



MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„The Ambiguous Effects of Penalties on Tax Compliance -
A Systematic Literature Review“

verfasst von / submitted by

Julian Rusche, BSc BA

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Master of Science (MSc)

Wien, 2018 / Vienna 2018

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

A 066 840

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Masterstudium Psychologie

Betreut von / Supervisor:

Univ.-Prof. Dr. Erich Kirchler

Mitbetreut von / Co-Supervisor:

Mag. Matthias Kasper, PhD

Table of Contents

Abstract	3
Introduction	4
<i>Standard Model of Income Tax Evasion</i>	5
<i>Slippery Slope Framework</i>	6
<i>Deterrence effect of penalties</i>	9
<i>Research Question</i>	12
Method	13
<i>Systematic Quantitative Approach</i>	13
<i>Selection criteria</i>	15
<i>PRISMA Statement</i>	15
<i>Coding procedure</i>	18
Results	19
<i>Quantitative Summary</i>	19
<i>Summary of qualitative conditions</i>	25
Discussion	28
<i>Limitations</i>	30
<i>Conclusion</i>	31
References	32
Appendix	40

Abstract

Penalties are the prevalent tool for policy-makers around the world, trying to decrease tax evasion. However, empirical studies testing the predictions of the established economic standard model of income tax evasion yielded ambiguous results regarding the deterrent effect of penalties in the tax context. More recent psychological models like the *Slippery Slope Framework* introduced the idea that fines in the tax context are subject to interpretation and depend on situational factors and noneconomic variables. The present study contributes to the current state of research by systematically and exhaustively searching for academic literature and summing up all empirical findings concerning the effect of penalties on tax compliance ($N=85$). Results are overall ambiguous, indicating, that evidence for the deterrent effect of penalties is at best weak. Further, we suggest that many of the counterintuitive findings of empirical studies in regards to penalties could be explained within the concepts of the *Slippery Slope Framework*.

Keywords: tax compliance, tax evasion, tax audit, penalty, fine, punishment

Introduction

„Our new Constitution is now established, and has an appearance that promises permanency; but in this world nothing can be said to be certain, except death and taxes“ is the original wording of the now famous quote that Benjamin Franklin, one of the Founding Fathers of the United States of America, wrote in a letter to French scientist Jean-Baptiste Leroy (Smyth, 1907, p.69).

With this choice of words Franklin implied that even those men and women who fled from the aristocratic governments in Europe and founded a new republic, a land of endless opportunities, had to pay their taxes, like all before and after them.

It seems like collecting taxes and reinvesting the revenue into public services, like infrastructure and education, is a necessary inconvenience for people to successfully live together in all modern day societies. As certain as taxes might be, they are not at all inevitable. Individuals might not like to pay taxes and engage in tax evasion by intentionally underreporting incomes, sales or wealth, or by overstating deductions or credits (Alm & Torgler, 2011). Not only does tax evasion reduce the government's ability to provide public services, it also increases national debt, therefore creating an even higher tax liability for honest taxpayers and future generations, that will likely lead to more discontent with the government and even less tax compliance (Cebula & Feige, 2012).

In practice tax-authorities around the world rely heavily on auditing taxpayers and punishing evaders with penalties, typically monetary fines in an attempt to increase compliance. They seem to act upon their conventional wisdom and presume that sanctions have a preventive effect that increases with its severity (Schwartz & Orleans, 1967). The assumption that audits and penalties effectively reduce tax evasion is supported by early theoretical models (Allingham & Sandmo, 1972; Yitzhaki, 1974). However, empirical evidence concerning the supposedly straightforward effect of these classical deterrents is inconclusive. While many studies did find that more severe sanctions are related to less tax evasion (e.g. Alm, Jackson, & Mckee, 1992; Alm, Sanchez, & De Juan, 1995), especially in recent years researchers seem to suggest, that this might not necessarily be the case (e.g. Alm, Clark, & Leibel, 2016; Ariel, 2012; Mendoza, Wielhouwer, & Kirchler, 2017). Furthermore, recent studies also suggest that a variety of noneconomic factors might be equally important in understanding tax compliance behavior (e.g. Feld & Frey, 2002; Hofmann, Hoelzl, & Kirchler, 2008). So while in practice penalties are in widespread use and their effectiveness seems practically undisputed, in the scientific community it is not clear whether they have

their intended deterring effect, as our knowledge of the reasons behind tax compliance behavior is still rather limited (Alm, McClelland, & Schulze, 1992).

Faced with this discrepancy this literature review aims to clarify the effect of penalties on tax compliance by exhaustively collecting and analyzing academic literature concerning the topic. In the following we examine two theoretical models that made major contributions to our understanding of tax compliance behavior and depict the current state of research concerning the effect of penalties on tax compliance. Subsequently, we derive our research questions from the presented literature. Then we explain the method that was used to conduct this literature review in a systematic and objective way. Next, we present the results, separated into two parts, a quantitative summary of the findings and a more explorative part. Finally, we discuss our findings and possible limitations of this study.

Standard Model of Income Tax Evasion

The *Standard Model of Income Tax Evasion* is somewhat an umbrella term that refers to early economic analyses of income tax compliance decisions. It is usually associated with the pioneering work of Allingham and Sandmo (1972), but papers from Srinivasan (1973) and Yitzhaki (1974) made major contributions as well. In the model taxpayers are assumed to conform to the Von Neumann-Morgenstern axioms for behavior under uncertainty. That means whether to pay taxes or to evade them is a decision under risk and decision-makers will always choose the extent of their evasion in a way that maximizes their expected utility (Allingham & Sandmo, 1972). Assuming nobody would voluntarily pay more taxes than they are required too, decision makers in the income tax context have two options; an agent can either declare his actual income, or declare less than his actual income. Allingham and Sandmo (1972) propose that this compliance decision depends on the four major factors: level of actual income, tax rates, audit probability and magnitude of fines. As for the deterrence variables audit-frequency and fine-rate, the model's predictions are very clear and unambiguous: Tax evasion decreases with an increase in audit probability and magnitude of fines. As for income and tax rate accurate predictions become increasingly difficult. The authors predict that evasion increases, when income increases, however that presumes that absolute risk aversion is decreasing (Myles, Hashimzade, & Tran-Nam, 2012). The relationship between tax rate and compliance is even more complex and depends on how penalties are imposed (Lee, 2015). In the Allingham and Sandmo model (1972) penalties are imposed on undeclared income. This is problematic, because it causes a substitution effect and an income effect that offset each other. Meaning that on the one hand higher marginal tax rates make evasion more attractive (substitution effect), but on the other hand higher tax rates

also lower the net income (income effect). So while it generally makes sense to evade more taxes, when tax rates are higher, assuming decreasing absolute risk aversion, also means that lower income makes taxpayers less willing to take risks, leading to the opposite conclusion that taxpayers will evade less, when tax rates increase (Borck, 2004). Therefore, the predictions regarding tax rate and evasion are unclear in the Allingham and Sandmo model (1972). In an attempt to clarify this, Yitzhaki (1974) modified the model by imposing penalties on taxes evaded not on undeclared income. He predicts that evasion will decrease with increasing tax rates. That seems counter-intuitive, but can be explained with the magnitude of fines rising simultaneously with the tax rates. The benefits from evading higher tax rates are therefore balanced out with the risk of higher fines, undoing the substitution effect. As the income effect still operates in Yitzhaki's model, it makes sense that a higher marginal tax rate results in less tax evasion, again assuming decreasing absolute risk aversion (Borck, 2004).

The publication of the standard model simulated a variety of tax research. However, researchers found empirical evidence for and against the predictions of all four main parameters of the standard model. These overall ambiguous results suggest that the economic approach, where all taxpayers are assumed to be *homines oeconomici*, is too narrow to fully explain the compliance behavior we observe in reality (Kirchler, Muehlbacher, Kastlunger, & Wahl, 2007).

One of the more notable shortcomings of the model is its inability to explain the rather larger number of taxpayers that are compliant. If we were to apply audit probabilities and fine rates observed in practice, the model predicts that all taxpayers should engage in tax evasion (Myles et al., 2012). Obviously that is not what we observe in reality as many people still pay their taxes. This phenomenon is also referred to as the compliance paradox. Consequently, researches in the 1990s started focusing on the question why so many people do not cheat in the tax contest (Torgler, 2003).

Slippery Slope Framework

One of the more recent and successful ideas that try to make sense of the compliance paradox is the so-called *Slippery Slope Framework* by Kirchler, Hoelzl and Wahl (2008). Their approach is based on distinguishing between enforced compliance and voluntary compliance. In the slippery slope model citizens who comply with tax laws, do so because they either calculate the cost for non-compliance as being too high (enforced) or because they feel obligated to do so as members of the community (voluntary). Which one of the two forms of compliance prevails in any given country depends on the interaction between taxpayers and

authorities. If authorities aim for transparent procedures and for respectful and supportive treatment of taxpayers, the authors speak of a synergistic climate. Voluntary compliance tends to be high in this “service and client” approach towards tax collection, as shown for example in Switzerland (Feld & Frey, 2002) or Australia (Holmes, 2011). On the other hand, if authorities assume taxpayers are cheaters and have to be monitored, citizens will feel persecuted and both parties will adopt a “cops and robbers” attitude. Voluntary compliance is negligible in such an antagonistic climate and the only way to collect taxes is to enforce them.

As the model assumes the interaction between tax authorities and taxpayers to be crucial, not only tax authorities regulations but also taxpayers’ attitudes towards the tax authorities are decisive factors for compliance behavior. Braithwaite (2003) defines five motivational postures to explain what attitudes citizens can have towards regulatory systems. “Commitment” and “Capitulation” are two postures that result out of a positive relationship between citizens and authorities. Committed taxpayers believe collecting tax revenue and reinvesting it into public goods is fair and feel morally obliged to pay their taxes. Capitulated taxpayers acknowledge that tax authorities are a legitimate power that serves to reach collective goals and cooperate with them.

A negative relationship between citizens and authorities can result in defiance behavior, that Braithwaite describes in the motivational postures “Resistance”, “Disengagement” and “Game-playing”. Resistant taxpayers are suspicious towards tax authorities intentions and doubt their authorities. Taxpayers with that motivational posture might not be compliant, however they are a healthy part of the tax system in total, that helps the system evolve over time. On the other hand, disengaged taxpayers are not beneficial for the system at all. They keep a great distance from authorities and general oppose the law. Finally, game-playing taxpayers compete with the tax authorities and try to exploit laws and loopholes in order to maximize their own utility.

Postures of commitment and capitulation are found to be linked positively to voluntary compliance. However, taxpayers whose compliance is enforced tend to not perceive tax authorities as a legitimate power. Therefore enforced compliance is positively correlated with the defiance postures “Resistance”, “Disengagement” and “Game-playing” and negatively correlated with the deference postures “Commitment” and “Capitulation” (Kirchler & Wahl, 2010).

The dynamics of the *Slippery Slope Framework* can be best described in the two dimensions “power of authorities” and “trust in authorities”. Power in this setting is the authorities’ ability to prosecute and punish tax evaders. Trust is defined by whether taxpayers

perceive the authorities as trustworthy and benevolent or not. The framework assumes that the two dimensions moderate each other as described by Kirchler et al. (2008):

The power and trust dimensions moderate each other. Variations in trust matter most when power is low; when power is at its maximum, variations in trust are irrelevant because authorities can enforce maximum compliance. Conversely, variations in power matter most when trust is low; however, when trust is at its maximum, variations in power do not matter because citizens contribute their share regardless. (p. 213)

As for the deterrence variables audit-frequency and fine-rate, that were positively related to compliance in the standard model, they are now subject to interpretation. High fines and frequent audits might increase compliance in an antagonistic climate, as they heighten authorities' power, whereas in a synergistic climate they might be construed as unfair or an act of distrust from the government. Subsequently the trust dimension will subside and so will voluntary compliance and compliance in total. Respectively, tax policies based on deterrence can be effective in climates where trust is low. However, in a climate of trust noneconomic parameters, like social norms, tax morale, tax knowledge or perceived fairness of the tax system might have a significant impact on tax compliance as well (Kirchler et al., 2008).

Multiple empirical studies found evidence supporting the predictions of the *Slippery Slope Framework* that trust in the government increases voluntary tax compliance (e.g. Kogler, Muehlbacher, & Kirchler, 2015; Wahl, Kastlunger, & Kirchler, 2010). Kaplanoglou and Rapanos (2015) found, that a high level of power generally increases enforced compliance while it does not have any effect on voluntary compliance, as long as trust is on a certain level. In low trust climates, however, higher power might decrease voluntary compliance even further.

The model also came as part of a new wave of research, investigating tax compliance behavior from a more psychological perspective, as the standard model seems to lose validity as soon as taxpayers have different motivational postures than just maximizing their utility (Kirchler et al., 2007). Several researchers showed relations between noneconomic factors and tax compliance. For example priming taxpayers to feel sympathy or empathy makes them less likely to evade (Calvet -Christian & Alm, 2014). Also social norms (Wenzel, 2004), personal norms (Jimenez & Iyer, 2016), political orientation (Wahlund, 1992) and religious beliefs (Torgler, 2006) have been shown to be related to tax compliance behavior. As for deterrence variables, contrary to prior beliefs, increasing the frequency of audits does not always lead to more compliance (Mendoza et al., 2017). For example, auditing already

compliant citizens can even lead to a decrease in tax compliance (Beer, Kasper, Kirchler, & Erard, 2015). Further, Cullis, Jones and Soliman (2012) found that if enforcement variables are considered to be excessive, they will lose their deterrence effect and will not lead to less tax evasion. This could be explained with the concept of retributive justice: Kirchler et al. (2008) propose that the perceived appropriateness of sanctions is part of the overall perceived fairness towards the tax system and therefore impacts compliance behavior.

These results can be explained within the *Slippery Slope Framework* and suggest that the “tax climate” of a country is equally important for its citizens compliance, as the actual tax policies themselves.

Deterrence effect of penalties

The *Slippery Slope Framework* creates a framework that incorporates noneconomic factors and the four main factors from the standard model in order to better describe real world tax compliance behavior. However in regards to penalties, it is still disputed, what exactly their effect on tax compliance is. Regardless, in practice penalties are commonly used in an attempt to deter people from evading taxes. Enforcement policies might even rely on them too heavily at times: While other deterrence strategies, like more frequent-audits, are costly, penalties actually provide a form of revenue. Therefore, governments tend to raise penalties before investing in other strategies that could increase tax compliance (Stella, 1991). In theory, governments could even set penalty-rates so high, that they would increase the state revenue, without actually increasing tax compliance (Andreoni, 1992). These findings can be attributed to the fact, that in practice penalties for tax evasion are mostly monetary fines, which do not use up social resources; On the contrary they are thought to increase social welfare and should be used whenever feasible (Becker, 1968).

In Austria, for example, the amount of a fine is calculated based on the following circumstances: The amount that was evaded, whether someone evaded for the first time or repeatedly, and whether it was done with intention or by accident. Usually the amount evaded has to be paid back together with a fine of 100% of the amount evaded (Torgler & Schneider, 2005). Fines can go up to 200% of the evaded amount or up to two years imprisonment. Longer prison terms only occur in serious cases of fraud, for example when falsified documents were used or when fictitious business structures were reported to the authorities (Schachner & Loukota, 2017). However, in practice sentences are often more lenient with fines typically being around 40% of the evaded amount (Kirchler et al., 2007).

Given the importance of penalties for real world policymaking, researchers naturally explored the topic in more detail theoretically and with empirical studies. The majority of this

research focuses on individual income tax evasion. A variety of more recent theoretical models come to conclusions that are in accordance with the standard model. Cho, Linn and Nakibullah (1996) argue, that more audits and higher penalties have a positive effect on compliance as they decrease incentives for people to evade. Chen (2003) and Levaggi and Menoncin (2016) also predict that tax evasion decreases with increasing penalties and Yoon et al. (2011) report that higher penalties are beneficial, as they lead to less aggressive tax reporting in general.

Other models suggest that the relation between penalties and compliance is not always positive. In practice penalties might be most effective at a non-maximum, as higher fines would induce more regulatory errors, like more bribes and more corrupt auditors and therefore decrease compliance overall (Bose, 1995). Fulton and Cule (2009) elaborate on this idea, pointing out that in a system, where corruption and bribes are already high, an increase in penalties would even result in an increase of tax evasion. Alm, Martinez-Vazquez and McClellan (2016) found empirical evidence for the positive relation between corruption and tax evasion, supporting the notion that classical deterrence strategies can only work properly in honest tax administrations.

Taking into account social norms, Davis, Hecht and Perkins (2003) argue that raising the penalty-rate has a rather small effect on initially noncompliant populations until it becomes sufficiently harsh, where they predict a sudden change to very high compliance levels in that population. Korobow, Johnson and Axtell (2007) show that for taxpayers, who know about their neighbors' compliance behavior, an increase in the penalty-rate leads to more under-reporting, while for taxpayers who do not know about their neighbors' actions, higher penalties have the intended deterring effect. Finally, assuming that collaborative tax evasion becomes the social norm, higher penalty-rates will result in a snowball effect, that further strengthens resistance towards the tax system and only leads to more tax evasion (Chang & Lai, 2004).

Considering the variety of predictions observed in theoretical studies, it does not come as a surprise, that empirical evidence for the effect of penalties on tax compliance is ambiguous as well.

Multiple experimental studies report a positive relation between severity of penalties and tax compliance (e.g. Beck, Davis, & Jung, 1991; Bruttel & Friehe, 2014). However, the positive effect is sometimes very small and almost negligible (Alm, Jackson, & McKee, 1992). It is also disputed, whether the fine-rate has a bigger effect on compliance than the audit-frequency, or whether it is the other way around. The former is supported by findings in

experiments from Friedland, Maital and Rutenberg (1978) and from Park and Hyun (2003). However, in a subsequent experiment Nehemia Friedland (1982) also found evidence for the latter.

Another common finding in experiments is, that penalties do not have any deterring effect on tax evasion at all (e.g. Hite & McGill, 1992; Spicer & Lundstedt, 1976; Webley & Halstead, 1986; Webley, Robben, Elffers, & Hessing, 1991). Empirical studies reporting a negative relation between penalties and compliance seem to be the exception, but they do occur. For example, Coricelli, Joffily, Montmarquette and Claire (2010) conducted a tax experiment, where higher fines led to more tax evasion in subsequent rounds.

Studies that analyze real world aggregate data report similar results. Multiple authors show that greater penalties are correlated with less evasion, higher reported income and smaller underground economies (Cebula, 1997, 2011; Crane & Nourzad, 1985; Lesnik, Kracun, & Jagric, 2014). Again effect sizes might be rather small, for example Feld and Larsen (2012) also report, that larger fines result in less undeclared work, however that effect is not significant. Just like with experimental studies, plenty of studies analyzing aggregate data find that penalties have no deterring effect at all (Alm, Clark, et al., 2016; Mehlkop & Becker, 2004; Pommerehne & Weck-Hannemann, 1996). Negative correlations between penalties and compliance seem to be the exception, but do also occur: For example Alm, Bahl and Murray (1990) found that higher penalties are associated with less reported income in Jamaica.

The overall effect of penalties on compliance has been summed up in several meta-analyses and literature reviews. All reviews support the standard model's predictions regarding the positive direction of the effect. Blackwell (2007) states, that there is strong evidence for the positive relation between penalties and compliance. Alm and McKee (1998) conclude, that raising penalties marginally increases compliance, however the compliance decision depends on a variety of non-economic factors as well. Bărbuță-Mișu (2011) finds that, although effects tend to be small, most studies report a positive relation between raising penalties and tax compliance. The authors also note, that very high penalties might result in taxpayers perceiving the tax system as unjust, which can result in more tax avoidance or even tax evasion behavior. Other reviews come to similar conclusions, stating that the evidence for a positive relation between fines and compliance is weak and the effect is at best very small, if not negligible (Andreoni, Erard, & Feinstein, 1998; Fischer, Wartick, & Mark, 1992; Kirchler et al., 2007).

To sum up, different theoretical models come up with different conclusions as to which variables can impact tax compliance behavior. From an economic point of view the implications for penalties are straightforward: More severe penalties will result in more tax compliance. Empirical studies seem to find overall support for the direction of this effect, however it is often very weak or ambiguous. In an attempt to understand why tax compliance is as high as we observe it to be, even though evading would be the rational decision for monetarily driven taxpayers, more noneconomic variables have been incorporated into compliance theories. Subsequent theoretical analyzes as well as empirical studies found even more ambiguous results concerning the effect of penalties on tax compliance. It seems as if the effectiveness of enforcement strategies is connected to the interaction of taxpayers and tax authorities. A variety of new factors, such as social norms, perceived fairness, religious beliefs and corruption, have been shown to influence tax compliance behavior and thus might moderate the effect of penalties on compliance (e.g. Calvet Christian & Alm, 2014; Torgler, 2006; Wenzel, 2004). Regardless of the theoretical complexity, penalties are extensively utilized as deterrence for tax evasion in practice and therefore should be continued to be studied.

Research Question

Conclusions drawn from both theoretical models as well as empirical studies regarding the effect of penalties on compliance can be best described as overall ambiguous. This has also been reported by prior meta-analyses and literature reviews. (Alm & McKee, 1998; Andreoni et al., 1998; Bărbuță-Mișu, 2011; Blackwell, 2007; Fischer et al., 1992; Kirchler et al., 2007). While these studies thoroughly investigated a variety of results, they did not collect their analyzed literature in a systematic way. The goal of this master's thesis is to contribute to the current state of research by systematically and exhaustively searching for academic literature and summing up all empirical findings concerning the effect of penalties on tax compliance. From that we derived the research question:

What are the effects of penalties on tax compliance?

We did not formulate specific hypotheses. Instead we defined a set of goals for our study, with the underlying purpose of answering the research question: Firstly, we intend to produce a structured quantitative summary of the topic by systematically analyzing existing academic literature. Secondly, we intend to categorize under which conditions evidence for effects of penalties and fines can be found. And finally we want to identify possible research gaps.

We approached these goals by collecting a literature sample consisting of academic studies published in peer-reviewed journals. We wanted to collect all papers published in academic journals that report empirical evidence regarding the research question. Whether that evidence was gathered through experiments, surveys or aggregate data did not matter. However, as we intend to summarize empirical evidence, we had to exclude purely theoretical papers and reviews and books, that report on preexisting findings. We discuss the exact inclusion criteria for papers below, when we proceed with describing the method that was used to collect and analyze literature in a systematic way. Then we present a quantitative summary of the findings and discuss possible requirements for penalties to have a deterring effect on evasion in a more exploratory approach.

Method

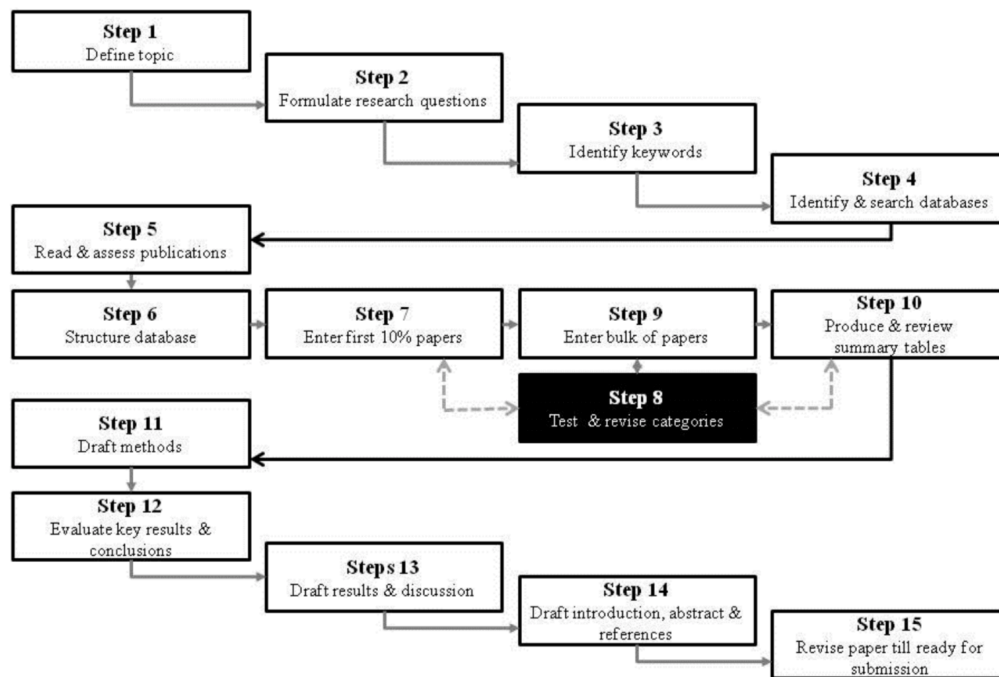
To assure the reproducibility of this systematic literature review, the methodical approach described by the *PRISMA Statement* (Liberati et al., 2009) and the *Systematic Quantitative Approach* (SQA) from Pickering and Byrne (2013) were followed closely while conducting this study.

Systematic Quantitative Approach

Pickering and Byrne (2013) argue that most scientific literature reviews follow a rather narrative style. Therefore, researchers producing high quality literature reviews need to be talented writers in addition to already having some expertise in the field they are reviewing. Not only is this approach prone to subjective tendencies, it also discourages researchers with little expertise in their field of interest to tackle literature reviews. However, young researchers might benefit from producing systematic literature reviews early in their career, as it can highlight research gaps and identify research questions for future studies in a particular field. Pickering and Byrne offer a new method to produce reviews, using what they call a *systematic quantitative assessment technique*. The method, also called *Systematic Quantitative Approach* (SQA), is best described in 15 consecutive stages (Figure 1) defining the process of collecting, mapping and analyzing literature. Following the SQA will result in a comprehensively structured and most importantly a reproducible literature review (Pickering & Byrne, 2013).

Figure 1

Fifteen stages in undertaking systematic quantitative literature reviews



Note. Figure 1: Fifteen stages in undertaking systematic quantitative literature reviews. Adopted from “The benefits of publishing systematic quantitative literature reviews for PhD candidates and other early-career researchers.” by Pickering and Byrne, 2013, *Higher Education Research and Development*, p. 10.

As for this study Step 1 (Defining topic) and Step 2 (Formulate research questions) have already been discussed. For Step 4 (Identify keywords) and Step 5 (Identify and search databases) we used an approach quite similar to the method reported by Hofmann, Voracek, Bock and Kirchler (2017) in their meta-analysis *Tax compliance across sociodemographic categories: Meta-analyses of survey studies in 111 countries*. We started with an informal literature search including a cited reference search of an influential paper from Alm, Jackson and McKee, (1992). The informal search identified 23 papers published between 1976 and 2016 that discussed the effects of penalties on tax compliance and also provided keywords. The most common keywords used were: *Evasion, Tax Compliance, Tax Evasion, Audit, Fines, Penalty, Punishment* and *Tax audits*. Following those findings, we established the following six terms to be searched for in the systematic literature search: “Tax + compliance”, “Tax + evasion”, “Tax + audit”, “Tax + punishment”, “Tax + fine” and “Tax + penalt*”. As for identifying databases we decided to adopt the following five scientific databases, that were also used in the Hofmann et al. meta-analysis (2017): *ISI Web of Science (SSCI)*, *PsycARTICLES*, *PsycInfo*, *PSYINDEXplus* and *EconLit*. We used *ISI Web of Science (SSCI)*, as it offers a broad coverage of all sciences. We chose *EconLit* to obtain a deeper coverage of

economic studies and the remaining three databases to obtain a deeper coverage of psychological publications. This way we wanted to ensure to collect relevant academic literature as exhaustively as possible.

Selection criteria

Before conducting the systematic literature search we established a set of criteria that defines whether certain studies were to be included or excluded in the review: First, to be included studies had to provide empirical information on tax compliance regarding penalties. In other words, the effect of an independent variable “penalties” on a dependent variable “tax compliance” had to be measured or estimated empirically. Second, we only incorporated studies in the review, if their understanding of the dependent variable “tax compliance” corresponded with either attitudes towards tax compliance, behavioral intentions to comply with tax law, or actual tax evasion and honest tax payments. Third, we did not restrict the inclusion of studies based on the method they used, as we intended to compare different methodical approaches in the analysis. However, we did control for the type of study. And finally, we did not limit the search to studies researching a specific type of tax (for example income tax or value added tax) or using a specific type of penalty (for example monetary fines or prison terms) as an independent variable, because we wanted to create an overview of the existing academic literature as complete as possible. It should be noted, that we defined “existing academic literature” as papers published in peer-reviewed journals. We excluded dissertations and books from the review.

We then proceeded with the systematic literature search using the predetermined key words. *ISI Web of Science*, *PsycARTICLES* and *PsycInfo* were researched in December 2017, *PSYNDExplus* and *EconLit* in February 2018. The PRISMA Statement flow chart (Liberati et al., 2009) was used to keep track of the number of results of the literature search and is depicted in Figure 2 .

PRISMA Statement

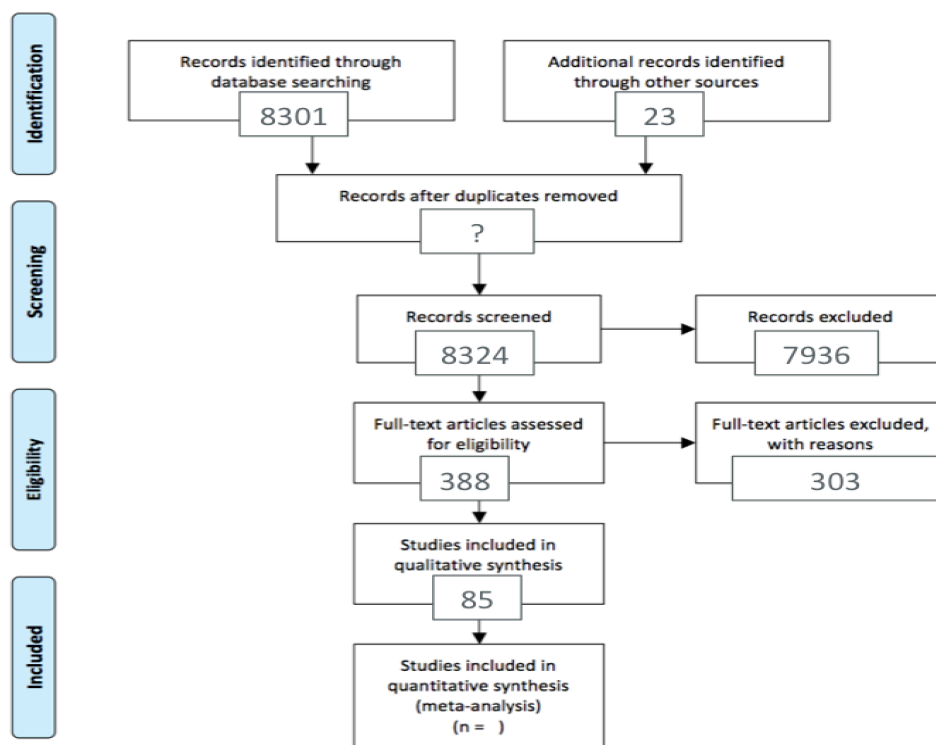
PRISMA is an abbreviation for “preferred reporting items for systematic reports and meta-analyses (Moher, Liberati, Tetzlaff, & Altman, 2009). It originated in health care but is applicable for other scientific fields for the most parts. The PRISMA Statement consists of two major parts: A 27 item checklist function as a quality guideline for the writing process of systematic reviews and meta-analyses and a 4-level flow chart which helps illustrate the literature collecting process in a comprehensible way. The authors argue that following the PRISMA Statement will result in a more transparent reporting of systematic reviews, hence preventing subjective tendencies and biases. The PRISMA 27 item checklist with corresponding page

numbers for this thesis can be seen in the Appendix. At this point it should be noted, that we did not follow the exact chronological order that the checklist suggests, as we structured our thesis more like a psychological paper. Also, we disregarded several points of the checklist that we deemed to health care specific for our means of writing a systematic literature review as a master's thesis. For example, item 19 suggests that authors should analyze the risk of bias for each individual study. We did not do so as we assumed to reach a certain quality standard by only incorporating articles from peer-reviewed journals into the review. Further we did not calculate a meta-analysis as our research question did not demand this and we had no funding available (items 21 and 27).

The PRISMA flow-chart for this study is shown in figure 2. We got 4089 results after we input our predetermined keywords into the *ISI Web of Science* search engine. *PsycARTICLES* turned out 59 results, *PsycInfo* 627 results, *PSYNDEXplus* 132 results and *EconLit* 3394 results, adding up to a total of 8301 records that have been identified through database searching. An additional 23 records have been identified before in the informal literature search at the start of our study. We did not keep track of duplicates found across the five databases, as we made use of the different options for organizing literature that the individual platforms provided. Also we did not see what the additional value was in knowing how many of the search results overlap across the five databases and therefore did not manually record that. Hence, we could not fill out the second box of the PRISMA Statement flow-chart and proceeded to the next step, screening our 8324 results. Applying the exclusion and inclusion criteria showed that 7936 of these records were not relevant for our research question and had to be dismissed. To help organize the remaining 388 articles we downloaded them in PDF-format and inserted them into the free reference manager program *Mendeley*. In accordance with Step 5 (Read & assess publications) of the *Systematic Quantitative Approach*, we then went on with reading those articles and assessing them in more detail, once again applying the inclusion and exclusion criteria. We had to exclude another 303 articles, although they seemed relevant for the research question at first. Why these articles had to be excluded, can be best explained by categorizing them into four groups. Firstly, it turned out that a large number of articles, especially those from economic journals, were describing and discussing theoretical models. A large number of these papers offered mathematical analyses of the effect of penalties on tax compliance in a variety of settings, but did not provide any empirical evidence and had to be excluded. Quite frequently, this was not apparent from only reading the title and the abstract of said articles. For example, the famous Allingham and Sandmo paper (1972) is not part of the actual review for this reason. Secondly, many of the experimental studies in our sample kept the penalty-rate constant in their tax

experiments, making it impossible to investigate its effect on tax compliance. Researchers most likely do that to save resources and focus on other variables of interest in the tax context (Alm, Jackson, & McKee, 2009). Thirdly, another group of studies had to be excluded, because they combined the penalty-rate with other variables to form a more general “deterrence” or “enforcement” variable. For example, most experiments that research the *Slippery Slope Framework* empirically, merge the audit-rate and the penalty-rate into a variable called *coercive-power* (Hofmann, Gangl, & Kirchler, 2014). Again, that makes it impossible to determine any effects that penalties alone could have on tax compliance. In general the exact qualities of dependent and independent variables had to be assessed in more detail causing the high number of papers being excluded at this stage of the literature search process. Finally, we also excluded all meta-analyses. Although it was convenient to collect them in the first place and countercheck, whether we included the same studies as already established researchers. However meta-analyses do not contribute any new information, especially considering, that the papers they analyze are most likely already incorporated in this review. Eventually, 85 studies were included in the analysis.

Figure 2
PRISMA 4-level flowchart



Note. Graphic taken from Liberati et al. (2009)

Coding procedure

We went on with coding the 85 remaining studies closely following Steps 6, 7, 8 and 9 from Pickering and Byrnes' *Systematic Quantitative Approach* (2013). As suggested in Step 6 (Structure database) we created our own database to organize the data obtained from the collected literature. We worked with a *Microsoft Excel* table, with each row representing a study and the columns being used to categorize collected data. At first we registered the following eight categories: Title, authors, year of publication, country in which the study was conducted, method, effect, conditions under which an effect was found/not found and an informal category summing up the findings of the study. The method used was coded as either an experiment (1), a survey (2) or as aggregate data (3). The effect a study found was coded as either no effect (0), a positive effect (1) or a negative effect (2). To clarify, a positive effect in this scenario means a study found that an increase in the independent variable penalty led to an increase in the dependent variable tax compliance, or respectively, a decrease in the dependent variable tax evasion. The last two categories collected qualitative data and were used in an explorative manner. The goal was to find similarities across studies, which indicate under what conditions penalties influence tax compliance. We used a variety of colors to mark papers with similarities in these categories.

As described in Step 7 (Enter first 10% papers) we started with entering a small sample of articles into the database. Some of these studies showed ambiguous results regarding the direction of the effect of penalties on tax compliance. Others came to different conclusions in different experimental settings, sometimes being unable to reproduce significant effects. That made it difficult to always assign one of the three values of the effect category to each paper. Therefore, we introduced the additional category "dependent effect" as a dummy variable. If studies found evidence for ambiguous effects between penalties and tax compliance, or if an effect only showed under certain conditions, that paper was coded with 1 in the dependent effect category. If studies found non-ambiguous effects they were coded with 0.

Further, we noticed that a considerable amount of papers also analyzed the effect of other factors that were known to have an impact on tax compliance decisions. Often the effects of changing audit frequencies or tax-rates, or appealing to social norms were compared to the effect of penalties. We added a tenth category "relative effect" and used it only for articles that directly compare at least one of those effects with the effect of penalties. The relative effect category was also a dummy variable. If a study found that the effect of

penalties on tax compliance was relatively smaller than the effect of another variable, it was coded 0. If the opposite was the case, it was coded 1.

After revising our categories as suggested in Step 8 (Test & revise categories) we continued with Step 9 (Enter bulk of papers) and entered all 85 papers into our database. The outcome of Step 10 (Produce & review summary tables) will be presented in the Results section. Step 12 (Evaluating key results & conclusion) was conducted in form of a presentation of our findings during a master's seminar at the Economic Psychology Department of the University of Vienna, followed by a discussion with the class and the department's staff including Prof. Dr. Erich Kirchler. The remaining four steps of Pickering and Byrnes' *Systematic Quantitative Approach* (2013) apply to the writing process itself. We followed the chronological order suggested by the authors and started writing by drafting the method section first (Step 11). Then we continued by drafting the results and discussion section (Step 13) and finally the introduction and the abstract (Step 14). As for Step 15 (Revise paper till ready for submission) our working paper was firstly peer-reviewed by co-students and revised. Consequently it was handed in to our supervisor Matthias Kasper, Ph.D., reviewed a second time and revised again before we submitted it to Prof. Kirchler.

Results

The 85 papers included in the final analysis range from 1967 to 2017 and are evenly distributed over time. 41% of the studies have been conducted in European countries, 40% in the U.S.A., 7% in Australia and New Zealand and the remaining 12% in South America, Africa, East Asia and India. As expected the majority of studies (87%) investigated the effect of penalties in regards to income tax evasion. Only three studies researched value added tax (VAT), two studies property tax, one study road fund tax and one study examined use tax. Four experimental studies conducted compliance games or social dilemma games without specifying the type of tax they were researching. Exact frequencies as well as a list of all included papers can be seen in the Appendix.

We first report a quantitative summary, describing studies by their applied method and the effects found, followed by an exploratory part, where we describe conditions under which penalties seem to have a more clear-cut effect.

Quantitative Summary

First, we categorized our data by method used and the direction of the effect that was found. This approach and the tables used to present these results were adopted from

Muehlbacher and Kirchler (2016). Table 1 shows the total frequencies of studies that found a positive effect, negative effect or no effect at all using the three different methodical approaches.

Table 1
Effect of penalties on compliance

Method	n	Positive	Zero	Negative
Experiment	49	29 (59%)	17 (35%)	3 (6%)
Survey	16	8 (50%)	7 (44%)	1 (6%)
Aggregate data	20	7 (35%)	9 (45%)	4 (20%)
Total	85	44 (52%)	33 (39%)	8 (9%)

Note. Frequencies indicate how many studies found a positive, negative or no effect at all using different methods.

With 52%, the majority of studies supported the intended positive effect of penalties on tax compliance. 39% found no effect at all and 9% reported a negative relation. These numbers are in line with previous meta-analyses. While a positive relation has been found most frequently, almost half of the studies did not report any effect or even showed a negative effect, supporting the notion that evidence for the deterring effect of penalties is weak, if not negligible.

Most of the studies conducted experiments, followed by analyzing aggregate data and finally conducting surveys. The external validity of tax experiments is often criticized. Compliance decisions in tax experiments are thought to resemble gambling behavior rather than real tax paying behavior. Other methods supposedly produce results with higher external validity but lower internal validity (Muehlbacher & Kirchler, 2016). As for the directional effect of penalties in this study, however, the three different methods produced very similar results. Both experiments and surveys yielded slightly more positive effects than no effects and very rarely showed negative effects. The distribution is shifted a little bit for analyses of aggregate data, as these studies found penalties to have no effect more often than a positive effect and reported the most negative effects in relative and absolute numbers. However, we do not want to over-interpret these numbers, as sample sizes are still rather small.

As described in the method section, we found that many studies reported rather ambiguous effects, which only showed in very specific situations. We introduced a dummy variable and coded all papers with “1” (effect dependent), if they reported an effect that only

showed in specific samples or under specific circumstances. For better illustration, we created a fourth directional effect and called it “ambiguous effect”. All the studies that had a positive coding in this dummy variable are now presented as “ambiguous”, regardless of what directional effect they had initially.

To better illustrate how we defined “ambiguous effect” we present a set of examples from the collected literature:

Bayer and Sutter (2009) found that raising the penalty-rate from 25% - 220% in their tax experiment had a small positive effect on tax compliance. But raising the penalty-rate from 25% - 100% or from 100% - 220% had no significant effect.

- We rated this a *positive effect* in Table 1 and an *ambiguous effect* in Table 2 (experimental study)

Iyer, Reckers and Sanders (2010) found that raising penalty-awareness for businesses with decreasing revenue had a positive effect on tax compliance. However, they did not find this effect for businesses with increasing revenue.

- We rated this a *positive effect* in Table 1 and an *ambiguous effect* in Table 2 (experimental study)

Verboon and Dijke (2011) reported that the moral evaluation of procedural justices mediates the effect of penalties on tax compliance. They found that the severity of sanction increases tax compliance in a system that is considered fair. However, this effect was not significant in a system that was considered unfair.

- We rated this a *positive effect* in Table 1 and an *ambiguous effect* in Table 2 (experimental study)

Verboon and Dijke (2007) reported the results of two separate surveys in their paper. In the first survey they showed a positive relation between severity of penalties and intended tax compliance, in the second survey they found no significant effect regarding penalties.

- We rated this a *positive effect* in Table 1 and an *ambiguous effect* in Table 2 (survey study)

Cullis et al. (2012) found that overall raising penalties and audits result in spite effects and have no significant effect on compliance. However there seems to be an “optimal range” where audit frequency is low and raising penalties lead to more compliance, until penalties are too high and lead to spite effects again!

- We rated this *no effect* in Table 1 and an *ambiguous effect* in Table 2 (experimental study)

Devos (2009) reported results from a paper-pen survey and from an interview study. The study found no significant effect from the perceived severity on tax compliance. However, in the interviews penalized evaders stated, that higher penalties made them more compliant.

- We rated this *no effect* in Table 1 and an *ambiguous effect* in Table 2 (survey study)

Bergman and Nevarez (2006) analyzed aggregate data from Argentina and Chile and found that larger fines are positively correlated with post-audit noncompliance, indicating that penalties have the undesired effect of reducing compliance. However, due to their sample, the authors did not analyze, whether this effect also shows when taxpayers were not audited. So while larger fines result in larger bomb-crater effects in Chile and Argentina, we cannot assume they have a negative deterring effect.

- We rated this a *negative effect* in Table 1 and an *ambiguous effect* in Table 2 (aggregate data study)

Table 2 shows the total frequencies of studies that found a positive effect, negative effect, no effect or an ambiguous effect using the three different methodical approaches.

Table 2
Effect of penalties on compliance

Method	n	Positive	Zero	Negative	Ambiguous
Experiment	49	16 (33%)	10 (20%)	2 (4%)	21 (43%)
Survey	16	4 (25%)	4 (25%)	1 (6%)	7 (44%)
Aggregate data	20	6 (30%)	7 (35%)	2 (10%)	5 (25%)
Total	85	26 (31%)	21 (25%)	5 (6%)	33 (39%)

Note. Frequencies indicate how many studies found a positive, negative, no effect or an ambiguous effect using different methods.

Introducing the additional direction “ambiguous” changed the distribution of effects. The most frequently reported effect is now the newly defined ambiguous effect, accounting for 39% of all studies. Total frequencies decreased for all other effects in all methodical approaches except for the one negative result reported from a survey study. Positive results still occur the second most often, however they also underwent the largest decrease in total numbers and percentagewise, dropping from 52% to 31%. Again, this decrease supports the notion, that evidence for the positive relation between penalties and tax compliance is usually weak. The total frequency of studies not finding any effect and the number of studies reporting negative effects dropped as well, losing about one third each to the ambiguous category.

Experimental studies yielded 43% ambiguous effects and surveys 44%. The distribution of studies that analyze aggregate data differs from the other two methods, as positive effects and no effects outweigh ambiguous effects here. Considering the small sample sizes, it does not seem like overall a certain method is more prone to yield ambiguous results than the others.

Finally, we identified 17 studies, that supported the intended deterring effect of penalties and directly compared that effect with the effect of the audit-frequency, the tax rate, or appealing to someone’s social norms or emotions. As described in the method section, these studies were coded with the additional dummy variable “relative effect” that recorded, whether the effect of penalties was relatively greater or smaller than the effect of the other variables.

To better illustrate how we defined “relative effect” we present a set of examples from the collected literature:

Cebula (2011) analyzed IRS data and found that a rise in the audit-rate of 1% reduces tax evasion by 18%, whereas a 1% increase in the penalty-rate reduces personal income tax evasion by 3%.

- We rated this a *smaller effect* of penalties compared to audit-frequency

Grasmick and Bursik (1990) found that feeling guilty for having committed tax evasion, as it is not accepted behavior according to someone's social norm, has a bigger effect on deterring it than the severity of legal sanctions.

- We rated this a *smaller effect* of penalties compared to appealing to social norms or emotions

Table 3 shows the total frequencies of studies that found relative effects.

Table 3

Relative effect of penalties on tax compliance

	n	Greater effect	Smaller effect
Audit-frequency	11	5 (45%)	6 (55%)
Tax-rate	2	-	2 (100%)
Appealing to social norms/emotions	4	1 (25%)	3 (75%)
Total	17	6(35%)	11(65%)

Note. Frequencies indicate how many studies found the effect of penalties on compliance to be relatively greater or smaller than the effect of audit-frequency, tax-rate or appealing to social norms or emotions

Overall 65% of these studies found the effect of penalties to be smaller than the effect of other variables. Audit-frequency was most often compared to penalty-rate, with support for either variable being almost even.

With the exception for one, all other studies agree that tax-rate and social norms are more important factors in explaining tax compliance behavior. However, we should consider the small sample sizes: a variable like tax rate, that is associated with rather ambiguous theoretical predictions, is probably underrepresented with only two studies comparing it to penalties. Also authors investigating a variety of factors might not treat factors like social

norms equivalently to audits and fines and the greater/smaller effects might be the result of a specific research design, rather than a true comparison.

Summary of qualitative conditions

In the following we sum up several conditions under which penalties seem to have the intended positive effect, or respectively the negative effect. Understanding conditional effects could help explain why empirical results are so diverse. We created these conditions by collecting analogies in our literature sample in a rather exploratory manner. While the scientific validity of these findings is questionable, they might be useful in identifying topics for possible future research.

Negative effects and locality. The first condition concerns the distribution of negative effects over the countries where these studies have been conducted. A total of eight articles reported a negative relation between penalties and tax compliance. These studies analyzed data collected in the following 9 countries: Argentina, Chile, Costa Rica, France, Italy, Jamaica, Switzerland, Tanzania and the U.S.A. While 88% of all studies were conducted in the “western world” (Europe, U.S.A., Australia), more than half of the negative effects were found in lower-income countries. Those studies were often the only studies conducted in these particular countries. For example, the only study that analyzed aggregate data from Jamaica found that an increase in penalties results in a decrease in the total reported tax base, suggesting that Jamaicans are not at all deterred by higher penalties but instead underreport more income (Alm, Bahl, et al., 1990). The one experiment that was conducted in Costa Rica found that the penalty-rate has a negative impact on tax compliance (Torgler & Schaltegger, 2005) and a survey from Tanzania reported that an increase in perceived severity of the sanction leads to more resistance from citizens towards the tax authorities, going so far that tax collectors have to be accompanied by militia in order not to be violently attacked while doing their job in rural areas (Fjeldstad & Semboja, 2001).

Additionally, other lower-income countries, like Bulgaria, Egypt, Ethiopia and India solely yielded studies that reported no effect at all or very ambiguous effects pointing to further possible negative relations between penalties and compliance. For example Soliman, Jones and Cullis (2014) conducted an experimental study in Cairo and found that when the audit rate is low, raising the penalty-rate from two-times the evaded amount to eight-times the evaded amount increased compliance, however raising it further to 15-times the evaded amount actually decreased compliance again. Also a survey in Ethiopia showed, that penalties are negatively correlated to compliance, however that effect was not significant (Engida & Baisa, 2014).

While sample sizes are too small to draw any scientifically valid conclusions, these findings seem to suggest, that the majority of studies conducted in poor countries show a negative effect or no effect at all. This could perhaps be explained within the *Slippery Slope Framework*: assuming that trust as well as power is low in these countries, raising penalties might worsen the relation between taxpayers and authorities. Higher fines might be considered a sign of distrust towards taxpayers, while not actually improving the capability of a given government to enforce their policies, therefore lowering trust while not increasing power. A better approach in said climates could be to focus on improving auditing strategies to increase power or to invest into public goods and make government spending more transparent in an attempt to increase trust. While we attempted to collect all peer-reviewed papers investigating the effect of penalties on tax compliance in this study, studies from third-world countries are still underrepresented and these conclusions are highly speculative. Future research is needed to determine how different penalties and auditing schemes effect tax compliance in poor countries.

Penalty awareness and businesses with decreasing revenue. Raising penalty awareness by sending details about the sanction severity when caught evading taxes to business owners did not result in more tax compliance overall, however it did result in more compliance from businesses that were struggling with decreasing revenue (Sanders, Reckers, & Iyer, 2008). In a follow up experiment by the same authors, the experimental manipulation of raising penalty awareness worked better, as they reported a positive relation between informing businesses about penalties and tax compliance. But they still found that the manipulation worked best for businesses with decreasing revenue, while it did not work for businesses with increasing revenue (Iyer et al., 2010).

Penalties and high-income taxpayers. Ali, Cecil and Knoblett (2001) analyzed aggregate data from the IRS and found, that both penalties and audits are more effective in deterring tax evasion for people with a high income. Further, Witte and Woodbury (1985) showed, that the severity of sanctions is only significantly related to tax compliance for high-level income taxpayers. This gave rise to the idea of adjusting the severity of a penalty to the evaders income, which has been shown to decrease the intend to evade for high-level income taxpayers (Muehlbacher, Kirchler, & Hoelzl, 2007).

The standard model predicts that taxpayers with more income will evade more taxes.

Empirical results suggest that raising penalties is an effective deterrence strategy against this specific kind of tax evasion.

Penalties and perceived fairness. The perceived fairness of tax authorities and of the policies they implement is part of the procedural justice and determines voluntary tax compliance in the *Slippery Slope Framework* (Kirchler et al., 2008). Moser, Evans III and Kim (1995) showed that the perceived fairness of the tax rate is a bigger factor in explaining tax compliance than the rate itself. There seems to be some sort of reluctance to comply with tax laws, when tax policies are considered unfair (Murphy, 2003). Especially evaders who have already been caught might react to their prosecution with evading even more taxes, when penalties are considered unfair (Murphy, 2005). Empirical studies found that more severe penalties increase compliance in a system that is considered fair, but not in a system that is considered unfair (Verboon & van Dijke, 2007). Also penalties that are considered to be unfair do not increase compliance with VAT, while penalties that are considered fair do (Adams & Webley, 2001). These findings could perhaps be explained within the *Slippery Slope Framework* and need further evaluation.

Penalties with high audit probabilities. Alm, Jackson and McKee (1992) pointed out, that the penalty-rate and the audit-frequency are dependent from each other. If the probability for an audit is very low, raising penalties will only affect very risk-averse taxpayers. Vice versa if penalties are very low, frequent audits will not have the intended deterring effect. Theoretically the deterring effect of penalties should be higher when detection of evasion is probable (Stella, 1991). Alm, Sanchez and De Juan (1995) found empirical evidence supporting this notion, showing that higher fines only result in more compliance when the audit-frequency is high too. Collins and Plumlee (1991) found that higher penalties increase compliance, when audits are likely or at least constant, however, when the probability for an audit is low, higher penalties might even lead to more tax evasion. In a more recent study, Cullis et al. (2012) found that penalties might be most effective, when the audit probability is low and fines are not too high, as the combination of severe sanctions and frequent audits caused spite effects and did not increase compliance.

Penalties after tax amnesties. An experiment from Alm, McKee and Beck (1990) showed, that tax compliance drops after a tax amnesty. This phenomenon can be counteracted by stronger enforcement efforts. Specifically, higher penalties had the desired effect of increasing compliance after the amnesty. The authors argue, that raising deterrence variables after an amnesty has a bigger effect on compliance, than raising deterrence variables without an amnesty. Analyzing IRS data showed that if penalties were increased after tax amnesties, the long term net revenue loss got smaller, therefore implying that taxpayers were more compliant (Luitel & Sobel, 2007). However, Torgler and Schaltegger (2005) reported

evidence contrary to these findings. Experiments in Costa Rica and in Switzerland showed that higher enforcement after the amnesty in general had no effect and raising penalties specifically had a small negative effect on tax compliance.

Discussion

In our study, we aimed to create a systematic literature review, investigating the effects of penalties on tax compliance. To improve on previous literature reviews and meta-analyses researching this topic, we focused extensively on the process of systematically and exhaustively collecting literature. We followed the *Systematic Quantitative Approach* (Pickering & Byrne, 2013) and the *PRISMA Statement* (Moher et al., 2009) to collect and analyze literature in a reproducible way. We then produced a structured quantitative summary of the topic and presented our findings in tables adopted from Muehlbacher and Kirchler (2016).

Findings presented in Table 1 show that slightly more than half of the papers we collected report that penalties have the intended positive effect on tax compliance, supporting the predictions of the standard model. However, almost half of the studies did not find any effect at all or even reported that higher penalties have a negative relation with tax compliance. The variance in these findings seems to be in line with previous reviews. We assume that penalties do increase compliance, however this effect is probably small or even negligible as suggested by prior research (e.g. Alm & McKee, 1998; Bărbuță-Mișu, 2011; Kirchler et al., 2007). How ambiguous empirical results regarding this effect can become, is displayed in Table 2. As mentioned in the method section, we introduced a dummy variable to code papers that showed certain results only in specific situations, but were overall ambiguous. After introducing the fourth category "ambiguous effect", we observed that the majority of results could be described as such an effect. Comparing Table 1 and Table 2 shows that the largest amount of studies categorized as reporting an ambiguous effect, were previously categorized as reporting a positive effect. This drop in positive results is observed over all three methodical approaches and further supports the notion, that empirical evidence for the deterring effect of penalties in the tax context is weak (e.g. Alm & McKee, 1998; Kirchler et al., 2007). Further, Table 3 shows that even when studies report a positive relation between penalties and tax compliance, other variables, like the audit-frequency, the tax-rate and noneconomic variables like social norms, are reported to have a bigger impact on

compliance behavior than the actual penalty. This also suggests that in practice penalties might have less impact on compliance behavior, than predicted by the standard model.

However, about one third of the studies, which were initially categorized as reporting no or a negative effect, were later categorized as reporting an ambiguous effect as well. Therefore, we conclude that empirical evidence against the predictions of the standard model is similarly weak, but also less frequent than evidence for it.

As the deterring effect of penalties seems to be highly dependent on situational factors, we attempted to identify conditions under which it can be predicted more reliably. Unfortunately, our literature sample and the initial study design only allowed us to investigate these conditions in an exploratory manner.

We noticed that a majority of studies, which were conducted in relatively poor countries, yielded results that did not show any effect or even a negative effect of penalties on tax compliance. This is contradictory to the predictions of the standard model but could perhaps be explained within the *Slippery Slope Framework*. Low-income countries might be prone to climates where both the level of power and the level of trust are so low, that the increase of power by raising the penalty rate becomes negligible, as it doesn't actually increase a given governments capability of enforcing its policies. Therefore citizens might not believe authorities can detect evasion and collect the higher fines accordingly. Additionally, countries with improperly functioning tax systems might face high levels of corruption and bribery; both factors that can further distort the effect of penalties and lead to the negative effects we observed (Fulton & Cule, 2009; Singh, 2008).

Finally, higher penalties are usually more effective in deterring evasion if audits are also frequent. This is considered a direct consequence of the relation between penalties and audits and has been supported theoretically (Alm, et al., 1992) and by our findings. However, there seems to be a threshold where audit-frequency and penalty-rate get too high and will not increase compliance anymore. Cullis et al. (2012) argued that the combination of high fines and frequent audits leads to spite effects. Stronger deterrence might also provide more incentives to evade for people, who have already been caught evading. Higher punishments can increase the bomb-crater effect and induce tax-evaders to misinterpret the chance of being detected again and try to make up for their loss with evading more (Kastlunger, Kirchler, Mittone, & Pitters, 2009). More severe punishments can also lead to stigmatization and resentment for caught evaders, again inducing them to comply less (Murphy, 2008).

Our literature sample suggests, that simply raising audits and penalties doesn't directly translate to more tax compliance. It seems the effect is moderated by a variety of situational factors. Keeping in mind that a large number of taxpayers are compliant anyway, authorities might be better off focusing on choosing the right taxpayers to audit, instead of increasing deterrence variables for the general public. Obviously creating and executing intricate auditing schemes is difficult and expensive. Future research could focus on how to efficiently identify and audit tax evaders.

Limitations

As this study is systematically analyzing peer-reviewed literature, it is prone to the publication bias. Considering, that researchers are supposedly likely to not publish their studies, when they did not find any effect at all, it is possible that this type of papers are underrepresented in this review. This might also be valid for papers, which find a negative relation between penalties and compliance contrary to theoretical assumptions. However, as most tax researchers are studying effects of various variables, it is plausible that they will find significant results for at least some of them and publish anyway. This seems especially true as we did find plenty of papers reporting no effect concerning the penalty-rate.

However, the same logic would suggest that ambiguous effects are also underreported. It seems likely, that studies that focus on investigating only one factor of compliance and find ambiguous results won't be published. Whereas studies that research a variety of specific factors and interactions might find significant results for some of those factors and are published, even though conclusions concerning the other factors are ambiguous. Furthermore it is also possible, that authors include significant results in their paper, while intentionally leaving out insignificant ones (Franco, Malhotra, & Simonovits, 2016). Therefore while ambiguous results are already the most common in our sample, they might still be underrepresented because of this kind of selective reporting.

Further, findings might be biased, as operationalizing tax variables is anything but easy. Especially tax experiments face the problem of low external validity. Researchers usually try to increase external validity by offering monetary incentives to participants. Unfortunately, that often provides windfall endowments and results in participants engaging in gambling rather than in making tax decisions (Kirchler et al., 2007). As for survey studies, they might not find any deterrent effect from raising the severity of penalties, if they fail to control for individual differences in the personal consequences of being punished with a specific sanction. Klepper and Nagin (1989) reported, that a non-zero chance for criminal persecution was an absolute deterrent for their middle-class sample,

making any change in severity of formal sanctions obsolete as a deterrent variable if there was a chance for criminal persecution.

Conclusion

We intended to answer our research question “What are the effects of penalties on tax compliance?” by following three propositions. Firstly, including a larger quantity of papers did not clear up the effect of penalties on compliance in comparison to previous reviews. On the contrary, we found more studies showing ambiguous results and even some reporting a negative relation. However, overall our quantitative summary yielded results in line with prior reviews: a slim majority of studies reported a positive relation between penalties and compliance, but many of these results can be regarded as ambiguous. Therefore we concluded that evidence for the effect of penalties is rather weak. Secondly, we identified several conditions under which penalties do have a more clear-cut effect. These findings imply that penalties are more likely to have the intended deterring effect in what is defined as a synergistic climate within the *Slippery Slope Framework*. However, this assumption is highly speculative and leads directly to our third proposition “identify research gaps”: It seems that while general assumptions from the *Slippery Slope Framework* have been supported empirically (e.g. Kastlunger, Muehlbacher, Kirchler, & Mittone, 2011; Muehlbacher, Kirchler, & Schwarzenberger, 2011), not many studies investigate specific deterrence variables in different socio-economic and political systems. We suggest that future research focuses on whether raising penalties has different effects in systems that can be defined as synergistic or antagonistic according to the *Slippery Slope Framework*.

References

- Adams, C., & Webley, P. (2001). Small business owners' attitudes on VAT compliance in the UK. *Journal of Economic Psychology*, 22(2), 195–216.
- Ali, M. M., Cecil, W. H., & Knoblett, J. A. (2001). The Effects of Tax Rates and Enforcement Policies on Taxpayer Compliance : A Study of Self-Employed Taxpayers. *Atlantic Economic Journal*, 29(2), 186–202.
- Allingham, M. G., & Sandmo, A. (1972). Income Tax Evasion: A Theoretical Analysis. *Journal of Public Economics*, 1, 323–338.
- Alm, J., Bahl, R., & Murray, M. N. (1990). Tax Structure and Tax Compliance. *The Review of Economics and Statistics*, 72(4), 603–613.
- Alm, J., Clark, J., & Leibel, K. (2016). Enforcement , Socioeconomic Diversity , and Tax Filing Compliance in the United States. *Southern Economic Journal*, 82(3), 725–747.
- Alm, J., Jackson, B., & Mckee, M. (1992a). Institutional Uncertainty and Taxpayer Compliance. *The American Economic Review*, 82(4), 1018–1026.
- Alm, J., Jackson, B. R., & Mckee, M. (1992b). ESstimating the Determinants of Taxpayer Compliance with Experimental Data. *National Tax Journal*, 45(1), 107–114.
- Alm, J., Jackson, B. R., & Mckee, M. (2009). Getting the word out : Enforcement information dissemination and compliance behavior. *Journal of Public Economics*, 93(3–4), 392–402.
- Alm, J., Martinez-Vazquez, J., & McClellan, C. (2016). Corruption and firm tax evasion. *Journal of Economic Behavior and Organization*, 124, 146–163.
- Alm, J., McClelland, G., & Schulze, W. (1992). Why Do People Pay Taxes ? *Journal of Public Economics*, 48, 21–38.
- Alm, J., & McKee, M. (1998). Extending the lessons of laboratory experiments on tax compliance to managerial and decision economics. *Managerial and Decision Economics*, 19(4–5), 259–275.
- Alm, J., Mckee, M., & Beck, W. (1990). Amazing Grace: Tax Amnesties and Compliance. *National Tax Journal*, 43(1), 23–37.
- Alm, J., Sanchez, I., & De Juan, A. (1995). Economic & Noneconomic Factors in Tax Compliance. *Kyklos*, 48(1), 3–18.
- Alm, J., & Torgler, B. (2011). Do Ethics Matter ? Tax Compliance and Morality. *Journal of Business Ethics*, 635–651.
- Andreoni, J. (1992). IRS as loan shark Tax compliance with borrowing constraints. *Journal of Public Economics*, 49, 35–46.

- Andreoni, J., Erard, B., & Feinstein, J. (1998). Tax Compliance. *Journal of Economics*, 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.
- Bărbuță-Mișu, N. (2011). A Review of Factors for Tax Compliance. *Economics and Applied Informatics*, 17(1), 69–76.
- Bayer, R., & Sutter, M. (2009). The excess burden of tax evasion — An experimental detection – concealment contest. *European Economic Review*, 53, 527–543.
- Beck, P. J., Davis, J. S., & Jung, W.-O. (1991). Experimental Evidence on Taxpayer Reporting Under Uncertainty. *The Accounting Review*, 66(3), 535–558.
- Becker, G. S. (1968). *Crime and Punishment: An Economic Approach*. *Journal of Political Economy*, 76, 1-54.
- Beer, S., Kasper, M., Kirchler, E., & Erard, B. (2015). Audit Impact Study 1, 2.
- Bergman, M., & Nevarez, A. (2006). Do Audits Enhance Compliance ? An Empirical Assessment of VAT Enforcement. *National Tax Journal*, 59, 817–832.
- Blackwell, C. (2007). A meta-analysis of tax compliance experiments. *Annual Meeting of the Public Choice Society*, (December), 7-24.
- Borck, R. (2004). Income tax evasion and the penalty structure. *Economics Bulletin*, 8(1), 1-9.
- Bose, P. (1995). Regulatory errors, optimal fines and the level of compliance. *Journal of Public Economics*, 56, 475–484.
- Braithwaite, V. (2003). Dancing with tax authorities: Motivational postures and non-compliant actions. In *Taxing Democracy: Understanding Tax Avoidance and Evasion* (pp. 15–39). Aldershot: Ashgate Publishing Ltd.
- Bruttel, L., & Friehe, T. (2014). On the path dependence of tax compliance. *European Economic Review*, 65, 90–107.
- Calvet Christian, R., & Alm, J. (2014). Empathy, sympathy, and tax compliance. *Journal of Economic Psychology*, 40, 62–82.
- Cebula, R. J. (1997). An Empirical Analysis of the Impact of Government Tax and Auditing Policies on the Size of the Underground Economy : The case of the United States 1973-94. *American Journal of Economics and Sociology*, 56(2), 173–185.
- Cebula, R. J. (2011). An Empirical Analysis of Determinants of Recent Federal Personal Income Tax Evasion in the U.S. *Journal of Public Finance and Public Choice*, 29(1–3), 33–54.
- Cebula, R. J., & Feige, E. L. (2012). America ’ s unreported economy : measuring the size ,

- growth and determinants of income tax evasion in the U . S . *Crime, Law & Social Change*, 57, 265–285.
- Chang, B. J., & Lai, C. (2004). Collaborative tax evasion and social norms : why deterrence does not work. *Oxford Economic Papers*, 56, 344–368.
- Chen, B. (2003). Tax evasion in a model of endogenous growth. *Review of Economic Dynamics*, 6, 381–403.
- Cho, J. M., Linn, S. C., & Nakibullah, A. (1996). Tax evasion with psychic costs and penalty renegotiation. *Southern Economic Journal*, 63(1), 172–190.
- Collins, J. H., & Plumlee, R. D. (1991). The Taxpayer ’ s Labor and Reporting Decision : The Effect of Audit Schemes. *The Accounting Review*, 66(3), 559–576.
- Coricelli, G., Joffily, M., Montmarquette, C., & Claire, M. (2010). Cheating , emotions , and rationality : an experiment on tax evasion. *Experimental Economics*, 13(2), 226–247.
- Crane, S. E., & Nourzad, F. (1985). Time Value of Money and Income-Tax Evasion Under Risk-Averse Behavior - Theoretical-Analysis and Empirical-Evidence. *Public Finance-Finances Publiques*, 40(3), 381–394.
- Cullis, J., Jones, P., & Soliman, A. (2012). “Spite effects” in tax evasion experiments. *Journal of Socio-Economics*, 41(4), 418–423.
- Davis, J. S., Hecht, G., & Perkins, J. D. (2003). Social behaviours, enforcement and tax compliance dynamics. *The Accounting Review*, 78(1), 39–69.
- Devos, K. (2009). An Investigation Into Australian Personal Tax Evaders- Their Attitudes Towards Compliance And The Penalties For Non-Compliance. *Revenue Law Journal*, 19(1), 1-49.
- Engida, T. G., & Baisa, G. A. (2014). Factors influencing taxpayers’ Compliance with the tax system: An empirical study in Mekelle City, Ethiopia. *eJournal of Tax Research*, 12(2), 433–452.
- Feld, L. P., & Frey, B. S. (2002). Trust breeds trust: How taxpayers are treated. *Economics of Governance*, 3(2), 87–99.
- Feld, L. P., & Larsen, C. (2012). Self-perceptions , government policies and tax compliance in Germany. *International Tax and Public Finance*, 19, 78–103.
- Fischer, C. M., Wartick, M., & Mark, M. M. (1992). Detection probability and taxpayer compliance : A review of the literature. *Journal of Accounting Literature*, 11(1a), 1–46.
- Fjeldstad, O., & Semboja, J. (2001). Why People Pay Taxes : The Case of the Development Levy in Tanzania. *World Development*, 29(12), 2059–2074.
- Franco, A., Malhotra, N., & Simonovits, G. (2016). Underreporting in Psychology

- Experiments: Evidence From a Study Registry. *Social Psychological and Personality Science*, 7(1), 8–12.
- Friedland, N. (1982). A Note on Tax Evasion as a Function of the Quality of Information about the Magnitude and Credibility of Threatened Fines: Some Preliminary Research. *Journal of Applied Social Psychology*, 12(1), 54–59.
- Friedland, N., Maital, S., & Rutenberg, A. (1978). A Simulation Study of Income Tax Evasion. *Journal of Public Economics*, 10, 107–116.
- Fulton, M., & Cule, M. (2009). Business culture and tax evasion : Why corruption and the unofficial economy can persist. *Journal of Economic Behavior & Organization*, 72, 811–822.
- Grasmick, H., & Bursik, R. (1990). Conscience, Significant Others, AND rational Choice: Extending The Deterrence Model. *Law & Society Review*, 24(3), 837–861.
- Hite, P. A., & McGill, G. A. (1992). An examination of taxpayer preference for aggressive tax advice. *National Tax Journal*, 45, 389–403.
- Hofmann, E., Gangl, K., & Kirchler, E. (2014). Enhancing Tax Compliance through Coercive and Legitimate Power of Tax Authorities by Concurrently Diminishing or Facilitating Trust in Tax Authorities. *Law and Policy*, 36(3), 290–313.
- Hofmann, E., Hoelzl, E., & Kirchler, E. (2008). Preconditions of Voluntary Tax Compliance. *Journal of Psychology*, 216(4), 209–217.
- Hofmann, E., Voracek, M., Bock, C., & Kirchler, E. (2017). Tax compliance across sociodemographic categories : Meta- analyses of survey studies in 111 countries. *Journal of Economic Psychology*, 62, 63–71.
- Holmes, R. Y. (2011). Forcing Cooperation: A Strategy for Improving Tax Compliance. *University of Cincinnati Law Review*, 79(4), 1415–1459.
- Iyer, G. S., Reckers, P. M. J., & Sanders, D. L. (2010). Increasing Tax Compliance in Washington State : A Field Experiment. *National Tax Journal*, 63(March), 7–32.
- Jimenez, P., & Iyer, G. S. (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.
- Kaplanoglou, G., & Rapanos, V. T. (2015). Why do people evade taxes? New experimental evidence from Greece. *Journal of Behavioral and Experimental Economics*, 56, 21–32.
- Kastlunger, B., Kirchler, E., Mittone, L., & Pitters, J. (2009). Sequences of audits , tax compliance , and taxpaying strategies. *Journal of Economic Psychology*, 30(3), 405–418.
- Kastlunger, B., Muehlbacher, S., Kirchler, E., & Mittone, L. (2011). What goes around comes

- around? Experimental evidence of the effect of rewards on tax compliance. *Public Finance Review*, 39(1), 150–167.
- Kirchler, E., Hoelzl, E., & Wahl, I. (2008). Enforced versus voluntary tax compliance : The “slippery slope” framework. *Journal of Economic Psychology*, 29, 210–225.
- Kirchler, E., Muehlbacher, S., Kastlunger, B., & Wahl, I. (2007). *Why pay Taxes? A review of tax compliance decisions*.
- Kirchler, E., & Wahl, I. (2010). Tax Compliance Inventory : TAX-I Voluntary tax compliance , enforced tax compliance , tax avoidance , and tax evasion. *Journal of Economic Psychology*, 31(August 2016), 331–346.
- Klepper, S., & Nagin, D. (1989). Tax Compliance and Perceptions of the Risks of Detection and Criminal Prosecution. *Law and Society Review*, 23(2), 209–240.
- Kogler, C., Muehlbacher, S., & Kirchler, E. (2015). Testing the “slippery slope framework” among self-employed taxpayers. *Economics of Governance*, 16(2), 125–142.
- Korobow, A., Johnson, C., & Axtell, R. (2007). An Agent–Based Model of Tax Compliance with Social Networks. *National Tax Journal*, 60(3), 589–610.
- Lee, K. (2015). Timing of Penalties , Tax Rates , and Tax Evasion. *Public Finance Analysis*, 71(1), 37-52.
- Lesnik, T., Kracun, D., & Jagric, T. (2014). Tax Compliance and Corporate Income Tax – The Case of Slovenia. *Lex Localis Journal of Local Self-Government*, 12(4), 793–811.
- Levaggi, R., & Menoncin, F. (2016). Optimal dynamic tax evasion : A portfolio approach. *Journal of Economic Behavior and Organization*, 124, 115–129.
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., & John, P. A. (2009). Research method & reporting the PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions : explanation and elaboration. *PLoS Med* 6(7).
- Luitel, H. S., & Sobel, R. (2007). The Revenue Impact of Repeated Tax Amnesties. *Public Budgeting & Finance*, 27(3), 19–38.
- Mehlkop, G., & Becker, R. (2004). Soziale Schichtung und Delinquenz. Eine empirische Anwendung eines Rational Choice-Ansatzes mit Hilfe von Querschnittsdaten des ALLBUS 1990 und 2000. *Koelner Zeitschrift Fuer Soziologie Und Sozialpsychologie*, 56(1), 95-126.
- Mendoza, J. P., Wielhouwer, J. L., & Kirchler, E. (2017). The backfiring effect of auditing on tax compliance. *Journal of Economic Psychology*, 62, 284–294.
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for

- systematic reviews and meta-analyses : the PRISMA statement, *PLoS Med* 6(7).
- Moser, D. V, Evans III, J. H., & Kim, C. K. (1995). Moser The Effects of Horizontal and Exchange Inequity on Tax Reporting Decisions. *The Accounting Review*, 70(4), 619–634.
- Muehlbacher, S., & Kirchler, E. (2016). Taxperiments : About the external validity of laboratory experiments in tax compliance. *Die Betriebswirtschaft*, (February), 7-19.
- Muehlbacher, S., Kirchler, E., & Hoelzl, E. (2007). Steuerhinterziehung und die Berücksichtigung des Einkommens in der Strafbemessung. *Wirtschaftspsychologie*, 4, 116–121.
- Muehlbacher, S., Kirchler, E., & Schwarzenberger, H. (2011). Voluntary versus enforced tax compliance : empirical evidence for the “ slippery slope ” framework. *Eur J Law Econ*, 89–97.
- Murphy, K. (2003). Procedural justice and tax compliance. *Australian Journal of Social Issues*, 38, 379–408.
- Murphy, K. (2005). Regulating More Effectively : The Relationship between Procedural Justice , Legitimacy , and Tax Non-compliance. *Journal of Law and Society*, 32(4), 562–589.
- Murphy, K. (2008). Enforcing Tax Compliance : To Punish or Persuade ? *Economic Analysis and Policy*, 38(1), 113–135.
- Myles, G. D., Hashimzade, N., & Tran-Nam, B. (2012). Applications of Behavioural Economics to Tax Evasion. *Journal of Economic Surveys*, 27(5), 941–977.
- Park, C., & Hyun, J. K. (2003). Examining the determinants of tax compliance by experimental data : a case of Korea. *Journal of Policy Modeling*, 25, 673–684.
- Pickering, C., & Byrne, J. A. (2013). The benefits of publishing systematic quantitative literature reviews for PhD candidates and other early-career researchers. *Higher Education Research and Development*, (December), 1–24.
- Pommerehne, W. W., & Weck-Hannemann, H. (1996). Tax rates, tax administration and income tax evasion in Switzerland. *Public Choice*, 88, 161–170.
- Sanders, D. L., Reckers, P. M. J., & Iyer, G. S. (2008). Influence of Accountability and Penalty Awareness on Tax Compliance. *The Journal of the American Taxation Association*, 30(2), 1–20.
- Schachner, G., Loukota, W. (o. J.). Austria Tax Controversy. Retrieved from <https://gettingthedealthrough.com/area/59/jurisdiction/25/tax-controversy-2018-austria>
- Schwartz, R. D., & Orleans, S. (1967). On Legal Sanctions. *University of Chicago Law*

- Review*, 34, 274–300.
- Singh, K. J. (2008). On tax evaders and corrupt auditors. *Journal of International Trade & Economic Development*, 17(1), 37–67.
- Smyth, Albert Henry (1907). *The Writings of Benjamin Franklin*, Vol. X (1789-1790). New York: MacMillan.
- Soliman, A., Jones, P., & Cullis, J. (2014). Learning in experiments : Dynamic interaction of policy variables designed to deter tax evasion. *Journal of Economic Psychology*, 40, 175–186.
- Spicer, M. W., & Lundstedt, S. B. (1976). Understanding tax evasion. *Public Finance*, 21(2), 295–305.
- Stella, P. (1991). An Economic Analysis of Tax Amnesties. *Journal of Public Economics*, 46(3), 383–423.
- Torgler, B. (2003). To evade taxes or not evade: That is the question. *Journal of Socio-Economics*, 32(3), 283–302.
- Torgler, B. (2006). The importance of faith : Tax morale and religiosity. *Journal of Economic Behavior and Organization*, 61, 81–109.
- Torgler, B., & Schaltegger, C. A. (2005). Tax amnesties and political participation. *Public Finance Review*, 33(3), 403–431.
- Torgler, B., & Schneider, F. (2005). Attitudes towards paying taxes in Austria: An empirical analysis. *Empirica*, 32(2), 231–250.
- Verboon, P., & Dijke, M. Van. (2011). When do severe sanctions enhance compliance ? The role of procedural fairness. *Journal of Economic Psychology*, 32(1), 120–130.
- Verboon, P., & van Dijke, M. (2007). A self-interest analysis of justice and tax compliance: How distributive justice moderates the effect of outcome favorability. *Journal of Economic Psychology*, 28(6), 704–727.
- Wahlund, R. (1992). Tax Changes and Economic-Behavior - the Case of Tax Evasion. *Journal of Economic Psychology*, 13(4), 657–677.
- Webley, P., & Halstead, S. (1986). Tax evasion on the micro: Significant simulations or expedient experiments? *Journal of Interdisciplinary Economics*, 1(2), 87–100.
- Webley, P., Robben, H., Elffers, H., & Hessing, D. (1991). *Tax evasion An experimental approach*. Cambridge: Cambridge University Press.
- Wenzel, M. (2004). The Social Side of Sanctions : Personal and Social Norms as Moderators of Deterrence. *Law and Human Behavior*, 28(5), 547–567.
- Witte, A. D., & Woodbury, D. F. (1985). The Effect of Tax Laws and Tax Administration on

Tax Compliance: The Case of The U.S. Individual Income Tax. *National Tax Journal*, 38(1), 1-14.

Yitzhaki, S. (1974). A Note on Income Tax Evasion: A Theoretical Analysis. *Journal of Public Economics*, 3(2), 201–202.

Yoon, S., Yoo, S., Kim, J., Yoon, S., Yoo, S., & Kim, J. (2011). Ambiguity , Audit Errors , and Tax Compliance. *Asia-Pacific Journal of Accounting & Economics*, 18(1), 10–26.

Appendix

Abstract English

Penalties are the prevalent tool for policy-makers around the world, trying to decrease tax evasion. However, empirical studies testing the predictions of the established economic standard model of income tax evasion yielded ambiguous results in regard to the deterrent effect of penalties in the tax contest. More recent psychological models like the *Slippery Slope Framework* introduced the idea that penalties in the tax context are subject to interpretation and depend on situational factors and noneconomic variables. The present study contributes to the current state of research by systematically and exhaustively searching for academic literature and summing up all empirical findings concerning the effect of penalties on tax compliance (N=85). Results are overall ambiguous, indicating, that evidence for the deterrent effect of penalties is at best weak. Further, we suggest that many of the counterintuitive findings of empirical studies in regards to penalties could be explained within the concepts of the *Slippery Slope Framework*.

Abstract Deutsch

Strafen sind ein weitverbreitetes Mittel der Steuerpolitik im Kampf gegen Steuerhinterziehung. Das ökonomische Standardmodell der Steuerehrlichkeit postuliert, dass höhere Strafen zu weniger Hinterziehung führen. Empirische Forschung kam zu uneindeutigen Ergebnissen, was diese Vorhersage betrifft. Neuere psychologische Modelle, wie das *Slippery Slope Framework* führten die Idee ein, dass Strafen im Steuerkontext Interpretationssache sind und von verschiedenen nicht-ökonomischen Kontextvariablen abhängen. Diese Studie erweitert den Forschungsstand, indem sie systematisch und erschöpfend die akademische Literatur über Strafen im Steuerkontext sucht und zusammenfasst (N=85). Ergebnisse sind im allgemeinen ambivalent und deuten an, dass Hinweise für die abschreckende Wirkung von Strafen schwach sind. Weiter schlagen wir vor, dass viele von den nicht intuitiven Resultaten aus empirischen Studien mit Hilfe des *Slippery Slope Frameworks* erklärt werden können.

Papers included in the sample

EXPERIMENTAL STUDIES

Title	Authors	Year	Effect
On Legal Sanctions	Schwartz, Orleans	1967	Positive/ambiguous
Understanding tax evasion	Spicer, Lundstedt	1976	No effect
A Simulation Study of Income Tax Evasion	Friedland, Maital, Rutenberg	1978	Positive
A Note on Tax Evasion as a Function of the Quality of Information about the Magnitude and Credibility of Threatened Fines: Some Preliminary Research	Friedland	1982	Positive
Components of Reward in Social Dilemmas	Komorita, Barth	1985	No effect/ambiguous
Tax evasion on the micro: Significant simulations or expedient experiments?	Webley, Halstead	1986	No effect
Extending the lessons of laboratory experiments on tax compliance to managerial and decision economics	Baldry	1987	No effect
Taxpayer attitudes toward tax audit risk	Chang, Nichols, Schultz	1987	Positive
Amazing Grace : Tax Amnesties and Compliance	Alm, McKee, Beck	1990	Positive/ambiguous
Tax evasion An experimental approach	Webley, Robben, Elffers	1991	No effect
Experimental Evidence on Taxpayer Reporting Under Uncertainty	Beck, Davis, Jung	1991	Positive
The Taxpayer ' s Labor and Reporting Decision : The Effect of Audit Schemes	Collins, Plumlee	1991	Positive/ambiguous
Beyond Deterrence : Behavioral Decision Theory and Tax Compliance	Scholz, Casey	1991	Positive
Institutional Uncertainty and Taxpayer Compliance	Alm, Jackson, McKee	1992	No effect/ambiguous
An examination of taxpayer preference for aggressive tax advice	Hite, McGill	1992	No effect
Estimating the Determants of Taxpayer Compliance with Experimental Data	Alm, Jackson, McKee	1992	Positive
An examination of tax practitioner decisions: The role of preparer sanctions and framing effects associated with client condition	Newberry, Reckers, Wyndelts	1993	Positive/ambiguous
Economic & Noneconomic Factors in Tax Compliance	Alm, Sanchez, De Juan	1995	Positive/ambiguous
Taxpayers' preference for type of advice from tax practitioner: A preliminary examination	Tan	1999	No effect
Examining the determinants of tax compliance by experimental data : a case of Korea	Park, Hyun	2003	Positive

Punishment in a Regulatory Setting : Experimental Evidence from the VCM	Anderson, Stafford	2003	Positive
Tax amnesties and political participation	Torgler, Schaltegger	2005	Negative
Tax compliance and obedience to authority at home and in the lab : A new experimental approach	Cadsby, Maynes, Trivedi	2006	No effect
Misperception of chance and loss repair : On the dynamics of tax compliance	Maciejovsky, Kirchler, Schwarzenberger	2007	Positive/ambiguous
Persuasive Communications : Tax Compliance Enforcement Strategies for Sole Proprietors	Hasseldine, Hite, James	2007	Positive/ambiguous
Influence of Accountability and Penalty Awareness on Tax Compliance	Sander, Reckers, Iyer	2008	No effect/ambiguous
The excess burden of tax evasion — An experimental detection – concealment contest	Bayer, Sutter	2009	Positive/ambiguous
Sequences of audits , tax compliance , and taxpaying strategies	Kastlunger, Kirchler, Mittone	2009	Positive/ambiguous
Increasing Tax Compliance In Washington State : A Field Experiment	Iyer, Reckers, Sanders	2010	Positive/ambiguous
Cheating , emotions , and rationality : an experiment on tax evasion	Coricelli, Joffily, Montemarquette	2010	Negative
Fighting income tax evasion with positive rewards	Bazart, Pickhardt	2011	Positive/ambiguous
When do severe sanctions enhance compliance ? The role of procedural fairness	Verboon, Van Dijke	2011	Positive/ambiguous
Public goods and voting on formal sanction schemes	Putterman, Tyran, Kamei	2011	Positive
Rationality Versus Emotions : The Case of Tax Ethics and Compliance	Maciejovsky, Kirchler, Schwarzenberger	2012	No effect/ambiguous
'Spite effects' in tax evasion experiments	Cullis, Jones, Soliman	2012	No effect/ambiguous
Deterrence and moral persuasion effects on corporate tax compliance: Findings from a randomized controlled trial	Ariel	2012	No effect
Certainty of Punishment versus Severity of Punishment : An Experimental Investigation	Friesen	2012	Positive
Learning in experiments : Dynamic interaction of policy variables designed to deter tax evasion	Soliman, Jones, Cullis	2014	No effect/ambiguous
On the path dependence of tax compliance	Bruttel, Friehe	2014	Positive
Empathy, sympathy, and tax compliance	Alm, Calvet Christian	2014	Positive
Tax compliance under horizontal and vertical equity conditions : An experimental	Finocchiaro, Ilde	2014	Negative/ambiguous

approach			
Does the Sole Description of a Tax Authority Affect Tax Evasion ? - The Impact of Described Coercive and Legitimate Power	Hartl, Hoffmann, Gnagl	2015	No effect
Factors Influencing Compliance Behavior in a Tax Laboratory Experiment	Spalek	2015	No effect
Tax evasion and social information : an experiment in Belgium , France , and the Netherlands	Lefebvre, Pestieau, Riedl, Villeval	2015	No effect/ambiguous
The effect of competition on tax compliance : The role of audit rules and shame	Casagrande	2015	Positive/ambiguous
Tax compliance and enforcement in the pampas evidence from a field experiment	Castro, Scartascini	2015	Positive
Do students behave like real taxpayers in the lab ? Evidence from a real effort tax compliance experiment	Choo, Fonseca, Myles	2016	Positive
The determinants of compliance with environmental tax : Behavioural study motivated by the case of Indonesia	Iskandar	2016	Positive
Gender differences in deception behaviour – the role of the counterpart Gender differences in deception behaviour – the role of the counterpart	Lohse, Qari	2017	Positive

SURVEY STUDIES

Title	Authors	Year	Effect
Tax Compliance and Perceptions of the Risks of Detection and Criminal Prosecution	Klepper, Nagin	1989	Positive
General prevention of tax evasion: A factorial survey approach	Thurman	1989	Positive
Conscience, Significant Others, AND rational Choice: Extending The Deterrence Model	Grasmick, Bursik	1990	Positive
Characteristics of New Zealand Tax Evaders: a Note	Hasseldine, Kaplan, Fuller	1994	No effect
Deterring economic crimes : The case of tax evasion	Varma, Doob	1998	No effect
Small business owners' attitudes on VAT compliance in the UK	Adams, Webley	2001	No effect/ambiguous
Why People Pay Taxes : The Case of the Development Levy in Tanzania	Fjeldstad, Semboja	2001	Negative
The Social Side of Sanctions : Personal and Social Norms as Moderators of Deterrence	Wenzel	2004	Positive/ambiguous
Steuerhinterziehung und die Berücksichtigung des Einkommens in der Strafbemessung	Muehlbacher, Kirchler, Hoelzl	2007	Positive/ambiguous
A self-interest analysis of justice and tax compliance: How distributive justice moderates the effect of outcome favorability	Verboon, Van Dijke	2007	Positive/ambiguous
An Investigation Into Australian Personal Tax Evaders- Their Attitudes Towards	Devos	2009	No effect/ambiguous

Compliance And The Penalties For Non-Compliance			
Factors affecting tax compliance behaviour in self assessment system	Palil, Mustapha	2011	Positive
Factors influencing taxpayers' Compliance with the tax system: An empirical study in Mekelle City, Ethiopia	Engida, Baisa	2014	No effect
The impact of threat of punishment on tax compliance and non- compliance attitudes in Malaysia	Mohdali, Isa, Hana	2014	No effect
Do Penalties and Enforcement Measures Make Taxpayers More Compliant?—The View of Australian Tax Evaders	Devos	2014	Positive/ambiguous
A Comparative Study of Compliant and Non- Compliant Individual Taxpayers in Ethiopia evidence from Amhara region	Oghenekaro, Olakunle	2015	No effect/ambiguous

AGGREGATE DATA

Title	Authors	Year	Effect
The Effect of Tax Laws And Tax Administration on Tax Cpmpliance: The Case of the U.S. Individual Income Tax	Witte, Woodbury	1985	No effect/ambiguous
Time Value of Money and Income-Tax Evasion Under Risk-Averse Behavior - Theoretical-Analysis and Empirical-Evidence	Crane, Nourzad	1985	Positive
Einkommensteuerhinterziehung in der Schweiz: Eine empirische Analyse	Pommerehne, Weck-Hannemann	1989	No effect
Tax Structure and Tax Compliance	Alm, Bahl, Murray	1990	Negative
Income Tax Compliance in India : An Empirical Analysis	Das-Gupta, Lahiri	1995	No effect
Tax rates, tax administration and income tax evasion in Switzerland	Pommerehne, Weck-Hannemann	1996	No effect
An Empirical Analysis of the Impact of Government Tax and Auditing Policies on the Size of the Underground Economy : The case of the United States 1973-94	Cebula	1997	Positive
The Effects of Tax Rates and Enforcement Policies on Taxpayer Compliance : A Study of Self-Employed Taxpayers	Ali, Cecile, Knoblett	2001	Positive/ambiguous
Soziale Schichtung und Delinquenz. Eine empirische Anwendung eines Rational Choice-Ansatzes mit Hilfe von Querschnittsdaten des ALLBUS 1990 und 2000	Mehlkop, Becker	2004	No effect
Road Fund Tax Compliance: An Analysis of Enhancement Strategies	Egger, Hackbart	2005	No effect
Do Audits Enhance Compliance ? An Empirical Assessment of VAT Enforcement	Bergman, Nevarez	2006	Negative/ambiguous
The Revenue Impact of Repeated Tax Amnesties	Luitel, Sobel	2007	Positive

Criminal Investigation Enforcement Activities and Taxpayer Noncompliance	Dubbin	2007	Negative/ambiguous
What Shapes Attitudes Toward Paying Taxes ? Evidence from Multicultural European Countries	Torgler, Schneider	2007	Negative
An Empirical Analysis of Determinants of Recent Federal Personal Income Tax Evasion in the U.S.	Cebula	2011	Positive
Self-perceptions , government policies and tax compliance in Germany	Feld, Larsen	2012	No effect/ambiguous
New and Current Evidence on Determinants of Aggregate Federal Personal Income Tax Evasion in the United States	Cebula	2013	Positive
Tax Compliance and Corporate Income Tax – The Case of Slovenia	Lesnik, Kracun, Jagric	2014	Positive
Beyond a Deterrence Approach towards the Undeclared Economy : Some Lessons from Bulgaria	Williams, Franic, Williams	2016	No effect
Enforcement , Socioeconomic Diversity , and Tax Filing Compliance in the United States	Alm, Clark, Leibel	2016	No effect

PRISMA 27-point checklist



PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4-12
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	12
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	-
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	15
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	15
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	16
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	15
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	18
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	18
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	15
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	20
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	19



PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	31
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	21-25
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	18
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	41-44
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	-
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	-
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	-
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	31
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	26 - 29
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	29 - 31
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	31
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	32
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	-