

Lohse, Tim; Simon, Sven A.

Working Paper

Compliance in Teams - Implications of Joint Decisions and Shared Consequences

CESifo Working Paper, No. 7807

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Lohse, Tim; Simon, Sven A. (2019) : Compliance in Teams - Implications of Joint Decisions and Shared Consequences, CESifo Working Paper, No. 7807, Center for Economic Studies and Ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/207198>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Compliance in Teams – Implications of Joint Decisions and Shared Consequences

Tim Lohse, Sven A. Simon

Impressum:

CESifo Working Papers

ISSN 2364-1428 (electronic version)

Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH

The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute

Poschingerstr. 5, 81679 Munich, Germany

Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de

Editor: Clemens Fuest

www.cesifo-group.org/wp

An electronic version of the paper may be downloaded

- from the SSRN website: www.SSRN.com
- from the RePEc website: www.RePEc.org
- from the CESifo website: www.CESifo-group.org/wp

Compliance in Teams – Implications of Joint Decisions and Shared Consequences

Abstract

In today's business environment, team work is omnipresent. But might teams be more prone toward non-compliance with laws and regulations than single individuals despite imminent negative consequences of uncovering misconduct? The recent prevalence of corporate delinquencies gives rise to this concern. In our laboratory experiment, we investigate the determinants of teams' compliance behavior. In particular, we disentangle the effect of deciding jointly as a team of two from sharing the economic consequences among both team members. Our findings provide evidence that teams are substantially less compliant than individuals are. This drop in compliance is driven by the joint, rather than the individual, liability of team members. In contrast, whether subjects make their decisions alone or together does not influence the overall compliance rate. When coordinating their compliance decision teams predominately discuss the risk of getting caught in an audit, and team decision-making is characterized by behavioral spillovers between team members. Holding each team member fully liable is a promising means to deter them from going astray.

JEL-Codes: C920, D910, K420.

Keywords: compliance, lying, team decision, shared liability, audit, communication, laboratory experiment.

Tim Lohse
Berlin School of Economics and Law
Badensche Straße 50-51
Germany – 10825 Berlin
tim.lohse@hwr-berlin.de

*Sven A. Simon**
Max Planck Institute for Tax Law and
Public Finance, Marstallplatz 1
Germany – 80539 Munich
sven.simon@tax.mpg.de

*corresponding author

August 5, 2019

1 Introduction

To comply, or not to comply: that is a question potentially worth billions of dollars as the dividend arbitrage trading scheme ‘CumEx’ or the Volkswagen emissions scandal show. These recent violations of tax laws and pollution standards, on the one hand, and the strong reliance by companies on team decisions, on the other hand, suggest a dark side of team work: despite enormous fines levied on companies, the well-documented deterring effect of audits seems to lose its bite when teams make compliance decisions. One crucial difference with an individual compliance problem is the involvement of several agents or parties. In order to run the highly complex share deals of the CumEx tax fraud scheme (which had no economic purpose other than to receive tax reimbursements for taxes that were never paid) employees both within, as well as between, banks, stock traders, and law firms had to cooperate. In the case of Volkswagen, a team of engineers and managers decided to employ manipulative software to mask the actual emissions level of their diesel engines. Apart from this team decision-making, there is a second crucial determinant of corporate non-compliance. The concomitant economic consequences – the gains from undetected wrongdoing or the losses in the case an audit reveals the misconduct – oftentimes do not arise to the responsible decision-makers alone. Typically, the entire company and potentially also third parties can be affected, such as shareholders via dividend payments and stock prices or customers via quality of goods and services. Thus, compliance decisions within firms might also hinge on the fact that the economic liability is shared among several agents.

This paper studies what determines teams’ compliance behavior. We examine whether dyads as the smallest but yet important size of a team are more prone toward non-compliance than single individuals. Firms, public administrations or the military frequently rely on dyads in the form of dual control in the course of which two persons have to agree on a certain decision. If so, dishonest behavior in teams could either be driven by the joint decision-making of several people. Alternatively, sharing the monetary consequences among the team members might explain corporate delinquencies, too. The separation of these two dimensions also allows for an investigation of how non-compliance by teams can effectively be deterred. In particular, we ask under which conditions an auditing mechanism is suited to address non-compliance by teams. In order to answer these research questions, we run a computerized compliance experiment. Participants first

earn a low or a high individual income, and then have the choice to report either a low or a high income with the latter being subject to a deduction. Non-compliance, i.e., falsely claiming to have a low income although the true income is high, and thereby evading the deduction, can pay off. But an exogenous audit may apply, revealing the underreporting and leading to a fine. Our incentivized laboratory experiment is based on a 2-by-2 design: along the first dimension, subjects either decide alone or in a team on their income reports. The two team members may coordinate before their decision-making via a chat. Along the second dimension, subjects are liable for the economic consequences either alone or in a team, i.e., team members either share their final earnings or not. Both dimensions are crucial but as the above mentioned real-world examples reveal they are also intertwined. Therefore, a laboratory experiment is particularly well suited to shed light on this matter.

Our results are as follows. Dyads are significantly more non-compliant than single individuals. Joint decision-making per se contributes very little to this increase in non-compliance and we find no significant effect along this dimension. In contrast, it is the shared (rather than individual) liability which leads to significantly more misreporting and accounts for most of the increased dishonesty of teams. This effect of shared liability is robust and independent of whether decisions are made on an individual basis or in a team. Hence, our findings suggest that holding each team member fully (rather than only partially) liable might effectively restore compliance in teams. Further results indicate that teams successfully coordinate by making their decisions consensually. In teams whose members decide together, we find evidence for behavioral spillovers between team members. This leads to more consequential declaration behavior. More often teams tend to be dishonest, or sometimes honest, for both declarations. Quantitatively, most important motivations for (non-)compliance in the team chat protocols include the risk dimension of the compliance task, the monetary consequences of the report, and the concept of being honest.

The remainder of the paper is organized as follows. In the next section we summarize the related literature. Section 3 explains the experimental design and our hypotheses. Section 4 provides an analysis of the experimental data, while section 5 discusses the experimental results, and section 6 concludes.

2 Related Literature

As the compliance decision is a specific case of the more general question of whether or not to be honest, our study contributes to the literature on deception in teams. This literature focuses on motivations for (dis)honesty but typically abstracts from an auditing mechanism. Whereas Sutter (2009) shows that teams are as deceptive as individuals but engage in more sophisticated lying, Cohen et al. (2009) argue that groups might cheat even more. Findings by Weisel and Shalvi (2015) in a setting of a sequential dice roll (Fischbacher and Föllmi-Heusi 2013) indicate more dishonest reporting in teams compared to individuals, but also stress the crucial role of aligned incentives. Closest to our setting, Kocher et al. (2017) conduct a variation of the dice-rolling experiment and confirm higher misreporting in teams. Communication in teams, as implemented by a chat, turns out to be crucial for coordination.¹ Inherent to team decisions is also the question of who benefits and who loses from a dishonest report. There is convincing evidence that subjects lie more often if others benefit (Wiltermuth 2011, Gino et al. 2013), but they lie less often when it comes at the expense of oneself or of others (“altruistic white lies” or “selfish black lies,” Erat and Gneezy 2012). Conrads et al. (2013) investigate team incentives in the absence of team decisions and find that equally shared earnings from individual reports increase dishonesty compared to individual earnings.

Our focus of compliance behavior in teams is insightful and contributes to the existing literature in at least three ways. First, compliance problems differ from mere cheating opportunities, leading to different behavioral predictions. Besides the intrinsic costs of lying (e.g., Abeler et al. 2014, Gneezy et al. 2018), the possibility of an audit introduces extrinsic costs such as monetary fines, and dishonest reporting imposes a (monetary) risk for the team partner.² Hence, our approach of potentially sharing the economic consequences goes beyond the literature studying lying in teams and the associated aspects of positive and negative externalities to other team members (such as pro-social lying, cf. e.g. Bizziou-van-Pol et al. 2015). Moreover, regret or shame may be more

¹However, Barr and Michailidou (2017) show that the introduction of a focal point in the form of a bonus leads to the coordination of reports even in the absence of interaction. Other aspects of dishonest decision-making in teams include the team formation process (Chytilova and Korbil 2014) and the gender composition of the team (Muehlheusser et al. 2015).

²Our approach is also promising given the inconclusive findings from the literature on individual versus group risk taking behavior. Baker et al. (2008) find groups to be risk neutral; Shupp and Williams (2008) that they are more risk loving, and Masclet et al. (2009) that they are more risk averse.

pronounced in a compliance framework, which makes the team decision process richer as compared to the cheating problem. Second, including both the liability dimension and the decision-making dimension allows us to identify and to quantify the importance of each of the two dimensions. This is new to the literature. In light of the real world examples of non-compliance sketched above, our findings provide answers to important management and policy questions such as if, and under which circumstances, audits might be able to restore honesty in team decisions. More specifically, do companies need to take a setback and rely more on the decisions of single employees rather than teams to prevent dishonesty? Should law enforcement hold both parties of a joint non-compliance decision (partially) liable, or rather focus on punishing one party to the full extent? Third, the analysis of the chat protocol enables us to identify the underlying motivations of the compliance decision in a team. Based on that, we can draw conclusions on the relationship of the compliance problem in our setting and the cheating problems studied by the deception literature so far.

More generally, we also add to the literature on tax compliance. Only a few papers experimentally study the role of social interactions within a group for tax evasion. Fairness seems to play an important role, but not social conformity (Fortin et al. 2007). Publicly exposing the deceivers in the laboratory deters evasion (Coricelli et al. 2010) and can have both a negative contagion effect and a positive shaming effect (Laury and Wallace 2005). As opposed to individual tax compliance, Abraham et al. (2017) show that joint tax compliance is positively affected by social norms. An inherent feature of team decision-making is also the transmission of information by taxpayer communication about the audit rate and compliance behavior of others, as studied by Alm et al. (2009, 2017). We contribute to the experimental compliance literature by explicitly studying differences in the compliance behavior between individuals and teams. Our differentiation between the collective decision and the shared economic consequences among the team members is new to the literature and promises important insights for a better understanding of non-compliance and tax evasion. Importantly, we provide evidence that it is the shared liability that boosts non-compliance and suggest that enforcement policies aiming at this dimension might be particularly promising.

3 The Experiment

3.1 Design of the laboratory experiment

Overview. We use an experimental compliance framework with 10 periods. A period commences with an income-generating stage: each subject i earns with probability 0.2 a low income Y^l (400 Experimental Currency Units [ECU] = EUR 8) and with probability 0.8 a high income Y^h (ECU 1000 = EUR 20).³ This binary income structure has two advantages: first, it facilitates the computation of outcomes and ensures a clear-cut compliance decision; and second, existence of the low income secures the credibility of declarations of a low income. While no deduction applies to a low income, subjects are supposed to pay a deduction Δ of ECU 400 on a high income. The deduction returns to the laboratory. For declarations of a high income, the deduction of ECU 400 is subtracted automatically from the income without further investigation. In contrast, declarations of a low income are independently from each other audited by the computer with a fixed probability $p = 0.5$.⁴ In the case of no audit or a truthful declaration, no deductions or fines apply. However, if an audit reveals non-compliance, i.e., if a high income has been falsely declared as low, subjects must remargin the deduction of ECU 400 topped up by a fine Θ of ECU 200. At the end of each period, subjects receive an overview of their earnings of this period (and of their partners' reporting and earnings in T2 to T4; see below). The unique best choice for subjects with a low income is to be honest about their income. Therefore, most of the subsequent analysis focuses on subjects with a high income since they face a trade-off of whether to be honest or not.

The experiment follows an innovative 2-by-2 between-subjects design as illustrated in Figure 1: first, we vary whether subjects decide on their declared income individually or jointly in a team; and second, we vary whether subjects are liable individually or collectively, i.e., on an individual level to the full extent or on a team level to proportional amounts.

Figure 2 shows the five stages of the decision process in a period of the experiment. In all treat-

³This stage is implemented as a real-effort task that involves counting numbers on the computer screen to generate income $Y_i \in \{Y^l, Y^h\}$ (Abeler et al. 2011). With probability 0.2 subjects have to count a few numbers and generate income Y^l , while with probability 0.8, they have to count many numbers and generate income Y^h . After the successful completion of the task, subjects received the respective pre-determined income.

⁴We have chosen a random audit because it is the easiest to understand. Recent literature studies alternatives, such as endogenous audit mechanisms where the audit probability increases with the estimated degree of underreporting (Gilpatric et al. 2011; Cason et al. 2016) or subjects' perceived untrustworthiness (Konrad et al. 2017). Our parametrization of incomes, deduction and fine borrows from Konrad et al. (2017).

1 st Dimension 2 nd Dimension	Individual Decision	Joint Decision
Indiv. Liability	Treatment 1 (T1) Subjects = 56 Observations = 428	Treatment 3 (T3) Subjects = 78 Observations = 632
Shared Liability	Treatment 2 (T2) Subjects = 52 Observations = 450	Treatment 4 (T4) Subjects = 82 Observations = 656

Figure 1: Overview of the experiment

ments with joint decision-making or shared liability (T2 to T4), the computer randomly matches one subject with another to form a team in each period after the income-generating stage (stage 1). The crucial behavioral differences we are studying primarily depend on whether a person is acting alone or not. We define a team as a dyad (cf. Conrads 2013, Muelheusser et al 2015, or Weisel and Shalvi 2015) being the smallest but yet important size of a group. As the main focus of our experiment is to isolate the effect of joint decision-making from the effects of shared liability, we abstract from team dynamics over time and use a pre-announced absolute stranger matching protocol. In addition, we ensure that matched team members have the same income, which leads to an equal playing field for both partners. The first stage is followed by an information stage (stage 2) where subjects learn of their income(s). Subjects in Treatment 3 and 4 then have the possibility to coordinate on their reports (stage 3), while subjects in Treatment 1 and 2 skip this step and enter the reporting stage directly (stage 4). For procedural reasons that will be explained below, the differentiation between “reports” - a subject’s own income statement in stage 4 - and “declarations” - the finally declared income which may be subject to an audit - is important to our setting.

First dimension: Decision-making. In treatments with individual decision-making (T1 and T2), subject i ’s only task is to make a report R_i of her own income Y_i with $R_i \in \{h, l\}$. This report automatically determines the declaration $D_i \in \{h, l\}$ of her income as $D_i = R_i$. In treatments with joint decision-making (T3 and T4), the computer forms a team by randomly matching a subject i with a partner j right after the income-generating stage. Team members i and j are mutually

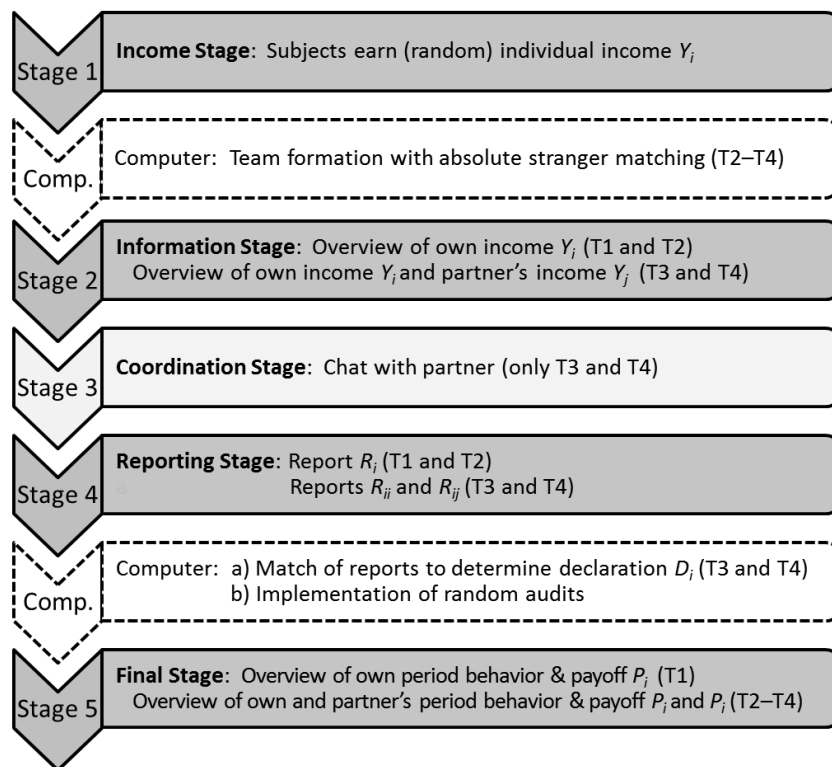


Figure 2: Procedure of one period in the experiment

informed about their incomes Y_i and Y_j . Each team member must then make two reports: one report of her own income and one report of the income of her partner. Hence, i reports R_{ii} and R_{ij} , whereas j reports R_{jj} and R_{ji} . In order to allow for a coordination on the reports before submitting them, team members enter a coordination stage (stage 3) and have the opportunity to exchange messages in a chatbox for 120 seconds. Then, both team members simultaneously make their reports on both incomes. For subject i the two reports R_{ii} and R_{ji} determine the declaration D_i of subject i 's income in the following way: if $R_{ii} = R_{ji}$, i.e., if the own income report matches with the partner's report on one's own income, this matched report is registered as declaration D_i of subject i : $D_i = R_{ji} = R_{ii}$. However, if $R_{ii} \neq R_{ji}$, the true income of subject i is taken as i 's declaration: $D_i = Y_i$ (and is never subject to an audit if $Y_i = Y^h$). Therefore, a dishonest declaration is only feasible if both team partners unanimously agree on being non-compliant.⁵ The described procedure applies vice versa to team partner j 's declaration and, thereby, implies the joint decision-making. Finally, a declaration of a low income $D_i = Y^l$ is audited independently of the team partner's declaration of a low income $D_j = Y^l$ with probability $p = 0.5$, even in the case of joint decision-making.

Second dimension: Liability. The second dimension varies to whom the payoff consequences of a declaration accrue. In treatments with individual liability (T1 and T3), subject i 's period payoff P_i consists only of the earnings E_i from the declaration of the own income including the potential deduction and fine. There are two possible earnings for a subject with a high income: first, in the case of an honest declaration $D_i = h$ earnings are $E_i = Y^h - \Delta$; second, in the case of a dishonest declaration $D_i = l$ subject i faces a potential audit with probability $1/2$ and expected earnings are $E_i = Y^h - \frac{1}{2}(\Delta + \Theta)$. In Treatment 1 each subject is self-responsible (with $D_i = R_i$) and the treatment resembles the baseline setting in standard compliance games. In Treatment 3, individual liability means that if one team member conspires to underreport the other team member's income, she herself neither benefits from it nor would she face any fines should an audit reveal the underreporting. To be more precise, a dishonest report R_{ji} by subject j about partner

⁵Two alternative designs for the default mechanism of non-matching reports are conceivable but inferior to our approach. In case an audit is automatically triggered, the best response to a team partner's report would be to state the same, i.e., comply if she complies and do not comply if she does not comply. In contrast, it is the weakly dominant strategy to base the report on one's true preference for our coordination mechanism. In case non-matching reports lead to the implementation of only one of the reports with equal probability, incentives to reach an agreement on dishonest reports would be reduced and joint decision-making would essentially be incomplete.

i 's income Y_i does not impact subject j 's earnings.

In treatments with shared liability (T2 and T4), both team members mutually share their earnings. A subject i 's period payoff is given by $P_i = \frac{1}{2}(E_i + E_j)$ with E_i and E_j denoting i 's and j 's earnings including potential deductions and fines, respectively. Thus, both team members share the potential monetary benefits of non-compliance. But in case the random and independent audits target D_i or/and D_j they also share the potential monetary costs of a dishonest declaration. This procedure keeps the range of potential earnings comparable to treatments without shared liability and does not change the total stake size of the individual reporting decision.

Implementation and payoffs. We conducted the experiment in July 2017 at the econlab Munich with predominantly local students. Fourteen sessions took place, each lasting for up to 100 minutes. In each session there were up to 24 subjects, with a total number of 268 subjects (average age 22.6; 51 percent female participants). Subjects were recruited using ORSEE (Greiner 2015) and the experiment was programmed and implemented in z-Tree (Fischbacher 2007). Upon arrival at the lab, subjects were seated randomly at individual cubicles. At each cubicle, subjects found a hard copy of the experimental instructions that were also read aloud by the same instructor before the start of each session. Subjects then answered some trial questions on the experimental setup to ensure a proper understanding of the instructions. After the successful completion, the main experiment with 10 periods started. One period was randomly chosen for the payoff. Following the final period, subjects first answered a short questionnaire on the experiment. They subsequently took part in several incentivized post-experimental tests that allowed for additional earnings. First of all, we conducted Holt and Laury's (2002) risk elicitation task to control for subjects' risk preferences in the compliance decision. Secondly, we used the Cognitive Reflection Test (Frederick 2005) to identify impulsive and reflective decision-making which might be related to the subjects' ability to find the economically optimal solution to the compliance problem. Thirdly, in order to capture aspects such as prosocial or competitive behavior, subjects participated in a variant of Murphy et al.'s (2011) test for social value orientation. This might be related to subjects' willingness to make a dishonest report when their partner is affected. Each session concluded with a socio-economic questionnaire and an overview of the subject's total earnings. Subjects earned an average of EUR 22 (min. EUR 9, max. EUR 31) including a show-up fee of EUR 6 and earnings from the post-experimental tests up to EUR 3.90.

3.2 Conceptual framework

In our framework the compliance decision rests on two factors. First, we include the utility from the monetary payoff of non-compliance. Non-compliance has the same expected payoff along all dimension, but shared liability leads to a more condensed distribution. This is relevant when a subject's compliance decision is affected by her risk preference. Second, subjects may incur psychological costs from an untruthful report which may very well depend on the specific experimental treatment.⁶ Along the first dimension, L_i^{ID} and L_i^{JD} denote subject i 's lying costs for individual and joint decision-making, respectively. Following the rich experimental evidence for an erosion of social norms in groups (Gino et al. 2009, Weisel and Shalvi 2015, Kocher et al. 2017) and given that decisions in teams allow individuals to disguise the own responsibility for a dishonest report (Conrads et al. 2013), we assume $L_i^{ID} > L_i^{JD}$. Lying costs can also occur along the second dimension and depend on whether subject i is individually liable (L_i^{IL}) or shares the economic consequences with the team partner (L_i^{SL}). Since sharing the potential benefits of a dishonest report has been shown to increase misreporting (Wiltermuth 2011, Gino et al. 2013) as it is essentially a Pareto white lie (Erat and Gneezy 2012), it is reasonable to suppose that $L_i^{IL} > L_i^{SL}$.

We focus on identical compliance behavior by both team partners and define $\tilde{Y} := Y^h - \frac{1}{2}(\Delta + \theta)$. Recall that the random audit mechanism targets each declaration D_i separately. We assume $u(\cdot)$ to be strictly concave and increasing in the earnings P_i . The treatment specific spread of expected payoffs (in utility terms) of subject i for non-compliance minus compliance read:

	<i>Non-Compliance</i>	<i>Compliance</i>
T1:	$(1-p)u\left(\tilde{Y} + \frac{1}{2}(\Delta + \theta)\right) + pu\left(\tilde{Y} - \frac{1}{2}(\Delta + \theta)\right) - (L_i^{ID} + L_i^{IL})$	$- u(Y^h - \Delta)$ (1)
T2:	$(1-p)^2u\left(\tilde{Y} + \frac{1}{2}(\Delta + \theta)\right) + 2p(1-p)u\left(\tilde{Y}\right) + p^2u\left(\tilde{Y} - \frac{1}{2}(\Delta + \theta)\right) - (L_i^{ID} + L_i^{SL})$	$- u(Y^h - \Delta)$ (2)
T3:	$(1-p)u\left(\tilde{Y} + \frac{1}{2}(\Delta + \theta)\right) + pu\left(\tilde{Y} - \frac{1}{2}(\Delta + \theta)\right) - (L_i^{JD} + L_i^{IL})$	$- u(Y^h - \Delta)$ (3)
T4:	$(1-p)^2u\left(\tilde{Y} + \frac{1}{2}(\Delta + \theta)\right) + 2p(1-p)u\left(\tilde{Y}\right) + p^2u\left(\tilde{Y} - \frac{1}{2}(\Delta + \theta)\right) - (L_i^{JD} + L_i^{SL})$	$- u(Y^h - \Delta)$ (4)

Aspects of risk aversion due to the monetary punishment in case of a detected misconduct on the one hand and strictly positive lying costs on the other hand may lead to negative signs of some or even all spreads. In such a case subject i would comply.⁷ The subsequent behavioral predictions

⁶Given the constant size of a lie, we confine our framework to fixed lying costs. Without loss of generality, the lying costs are denoted separately according to the two dimensions of the experiment. This approach provides a clearer picture of the different costs at work.

⁷For a risk-neutral subject without any moral concerns (i.e. lying costs of zero), the spreads as given by (1) to (4)

about the effect of our two dimensions on compliance behavior are based on a comparison of the spreads across treatments.

Hypothesis 1 *Joint rather than individual decision-making leads to less compliance.*

Proof. Comparing the spreads of T3 and T1, the claim follows directly from the positive sign of (3)–(1):

$$\begin{aligned} & (1-p)u\left(\tilde{Y} + \frac{1}{2}(\Delta + \theta)\right) + pu\left(\tilde{Y} - \frac{1}{2}(\Delta + \theta)\right) - (L_i^{JD} + L_i^{IL}) - u(Y^h - \Delta) \\ & - \left[(1-p)u\left(\tilde{Y} + \frac{1}{2}(\Delta + \theta)\right) + pu\left(\tilde{Y} - \frac{1}{2}(\Delta + \theta)\right) - (L_i^{ID} + L_i^{IL}) - u(Y^h - \Delta) \right] \\ & = L_i^{ID} - L_i^{JD} > 0 \end{aligned}$$

and, in an analogous way, from the positive sign of (4)–(2). ■

Beyond this model based prediction, further behavioral considerations can point in the same direction as Hypothesis 1. First, by aggregating individual capabilities, teams enjoy higher cognitive skills (Kocher and Sutter 2005, Charness et al. 2007, Feri et al. 2010, Sutter et al. 2013, among others), act with more self-interest (Bornstein et al. 2004, Kugler et al. 2007), and, therefore, might find the economically optimal solution to the compliance problem more easily. Second, teams might have access to richer information than individuals, in particular they might update on the compliance behavior of others and the tightness of the audit regime. Nevertheless, this effect is potentially limited since audit probabilities are common knowledge and subjects receive feedback on their own behavior as well as their partner’s behavior at the end of the period.

Evidence for Hypothesis 1 would cast a severe shadow on the concept of team work and joint decision-making in general and would provide a potential explanation for the recent corporate scandals of non-compliance. An impending implication for the design of governance rules would be to strengthen the role of individual decision-making and, thereby, possibly bringing back deep hierarchies with lonely leaders.

However, the prediction of Hypothesis 1 should be taken with a grain of salt. Following the literature, we assume $L_i^{ID} > L_i^{JD}$. But in T4 if partner j is devoted to honesty while i isn’t, the latter would need to convince the former of joint non-compliance. Since team partners can

are strictly positive and all equal, suggesting that the subject would underreport in all treatments.

always enforce an honest report, only deceiving subjects have an incentive to convince the partner to give a dishonest report, too. Thus, arguments in favor of dishonesty might be more frequent and cause spillovers for honest subjects. But this might be quite cumbersome and drive up i 's psychological costs of joint decision-making. If that is the case, the inequality of the associated lying costs could become less strict or even reverse, and Hypothesis 1 might not materialize (sharply).

Hypothesis 2 *Shared rather than individual liability leads to less compliance.*

Proof. Given individual decision-making, a positive sign of the difference (2)–(1) would indicate that the spread under shared liability is larger than under individual liability, and, thereby, imply less compliance in T2 than in T1:

$$\begin{aligned}
& (1-p)^2 u\left(\tilde{Y} + \frac{1}{2}(\Delta + \theta)\right) + 2p(1-p)u\left(\tilde{Y}\right) + p^2 u\left(\tilde{Y} - \frac{1}{2}(\Delta + \theta)\right) - (L_i^{ID} + L_i^{SL}) - u(Y^h - \Delta) \\
& - \left[(1-p)u\left(\tilde{Y} + \frac{1}{2}(\Delta + \theta)\right) + pu\left(\tilde{Y} - \frac{1}{2}(\Delta + \theta)\right) - (L_i^{ID} + L_i^{IL}) - u(Y^h - \Delta) \right] \\
& = [(1-p)^2 - (1-p)] u\left(\tilde{Y} + \frac{1}{2}(\Delta + \theta)\right) + 2p(1-p)u\left(\tilde{Y}\right) + (p^2 - p)u\left(\tilde{Y} - \frac{1}{2}(\Delta + \theta)\right) - L_i^{SL} + L_i^{IL} \\
& = (p^2 - p) \left[u\left(\tilde{Y} + \frac{1}{2}(\Delta + \theta)\right) + u\left(\tilde{Y} - \frac{1}{2}(\Delta + \theta)\right) \right] + p(1-p)2u\left(\tilde{Y}\right) + L_i^{IL} - L_i^{SL} \\
& = p(1-p) \left[2u(\tilde{Y}) - \left[u\left(\tilde{Y} + \frac{1}{2}(\Delta + \theta)\right) + u\left(\tilde{Y} - \frac{1}{2}(\Delta + \theta)\right) \right] \right] + L_i^{IL} - L_i^{SL} > 0
\end{aligned}$$

The first two factors are positive, and the term in square brackets is strictly positive by Jensen's inequality. Since $L_i^{IL} > L_i^{SL}$, the sign of (2)–(1) is strictly positive. In an analogous way, the same holds for the sign of (4)–(3) when comparing T4 to T3. ■

Beyond our framework, other aspects support Hypothesis 2. Sharing the potential fine may increase non-compliance because the per capita size of the loss per declaration is smaller and, therefore, appears individually less threatening. Recent results by Andersson et al. (2016) also support Hypothesis 2. They find more risk-taking for decisions made on the behalf of others.

Evidence for Hypothesis 2 would not question working in teams or flat hierarchies in general, but allude to a more subtle source of non-compliance. Sharing or even fully externalizing the costs of one's misconduct would provide exactly the wrong incentives since it would lead to more non-compliance.

4 Experimental Results

4.1 The report of the own income

We start our analysis with subjects' reports R_i (Treatment 1 and 2) and R_{ii} (Treatment 3 and 4) of their own income Y_i . This measure of compliance is directly comparable across all treatments.⁸ In the following, we focus on potential deceivers, i.e., subjects with a high income Y^h . Only they face a trade-off between reporting honestly or evading the deduction and reporting dishonestly. Subjects with a low income Y^l have no incentive to cheat by overreporting their income as the deduction Δ would be due. Thus, they are excluded for the respective period. This leaves us with 2,166 observations from 268 subjects over 10 periods.

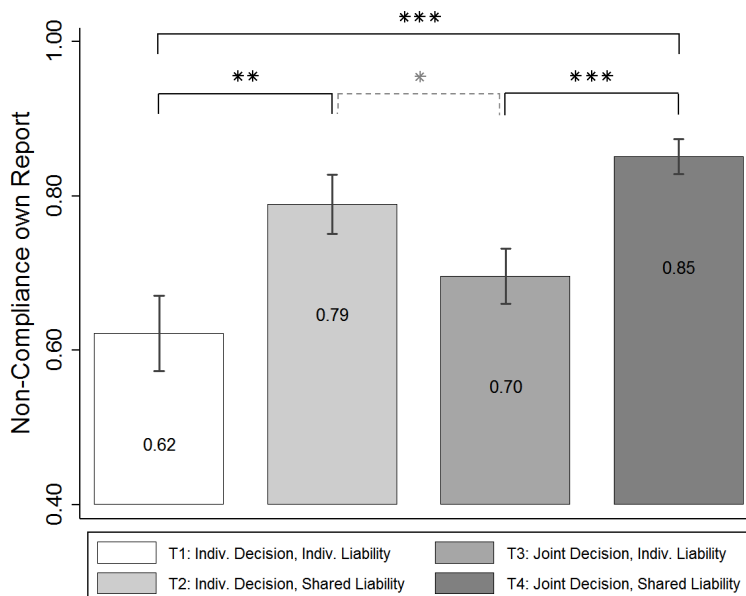


Figure 3: Misreporting of own income by treatment. Error bars display mean \pm SEM. Brackets indicate significant treatment differences, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure 3 displays average non-compliance for the report of the own income in each treatment.⁹

A comparison of Treatment 1 with individual liability and decision-making (far-left bar) to Treat-

⁸ As outlined in section 3.1, the structure of the declaration D_i of income Y_i is different along the decision-making dimension. Therefore, the declaration is discussed in a later stage of the analysis.

⁹ (Non-)Compliance is computed as follows: we first calculate the compliance rate of each subject over the respective periods. This leaves us with one observation per subject for non-parametric tests. A complementary regression analysis accounts for the panel structure of our data instead. Non-Compliance per treatment, as displayed in Figure 3, is the average over individual compliance rates. The same procedure is applied to the subsequent analysis of compliance for the report of the partner's income and the final declaration.

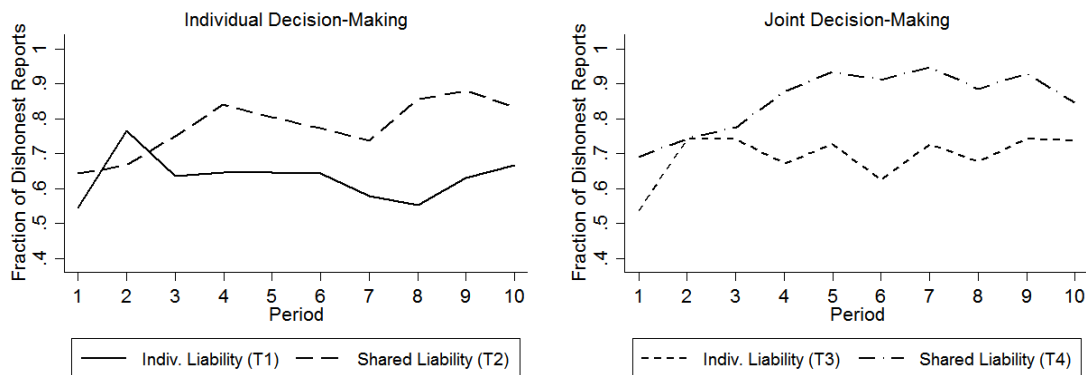


Figure 4: Misreporting of own income over periods (as a fraction of dishonest reports to potential dishonest reports)

ment 4 with shared liability and decision-making (far-right bar) reveals that the fraction of dishonest reports significantly increases from 62 percent to 85 percent (Wilcoxon-Mann-Whitney test: $p < 0.01$). This finding is in line with earlier results of the deception literature that finds more deception for teams/groups (Weisel and Shalvi 2015, Kocher et al. 2017). Exploring the intermediate treatments refines the picture: joint decision-making has only a small effect and increases non-compliance by 7 percentage points on average (from 0.62 to 0.70 in T1 vs. T3, and from 0.79 to 0.85 in T2 vs. T4). These differences are not statistically significant (Wilcoxon-Mann-Whitney test: $p = 0.29$ and $p = 0.41$). In contrast, effects for the shared liability dimension are clear-cut. Moving from individual to shared costs and benefits raises the fraction of dishonest reports significantly by 16 percentage points on average (from 0.62 to 0.79 in T1 vs. T2 and from 0.70 to 0.85 in T3 vs. T4, Wilcoxon-Mann-Whitney test: $p = 0.02$ and $p < 0.01$, respectively). Importantly, this increase in non-compliance is nearly identical along the decision-making dimension. Hence, sharing the economic consequences drives non-compliance independent of whether the decision was made individually or as a team.¹⁰

Figure 4 compares compliance behavior over periods. The left side displays the evolution of non-compliance for individual and shared liability conditional on individual decision-making (Treatment 1 and 2). The right side displays this evolution conditional on joint decision-making (Treatment 3 and 4). The insights from aggregate misreporting above are in line with these observations over

¹⁰As of the variation of both dimensions, a comparison of the difference between Treatment 2 and 3 (Wilcoxon-Mann-Whitney test: $p < 0.08$) is not meaningful and is hard to interpret. Hence, we abstract from this comparison.

time: a 'compliance gap' of 10 to 20 percentage points along the liability dimension emerges from period three onwards. Both graphs exhibit the same pattern over periods. Compliance behavior still fluctuates in the first three periods while the rate of dishonest reports increases in three out of four treatments. Subsequently, the compliance rate converges to a treatment-specific level.¹¹ We cannot observe strong reversals in the compliance behavior or end-round effects. Therefore, compliance behavior is not the outcome of some specific outlier periods but it is a stable pattern. Comparing the evolution in Treatment 1 to Treatment 3 and Treatment 2 to Treatment 4 reveals that the effects for decision-making are weaker. There seems to be no stable 'compliance gap' along this dimension (compare to Figure A1 in the appendix). The graphs also reveal no evidence of an interaction of the liability and the decision-making dimension, since we observe that the liability effect for both individual and team decision-making are fairly similar.

Table 1 confirms the findings in a multivariate analysis. The dependent variable is the misreporting of the own income (coded as a dummy variable). We show effects for a random effects probit regression that takes account of the panel data structure. Column (1) is a baseline specification and includes dimension dummies and a set of basic socio-economic control variables.¹² Joint decision-making has no significant effect on non-compliance, and we cannot confirm Hypothesis 1. In line with Hypothesis 2, shared liability significantly increases the probability of giving a dishonest report. Column (2) indicates the results of a specification that includes an interaction term between both treatment dimensions. Results from column (1) are validated, but we find no evidence of an interaction effect: the coefficient on the interaction term is small and not significant. The average marginal effect of shared liability in both specifications amounts to 12 percentage points and highlights the economic relevance of this dimension, while the non-significant average marginal effects of joint decision-making and the interaction term are smaller (4–5 percentage points and 1 percentage point, respectively). Hence, the two baseline regressions only point toward a negative effect of shared liability on compliance.

¹¹Although we ran pre-experimental practice questions, the most probable explanation is the learning of subjects. Since the shared liability treatments are slightly more complicated than individual liability treatments, convergence takes one to two periods longer. Nevertheless, the between-treatment differences in compliance behavior in period 1 are already broadly in line with following periods, with the difference between T3 and T4 even being significant (χ^2 -Test: $p = 0.08$).

¹²Control variables include gender, age, number of siblings, and a dummy for students of economics and business administration. In line with earlier literature, we find that women are less likely to misreport. Moreover, age is positively correlated to compliance.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Joint Decision	0.233 (0.166)	0.219 (0.229)	0.215 (0.224)	0.203 (0.226)	0.237 (0.230)	0.213 (0.234)
Shared Liability	0.594*** (0.163)	0.575** (0.254)	0.554** (0.248)	0.533** (0.250)	0.612** (0.256)	0.589** (0.259)
Joint Dec. x Shared Liab.		0.0306 (0.328)	0.017 (0.320)	0.035 (0.322)	-0.010 (0.330)	0.0490 (0.335)
Risk Attitude			0.122*** (0.041)			
Cog. Refl. Test				0.243*** (0.073)		
Soc. Val. Ori.					-0.007 (0.006)	
Early Period						-0.439*** (0.079)
Socio-Eco. Controls	YES	YES	YES	YES	YES	YES
Constant	1.202 (0.740)	1.208 (0.743)	0.917 (0.729)	1.025 (0.731)	1.272* (0.744)	1.357* (0.759)
Observations	2,166	2,166	2,166	2,166	2,166	2,166
Number of ID	268	268	268	268	268	268

Notes: The table presents results of a random effects probit specification with dishonest reporting (binary variable) as the dependent variable. "Joint Decision" is a dummy variable that indicates treatments with joint decision-making, "Shared Liability" is a dummy variable that indicates treatments with shared payoffs, and "Joint Dec. x Shared Liab." is the interaction between both dummies. The reference group is Treatment 1 with individual payoffs and individual decision-making. "Risk Attitude" refers to the number of risky decisions (0 - 10) in Holt and Laury (2002). "Cog. Refl. Test" is the number of correct answers (0-3) in the Cognitive Reflection Test (Frederick 2005). "Soc. Val. Ori." indicates Murphy's et al. (2011) measure of distributional preferences (-16.26 - 63.39). "Early Period" is a dummy variable for periods 1 - 3. Standard errors are given in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 1: Multivariate analysis of dishonest reports of the own income

The alternative specifications in Table 1 (columns 3 to 6) include post-experimental tests and further robustness checks. Importantly, the main results from the baseline specifications (1) and (2) are robust. Since the audit introduces a risk dimension in the compliance decision, we control for subjects' risk preferences (Holt and Laury 2002) in column (3). In line with the previous literature, we find that being less risk-averse (i.e., choosing the risky lottery more often) is positively correlated to higher non-compliance. Cognitive ability, as measured by Frederick (2005), might influence the compliance decision as more impulsive subjects may over- or underestimate potential costs. We find that reflective thinking is positively correlated to a higher probability of giving a dishonest report (column (4)). As outlined in section 3.1, social preferences might also alter the compliance

decision. In column (5), social preferences as measured by the Social Value Orientation (Murphy et al. 2011) are not significantly related to the compliance decision. A final robustness check includes a dummy for the first three periods (compare to the discussion of Figure 4). The negative and highly significant coefficient in column (6) indicates that subjects are less dishonest in early periods, which could be due to the initial learning of the game.¹³ However, controlling for early periods has no influence on our main results. We summarize the findings above in

Result 1 *For the report of the own income*

- (i) *joint decision-making has no significant effect on the compliance rate;*
- (ii) *shared liability significantly increases non-compliance.*

4.2 The report of the partner’s income and the final declaration

Besides the report for their own income R_{ii} , subjects are required to make a report R_{ij} for their partner’s income Y_j in treatments with joint decision-making (Treatment 3 and 4). The procedure leaves us with two reports per subject in these treatments: a report for the own income and a report for the income of their partner. Each income report is then matched to the report of the partner for the respective income. If both reports match, the report is implemented as the declaration of a subject’s income. Vice versa, if the reports do not match, a truthful declaration is automatically applied. A dishonest declaration therefore requires some coordination within a team.

Figure 5 displays the non-compliance rate for the reports on the partner’s income. Aggregate non-compliance is comparable to the report of the own income. In Treatment 3 with joint decision-making and individual liability, 70 percent of reports for the partner’s income are dishonest. This fraction increases to 84 percent in Treatment 4 with joint decision-making and shared liability (Wilcoxon-Mann-Whitney test: $p < 0.01$).¹⁴ Moreover, the vast majority of teams successfully coordinate on the reports. Reports match with 95 percent probability in Treatment 3 and with 92 percent probability in Treatment 4. In the few cases of non-matching reports, a subject’s partner

¹³As an additional robustness check, we include lagged audits in our regression analysis. In line with earlier literature (Mittone 2006), we find that lagged audits on the declaration of the own income increase non-compliance ($p < 0.01$). In contrast, we find no significant effect for lagged audits on the declaration of the partner’s income ($p = 0.37$).

¹⁴The (nearly) perfect match of compliance for the report of the partner’s income (Figure 5) with compliance for the report of the own income (Figure 3) is not explained by subjects necessarily choosing the same report for their own income R_{ii} and the income of their partner R_{ij} . Only 72 percent do so in Treatment 3, but 91 percent do so in Treatment 4.

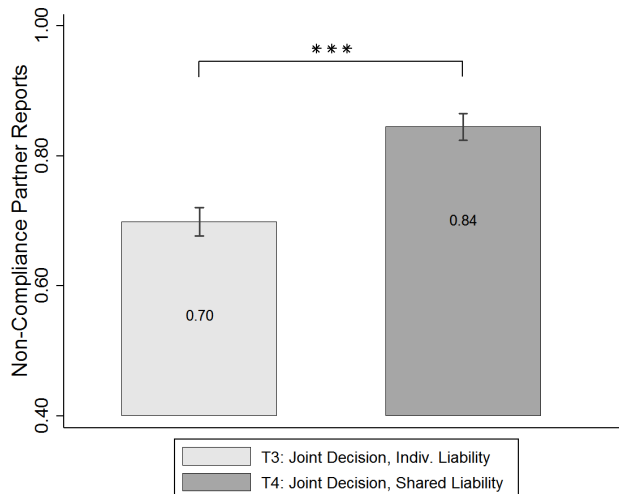


Figure 5: Misreporting of partner’s income. Error bars display mean \pm SEM. Brackets indicate significant treatment differences, *** $p < 0.01$.

does not systematically deviate in one direction. The probability of a subject making a dishonest report and her partner making an honest report on the subject’s income is roughly comparable to the opposite case of a subject making an honest report and her partner making a dishonest report on the respective income.

Result 2 *Subjects are willing to make a dishonest report on behalf of their partner, even if they are not involved in the economic consequences or are honest about their own income report.*

An analysis of the final declaration of incomes reveals that the fraction of dishonest declarations is almost identical to the fraction of dishonest reports. For treatments with individual decision-making, the fraction of dishonest reports equals the fraction of dishonest declarations by design. For treatments with joint decision-making, the declaration mechanism itself may increase compliance: if subjects disagree and state different reports for an income, the mechanism automatically leads to a truthful declaration. Due to high coordination within teams, the tendency toward compliance is very moderate in size but decreases the effect of joint decision-making even further. Non-compliance for declarations amounts to 67 percent in Treatment 3 and 81 percent in Treatment 4 (as compared to 70 percent and 85 percent dishonest reports, respectively).¹⁵ The results of a multivariate

¹⁵ A Wilcoxon-Mann-Whitney test confirms the earlier findings for the report of the own income: the differences along the decision-making dimension are not significant (T1 vs. T3 and T2 vs. T4, $p = 0.63$ and $p = 0.54$), while joint liability significantly increases non-compliance (T1 vs. T2 and T3 vs. T4, $p = 0.02$ and $p < 0.01$).

analysis on the reporting behavior for the declarations are in line with the results on the reporting behavior for the reports (compare to Table A.2 in the appendix). We summarize our findings in

Result 3 *Subjects' declaration behavior parallels the previous findings for their reporting behavior.*

4.3 The declaration behavior within teams

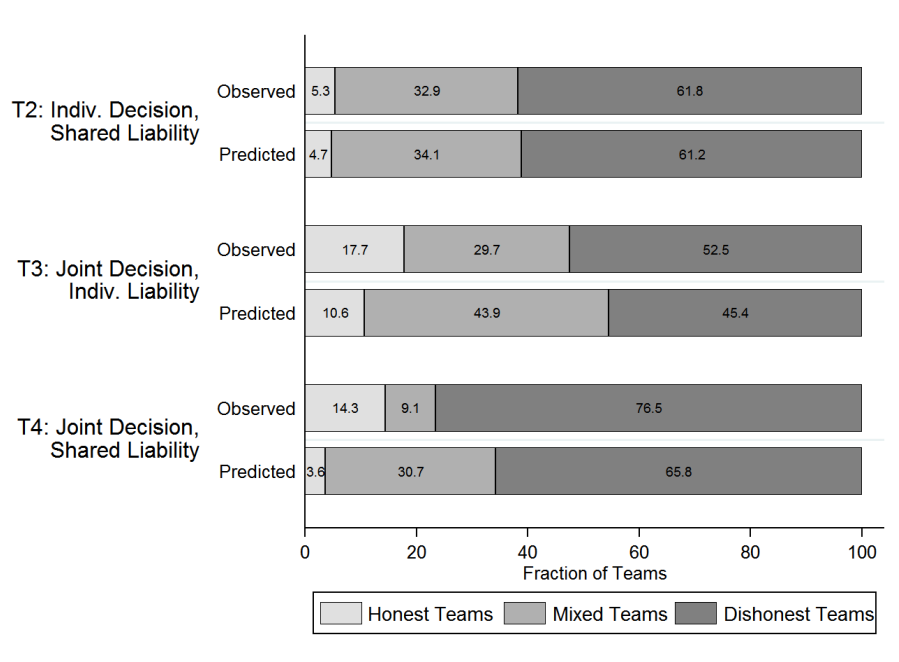


Figure 6: Composition of declarations within teams

Although joint decision-making has no effect on the aggregate compliance rate, it does influence the composition of declarations within teams. Figure 6 summarizes the two declarations of a team in Treatment 2 to 4.¹⁶ For each treatment, the upper bar ('Observed') displays the actual composition of declarations as observed in our data. On the lower bar ('Predicted'), we display a pseudo-prediction on the composition of declarations. This prediction assumes that coordination is not possible, and consequently, there are no behavioral spillovers between the team members. The composition of declarations follows a simple calculation based on the observed aggregate compliance rate in the respective treatment.¹⁷ Each bar is interpreted as follows: The slightly gray shaded area

¹⁶Since Treatment 1 (individual decision and liability) has no team characteristics, we do not include this treatment here. If the same matching protocol as for the other treatments is applied, the results for artificial teams read as follows: 14.0 percent of teams are classified as honest, 46.3 percent of teams are classified as mixed, and 39.7 percent of teams are classified as dishonest.

¹⁷In absence of coordination, the predicted fraction of honest teams is the square of the compliance rate (1 subtracted

on the very left ('honest teams') refers to the fraction of teams with two honest declarations, the one in the middle ('mixed teams') refers to the fraction of teams with one honest and one dishonest declaration, and the dark gray area on the far right ('dishonest teams') refers to the fraction of teams with two dishonest declarations.

A comparison of the observed and the predicted composition of declarations allows for detailed insights on the impact of joint decision-making. In Treatment 2, decision-making occurs individually and coordination among team members is not possible. The distribution of honest, mixed and dishonest teams closely follows the theoretically expected distribution based on the aggregate compliance rate. In contrast, both treatments with joint decision-making are characterized by a medium to high correlation of declarations within teams. Consequently, individual declarations of team members are not correlated.¹⁸ The positive correlations suggest that joint decision-making is marked by spillovers between team members which results in more consequential declaration behavior. In comparison to the predicted composition in absence of coordination, teams choose more often to be honest or to be dishonest for both declarations, but less often the intermediate case of mixed declarations. For Treatment 3 with joint decision-making and individual liability, the observed fraction of mixed declarations is 14 percentage points lower than predicted (29.7 percent vs. 43.9 percent), while both the fraction of two honest or two dishonest declarations is higher than predicted. For Treatment 4 with joint decision-making and shared liability, this effect is even more pronounced and the case of mixed declarations is almost crowded-out completely (9.1 percent vs. 30.7 percent). This pattern is also evident from a direct comparison of the observed composition of declarations between Treatment 2 and Treatment 4: moving from individual to joint decision-making leads to a reduction of mixed declarations, and a corresponding increase in mainly dishonest but also some honest declarations.¹⁹ Ex ante, one might rather expect a crowding-out of

by the non-compliance rate), the predicted fraction of mixed teams is two times the compliance rate multiplied by the non-compliance rate, and the predicted fraction of dishonest teams is the square of the non-compliance rate. By illustration, the calculation for Treatment 2 is the following: fraction of honest teams is $(1 - 0.79)^2 = 0.04$; fraction of mixed teams is $2 \cdot (1 - 0.79) \cdot 0.79 = 0.33$; fraction of dishonest teams is $0.79^2 = 0.62$.

¹⁸We measure the correlation of declarations by the phi coefficient, which has a similar interpretation to the Person correlation coefficient but is suited for binary variables. Calculation is based on a 2x2 contingency table of pairs of declarations in teams. Rows reflect the declaration of the own income and columns the declaration of the partner's income. A correlation of zero results when the probability in each of the four cells (honest, mixed and dishonest teams) equals to the product of marginal probabilities (the aggregate (non-)compliance rate). While the phi coefficient is not different from zero in Treatment 2 (corr=0.04, χ^2 -statistic: $p = 0.60$), the phi coefficient is positive in Treatment 3 and 4 (corr = 0.32 and corr = 0.70, χ^2 -statistic: $p < 0.01$, respectively).

¹⁹In contrast, a comparison of Treatment 3 and Treatment 4 reveals that the effect of shared liability is unidirectional: the decrease in the fraction of mixed teams leads to an increase in dishonest teams only.

the case of two honest declarations, as the coordination on mixed declarations might be attractive for risk-averse subjects in this treatment. Instead, subjects seem to adapt to their team partner’s compliance behavior.

In summary, the compliance behavior under joint decision-making is potentially two-directional: some teams may be less compliant, and some teams may be more compliant. As the magnitude of both directions is roughly comparable and cancels out, this might explain why joint decision-making has no effect on the aggregate compliance rate.

Result 4 *Joint decision-making is characterized by behavioral spillovers between team members. As a consequence, teams are prone to be either honest or dishonest.*

4.4 Evaluation of the team chats

The analysis of chat protocols allows for further insights into team decision-making in Treatment 3 and 4.²⁰ Less than 2 percent of teams abstain from using the chat. Usage of the chat leads to a successful coordination on the mutual income reports (99 percent in Treatment 3 and 96 percent in Treatment 4). The finding is in line with the low rate of non-matching reports (compare to section 4.2). In half of all chats, the chat is not only a mere instrument for coordination (e.g., typing in a single number as a proposed report), but it is also used for the exchange of arguments in favor or against non-compliance. In early periods this exchange of arguments is slightly more pronounced (57 percent) as compared to final periods (46 percent).

Conditional on subjects giving an honest or dishonest report, Figure 7 provides an overview of the frequency with which a particular argument is mentioned in the chat.²¹ “Money” refers to the frequency with which the monetary consequences of a report are addressed, while “Honesty” indicates discussions on being truthful. “Rules” refers to the experimental instructions, in particular that subjects were supposed to report their actual income. “Other” indicates discussions about the compliance behavior of other subjects, while “Prev. Audit” captures the experience with audits in

²⁰Inspired by Kocher et al. (2017), the evaluation of chats was carried out independently by five research assistants according to an analysis protocol (available upon request). For each chat, we take the median evaluation of graders.

²¹We focus on subjects with a high income that used the chat. For honest reporting, this leaves us with 189 observations in Treatment 3 and 91 observations in Treatment 4. For dishonest reporting, we have 439 observations in Treatment 3 and 551 observations in Treatment 4. There might be zero, one or multiple arguments per chat (hence, frequencies do not add up to 1), and some statements may fall into more than one category. This procedure allows for a direct assessment of the importance of arguments in the respective four cases.

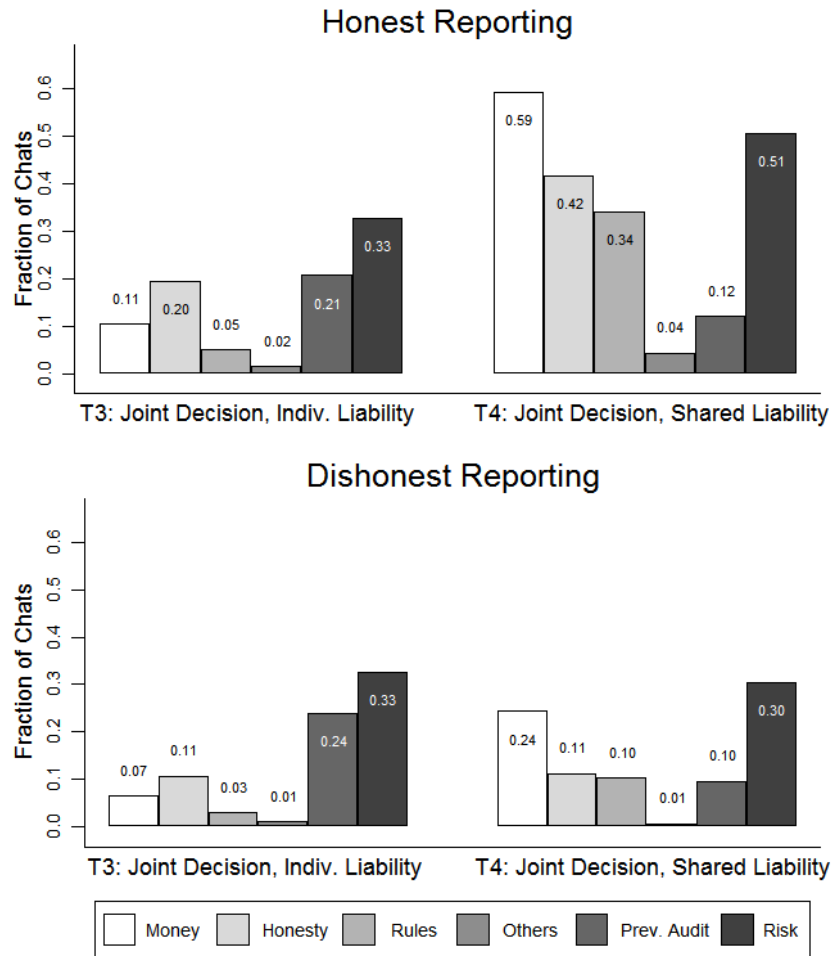


Figure 7: Frequency of arguments mentioned in chat

previous periods. Finally, “Risk” refers to the risk dimension of the compliance task.

A comparison between honest and dishonest reporting reveals that arguments are used more often for honest reporting. In quantitative terms, the risk dimension is the most important argument,²² followed by the monetary consequences of a report, the concept of being honest, and experience with previous audits. Adherence to rules and the behavior of other subjects are mentioned only in a few cases (except for honest subjects in Treatment 4). Overall, the monetary consequences of a report and the adherence to rules are more prominent as arguments for shared liability (Treatment 4), while the reference to previous audits is more pronounced for individual liability (Treatment 3). For dishonest reporting, we do not observe strong treatment differences, while there are evident differences for honest reporting. Both in terms of the frequency and the distribution of arguments, honest reporting in Treatment 4 is different compared to the other cases and seems to be the most controversial option. Reasoning relies more heavily not only on monetary consequences, but also on the concept of being honest and on compliance with the rules.

As we partially focus on categories of arguments comparable to Kocher et al. (2017), we are able to relate our compliance task to their computerized deception problem. This gives insights into the similarities and differences of compliance and deception problems. We find differences mainly in two aspects: first, the frequency of argumentation is higher for honest reporting (and not for dishonest reporting) in the compliance task. Second, the most frequent argument for dishonest reporting is the risk dimension, as compared to the monetary consequences in the deception problem. However, monetary consequences are also relevant for the compliance problem, though mostly for honest reporting. Moreover, both the discussion of monetary consequences and of risk considerations are related to the economic consequences of a report, and therefore indicate a similarity rather than a difference between both problems. Both settings share that the concept of being honest is identified for an intermediate number of chats. Adherence to rules and the behavior of others are only of minor importance for both settings. In summary, the introduction of an audit shifts attention toward the risk dimension and the audit itself, while other motivations, such as the monetary consequences or the concept of being honest, remain important for the compliance task.

²²We do not find a correlation between the frequency of the argument “Risk” in the chat and risk aversion as measured by Holt and Laury (2002). The possibility of playing a mixed lottery by making one honest and one dishonest declaration is only mentioned in 8 percent of the chats in Treatment 4, indicating the limited importance of this option.

Since the design of this study does not focus on the investigation of gender differences, we abstract from gender effects in our main analysis (for an explicit study of gender differences, see Lohse and Qari 2019, Muehlheusser et al. 2015). Concerning the composition of declarations within teams (previous section), female and mixed teams tend to be more honest on average, but exhibit a similar behavioral pattern as compared to male teams. An exception is Treatment 3 with less honest pairs of declarations in male teams, which might indicate that individual liability rules are not as effective for male teams. In the chat, male teams tend to exchange more arguments and the argumentation relies more heavily on the risk dimension and the monetary consequences.

Result 5 *The vast majority of subjects use the chat to coordinate on the mutual reports. The most important arguments are the risk dimension of the compliance task, monetary consequences, and the concept of being honest.*

5 Discussion

Joint decision-making in our setting is characterized by two main components: first, a coordination phase that enables the decision-formation via the chat, and second, the implementation of the decision via the mutual reporting of incomes. Both components are crucial: in the absence of mutual reporting, the coordination phase lacks the need to reach an agreement. Essentially, it is reduced to cheap talk and allows at most for an update on the other’s behavior. In the absence of any prior coordination, simultaneous mutual reporting may lead to non-matching reports by design. Coordination could also be feasible without chat, e.g., by the mutual signaling of reporting preferences or by sequential reporting. However, such coordination mechanisms leave subjects with very little influence on the partner and might not reflect joint decision-making appropriately. Finally, we ensure that a dishonest declaration is only feasible if both team partners agree on being non-compliant. This reflects situations outside the laboratory in which honest subjects are in a better bargaining position since an honest report is typically a feasible outside option.

The small insignificant increase of dishonest reporting along the decision-making dimension is good news and bad news at the same time. On the one hand, the deception literature’s doom assessment of team behavior as rather deceptive does not entirely carry over to the more complex framework of compliance. While the rules of the decision situation and the possibility to deceive

remain somewhat ambiguous in the standard settings used to study deception (such as the die role paradigm or the mind game), available strategies and potential consequences are clearly specified for the compliance task. Hence, the effect of higher cognitive skills in teams as well as the exchange of information within teams might only play a minor role. Although team decision-making might also allow us to sustain a positive self-image or reputation toward others by putting forward the excuse of having been convinced - or even having been put under pressure - by the team partner, delegating the (moral) responsibility for a dishonest report is only partially feasible. After the chat, each team member makes the two reports on her own and does not face direct peer pressure in the reporting situation. Moreover, honest reporting is always a costless outside option in monetary terms. In Treatment 3, partners do not benefit from dishonest reporting on behalf of their team partner and disguising the responsibility for a dishonest report of one's own income is particularly difficult. Finally, the role of intrinsic lying costs or social norms might be less pronounced in compliance tasks due to the presence of an audit. The reduction/erosion of these costs/norms in teams, as found in mere deception problems, might therefore have weaker implications. This is in line with the fairly stable compliance behavior over time.

On the other hand, a negative implication is that third-party reporting seems to be ineffective in reducing dishonest behavior when agents have the possibility to coordinate beforehand. The reporting mechanism in the joint decision treatments is a form of third-party reporting: successful dishonest declarations require not only a silent approval or non-objection, but an active statement by the team member. This condition leads neither to a large fraction of non-matching reports nor does it prevent successful evasion. In Treatment 3 with individual liability, a potential explanation might be reciprocal behavior: if both partners agree to mutually falsely report low incomes they may both benefit. However, we find that 30 percent of teams coordinate on one honest and one dishonest declaration. Although honest subjects are not dependent on their partner to achieve a truthful declaration, a substantial fraction of subjects are willing to actively assist their partner in being non-compliant. As long as honest subjects are not negatively affected by the non-compliance of their partners, they seem to be more committed to their partner than to being honest. Potential reasons include that honest subjects are afraid of disappointing their team partner or that breaking the agreement of giving a dishonest report on behalf of the partner outweighs the moral cost of making the dishonest report itself.

Unlike for joint decision-making, our results on shared liability are clear-cut: independent of the decision-making dimension, sharing the costs and benefits of a dishonest report decreases compliance. In quantitative terms, this dimension accounts for over two thirds of the difference in the compliance rate between Treatment 1 and Treatment 4. An explanation is that the concept of “White Lies” (Erat and Gneezy 2012) transfers to compliance problems, i.e., subjects are more likely to not comply if it benefits not only themselves on expectation but also their partners. The magnitude of the effect would be in line with previous findings from the deception literature.²³ In addition, despite sharing both the potential gains and losses from the compliance decision, the effect for the loss domain might be more pronounced and shared fines in turn less of a deterrent. Moreover, the more condensed distribution of final earnings for shared liability might induce subjects to not comply more often, although the expected value of non-compliance is the same under individual and shared liability. Another potential explanation is the reduction of the social risks related to non-compliance. The perfect correlation of final earnings of team partners might reduce social regret (Cooper and Rege 2011) or disutility from social comparisons (Friedl et al. 2014, Lahno and Serra-Garcia 2015).

Although shared liability increases non-compliance independent of the decision-making dimension, team decision-making seems to crowd out the intermediate case of mixed (one honest and one dishonest) declarations. This indicates that those honest subjects who support the non-compliance of their partner for individual liability might be more prone toward dishonesty for shared liability themselves. For example, their dishonest partners might convince them to deceive (dishonest partners have an incentive to do so for shared liability), or they might turn dishonest since sharing the consequences of the declaration of their own and their partner’s income reduces the risk on their final earnings. Therefore, at least for settings with a positive expected value of non-compliance, the combination of shared liability and joint decision-making is detrimental.

Implications arise to the design of the internal structures of companies or organizations. Importantly, team decision-making per se is not the major source of non-compliance in group settings. Hence, there is no need for firms to forgo the benefits of team work in order to induce more honest

²³Since the possibility of an audit in our setting is a major difference to the deception problem, a direct comparison should be interpreted cautiously. Erat and Gneezy (2012) find an increase of cheating of 13 percentage points (treatment T[10,0] vs. T[10,10]), while Wiltermuth (2011) reports an increase between 10 and 15 percentage points (treatment self-alone vs. self-and-other, study 2 and 3). Conrads et al. (2013) show that the average reported number in a dice-rolling experiment is inflated by 17 percent for teams (treatment Individual vs. Team).

decisions by its employees. Instead, our results suggest that it is sufficient to set liability rules right, meaning that each team member could be fully rather than just partially economically liable. This should deter misconduct *ex ante*. In the case of the Volkswagen emission scandal, for example, full individual liability might have prevented engineers and managers from committing their offenses in the first place. The recommendation of full individual liability does not only hold for teams, but also for compliance decisions by single employees that have potentially large external effects on the company: our findings from Treatment 2 show that shared liability substantially increases non-compliance even under individual decision-making. Famous rogue traders Nick Leeson (Barings Bank) and Jérôme Kerviel (Société Générale), for example, caused huge losses (and, in Leeson’s case, the collapse of the bank) in an attempt to increase company profits through trading activities which violated internal regulations. Importantly, while acting on their own, successful non-compliance would have been individually advantageous for them as well as for the entire company. However, these cases also show the limits of individual liability rules *ex post*. The losses caused by Leeson or Kerviel outweighed their individual economic capability by far. Ultimately, such settings call for tight internal monitoring of employees’ activities and external prosecution of misconduct for which the culprit could even face imprisonment. For other applications, such as joint tax evasion, individual liability is potentially a feasible policy: fines for unreported employment or evaded sales taxes might be imposed on either party to the full extent, and thereby deployed as an effective deterrence mechanism. In this sense, the United Kingdom issued new regulations to fight corporate tax evasion with shared liability of the involved parties in late 2017. In order to comply with the “Facilitation of Tax Evasion Offences,” firms have to assess risk across their business, including suppliers and customers. As reported by Bloomberg BNA (2017), British tax authorities argued that especially large corporations “can put pressure to comply on anyone looking to join their supply chain, requiring compliance up front before they can enter.” Companies that fail to prevent the facilitation of tax evasion are threatened with unlimited fines and a criminal record.

6 Conclusion

Team decisions are widespread, and firms heavily rely on team work. However, recent corporate scandals suggest that team decisions may also have a dark side: teams seem to be more dishonest than individual decision-makers. In this paper, we study the compliance decision of dyads as the smallest form of teams in a laboratory experiment. We ask whether the compliance behavior of teams is different from the compliance behavior of individuals and identify the effect of key differences to individual decisions. The first dimension of our 2-by-2 between-subjects design varies whether the compliance decision is made individually or in a team. Teams coordinate via a chat and each member subsequently reports both the own income and the income of the team member. The second dimension varies whether the monetary consequences of the (non-)compliance decision, in particular gains from undetected non-compliance and losses in the case of an audit, accrue only to the respective subject alone or if they are shared between both team members. This allows us to answer an even more important question, namely how non-compliance by teams can effectively be deterred. Specifically, we ask under which conditions an auditing mechanism is suited to address non-compliance by teams.

We find significantly more dishonesty in dyads as compared to individuals. This effect can mainly be attributed to the liability dimension: independent of whether the compliance decision is made on an individual basis or in teams, shared economic consequences among team members lead to significantly increased non-compliance. In contrast, the impact of team decision-making is small and we find no significant effect along this dimension. Hence, the increased non-compliance of teams is explained by shared liability rather than the team decision-making. Further results indicate that decisions in teams are characterized by a low rate of disagreement in terms of non-matching reports and a medium to high correlation of reporting behavior. Most important motivations for (non-)compliance include the risk dimension of the compliance task and the monetary consequences of a report. Our findings are corroborated by different robustness checks. An implication of these insights for firms or organizations in general, is that team structures do not necessarily lead to more non-compliance. Corporate misconduct as with Volkswagen in the diesel emission scandal can not necessarily be explained by group decision-making per se. Rather, spreading the economic consequences among several engineers and managers seems to be the cause. In order to restore

compliance within teams, it is advisable to prohibit sharing of consequences but to place an emphasis on the full liability of each individual team member.

Acknowledgments

We thank Jana Cahlíková, Zhixin Dai, Enrico Diecidue, Christoph Engel, Laszlo Goerke, Marco Kleine, Martin Kocher, Kai Konrad, Nadja Müller, Daniele Nosenzo, Klaus Schmidt, Simeon Schudy, Lisa Spantig, Matthias Sutter, Marie Claire Villeval, Ori Weisel, participants of the 2018 Thurgau Experimental Economics Meeting in Kreuzlingen, the fifth International Meeting on Experimental and Behavioral Social Sciences (IMEBESS) in Florence, the 2018 ESA World Meeting in Berlin, the ComplianceNet First Annual Conference 2018 in Irvine, the 2018 PET Conference in Hue, the 2018 IIPF Congress in Tampere, the 2018 annual congress of the German Economic Association in Freiburg, workshops in Bonn, Munich, Rome and Trier and our colleagues at the MPI and the Berlin School of Economics and Law for helpful comments. Moreover, we thank the econlab in Munich for providing the laboratory resources.

References

- [1] Abeler, Johannes, Anke Becker, and Armin Falk, 2014, Representative evidence on lying costs, *Journal of Public Economics* 113, 96–104.
- [2] Abeler, Johannes, Armin Falk, Lorenz Goette, and David Huffman, 2011, Reference points and effort provision, *American Economic Review* 101(2), 470–492.
- [3] Abraham, Martin, Kerstin Lorek, Friedemann Richter, and Matthias Wrede, 2017, Collusive tax evasion and social norms, *International Tax and Public Finance* 24(2), 179–197.
- [4] Alm, James, Kim M. Bloomquist, and Michael McKee, 2017, When You Know Your Neighbour Pays Taxes: Information, Peer Effects and Tax Compliance, *Fiscal Studies* 38(4), 587–613.
- [5] Alm, James, Betty R. Jackson, and Michael McKee, 2009, Getting the word out: Enforcement information dissemination and compliance behavior, *Journal of Public Economics* 93(3), 392–402.

- [6] Andersson, Ola, Hakan Holm, Jean-Robert Tyran, and Erik Wengström, 2016, Deciding for others reduces loss aversion, *Management Science* 62(1), 29–36.
- [7] Baker, Ronald J., Susan K. Laury, and Arlington W. Williams, 2008, Comparing small-group and individual behavior in lottery-choice experiments, *Southern Economic Journal* 72, 367–382.
- [8] Barr, Abigail, and Georgia Michailidou, 2017, Complicity without connection or communication, *Journal of Economic Behavior & Organization* 142, 1–10.
- [9] Biziou-van-Pol, Laura, Jana Haenen, Arianna, Novaro, Andrés Occhipinti Liberman, and Valerio Capraro, 2015, Does telling white lies signal pro-social preferences?, *Judgment and Decision Making* 10, 538–548.
- [10] Bloomberg BNA, 2017, Here’s How the U.K. Is Using Big Companies to Fight Tax Evasion, <https://www.bna.com/heres-uk-using-n73014470233/> [January 31, 2018].
- [11] Bornstein, Gary, Tamar Kugler, and Anthony Ziegelmeyer, 2004, Individual and group decisions in the centipede game: Are groups more “rational” players? *Journal of Experimental Social Psychology* 40(5), 599–605.
- [12] Cason, Timothy N., Lana Friesen, and Lata Gangadharan, 2016, Regulatory performance of audit tournaments and compliance observability, *European Economic Review* 85, 288–306.
- [13] Charness, Gary, Edi Karni, and Dan Levin, 2007, Individual and group decision making under risk: An experimental study of Bayesian updating and violations of first-order stochastic dominance, *Journal of Risk and Uncertainty* 35(2), 129–148.
- [14] Chytilova, Julie, and Vaclav Korbek, 2014, Individual and group cheating behavior: a field experiment with adolescents. No. 06/2014. IES Working Paper.
- [15] Cohen, Taya R., Brian C. Gunia, Sun Young Kim-Jun, and J. Keith Murnighan, 2009, Do groups lie more than individuals? Honesty and deception as a function of strategic self-interest, *Journal of Experimental Social Psychology* 45(6), 1321–1324.

- [16] Conrads, Julian, Bernd Irlenbusch, Rainer M. Rilke, and Gari Walkowitz, 2013, Lying and team incentives, *Journal of Economic Psychology* 34, 1–7.
- [17] Cooper, David J., and Mari Rege, 2011, Misery loves company: Social regret and social interaction effects in choices under risk and uncertainty, *Games and Economic Behavior* 73(1), 91-110.
- [18] Coricelli, Giorgio, Mateus Joffily, Claude Montmarquette, and Marie Claire Villeval, 2010, Cheating, emotions, and rationality: an experiment on tax evasion, *Experimental Economics* 13(2), 226–247.
- [19] Erat, Sanjiv, and Uri Gneezy, 2012, White lies, *Management Science* 58(4), 723–733.
- [20] Feri, Francesco, Bernd Irlenbusch, and Matthias Sutter, 2010, Gains from team-based coordination—arge-scale experimental evidence, *American Economic Review* 100(4), 1892–1912.
- [21] Fischbacher, Urs, 2007, z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10(2), 171–178.
- [22] Fischbacher, Urs, and Franziska Föllmi-Heusi, 2013, Lies in disguise—an experimental study on cheating, *Journal of the European Economic Association* 11(3), 525–547.
- [23] Fortin, Bernard, Guy Lacroix, and Marie-Claire Villeval, 2007, Tax evasion and social interactions, *Journal of Public Economics* 91(11-12), 2089–2112.
- [24] Frederick, Shane, 2005, Cognitve Reflection and Decision Making, *Journal of Economic Perspectives* 19(4), 25–42.
- [25] Friedl, Andreas, Katharina Lima De Miranda, and Ulrich Schmidt, 2014, Insurance demand and social comparison: An experimental analysis, *Journal of Risk and Uncertainty* 48(2), 97-109.
- [26] Gilpatric, Scott M., Christian A. Vossler, and Michael McKee, 2011, Regulatory enforcement with competitive endogenous audit mechanisms, *The RAND Journal of Economics* 42(2), 292–312.

- [27] Gino, Francesca, Shahar Ayal, and Dan Ariely, 2009, Contagion and differentiation in unethical behavior: The effect of one bad apple on the barrel, *Psychological Science* 20(3), 393–398.
- [28] Gino, Francesca, Shahar Ayal, and Dan Ariely, 2013, Self-serving altruism? The lure of unethical actions that benefit others, *Journal of Economic Behavior & Organization* 93, 285–292.
- [29] Gneezy, Uri, Agne Kajackaite, and Joel Sobel, 2018, Lying Aversion and the Size of the Lie, *American Economic Review* 108(2), 419–453.
- [30] Greiner, Ben, 2015, Subject pool recruitment procedures: organizing experiments with ORSEE, *Journal of the Economics Science Association* 1(1), 114–125.
- [31] Holt, Charles A., and Susan K. Laury, 2002, Risk aversion and incentive effects, *American Economic Review* 92(5), 1644–1655.
- [32] Kocher, Martin G., and Matthias Sutter, 2005, The decision maker matters: Individual versus group behaviour in experimental beauty-contest games, *Economic Journal* 115, 200–223.
- [33] Kocher, Martin G., Simeon Schudy, and Lisa Spantig, 2017, I lie? We lie! Why? Experimental evidence on a dishonesty shift in groups, *Management Science* 64(9), 3995–4008.
- [34] Konrad, Kai A., Tim Lohse, and Salmai Qari, 2017, Compliance with Endogenous Audit Probabilities, *Scandinavian Journal of Economics* 119, 821–850.
- [35] Kugler, Tamar, Gary Bornstein, Martin G. Kocher, and Matthias Sutter, 2007, Trust between individuals and groups: Groups are less trusting than individuals but just as trustworthy, *Journal of Economic Psychology* 28(6), 646–657.
- [36] Lahno, Amrei M., and Marta Serra-Garcia, 2015, Peer effects in risk taking: Envy or conformity?, *Journal of Risk and Uncertainty* 50(1), 73–95.
- [37] Laury, Susan, and Sally Wallace, 2005, Confidentiality and taxpayer compliance, *National Tax Journal* 58(3) 427–438.
- [38] Lohse, Tim, and Salmai Qari, 2019, Gender differences in face-to-face deceptive behavior. *Mimeo*.

- [39] Masclet, David, Natalie Colombier, Laurent Denant-Boemont, and Youenn Lohéac, 2009, Group and individual risk preferences: A lottery-choice experiment with self-employed and salaried workers, *Journal of Economic Behavior & Organization* 70, 470–484.
- [40] Mittone, Luigi, 2006, Dynamic behaviour in tax evasion: An experimental approach, *Journal of Socio-Economics* 35, 813–835.
- [41] Muehlheusser, Gerd, Andreas Roeder, and Niklas Wallmeier, 2015, Gender differences in honesty: Groups versus individuals, *Economics Letters* 128, 25–29.
- [42] Murphy, Ryan O., Kurt A. Ackermann, and Michel J. J. Handgraaf, 2011, Measuring social value orientation, *Judgement and Decision Making* 6(8), 771–781.
- [43] Shupp, Robert S., and Arlington W. Williams, 2008, Risk preference differentials of small groups and individuals, *The Economic Journal* 118, 258–283.
- [44] Sutter, Matthias, 2009, Deception through telling the truth?! Experimental evidence from individuals and teams, *Economic Journal* 119, 47–60.
- [45] Sutter, Matthias, Simon Czermak, and Francesco Feri, 2013, Strategic sophistication of individuals and teams. Experimental evidence, *European Economic Review* 64, 395–410.
- [46] Weisel, Ori, and Shaul Shalvi, 2015, The collaborative roots of corruption, *Proceedings of the National Academy of Sciences* 112(34), 10651–10656.
- [47] Wiltermuth, Scott S, 2011, Cheating more when the spoils are split, *Organizational Behavior and Human Decision Processes* 115(2), 157–168.

Appendix A.1

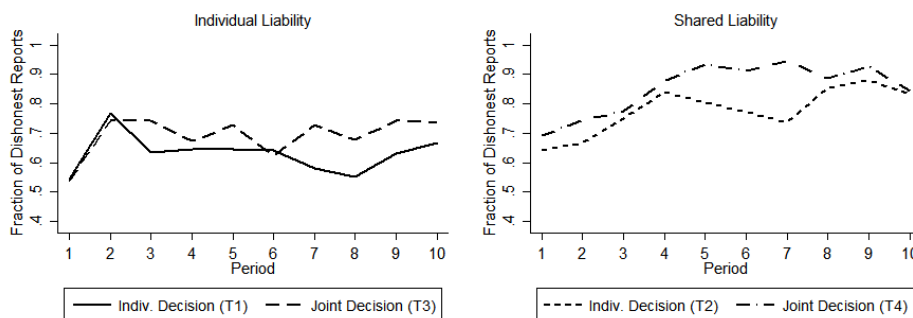


Figure A.1: Misreporting of own income over periods conditional on individual liability (left side) and shared liability (right side) as a fraction of dishonest reports to potential dishonest reports

Appendix A.2

VARIABLES	(1)	(2)	(3)	(4)
Joint Decision	0.024 (0.147)	0.074 (0.204)	0.069 (0.197)	0.064 (0.208)
Shared Liability	0.497*** (0.143)	0.560** (0.228)	0.540** (0.222)	0.575** (0.233)
Joint Dec. x Shared Liab.		-0.104 (0.291)	-0.136 (0.282)	-0.090 (0.297)
Early Period				-0.462*** (0.075)
Post Tests	NO	NO	YES	NO
Socio-Eco. Controls	YES	YES	YES	YES
Constant	1.082* (0.656)	1.060 (0.659)	0.708 (0.644)	1.217* (0.674)
Observations	2,166	2,166	2,166	2,166
Number of ID	268	268	268	268

Notes: The table presents results of a random effects probit specification with dishonest declarations (binary variable) as the dependent variable. "Joint Decision" is a dummy variable that indicates treatments with joint decision-making, "Shared Liability" is a dummy variable that indicates treatments with shared payoffs, and "Joint Dec. x Shared Liab." is the interaction between both dummies. The reference group is Treatment 1 with individual payoffs and individual decision-making. "Early Period" is a dummy variable for periods 1 - 3. "Post Tests" include control variables for risk aversion (Holt and Laury 2002), the Cognitive Reflection Test (Frederick 2005) and distributional preferences (Murphy et al. 2011). Standard errors are given in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.2: Multivariate analysis of the dishonest declarations

In Table A.2, we show the results of a random effects probit regression with dishonest declarations as the dependent variable (coded as a dummy variable). The results are broadly in line with our previous analysis of dishonest reports (Table 1). The two baseline specifications include dimension dummies (column 1) and an interaction term (column 2). Shared liability significantly increases the probability of a dishonest declaration. In contrast, we find no evidence of an effect of joint decision-making or an interaction between both dimensions. For declaration behavior, the coefficient for joint decision-making is even smaller than for reporting behavior. This is also reflected by the average marginal effects: the effect of shared liability is stable and between 11 to 13 percentage points, while the (non-significant) effect of joint decision-making is only 1–2 percentage points. Two alternative specifications confirm the robustness of our findings: specification (3) includes all post-experimental tests (Risk attitudes, Cognitive Reflection Test, and Social Value Orientation) while specification (4) controls for early periods (periods 1 to 3). As for reporting behavior, more risk-taking and better performance in the Cognitive Reflection Test is correlated to a higher non-compliance for declarations. In contrast, we do not find an effect for social preferences. Finally, the fraction of dishonest declarations is significantly smaller in early periods. As for the report of the own income, we find no evidence for Hypothesis 1, but we confirm Hypothesis 2.