives first ideas for the existence of overlaps and interdependencies. For example, higher fairness and lower complexity increase the non-monetary costs of tax evasion and consequently increase tax compliance. However, fairness and complexity are related to each other. It has been acknowledged that tax systems need complex rules to avoid tax evasion and tax avoidance (Picciotto, 2007). Assuming that part of the complexity of tax law arises from addressing concerns regarding fairness, Carnes and Cuccia (1996) find out that perceived complexity generally has a negative effect on perception of fairness. However, their results suggest also that individuals know about the necessity of complexity. Therefore, the justification of complexity differs systematically across different tax items. As a result, Carnes and Cuccia (1996) emphasize that compliance improves most by reforming areas that are less justifiable complex and these areas are not necessarily the most complex ones. Furthermore, such reductions in complexity increase trust in authorities and therefore increases tax compliance even further, following Kirchler et al. (2008). As individuals acknowledge justifiable complexity, Carnes and Cuccia (1996) highlight that perceived fairness is also improved by educating individuals on the purpose of complex tax law. Accordingly, Eriksen and Fallan (1996) observe that perceived fairness rises following an increase in tax knowledge. This is also supported experimentally by Wartick (1994). Finally, Eriksen and Fallan (1996) observe that attitudes towards tax evasion becomes stricter following an increase in tax knowledge. Furthermore, fairness is an important factor for trust and vice-versa, as pointed out by Job et al. (2007) and Jimenez and Iyer (2016).

[25] The earlier mentioned slippery-slope framework of Kirchler et al. (2008) circumvents this issue to some extent, as it aggregates "voluntary" compliance.

There is also evidence that participation rights (e.g. voting) increase procedural fairness and finally trust, following Wahl et al. (2010b) and Pommerehne and Weck-Hannemann (1996). Feld and Kirchgässner (2000) observed that citizens in direct democracies are better informed than in representative democracies. Arguably, more participation rights are therefore linked with better knowledge. Accordingly, Hug and Spörri (2011) find that allowing for referendums strengthens the link between trust and tax attitude.

4.5.2 Interdependencies between non-monetary and monetary costs of tax evasion

In an experimental study, Farrar et al. (2019) find out that detection moderates the relation between procedural fairness and tax compliance intentions. When detection probability is low, a high procedural fairness enhances compliance. However, when detection probability is high, the impact of procedural fairness is diminished. The detection probability sets a boundary for the effectiveness of fairness. Also public enforcement (and public shaming) could enhance perceived fairness (Braithwaite, 2003). The other way around, weak enforcement can decrease perceived fairness (also trust in tax authorities). However, audits and penalties perceived as unfair might lead to a negative tax attitude, following Wenzel and Thielmann (2006).

Olsen et al. (2018) find that enforcement induces negative emotions resulting in enforced compliance and increased readiness to evade. Therefore, high monetary costs of tax evasion might lower non-monetary costs of tax evasion. This crowding-out effect was already emphasized by Frey (1997a), who argues that monetary costs of tax evasion (in particular deterrence) may crowd out non-monetary costs of tax evasion. Early evidence is provided by Frey (1997b), who shows that more punitive enforcement crowd out tax compliance. However, research findings on the potential crowding-out of tax compliance are unclear. On the one hand, Dwenger et al. (2016) find in a field study in Germany in the context of a local church tax no crowding-out between deterrence and tax compliance. They utilize the fact that the local church tax relied on zero deterrence and incentivized compliance through deterrence or rewards. However, in a related study of Boyer et al. (2014), they find a crowding-out of tax compliance for weakly intrinsically motivated individuals in the context of a field study dealing with a local catholic church tax. Contrary, Filippin et al. (2013) even find that tax enforcement can have a positive effect on tax compliance.

More evidence on crowding-out is provided in general compliance research. Frey and Jegen (2001) provide a survey on intrinsic and extrinsic considerations and regarding crowding effects. Most prominent is a field study with day-care centers by Gneezy and Rustichini (2000). They introduced a monetary fine for late-coming parents and as a consequence the number of late-coming parents increased. Recently, Fochmann et al. (2020b) have observed in a laboratory experiment a crowding-out of intrinsic motivation by introducing audits in a simple compliance setting. Participants were asked to report a seen die roll and received a payoff in relation to the reported number. The introduction of an audit with a detection probability of zero percent increased the average reported die roll and the share of misreports. This result is significant in the context of penalties on misreporting and in the context of rewards on compliance. Translated to the framework of tax compliance developed in this study, their results suggest that monetary costs of tax evasion may crowd out non-monetary costs of tax evasion.

Conclusion 9: *There is evidence for interdependencies between sources of non-monetary costs of tax evasion. Therefore, policy makers implementing (or strengthening) one of these sources have to consider other sources as well. This recommendation holds also true for interdependencies between monetary and non-monetary costs of tax evasion.*

4.6 Concluding remarks on future research

Monetary and non-monetary costs of tax evasion are two sides of the same coin determining tax compliance of individuals. While research regarding monetary costs of tax evasion was most prominent over the last decades, recent research shed light on the importance of non-monetary costs of tax evasion. However, research regarding non-monetary costs of tax evasion lacks systemization. The study at hand is one approach tackling this problem. Furthermore, there is striking evidence that social norm, trust, fairness, participation rights, complexity and knowledge, patriotism, and moral emotions influence individuals in their decision on tax compliance. Accordingly, governments can expand their influence on these sources in order to enhance tax compliance. First ideas were presented in this paper. However, researchers and governments have to keep in mind complex interdependencies between these sources of non-monetary costs of tax evasion and between non-monetary and monetary costs of tax evasion.

Given all the information above, I emphasize four ideas for future research. First, research regarding complexity and knowledge, retributive fairness and procedural fairness are the most promising directions from a practical perspective. As shown in Figure 3, these variables are mainly allocated to the design of fiscal procedure. On the one hand, tax authorities can shape the design of these variables to some extent autonomously without a legislative process (e.g. in Germany). On the other hand, the design of fiscal procedure is more or less exclusively determined by considerations regarding taxation.[26] Consequently, research in the field of the design of fiscal procedure promises impact and offers the possibility for a dialogue between researchers and policymakers. Moreover, this connects with the service-paradigm that is frequently postulated in tax compliance research (Kirchler et al., 2008; Alm, 2012 and Alm, 2019).

Second, research dealing with tax culture lacks precision. It is most likely that a culture variable just measures for example differences in social norm or trust. The interpretation and the overall learning of culture effects is therefore limited. A more promising approach has to sharpen its focus and consider and analyze (at least) the range of variables described in this paper.

Third, tax compliance research has not solely to focus on specific sources of non-monetary costs of tax evasion. A promising field of research deals with mechanisms that enhance tax compliance, even though the mode of action is not precise. One recent example is research dealing with prefilling of tax returns. In a study that sets a novel focus on reporting of deductions rather than income, Fochmann et al. (2020a) find evidence that prefilled tax returns enhance tax compliance. In particular, prefilling of deductions increased tax compliance especially on items that were preferred by tax evaders. Following the authors, the increased compliance is due to increased non-monetary costs of tax evasion as there is no change in monetary costs of tax evasion in the experiment. Similar results can be observed for the prefilling of income in tax returns (Fochmann et al., 2018). Future research might identify the specific source of non-monetary costs of tax evasion that is causative for the positive influence of prefilling on tax compliance (e.g. it is conceivable that prefilling lowers complexity, increases procedural fairness and/or increases trust).

Fourth, researchers have to ensure that they measure variables without overlap. My systematization in Figure 3 shows variables that are clustered, e.g. distributive fairness,

[26] Contrary, governments face many other non-tax-related considerations if they for example debate about participation rights, that are allocated to the design of the institutional framework.

trust and participation rights. These variables are likely to measure the same phenomenon if experimental design lacks precision. The interdependencies presented in section 4.5 emphasize already this issue.

As a final remark and given the complexity of human behavior, there is no doubt that there will not be a single framework of tax compliance applicable to all individuals in every situation at all times. However, I strongly belief that the development of a reasonable systemization of factors influencing tax compliance is a key component for future research as an orientation and communication aid. Furthermore, being able to explain at least most of the puzzle of tax compliance is already a major achievement.

References

Allingham, M. G., and A. Sandmo. 1972. Income tax evasion: A theoretical analysis. *Journal of Public Economics* 1 (3-4):323-338.

Alm, J. 1991. A Perspective on the Experimental Analysis of Taxpayer Reporting. *The Accounting Review* 66 (3):577-593.

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

———. 2019. What Motivates Tax Compliance? *Journal of Economic Surveys* 33 (2):353-388.

Alm, J., K. M. Bloomquist, and M. McKee. 2015. On the External Validity of Laboratory Tax Compliance Experiments. *Economic Inquiry* 53 (2):1170-1186.

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

Alm, J., J. Clark, and K. Leibel. 2016. Enforcement, Socioeconomic Diversity, and Tax Filing Compliance in the United States. *Southern Economic Journal* 82 (3):725-747.

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

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

———. 1999. Changing the Social Norm of Tax Compliance by Voting. *Kyklos* 52 (2):141-171.

Alm, J., M. McKee, and W. Beck. 1990. Amazing Grace: Tax Amnesties and Compliance. *National Tax Journal* 43 (1):23-37.

Alm, J., I. Sanchez, and A. de Juan. 1995. Economic and Noneconomic Factors in Tax Compliance. *Kyklos* 48 (1):1-18.

Alm, J., W. D. Schulze, C. von Bose, and J. Yan. 2019. Appeals to Social Norms and Taxpayer Compliance. *Southern Economic Journal* 86 (2):1-29.

Alm, J., and B. Torgler. 2011. Do Ethics Matter? Tax Compliance and Morality. *Journal of Business Ethics* 101 (4):635-651.

Alstadsæter, A., W. Kopczuk, and K. Telle. 2019. Social networks and tax avoidance: evidence from a well-defined Norwegian tax shelter. *International Tax and Public Finance* 26 (6):1291-1328.

Andreoni, J., B. Erard, and J. Feinstein. 1998. Tax compliance. *Journal of Economic Literature* 36 (2):818-860.

Ariel, B. 2012. Deterrence and Moral Persuasion Effects on Corporate Tax Compliance: Findings from a Randomized Controlled Trial. *Criminology* 50 (1):27-69.

Bardach, E. 1989. Moral Suasion and Taxpayer Compliance. *Law & Policy* 11 (1):49-69.

Bazart, C., and A. Bonein. 2014. Reciprocal relationships in tax compliance decisions. *Journal of Economic Psychology* 40:83-102.

Becker, G. S. 1968. Crime and Punishment: An Economic Approach. *Journal of Political Economy* 76 (2):169-217.

Bergman, M. 2002. Who Pays for Social Policy? A Study on Taxes and Trust. *Journal of Social Policy* 31 (2):289-305.

Besley, T., A. Jensen, and T. Persson. 2019. Norms, enforcement, and tax evasion. In *NBER Working Paper Series: Working Paper 25575*: National Bureau of Economic Research.

Blumenthal, M., C. W. Christian, and J. Slemrod. 2001. Do Normative Appeals Affect Tax Compliance? Evidence From a Controlled Experiment in Minnesota. *National Tax Journal* 54 (1):125-138.

Bobek, D. D., A. M. Hageman, and C. F. Kelliher. 2013. Analyzing the Role of Social Norms in Tax Compliance Behavior. *Journal of Business Ethics* 115 (3):451-468.

Bott, K. M., A. W. Cappelen, E. Sorensen, and B. Tungodden. 2017. You've Got Mail: A Randomised Field Experiment on Tax Evasion. *SSRN Electronic Journal*.

Boyer, P., N. Dwenger, and J. Rincke. 2014. Do Taxes Crowd Out Intrinsic Motivation? Field-Experimental Evidence from Germany. In *Working Paper 2014 – 23*: Working Paper of the Max Planck Institute for Tax Law and Public Finance.

Braithwaite, V. 2003. Perceptions of Who's not Paying Their Fair Share. *Australian Journal of Social Issues (Australian Council of Social Service)* 38 (3):335-362.

Brehm, J. W. 1966. A theory of psychological reactance.

Carnes, G. A., and A. D. Cuccia. 1996. An analysis of the effect of tax complexity and its perceived justification on equity judgments. *The Journal of the American Taxation Association* 18 (2):40-56.

Casal, S., C. Kogler, L. Mittone, and E. Kirchler. 2016. Tax compliance depends on voice of taxpayers. *Journal of Economic Psychology* 56:141-150.

Castro, L., and C. Scartascini. 2015. Tax compliance and enforcement in the pampas evidence from a field experiment. *Journal of Economic Behavior & Organization* 116:65-82.

Chirico, M., R. Inman, C. Loeffler, J. MacDonald, and H. Sieg. 2019. Deterring Property Tax Delinquency in Philadelphia: An Experimental Evaluation of Nudge Strategies. *National Tax Journal* 72 (3):479-506.

Choo, C. Y. L., M. A. Fonseca, and G. D. Myles. 2016. Do students behave like real taxpayers in the lab? Evidence from a real effort tax compliance experiment. *Journal of Economic Behavior & Organization* 124:102-114.

Christian, C. R., and J. Alm. 2014. Empathy, sympathy, and tax compliance. *Journal of Economic Psychology* 40:62-82.

Christian, C. W., S. Gupta, and S. Lin. 1993. Determinants of Tax Preparer Usage: Evidence from Panel Data. *National Tax Journal* 46 (4):487-487.

Cialdini, R. B., and M. R. Trost. 1998. Social influence: Social norms, conformity and compliance. In *The handbook of social psychology, Vols. 1-2, 4th ed.* New York, NY, US: McGraw-Hill, 151-192.

Cinyabuguma, M., T. Page, and L. Putterman. 2005. Cooperation under the threat of expulsion in a public goods experiment. *Journal of Public Economics* 89 (8):1421-1435.

Clotfelter, C. T. 1983. Tax Evasion and Tax Rates: An Analysis of Individual Returns. *The Review of Economics and Statistics* 65 (3):363-373.

Coricelli, G., M. Joffily, C. Montmarquette, and M. C. Villeval. 2010. Cheating, emotions, and rationality: an experiment on tax evasion. *Experimental Economics* 13 (2):226-247.

Coricelli, G., E. Rusconi, and M.-C. Villeval. 2014. Tax evasion and emotions: An empirical test of re-integrative shaming theory. *Journal of Economic Psychology* 40:49-61.

Cowell, F. 1991. *Tax-evasion experiments: An economist's view*. edited by P. Webley, H. Robben, H. Elffers and D. J. Hessing: Cambridge University Press.

Cowell, F. A. 1992. Tax evasion and inequity. *Journal of Economic Psychology* 13 (4):521-543.

Cox, S. P., and R. J. Eger, III. 2006. Procedural Complexity of Tax Administration: The Road Fund Case. *Journal of Public Budgeting, Accounting & Financial Management* 18 (3):259-283.

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

Cullis, J., P. Jones, and A. Soliman. 2012b. 'Spite effects' in tax evasion experiments. *The Journal of Socio-Economics* 41 (4):418-423.

Cullis, J. G., and A. Lewis. 1997. Why people pay taxes: From a conventional economic model to a model of social convention. *Journal of Economic Psychology* 18 (2):305-321.

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

Dean, P., T. Keenan, and F. Kenney. 1980. Taxpayers' Attitudes to Income Tax Evasion: An Empirical Study. *British Tax Review* (1):28-44.

DeBacker, J., B. T. Heim, and A. Tran. 2015. Importing corruption culture from overseas: Evidence from corporate tax evasion in the United States. *Journal of Financial Economics* 117 (1):122-138.

Del Carpio, L. 2014. Are the Neighbors Cheating? Evidence from a Social Norm Experiment on Property Taxes in Peru.

Doerrenberg, P., and A. Peichl. 2018. Tax morale and the role of social norms and reciprocity: Evidence from a randomized survey experiment. In *ZA DP No. 11714*: Discussion Paper Series.

Doerrenberg, P., and J. Schmitz. 2015. Tax compliance and information provision–A field experiment with small firms. In *ZEW-Centre for European Economic Research Discussion Paper No. 15-028*.

Drago, F., F. Mengel, and C. Traxler. 2020. Compliance Behavior in Networks: Evidence from a Field Experiment. *American Economic Journal: Applied Economics* 12 (2):96-133.

Dubin, J. A., M. J. Graetz, M. A. Udell, and L. L. Wilde. 1992. The Demand for Tax Return Preparation Services. *The Review of Economics and Statistics* 74 (1):75-82.

Dulleck, U., J. Fooken, C. Newton, A. Ristl, M. Schaffner, and B. Torgler. 2016. Tax compliance and psychic costs: Behavioral experimental evidence using a physiological marker. *Journal of Public Economics* 134:9-18.

Dwenger, N., H. J. Kleven, I. Rasul, and J. Rincke. 2016. Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany. *American Economic Journal. Economic Policy* 8 (3):203-232.

Elster, J. 1989. Social Norms and Economic Theory. *Journal of Economic Perspectives* 3 (4):99-117.

Engel, C. 2011. Dictator games: a meta study. *Experimental Economics* 14 (4):583-610.

Engström, P., K. Nordblom, H. Ohlsson, and A. Persson. 2015. Tax Compliance and Loss Aversion. *American Economic Journal: Economic Policy* 7 (4):132-164.

Erard, B. 1993. Taxation with representation: An analysis of the role of tax practitioners in tax compliance. *Journal of Public Economics* 52 (2):163-197.

———. 1997. Self-selection with measurement errors A microeconometric analysis of the decision to seek tax assistance and its implications for tax compliance. *Journal of Econometrics* 81 (2):319-356.

Eriksen, K., and L. Fallan. 1996. Tax knowledge and attitudes towards taxation; A report on a quasi-experiment. *Journal of Economic Psychology* 17 (3):387-402.

Faizal, S. M., M. R. Palil, R. Maelah, and R. Ramli. 2017. Perception on justice, trust and tax compliance behavior in Malaysia. *Kasetsart Journal of Social Sciences* 38 (3):226-232.

Falk, A., and J. J. Heckman. 2009. Lab Experiments Are a Major Source of Knowledge in the Social Sciences. *Science* 326 (5952):535-538.

Falkinger, J. 1995. Tax evasion, consumption of public goods and fairness. *Journal of Economic Psychology* 16 (1):63-72.

Falkinger, J., and H. Walther. 1991. Rewards Versus Penalties: on a New Policy against Tax Evasion. *Public Finance Quarterly* 19 (1):67-79.

Farrar, J., S. E. Kaplan, and L. Thorne. 2019. The Effect of Interactional Fairness and Detection on Taxpayers' Compliance Intentions. *Journal of Business Ethics* 154 (1):167-180.

Feld, L. P., and B. S. Frey. 2002. Trust breeds trust: How taxpayers are treated. *Economics of Governance* 3 (2):87-99.

Feld, L. P., and G. Kirchgässner. 2000. Direct democracy, political culture, and the outcome of economic policy: a report on the Swiss experience. *European Journal of Political Economy* 16 (2):287-306.

Feld, L. P., and J.-R. Tyran. 2002. Tax Evasion and Voting: An Experimental Analysis. *Kyklos* 55 (2):197-221.

Fellner, G., R. Sausgruber, and C. Traxler. 2013. Testing Enforcement Strategies in the Field: Threat, Moral Appeal and Social Information. *Journal of the European Economic Association* 11 (3):634-660.

Filippin, A., C. V. Fiorio, and E. Viviano. 2013. The effect of tax enforcement on tax morale. *European Journal of Political Economy* 32:320-331.

Fishbein, M., and I. Ajzen. 1975. *Belief, attitude, intention and behaviour: An introduction to theory and research*. Vol. 27.

Fjeldstad, O.-H. 2004. What's trust got to do with it? Non-payment of service charges in local authorities in South Africa. *The Journal of Modern African Studies* 42 (4):539-562.

Fochmann, M., F. Hechtner, E. Kirchler, and P. Mohr. 2019. When happy people make society unhappy: Incidental emotions affect compliance behavior. *SSRN Electronic Journal*.

Fochmann, M., F. Hechtner, T. Kölle, and M. Overesch. 2020a. Combating Overreporting of Deductions in Tax Returns: Prefilling and Restricting the Deductibility of Expenditures. *SSRN Electronic Journal*.

Fochmann, M., T. Kölle, P. Mohr, and B. Rockenbach. 2020b. Trust Them, Threaten Them, or Lure Them? Effective Audit Systems to Promote Compliance. *SSRN Electronic Journal*.

Fochmann, M., N. Müller, and M. Overesch. 2018. Less cheating? The effects of prefilled forms on compliance behavior. *SSRN Electronic Journal*.

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

Frey, B. S. 1997a. A Constitution for Knaves Crowds out Civic Virtues. *The Economic Journal* 107 (443):1043-1053.

Frey, B. S. 1997b. *Not just for the money*: Edward Elgar.

Frey, B. S., and R. Jegen. 2001. Motivation Crowding Theory. *Journal of Economic Surveys* 15 (5):589-611.

Frey, B. S., and B. Torgler. 2007. Tax morale and conditional cooperation. *Journal of Comparative Economics* 35 (1):136-159.

Gangl, K., B. Torgler, and E. Kirchler. 2016. Patriotism's Impact on Cooperation with the State: An Experimental Study on Tax Compliance. *Political Psychology* 37 (6):867-881.

Gërxhani, K., and A. Schram. 2006. Tax evasion and income source: A comparative experimental study. *Journal of Economic Psychology* 27 (3):402-422.

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

Graetz, M. J., and L. L. Wilde. 1985. The economics of tax compliance: fact and fantasy. *National Tax Journal* 38 (3):355-363.

Groenland, E. A. G., and G. M. van Veldhoven. 1983. Tax evasion behavior: A psychological framework. *Journal of Economic Psychology* 3 (2):129-144.

Guala, F., and L. Mittone. 2010. How history and convention create norms: An experimental study. *Journal of Economic Psychology* 31 (4):749-756.

Halla, M. 2012. Tax morale and compliance behavior: First evidence on a causal link. *The BE Journal of Economic Analysis & Policy* 12 (1):1-25.

Hallsworth, M. 2014. The use of field experiments to increase tax compliance. *Oxford Review of Economic Policy* 30 (4):658-679.

Hallsworth, M., J. A. List, R. D. Metcalfe, and I. Vlaev. 2017. The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of Public Economics* 148:14-31.

Hammar, H., S. C. Jagers, and K. Nordblom. 2009. Perceived tax evasion and the importance of trust. *The Journal of Socio-Economics* 38 (2):238-245.

Hartner-Tiefenthaler, M., B. Kubicek, E. Kirchler, S. Rechberger, and M. Wenzel. 2012. Perceived Distributive Fairness of European Transfer Payments and EU-Taxes in Austria, the Czech Republic, and the United Kingdom. *Applied Psychology* 61 (3):454-478.

Hartner, M., S. Rechberger, E. Kirchler, and A. Schabmann. 2008. Procedural Fairness and Tax Compliance. *Economic Analysis and Policy* 38 (1):137-152.

Hashimzade, N., G. D. Myles, and B. Tran-Nam. 2013. Applications of Behavioural Economics to Tax Evasion. *Journal of Economic Surveys* 27 (5):941-977.

Hasseldine, J. 1998. Tax amnesties: an international review. *Bulletin for international fiscal documentation* 52:303-310.

Henderson, J. V., A. Storeygard, and D. N. Weil. 2012. Measuring Economic Growth from Outer Space. *American Economic Review* 102 (2):994-1028.

Hofmann, E., E. Hoelzl, and E. Kirchler. 2008. Preconditions of Voluntary Tax Compliance: Knowledge and Evaluation of Taxation, Norms, Fairness and Motivation to Cooperate. *Zeitschrift für Psychologie/Journal of Psychology* 216 (4):209-217.

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

Hunt, N. C., G. S. Iyer, and P. Jimenez. 2019. Election Outcome and Tax Compliance: The Role of Political Party Affiliation, Affect Balance, and Trust in Government. *Applied Psychology* 68 (2):341-372.

Jacquemet, N., S. Luchini, A. Malézieux, and J. F. Shogren. 2019. A Psychometric Investigation of the Personality Traits Underlying Individual Tax Morale. *B.E. Journal of Economic Analysis and Policy* 19 (3):1-25.

Jimenez, P., and G. S. Iyer. 2016. Tax compliance in a social setting: The influence of social norms, trust in government, and perceived fairness on taxpayer compliance. *Advances in Accounting* 34:17-26.

Job, J., A. Stout, and R. Smith. 2007. Culture Change in Three Taxation Administrations: From Command-and-Control to Responsive Regulation. *Law & Policy* 29 (1):84-101.

Jones, C. 1996. Mass-based income taxation: Creating a taxpaying culture, 1940-1952. In *Funding the Modern American State*, edited by W. E. Brownlee: New York: Cambridge University Press, 107-147.

Kang, S. W., and H. Rockoff. 2015. Capitalizing patriotism: the Liberty loans of World War I. *Financial History Review* 22 (1):45-78.

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

Kinsey, K. A., and H. G. Grasmick. 1993. Did the Tax Reform Act of 1986 Improve Compliance? Three Studies of Pre- and Post-TRA Compliance Attitudes. *Law & Policy* 15 (4):293-325.

Kinsey, K. A., H. G. Grasmick, and K. W. Smith. 1991. Framing Justice: Taxpayer Evaluations of Personal Tax Burdens. *Law & Society Review* 25 (4):845-873.

Kirchler, E. 2007. *The economic psychology of tax behaviour*: Cambridge University Press.

Kirchler, E., E. Hoelzl, and I. Wahl. 2008. Enforced versus voluntary tax compliance: The "slippery slope" framework. *Journal of Economic Psychology* 29 (2):210-225.

Kirchler, E., and B. Maciejovsky. 2001. Tax compliance within the context of gain and loss situations, expected and current asset position, and profession. *Journal of Economic Psychology* 22 (2):173-194.

Kirchler, E., A. Niemirowski, and A. Wearing. 2006. Shared subjective views, intent to cooperate and tax compliance: Similarities between Australian taxpayers and tax officers. *Journal of Economic Psychology* 27 (4):502-517.

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez. 2011. Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark. *Econometrica* 79 (3):651-692.

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

Konrad, K. A., and S. Qari. 2012. The Last Refuge of a Scoundrel? Patriotism and Tax Compliance. *Economica* 79 (315):516-533.

Kountouris, Y., and K. Remoundou. 2013. Is there a cultural component in tax morale? Evidence from immigrants in Europe. *Journal of Economic Behavior and Organization* 96:104-119.

Kuchumova, Y. 2017. The Optimal Deterrence of Tax Evasion: The Trade-off Between Information Reporting and Audits. *Journal of Public Economics* 145:162-180.

Lamberton, C., J.-E. De Neve, and M. I. Norton. 2018. The Power of Voice in Stimulating Morality: Eliciting Taxpayer Preferences Increases Tax Compliance. *Journal of Consumer Psychology* 28 (2):310-328.

Lefebvre, M., P. Pestieau, A. Riedl, and M. C. Villeval. 2015. Tax evasion and social information: an experiment in Belgium, France, and the Netherlands. *International Tax and Public Finance* 22 (3):401-425.

Long, J. E., and S. B. Caudill. 1987. The Usage and Benefits of Paid Tax Return Preparation. *National Tax Journal* 40 (1):35-35.

Lubian, D., and L. Zarri. 2011. Happiness and tax morale: An empirical analysis. *Journal of Economic Behavior & Organization* 80 (1):223-243.

Luttmer, E. F. P., and M. Singhal. 2014. Tax Morale. *Journal of Economic Perspectives* 28 (4):149-168.

Mascagni, G. 2018. From the Lab to the Field: A Review of Tax Experiments. *Journal of Economic Surveys* 32 (2):273-301.

Mascagni, G., C. Nell, and N. Monkam. 2017. One size does not fit all: a field experiment on the drivers of tax compliance and delivery methods in Rwanda. In *ICTD Working Paper 58*.

Mason, L. 2015. "I Disrespectfully Agree": The Differential Effects of Partisan Sorting on Social and Issue Polarization. *American Journal of Political Science* 59 (1):128-145.

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

Meiselman, B. S. 2018. Ghostbusting in Detroit: Evidence on nonfilers from a controlled field experiment. *Journal of Public Economics* 158:180-193.

Mendoza, J. P., J. L. Wielhouwer, and E. Kirchler. 2017. The backfiring effect of auditing on tax compliance. *Journal of Economic Psychology* 62:284-294.

Muehlbacher, S., B. Hartl, and E. Kirchler. 2017. Mental Accounting and Tax Compliance:Experimental Evidence for the Effect of Mental Segregation of Tax Due and Revenue on Compliance. *Public Finance Review* 45 (1):118-139.

Murphy, K. 2003. Procedural justice and tax compliance. *Australian Journal of Social Issues* 38 (3):379-408.

———. 2004. The Role of Trust in Nurturing Compliance: A Study of Accused Tax Avoiders. *Law and Human Behavior* 28 (2):187-209.

Nagel, H., L. Rosendahl Huber, M. Van Praag, and S. Goslinga. 2019. The effect of a tax training program on tax compliance and business outcomes of starting entrepreneurs: Evidence from a field experiment. *Journal of Business Venturing* 34 (2):261-283.

Niemirowski, P., S. Baldwin, and A. J. Wearing. 2003. Tax Related Behaviours, Beliefs, Attitudes and Values and Taxpayer Compliance in Australia. *Journal of Australian Taxation* 6 (1):132-165.

OECD. 2019. Tax Morale - What drives people and businesses to pay tax?, edited by O. Publishing. Paris: OECD Publishing, 1-68.

Olsen, J., M. Kasper, J. Enachescu, S. Benk, T. Budak, and E. Kirchler. 2018. Emotions and tax compliance among small business owners: An experimental survey. *International Review of Law and Economics* 56:42-52.

Ortega, D., and C. Scartascini. 2015. Don't blame the messenger: A Field Experiment on Delivery Methods for Increasing Tax Compliance. In *CAF – Working paper N° 2015/09*.

Orviska, M., and J. Hudson. 2002. Tax evasion, civic duty and the law abiding citizen. *European Journal of Political Economy* 19 (1):83-102.

Park, C.-G., and J. K. Hyun. 2003. Examining the determinants of tax compliance by experimental data: a case of Korea. *Journal of Policy Modeling* 25 (8):673-684.

Picciotto, S. 2007. Constructing Compliance: Game Playing, Tax Law, and the Regulatory State. *Law & Policy* 29 (1):11-30.

Pommerehne, W. W., and H. Weck-Hannemann. 1996. Tax rates, tax administration and income tax evasion in Switzerland. *Public Choice* 88 (1):161-170.

Rawlings, G. 2003. Cultural narratives of taxation and citizenship: Fairness, groups and globalisation. *Australian Journal of Social Issues* 38 (3):269-306.

Rees-Jones, A. 2014. Loss aversion motivates tax sheltering: Evidence from US tax returns. *SSRN Electronic Journal*.

Rheinberg, F., and S. Engeser. 2018. Motivation and action: Introduction and overview. In *Motivation and action*, edited by J. Heckhausen and H. Heckhausen: Springer, 423-450.

Richardson, G. 2008. The relationship between culture and tax evasion across countries: Additional evidence and extensions. *Journal of International Accounting, Auditing and Taxation* 17 (2):67-78.

Roberts, M. L., and P. A. Hite. 1994. Progressive Taxation, Fairness, and Compliance. *Law & Policy* 16 (1):27-48.

Rodriguez-Justicia, D., and B. Theilen. 2018. Education and tax morale. *Journal of Economic Psychology* 64:18-48.

Saad, N. 2014. Tax Knowledge, Tax Complexity and Tax Compliance: Taxpayers' View. *Procedia - Social and Behavioral Sciences* 109:1069-1075.

Sausgruber, R., and H. Winner. 2004. Steueramnestie abgesagt: Eine kluge Entscheidung? Empirische Evidenz aus OECD-Ländern. *Österreichische Steuer-Zeitung* 57:207-211.

Slemrod, J. 2007. Cheating Ourselves: The Economics of Tax Evasion. *Journal of Economic Perspectives* 21 (1):25-48.

———. 2016. Tax Compliance and Enforcement: New Research and Its Policy Implications. In *Ross School of Business Paper No. 1302*.

Slemrod, J., M. Blumenthal, and C. Christian. 2001. Taxpayer response to an increased probability of audit: evidence from a controlled experiment in Minnesota. *Journal of Public Economics* 79 (3):455-483.

Spicer, M. W., and L. A. Becker. 1980. Fiscal Inequity and Tax Evasion: An Experimental Approach. *National Tax Journal* 33 (2):171-175.

Taylor, N. 2003. Understanding taxpayer attitudes through understanding taxpayer identities. In *Working Paper No 14*: Centre for Tax System Integrity, 71-92.

The Economist. 2018. Taxation - What about representation? *The Economist*, December 1st 2018.

Torgler, B. 2002. Speaking to Theorists and Searching for Facts: Tax Morale and Tax Compliance in Experiments. *Journal of Economic Surveys* 16 (5):657-683.

———. 2003. Tax Morale, Rule-Governed Behaviour and Trust. *Constitutional Political Economy* 14 (2):119-140.

———. 2004. Moral Suasion: An Alternative Tax Policy Strategy? Evidence from a Controlled Field Experiment in Switzerland. *Economics of Governance* 5:235-253.

———. 2005. Tax morale and direct democracy. *European Journal of Political Economy* 21 (2):525-531.

———. 2007. *Tax compliance and tax morale: A theoretical and empirical analysis*: Edward Elgar Publishing.

———. 2013. A Field Experiment on Moral Suasion and Tax Compliance Focusing on Under-Declaration and Over-Deduction. *FinanzArchiv / Public Finance Analysis* 69 (4):393-411.

Torgler, B., and F. Schneider. 2005. Attitudes Towards Paying Taxes in Austria: An Empirical Analysis. *Empirica* 32 (2):231-250.

Trivedi, V. U., M. Shehata, and B. Lynn. 2003. Impact of Personal and Situational Factors on Taxpayer Compliance: An Experimental Analysis. *Journal of Business Ethics* 47 (3):175-197.

Tyler, T. R., and H. J. Smith. 1998. Social Justice and Social Movements. In *The handbook of social psychology*, edited by D. T. Gilbert, S. T. Fiske and G. Lindzey: Oxford University Press.

Verboon, P., and M. van Dijke. 2011. When do severe sanctions enhance compliance? The role of procedural fairness. *Journal of Economic Psychology* 32 (1):120-130.

Wahl, I., B. Kastlunger, and E. Kirchler. 2010a. Trust in Authorities and Power to Enforce Tax Compliance: An Empirical Analysis of the "Slippery Slope Framework". *Law & Policy* 32 (4):383-406.

Wahl, I., S. Muehlbacher, and E. Kirchler. 2010b. The Impact of Voting on Tax Payments. *Kyklos* 63 (1):144-158.

Wartick, M. L. 1994. Legislative justification and the perceived fairness of tax law changes: A referent cognitions theory approach. *The Journal of the American Taxation Association* 16 (2):106-106.

Weck-Hannemann, H., and W. W. Pommerehne. 1989. Einkommensteuerhinterziehung in der Schweiz: Eine empirische Analyse. *Swiss Journal of Economics and Statistics (SJES)* 125 (IV):515-556.

Wenzel, M. 2002. The Impact of Outcome Orientation and Justice Concerns on Tax Compliance: The Role of Taxpayers' Identity. *The Journal of applied psychology* 87:629-645.

———. 2003. Tax compliance and the psychology of justice: Mapping the field. In *Taxing Democracy: Understanding Tax Avoidance and Evasion*, 41-70.

———. 2004. An analysis of norm processes in tax compliance. *Journal of Economic Psychology* 25 (2):213-228.

———. 2007. The Multiplicity of Taxpayer Identities and Their Implications for Tax Ethics. *Law & Policy* 29 (1):31-50.

Wenzel, M., and I. Thielmann. 2006. Why We Punish in the Name of Justice: Just Desert versus Value Restoration and the Role of Social Identity. *Social Justice Research* 19 (4):450-470.

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

CHAPTER 5

Concluding remarks

This thesis expands the understanding of tax compliance of individuals. The three essays provide novel insights that might be especially relevant for governments combating tax evasion.

Chapter 2 studies the design of effective audit systems to promote compliance. In particular, it analyzes the effects of (1) increasing the detection probability of non-compliance, (2) monetary incentives to promote compliance (bonuses and penalties), and the interaction of (1) and (2). In conclusion, compliance decreases with audit systems that penalize non-compliance (or reward compliance) with a low detection probability compared to a situation without any audits. Only a penalty system with a high detection probability ensures higher compliance than without audits.

Chapter 3 analyzes three mechanisms that might affect tax compliance: 1) prefilling of deductions in tax returns, 2) restricting tax evasion opportunities by either disallowing or 3) limiting the deductibility of expenditures. In conclusion, prefilling reduces tax evasion compared to blank forms. Contrary, cutting the number of tax evasion opportunities by disallowing the deductibility of expenditure items is ineffective. In fact, individuals shift their tax evasion activities from the disallowed item to non-restricted items. However, this evasion-shift-effect seems to be avoidable by just limiting the deductibility of expenditures.

Chapter 4 reviews the non-monetary costs of tax evasion. Backed up with a theoretical framework of non-monetary costs of tax evasion, the chapter emphasizes the importance of non-monetary costs of tax evasion. In particular, it explains the influence of social norm, trust, fairness, participation rights, complexity and knowledge, patriotism, and moral emotions on tax compliance behavior.

In conclusion and if I were to give advice to governments combating tax evasion, I would stress three insights based on my research on tax compliance. First, governments should rethink about audits that are performed with a (known) low probability. Either governments should increase the audit probability or abolish these audits and utilize the intrinsic motivation of taxpayers. Second, governments should restrict the deductibility of expenditures with caution because of potential evasion-shift-effects. In fact, they should utilize prefilled tax returns. Third, governments should strengthen their influence on non-monetary costs of tax evasion (e.g. enhance the social norm of tax compliance, trust in tax authorities, fairness etc.). These factors have a measurable effect on tax compliance and are often forgotten compared to audit probability, fine and tax rate.

CURRICULUM VITAE

Tobias Kölle

Personal Data

Date of Birth	1991
Place of Birth	Munich
Citizenship	German

Education

PhD Student and Research Assistant — **Cologne**
Chair of Business Taxation, Prof. Dr. Michael Overesch — Apr 2018-Oct 2020
University of Cologne

Master of Science — **Cologne**
University of Cologne — Oct 2015-Mar 2018

Bachelor of Science — **Cologne**
University of Cologne — Apr 2012-Apr 2015

Allgemeine Hochschulreife (A-Levels) — **Idstein**
Pestalozzischule Idstein — Sep 2004-Jun 2011

Cologne, September 2020