Why we Cheat: Evidence from Tax Compliance Experiments *

Raymond M. Duch
raymond.duch@nuffield.ox.ac.uk
Nuffield College
University of Oxford

Hector Solaz
h.solaz@bham.ac.uk
Department of Economics
University of Birmingham

November 12, 2016

*Working Paper. Centre for Experimental Social Sciences, Nuffield College, University of Oxford. November, 2016.

Abstract

We conduct tax compliance experiments in which subjects earn real money, are subject to a tax, and can lie about their earnings. Performing better on incentivized real effort tasks results in more cheating. Cheating rises as the tax rate increases; but the higher levels of cheating by those who excelled at the tasks persists in high and low tax treatments. This correlation persists when earnings are associated with luck; the correlation persists in experimental treatments with more redistributive taxation. The correlation persists for choices made in other games by the same subjects : High performance types give less in a conventional Dictator Game and also cheat more in a classic die game in which they privately report the results of tossing a die.

1 Why We Cheat?

Individuals in the population, controlling for its costs and benefits, systematically lie or cheat (Gneezy, Rockenback and Serra-Garcia 2013, Cappelen, Sorensen and Tungodden 2013, Hurkens and Kartik 2009). This essay demonstrates that it is High Performance types who are more likely to cheat. Our evidence is based primarily on experimental games in which choices resemble the decision to comply with taxation. We provide important insights into tax cheating in particular but also into cheating behavior more generally.

It is generally accepted that cheating is costly – there is a considerable body of literature documenting the economic and political costs associated with cheating. An area in which the economic costs of cheating are particularly prominent is taxation. One of these claims is that tax cheating by the rich contributes to growing economic inequality (Zucman 2015, Piketty 2014, Atkinson 2015). This is a prominent issue, of course, because there are significant opportunities for cheating, particularly by the self-employed "rich", in virtually all tax regimes.[1]

Our effort to explain cheating behavior in the population is based on experimental games in which subjects make decisions that resemble the decision to cheat or not to cheat on one's taxes. The literature on tax compliance focuses on why individuals do not cheat. Benabou and Tirole (2011) make a persuasive case for how intrinsic motivation and reputation concerns in the population can inform the design of optimal incentives that depart from standard Pigou-Ramsey taxation. Non-standard preferences play an important role in tax compliance. As Luttmer and Singhal (2014) point out there is considerable evidence suggesting that intrin-

[1]The U.S. Internal Revenue Service audits less than two percent of tax returns and the penalties are relatively light (Posner 2000). Audit rates and penalties are even lower in most other countries.

sic motivations – or "tax morale" – contribute to explaining tax compliance.

While there is a recognition that intrinsic motivations play an important role in understanding tax compliance they are typically treated as being homogeneous in the population. This is probably not the case. We contend that High Performance types in the population are less motivated by these intrinsic payoffs. We present evidence indicating that High Performance types are more likely to cheat. Our evidence is based primarily on their choices in tax compliance experiments but we claim this generalizes to any situation in which there are opportunities to cheat.

2 Conjectures Regarding Cheating

Our contention is that High Performance types value cheating more than do Low Performance types. This section states our conjectures along with possible counterfactuals. We illustrate our conjecture with a tax declaration decision that is made under certainty. Failure to report ones full income to the tax authorities in this illustration does not provoke any penalties or any other form of social costs, such as ostracism. Following Allingham and Sandmo (1972), the taxpayer is expected to choose between two main strategies: (1) she may declare her actual income, or (2) she may declare less than her actual income. The utility of the taxpayer decreases with the amount declared because a percentage of it is deducted through taxes. Separately, the individual derives some pay-off from cheating that is simply determined by their Performance type.

The actual income of the taxpayer, i, is endogenously determined by her performance, P_i; and a wage component that is performance based, s. Only the taxpayer knows this income. Tax is levied at a constant rate, t, on declared income, which is determined by what proportion, c_i, of actual income the taxpayer declares. We can think of c_i as dichotomous (cheat or not cheat) or as some continuum of cheating that varies from 0 (no cheating) to 1 (full compliance). And we can think of P_i as indicating one's ability or productivity such that a higher value represents high productivity. The payoff from cheating is then determined by two terms: $\alpha[(c_i) * (P_i * s * t] + \beta[P_i * c_i]$.

The first square bracket captures the standard pay-off from cheating and has a weight of α: For any level of cheating, (c_i), the payoffs (from cheating) rise with $(P_i * s * t)$. This is the case because ones taxable income, $P_i * s$ increases with ones Performance type (more productive employees earn more money).

We conjecture that Performance type has an affect on the cheating decision that is independent of this cost term. The second term, $[P_i * c_i]$, captures this effect of tax-payer Performance type on cheating pay-offs. For any level of cheating, the pay-offs (from cheating) rise with one's Performance type. The importance of this second term is represented by β.

This representation of tax compliance is a significant simplification. Our formulation ignores all possible sources of uncertainty that are commonly associated with tax cheating. Since there are no audits in the experiments reported here, there is no possibility of receiving a penalty for lying about one's income.

Our conjecture regarding High Performance types is captured by the right-hand term $\beta(P_i * c_i)$. At any level of cheating $(c_i > 0)$, the pay-offs from cheating will be great for High, compared to, Low Performance types. A large weight, β, on the interaction term would lend credence to our argument. But Performance, P_i, also increases the pay-offs to cheating because it increases one's taxable income, $P_i * s$. The challenge is distinguishing the independent contribution of these two terms to the cheating decision – our experiments are designed to disentangle these two effects of Performance type on the pay-offs from cheating.

Performance. We argue that High Performance types are more likely to cheat. There are relatively few explicit efforts linking ability to cheating. An exception is Gill, Prowse and Vlassopoulos (2013) who find, in their online experiments with real effort tasks, that highly productive workers cheat more. Ariely (2012) finds that creativity is, while intelligence is not, correlated with cheating.

Our conjecture is consistent with a body of literature that links wealth, competitive success and power with greedy and unethical behaviour. Piff et al. (2012) implement seven experiments with convenience samples that support this correlation between wealth and unethical behaviour.[2] An analysis of shop-lifting behaviour found that incident rates are significantly higher for those in the highest income categories (Blanco et al. 2008). Under this line of reasoning the rich's antipathy towards tax compliance is based on more favorable attitudes toward greed and a predilection for unethical behavior. But we contend that its not wealth per se that explains unethical behavior; rather ability accounts for the rich's proclivity to engage in anti-social behavior.

A related literature suggests that competition breeds a sense of entitlement on the part of winners (Major 1994, Major and Testa 1989) that in turn facilitates dishonest behaviour by the winners. Recent experimental results (Schurr and Ritov 2016) indicate that winning a competition predicts dishonest behaviour. These

[2] Although in a subsequent letter to the editor Francis (2012) raised questions about the plausibility of the results (all seven experiments rejecting the null hypothesis) given that the observed Power of the tests in each of these experiments hovered around .5. Though subsequent replications of these experiments by Dubois, Rucker and Galinsky (2015) generated results that are consistent with the Piff et al. (2012) results.

results are consistent with our conjecture that individuals "who are likely to be winners" are more likely to cheat. Schurr and Ritov (2016) suggest that dishonest behaviour results from winning a competition in which there are identifiable losers. We contend, somewhat differently, that there are types in the population that identify themselves as "winners" or what we could call High Performance types. And High Performance types are more likely to cheat. Since they typically self-identify as High Performance types they do not require explicit competitions, that identify winners and losers, in order to exhibit dishonest behaviour.

The mechanism may be the sense of entitlement that is associated with identifying as a High Performer type. Experimental evidence suggests that a sense of entitlement is an important contributing factor to dishonest behaviour (Vincent and Kouchaki 2015, Major 1994, Major and Testa 1989). And performing well, or excelling in competition, appears to enhance these feelings of entitlement (Vincent and Kouchaki 2015, Major 1994, Major and Testa 1989). Entitlement may be the underlying mechanism that generates anti-social or dishonest behavior; being above average entitles one to be "above the law."

Finally, recent experimental findings suggest that relatively small incentives lead powerful subjects to make unethical decisions that harm less powerful subjects (Swanner and Beike 2015). When put in a position of high power and offered an incentive, the majority of subjects knowingly falsely informed on a confederate. The mechanism here may be that incentives cause powerful people to focus on their own rewards rather than the well-being of others.

These findings are consistent with our intuition that ability causes cheating. Individuals who are financially successful or who typically win at competition or have attained powerful positions tend to be High Performance types. We contend that its this self-perceived ability that leads to unethical behaviour. And its not a perception that necessarily needs to be primed – individuals know their performance type.

C_1: High performance reduces the percent of income declared for tax purposes.

Cost of Compliance. The expected costs of cheating matter. The intrinsic payoffs individuals might realise from tax compliance obviously come at an "extrinsic" cost. Individuals in the population, regardless of whether they are of high or low ability, share a similar, presumably downward sloping, demand function for "intrinsic" rewards. This is consistent with findings suggesting that charitable giving increases as its price decreases (Andreoni and Miller 2002, Gneezy, Meier and Rey-Biel 2011) but also experimental evidence that lying is sensitive to the costs associated with telling the truth (Gibson, Tanner and Wagner 2013). As the size of their tax obligation rises, which will be a function of income and tax rates, we should see cheating – the percentage of income undeclared – increase.

Our tax treatments are designed to help tease out the relative importance of ability versus costs in determining tax compliance: We observe low and high ability types at different prevailing tax rates; and for any compliance cost level, we observe how ability levels affect rates of cheating.

C_2: High income reduces the percent of income declared for tax purposes.

Winners versus Losers. Cheating behaviour may simply be triggered by discrete events that identify a winner or loser. Experimental evidence from Schurr and Ritov (2016) demonstrates that unethical behaviour by "winners" is triggered by competition that clearly identifies winners and losers. Our experimental design goes to considerable length to ensure we are not confounding "winning" with High

Performance type.

First, arguments about winners and unethical behaviour assume that individuals are informed about their location in the overall income distribution or about their success in competitions. One cannot "be" rich in this sense without perceiving that there are others who are poor, i.e., earning less than they do. Or one cannot be a "winner" in a competition without there being "losers" (Schurr and Ritov 2016). Its the fact that one knows one is rich or a winner that gives license to unethically or greedily behavior. In our experiments, subjects are not provided with this information which helps us isolate the effect of performance type as opposed to the effect of one's "winner" versus "loser" status.

Secondly, in our experiments we observe cheating behaviour over multiple rounds of the same tax compliance game. We can think of each round as a competition. And we observe the extent to which our High and Low Performance types exhibit consistent cheating behaviour. While its true that we expect High Performance types to do well on average, we will observe rounds in which they perform below average. If winning and losing in particular competitions affects cheating behaviour (as opposed to being affected by ones type) then our subjects' cheating behavior should respond to these deviations from average performance levels.

C_3: "Winners" declare lower percent of income for tax purposes.

Luck Trumps Ability. Some argue that tax compliance is conditioned on beliefs regarding returns to labor inputs (Alesina and Angeletos 2005). If rich taxpayers believe labor markets are efficient and reward ability they are more averse to taxation. Frank (2016) suggests that the failure to recognize the role of chance in determining success and income my increase one's reluctance to pay taxes. On the other hand, if opportunities for wealth are perceived to be more weakly correlated with ability and more likely to be determine by luck, status, class or government largess, the rich will be more tax compliant.

This argument implies that the impact of ability on cheating should be moderated when luck or status is perceived as an important determinant of success. There is evidence in the work on tax compliance to suggest that cheating is conditioned on the source of income or wealth (Durham, Manly and Ritsema 2014).[3]

We explore this claim by implementing two variations in the baseline compliance experiment: a "status" and a "shock" variant. Random assignment in both variants signals to some subjects that part of their earnings results from chance or luck. If the argument is correct, then cheating should be moderated for those High Performance types who benefit from chance.

C_4: When income results from luck (as opposed to ability) individuals declare a higher percent of income for tax purposes.

Fairness of the Tax Regime. A fifth conjecture is that there are social norms regarding the appropriateness or "fairness" of different tax regimes. Intrinsic motivation for tax compliance may vary according to perceived variations in "fairness." Besley, Jensen and Persson (2015) draw such a conclusion from their study of the UK poll tax. The 1990 introduction of the Poll Tax in the UK clearly induced a

[3]Generally, when experiments employ earned income they find a negative relationship between income and compliance (Alm and McKee 2006, Anderhub et al. 2001, Bradly 1987, Becker, Buchner and Sleeking 1987, Trivedi and O.Y.Chung 2006, Cherry, Kroll and Shogren 2005, Chan et al. 1999, 1996). Evidence from Dictator games suggest that windfall wealth tends to have an overall positive effect on contributions while in Dictator experiments with earned income subjects generally tend to be much less generous (Cherry, Frykblom and Shogren 2002, Hoffman, McCabe and Smith 1996).

negative shock in the intrinsic motivation to comply with taxation. Compliance dropped because taxpayers questioned the "fairness" of the tax being imposed. Cheating rises when individuals perceive the tax regime as unfair and vice-versa.

Fairness here refers to the manner in which taxes are raised and redistributed – the progressivity of the tax regime would be a case in point. Heterogeneity in intrinsic motivations or compliance could simply reflect variations in perceived fairness of the tax regime. Highly progressive tax regimes, for example, might create social norms that promote compliance. Our "Redistribute" treatment introduces progressive redistribution of the tax revenues collected in each group. If progressivity increases the intrinsic benefits of compliance for high ability types then we should see a moderation of the correlation between performance and cheating.

C_5: Under redistributive tax regimes individuals declare a higher percent of income for tax purposes.

Generalized Cheating Behaviour. Our conjecture is that 1) individuals in the population self-identify as High Performance types; and 2) their proclivity to cheat is not restricted to a particular domain (i.e., tax compliance). We designed treatments that explicitly test these two conjectures.

Performance on a real effort task (RET) allows us to identify those with high ability. In order to assess whether subjects self-identify as high versus low performance types we elicit, in an incentivized fashion, subjects' assessments of both how well they expect to perform and, after the fact, their relative performance on the RET compared to other group members. Our conjecture is that subjects will anticipate their high or low performance on the RET. Hence subjects recognise their ability without requiring any feedback on other group members' performance.

Second, the proclivity to cheat by High Performance types is not restricted to taxes. Accordingly we have included other opportunities for subjects to cheat. One of these efforts simply gauges the other-regarding preferences of the subjects. Prior to the RET, and before they make decisions in the tax compliance module, subjects play a conventional Dictator Game that allows us to compare the generosity of High versus Low Performance types. Subjects also have the opportunity to cheat at the end of the experiment when they toss a die to determine compensation for completing a questionnaire. After the tax compliance games, subjects play a version of the Fischbacher and Follmi-Heusi (2013) die tossing game that again allows us to compare the unethical behaviour of high versus low performance types. Our expectation is that Higher Performance types will lie more about their die toss than is the case low ability types.

C_6: Cheating by high performance types generalizes to decisions other than declared income for tax purposes.

Summary Our conjecture is that High Performance types cheat more than Low Performance types. Other factors might moderate this relationship. In particular, for both High and Low performance types, we might see cheating increase as the price of honest behavior, tax compliance in our experiments, increases.

Contextual factors might moderate the relationship between ability and cheating. High Performance types might cheat less when income is determined by luck or status. And the intrinsic benefits of compliance by High Performance types may be higher under highly progressive tax regime (again moderating this ability-cheating correlation). These possibilities are explored in the experimental design. High Performance types self-identify – a conjecture also explored in the experimental design. And we contend that this proclivity to cheat or engage in unethical behaviour by High Performance types is a general phenomenon. Accordingly we incorporate diverse decision making situations in which subjects have an opportunity to cheat.

3 Tax Compliance Experiments: Design

Tax Cheating. We employ real effort tax compliance experiments in order to understand cheating.[4] Our tax compliance experiments are designed to isolate in as neutral a context as possible the micro-foundations for cheating. It is explicitly a highly simplified tax regime and we make no claims regarding the external validity of the "stylised" regime itself. Individuals reveal their type by performing the RET. Their type explains subsequent cheating behavior. Subjects are randomly assigned to different versions of this "stylised" tax regime that help isolate the factors causing individuals to cheat.

This essay reports the results for five treatments designed to identify the factors causing cheating. Subjects are paid at the end of experiment, and do not receive feedback about earnings until the end of the experiment. Participants receive printed instructions at the beginning of each module, and instructions are read and explained aloud.

The tax treatments consist of ten rounds each. Table 1 summarises the treatments. Prior to the tax treatments, participants are randomly assigned to groups

[4] There is a considerable literature on tax compliance experiments (Slemrod 2007, Alm, Bloomquist and Mckee 2015). Typically these experiments have explored whether or not compliance is affected by perceived features of the tax system(Alm, Jackson and McKee 1992, Spicer and Becker 1980, Falkinger 1995, Cowell 1990) or changes in the tax system(Heinemann and Kocher 2013).

of four. In four of the treatments we follow a partner matching such that the composition of each group remains unchanged for the two tax treatment modules. In a fifth treatment, Sessions 15-20, participants are randomly assigned to groups of four after each of the ten rounds of the two tax modules.

Each round is divided in two stages. In the first stage subjects perform a real effort task. This task consist of computing a series of additions in one minute. Their Preliminary Gains depend on how many correct answers they provide, getting a set number of ECUs for each correct answer (in the Baseline Treatment this is 150 ECUs for each correct answer where 300 ECU = £1).

After subjects receive information concerning their Preliminary Gains, they are asked to declare these gains. A certain percentage or "tax" (that depends on the treatment) of these Declared Gains is then deducted from their Preliminary Gains.[5] These deductions are then divided amongst the members of the group (in most treatments the deductions are divided evenly amongst group members). Note that in each session the tax rate is consistent and it does not vary. The tax treatments are the following: 10%, 20%, 30%.[6]

We implemented five treatments designed to identify conditions under which subjects might vary their degree of tax cheating. In the first equal salary (*Baseline*) treatment subjects get the same payment for correct answers to the real effort test (10 pence). This represents the "ability" treatment in which salaries are strictly

[5]We explicitly avoid framing the game in terms of "taxes". Subjects are told that a deduction (rather than a "tax") would be applied to earnings. And while our neutral framing does not identify the deduction as a tax, we do explicitly tell subjects that there is a probability of verification of income (an "audit") that could result in a penalty. We believe this framing signals to subjects that compliance is encouraged. In the results reported here, subjects are informed that the probability of such an "audit" is zero. And they are told, accurately, that subsequent modules of the session could have non-zero audits. We only report the results here of the zero audit module.

[6]One of the Redistribute sessions had a tax rate of 40%.

Table 1: Summary of Tax Compliance Experimental Treatments

Session	Participants	Groups	Tax Rate	Treatment
1	24	6	10%	Baseline
2	24	6	20%	Baseline
3	24	6	30%	Baseline
4	24	6	10%	Status
5	12	3	20%	Status
6	16	4	20%	Status
7	20	5	30%	Status
8	24	6	10%	Redistribute
9	20	5	20%	Redistribute
10	20	5	30%	Redistribute
11	20	5	40%	Redistribute
12	16	4	10%	Shock
13	20	5	20%	Shock
14	20	5	30%	Shock
15	16	4	10%	Baseline Non-fixed
16	16	4	10%	Baseline Non-fixed
17	16	4	10%	Baseline Non-fixed
18	12	3	10%	Baseline Non-fixed
19	12	3	20%	Baseline Non-fixed
20	16	4	30%	Baseline Non-fixed

tied to performance. The second (*Status*) treatment consists of an inequality salary treatment in which two ("Low Status") subjects get 5 pence per correct answer and two ("High Status") subjects get 15 pence per correct answer. Random assignment determines those subjects earning higher returns to effort and hence introduces our notion of "luck" or "status" into the resulting income distribution.

In a third (*Redistribute*) treatment the two participants per group with the lower income (each round) receive 35% of the pooled deductions, while the two with higher income receive 15% – in case of ties on the number of additions computed (income), the division is decided at random. This represents the treatment in which the redistributive use of the tax revenues is the most aggressive.

Our fourth (*Shock*) treatment (Sessions 12-14) randomly assigns half of the subjects to a control treatment that resembles the *Baseline* treatment. Half of the subjects are randomly assigned to a "shock" treatment in which their earnings from the RET are incremented by 150 ECUs (50 pence). Subjects are not informed about their bonus until after they complete their RET and before they report their income. Subjects are presented with a breakdown of their earnings – gains associated with their performance in the RET and the portion associated with the bonus. These random assignments occur after each round is played.

Our final (*Baseline Non-fixed*) treatment (Sessions 15-20) resembles the Baseline treatment with the exception that subjects were randomly assigned to new groups after each round of play.

Subjective Ability. In Sessions 15-20 we elicit incentivised subjective assessments of ability in the RET. After completing the first set of practice rounds for the RET, but prior to the RET for the first round, we asked subjects to indicate the expected ranking (rank 1 through 4) of their first round performance relative to other members of their group. They were informed that they would earn 150 ECUs if they were correct. A similar question was then asked of each subject for two randomly selected rounds (excluding the first round). In these cases subjects were asked to rank their performance in the RET effort they had just completed.

Selfish Preferences. We measure the general other-regarding preferences of subjects in the first module of each session with a standard Dictator Game (Engel 2011). Subjects are asked to allocate an endowment of 1000 ECUs between them and another randomly selected participant in the room. Participants are informed that only half of them will receive the endowment, and the ones who receive the endowment will be randomly paired with those who don't. However, before the endowments are distributed and the pairing takes place, they may allocate the endowment between themselves and the other person as they wish if they were to receive the endowment.

Lying. In Sessions 15-20, we implemented a version of the Fischbacher and Follmi-Heusi (2013) die tossing game in order to compensate respondents for completing an attitudinal and demographic questionnaire. Subjects were asked to toss a die (in total privacy without any means for the experimenter to see the result) and report the result. Subsequent to their first toss they were given an opportunity to toss the die as many times as they wanted in order to ensure the die was fair. The number reported from the die toss translated to payoffs at a rate of 100 ECU per unit reported (where 300 ECU = £1) – so they could earn up to £2.

Subjects and Earnings. Subjects are informed that there is a certain probability that the Declared Gains are compared with the actual Preliminary Gains in order to verify these two amounts correspond. In the module results presented in this essay the probability of such an audit is 0%. At the end of each round participants are informed of their Preliminary and Declared gains; the amount they receive from the deductions in their group; and the earnings in the round. At the end of each module one of the ten rounds is chosen at random, and their earnings are based on their profit for that round. At the end of the experiment ECU earnings are converted at the exchange rate 300 ECUs = 1£. While the earnings are prepared participants answer a questionnaire, which consists on an Integrity Test, and a series of socio-demographic questions.

All of the sessions were conducted at CESS (Centre for Experimental Social Sciences), a research facility of Nuffield College, at the University of Oxford. Subjects are undergraduate and graduate students from Oxford. Some subjects had participated in previous experiments, but all of them were inexperienced in this particular type of experiment. No subject participated in more than one session of the study. On average, a session lasted around 90 minutes, including instructions and payment of subjects, and the average payment was around 17£. The experiment was computerized using ZTREE (Fischbacher 2007). A copy of the instructions can be found in the Appendix.

4 Tax Compliance Experiment: Results

4.1 Effort

We implemented treatments in which income from the additions RET is entirely determined by performance – each subject receives 10 pence per correct addition; a treatment in which a randomly assigned status (high/low) determines the subject's income associated with correct additions (5 versus 15 pence); and a treatment in which half the subjects are randomly assigned to a shock treatment whereby their income is incremented by 50 pence. The mean correct additions is similar across treatments: 12.2 in the baseline; 12.7 in the low status; 13.2 in the high status; and 12.8 in the shock treatment. There is no indication that effort was conditioned on the tax rates – the average correct additions for the 10%, 20%, and 30% tax rates, respectively, were 12.7, 11.8, and 12.1.[7]

[7]Figure 5 in the Appendix summarises the performance of the subjects in the different treatments.

4.2 Cheating

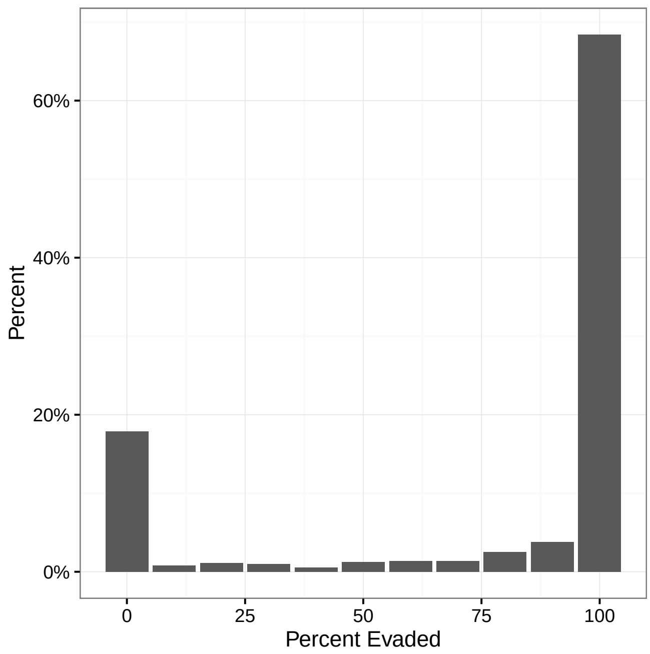
The outcome variable of interest is the amount of misreported earnings in this tax compliance game. The audit rate is zero and hence subjects were not penalized for cheating. Revenue collected from these taxes is (with the exception of one treatment) distributed equally amongst subjects and hence there are no social gains (or losses) associated with compliance. The equilibrium choice for all subjects is to report zero earnings.

Virtually all of the subjects in the experimental sessions cheated – 81 percent of the declarations under-reported the subjects' actual earnings.[8] Figure 1 summarises the subjects' cheating behaviour. On average, subjects' choices resemble standard compliance models (Allingham and Sandmo 1972): Subjects report about 20 percent of total income earned over the 10 sessions with zero audit.[9] There is variation in cheating. The left graph in Figure 1 presents the frequency of subjects' average ratio of non-declared to total earnings. About two-thirds of the subjects are cheating virtually in every round and about 10 percent never cheat.

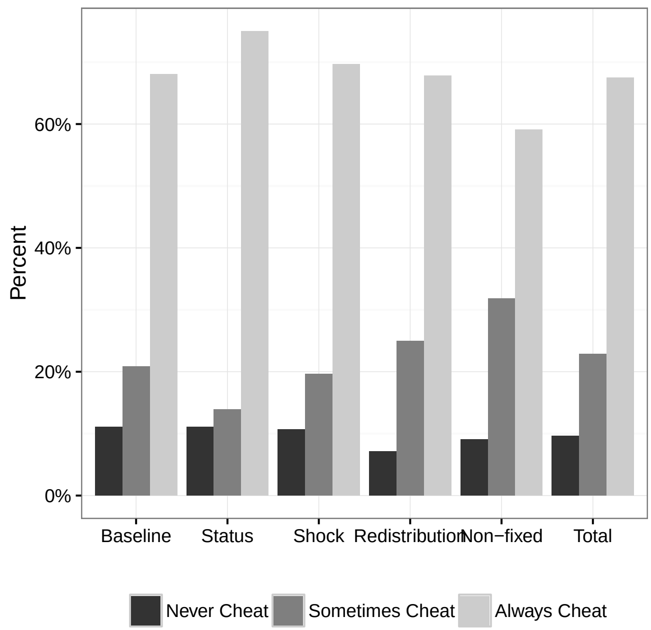
[8]This is almost precisely the percentage of individuals evading in the Dwenger et al. (2015) German local tax field experiment. And this suggests that the overwhelming majority of subjects conform to the Allingham and Sandmo (1972) theoretical framework.

[9]The average 20 percent compliance results are similar to those obtained by Alm, Bloomquist and Mckee (2015) in their tax experiments and benchmark closely to the compliance rates they report from their analysis of actual U.S. Internal Revenue Service data made available by the National Research Program initiative. They find that the average compliance rate for Sole proprietor Section C filers in the NPR data is 31 percent (24 percent in the data weighted to the population distribution). And in their experiments with student subjects, the compliance rate is 28 percent in the zero audit probability treatment.

Figure 1: Subjects' Non-declared to Total Earnings



(a) Frequency Average Non-declared to Total Earnings



(b) Frequency: Alway, Sometimes, Never Cheat

Subjects earn and report income over ten periods. We do not expect significant within subject variation with respect to cheating. Our principal conjecture is that compliance is shaped by one's performance type that does not vary significantly by subject over the ten rounds. Our expectation then is not unlike those of Kajackaite and Gneezy (2015), Gneezy, Rockenback and Serra-Garcia (2013) – subjects, given a particular treatment condition, will either sort into economic types and cheat most of the time or will rarely cheat.

The right graph in Figure 1 suggests this is a reasonable characterisation of the subjects' behaviour. Subjects are categorised as "Always Cheat" if in at least 9 out of 10 rounds they reported less than 15 percent of their earned income; as "Never Cheat" if in at least 9 out of 10 rounds they reported more than 85 percent of their earned income; and "Sometimes Cheat" if they fell between these two extremes. Two-thirds of the subjects are consistent cheaters using this categorisation strategy. And about 10 percent are serial non-cheaters. One-quarter of the subjects vary their cheating behaviour across the 10 rounds.

This result is consistent with other claims that individuals in the population sort into types that are highly prone to cheat or lie (subject to an assessment of costs and benefits) while others rarely cheat or lie (Kajackaite and Gneezy 2015, Gneezy, Rockenback and Serra-Garcia 2013). Our conjecture is that one's Performance type determines cheating proclivities.

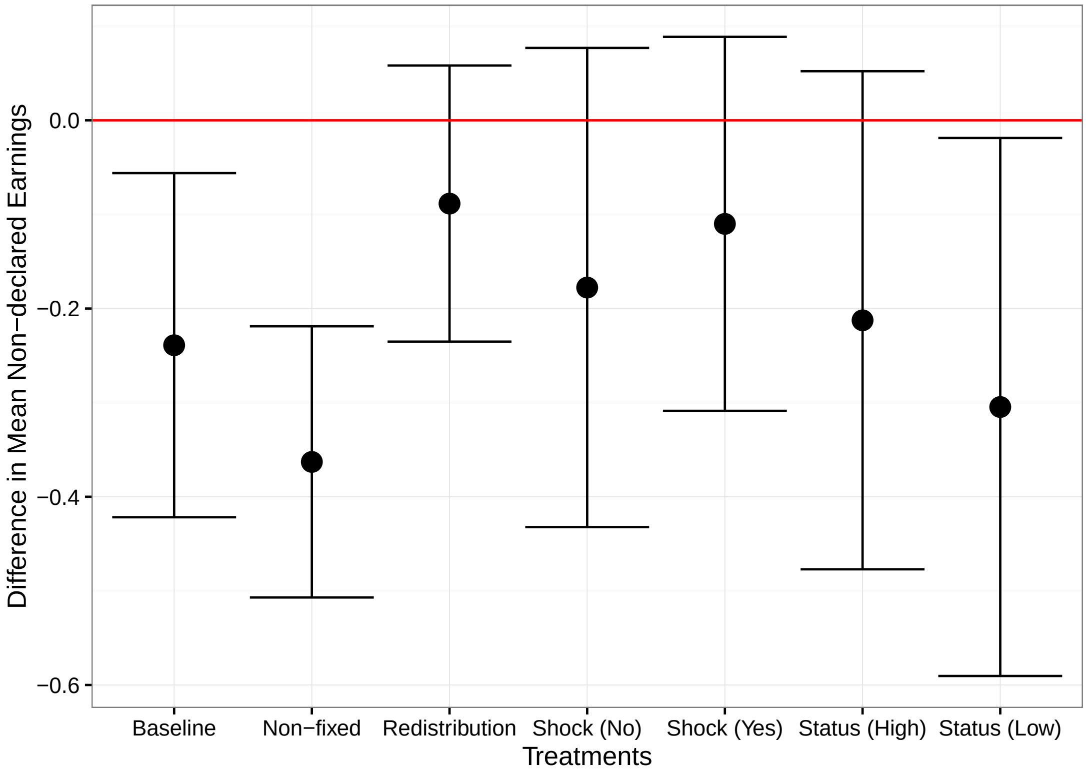
Figure 2 presents the difference in treatment effects between Low and High Performance types, along with their boot-strapped confidence intervals estimated with clustered standard errors. We define Low and High Performance types according to whether their average performance was above or below the median performance on the real effort tasks under each treatment (which is either 11 or

12 correct additions depending on the treatment). The point estimates represent the percent of earnings evaded by the High Performance types subtracted from those for the Low Performance types. The typical difference across these seven treatments is about -0.20.

Clearly performance type matters – High Performance types cheat more. And the results are quite robust to treatment: High Performance types cheat more regardless of treatment. In the Baseline treatment, there is a difference of more than -0.20 percent in cheating – High Performance types do not declare about 85 percent of their income while Low Performance types cheat at a rate of about 65 percent. Moreover, the confidence intervals calculated using clustered standard errors here do not include zero. This is consistent with our conjecture that the interaction term $(P_i * c_i)$ weighs heavily in the cheating calculus described earlier.[10]

[10] Risk preferences could affect cheating behavior. The tax compliance games resemble a public good game in that each subject's final earnings is affected by tax "contributions" by other members of the group. Given there is uncertainty as to the choices of other group members, risk aversion can reduce the contributions of subjects (Schechter 2007, Teyssier 2012). Risk preferences may be confounded with performance type in our tax compliance games. We explore the impact of risk preferences on our findings in the Online Appendix and conclude that the performance type effect we estimate is not confounded with risk preferences.

Figure 2: Difference in Mean Non-declared Earnings High-Low Performance Types



Cost of Compliance (α) versus Performance Type (β). The cheating calculus described earlier is driven both by the cost of compliance (which, for any performance type, is a function of salary and the prevailing tax rate) and whether one is a high or low performance type. Table 2 provides some insight into the relative magnitudes of α and β. Comparing the % Evaded columns for Low versus High Performance types confirms the Figure 2 results: High types consistently cheat more, regardless of the tax rate.

Table 2: Tax Rates, Average Cost of Compliance, and Cheating

	Low Ability		High Ability	
	Cost to Comply	% Evaded	Cost to Comply	% Evaded
10% Tax	150	65%	205	81%
20% Tax	260	62%	422	87%
30% Tax	398	78%	851	89%

Secondly, the % Evaded is clearly higher in the 30% tax rate compared to the 10% tax rate for both Low and High Ability types. And, the % Evaded for the High Ability types in the 30% tax treatment is higher than the % Evaded for the Low Ability types in the 30% tax treatment. This is consistent with the notion that both higher tax rates and higher performance levels contribute to cheating. Its reasonable to conclude that the tax component (t) of the cheating calculus is affecting the pay-offs from cheating as conjectured.

Surprisingly, the entire measure of the cost to comply term $((P_i * s * t))$ in the cheating calculus is at best very weakly correlated cheating. Within the Low and High ability types there clearly is a positive relationship between cost of complying and the % Evaded. But if we compare the 10% and 20% treatment average costs to comply figures from the Low Ability types, they are larger than the average cost to comply for High Ability types in the 10% tax treatment. Yet the % Evaded for the High Ability types in the 10% tax treatment is almost 20% higher than it is for these two Low Ability tax categories – in spite of having a lower average cost to comply. This at least suggests that Performance type might be a stronger driver of cheating than is the cost of complying.

Cheating in Context. Our experimental results suggest that the correlation between ability and cheating is very robust, although there is some evidence context can matter. In the *Status* sessions, subjects are randomly assigned to either a low or high wage rate. Again for both treatments we see that cheating is significantly higher for high as opposed to low performance types. The difference is higher for the those assigned to the low status treatment and with a confidence interval below zero.

In the *Shock* sessions, half the subjects randomly receive a bonus (*Shock*) after completing their RET. Differences for high versus low performance types is about -0.20 in the *No Shock* condition and roughly -0.10 for subjects in the *Shock* treatment. In the Redistribute sessions, we see a moderation in the difference between low and high performance types – it is just under -.0.10.

In our highly stylized tax regime, subjects interact in small groups of four and they retain the same matched partners throughout the 10 rounds reported in this essay. Features of this decision making resemble a standard public goods game. Accordingly, some subjects, particularly those who are net beneficiaries from the public good, may comply in an effort to encourage other members to reciprocate, i..e, contribute generously to the public good in subsequent rounds (Andreoni 1995, Fischbacher, Gchter and Fehr 2001, Sonnemans, Schram and Offerman 1999). This could explain higher compliance by low performers in our tax compliance games.

We explore this possibility by implementing six sessions (Sessions 15-20) in which subjects are randomly assigned to a new group after each round of the tax game. Figure 2 reports the difference between high-low performance types in these three non-fixed sessions. Cheating by high performance types compared to low performers is even greater in the non-fixed (Sessions 15-20), relative to the, matched baseline treatments (Sessions 1-3). Its unlikely that the higher levels of cheating by high ability types in our experiment are an artefact of the fixed matching design.

High Performance Types. We conjecture that individuals know their type and behave accordingly. Cheating by high ability types does not need to be primed necessarily by an external cue signalling their type. A number of features of our experiments allow us to test this conjecture.

Subjects here are not informed of the distribution of outcomes. Hence the difference in the cheating behaviour we observe is unlikely to result from subjects comparing their outcome in each round to the performance of other members of their group. Rather subjects are self-aware of their performance type. One indication of this is that subjects perform quite consistently over the multiple rounds of the addition real effort task. Recall we categorised subjects into High and Low Performance types based on whether they performed, on average, more or less than 11 correct additions (the overall mean). Rarely did each subject's performance deviate significantly from their performance "type" (either Low Performance types deviating significantly above the mean or High Performance types deviating significantly below the mean). As an illustration, we define low performance deviations as an outcome in which a Low Performance type performs more than 12 additions. And high performance deviations occur when a High Performance type only manages less than 10 correct additions. There are only 78 occurrences of the high performance deviations which represents less than 5 percent of high performer outcomes. And there are only 68 occurrences of low performance deviations which also represents less than 5 percent of low performer outcomes.

Cheating behaviour should be consistent within subjects – it should not fluctuate significantly and we would not expect it to respond to stochastic shocks in performance. This implies that, for any particular subject, cheating behaviour is not correlated with RET outcomes that deviate significantly from their overall performance levels. Its not the case, for example, that when a high performance type experiences an unexpectedly poor RET performance her cheating sharply declines. To test this, we calculated, for occurrences of both low and high performance deviations, the average change in percent of earnings evaded. In both cases, the average change was not significantly different from zero.[11]

Finally, as part of Sessions 15-20 we elicited, in an incentive compatible fashion, subjective assessment of ability on the RET. At the very outset of the tax compliance module, after subjects had become familiar with the RET (but before they began playing any of the 10 rounds), we asked subjects how they expected their performance to rank relative to the other three members of their group.

Table 3 presents a cross-tabulation of subject's rank with our categorization of

[11]Figure 7 in the Online Appendix provides a frequency plot of these deviations in cheating for the low and high performance deviation cases.

the subject into a low or high ability type. As expected, subjects who performed above average on the RET are much more likely to have anticipated being either ranked first or second in their group. Almost 60 percent of the High Performance types anticipated being ranked first and about 35 percent expected to be ranked second. Very few of the High Performance types anticipated being ranked third or fourth. On the other hand the Low Performance types were much more likely to anticipate performing poorly: half of the Low Performance types expected to rank third or fourth while only about 20 percent of them expected to rank first.

Table 3: Subjective Assessment of Performance on Real Effort Task (Sessions 15-20)

	Low Ability	High Ability
First Rank	21%	57%
Second Rank	26%	37%
Third Rank	37%	5%
Fourth Rank	15%	2%
Total	84	60

Our conjecture is that High and Low Performance types generally know their type. Their cheating behavior is not conditioned on a particular outcome (such as winning a competition or ranking first in a tournament). In our experiments subjects are never told about the performance of other group members – nevertheless the High Performance types cheat more than the Low Performers. And cheating behaviour is relatively constant across each round of the tax experiment; cheating does not respond to positive or negative shocks to subjects' typical performance on the RET. Finally, we explicitly ask subjects to assess their expected performance on the RET and there is a strong correlation between self assessed ability and performance on the RET. There is strong evidence to suggest that individuals recognise that they are either a Low or High Performance type and this determines how much they cheat in each round of the game.

Generalized Cheating Behavior. We conjecture that high ability types will be more likely to cheat in contexts other than this particular tax game. To test this we have subjects make decisions in two other games in which they have an opportunity to behave selfislly or unethically.

Prior to the tax compliance modules, subjects are randomly paired and one of the pair is randomly given the opportunity to share any portion of an endowment of 1000 ECUs with his or her partner. In this standard Dictator Game we expect that high ability types (who are identified in a subsequent tax compliance game) would be much less generous than the low ability types. Analysis of the amounts given in the Dictator Game clearly confirm our conjecture. The average offer, over all the sessions, by low performance types is 289 ECUs and by high performance types it is 197 ECUs. And the overall correlation between number of correct additions and offers in the Dictator Game is -0.22. Even prior to performing the RET, high performance types clearly distinguish themselves as being less generous.

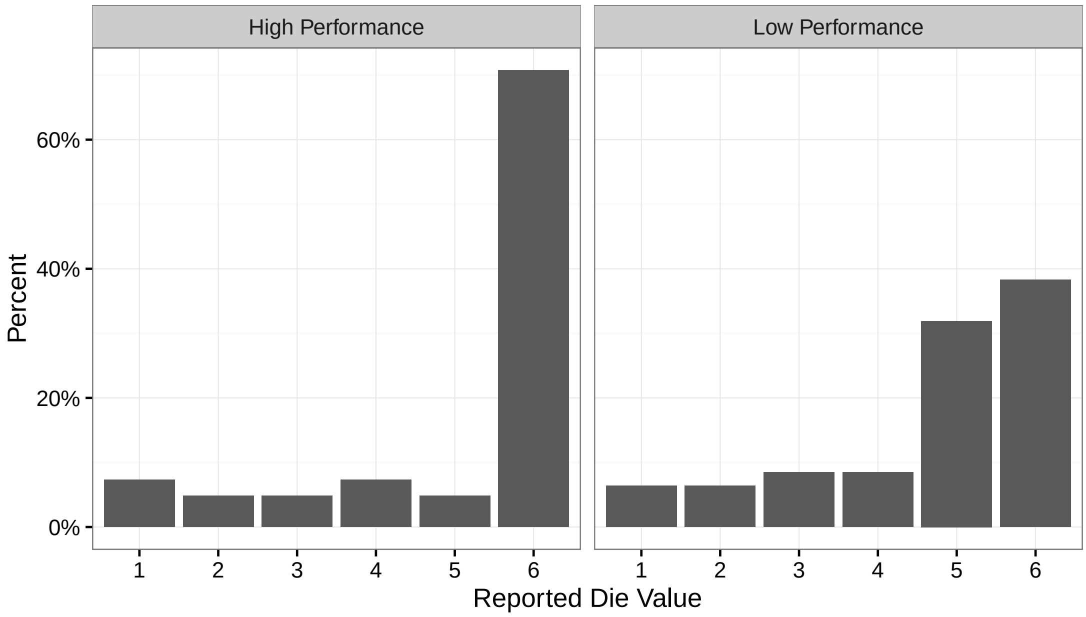
We implemented a more direct measure of cheating in Sessions 15-20. After subjects complete the tax modules they are asked to complete a questionnaire. Compensation for this effort is determined by their reporting the results of a completely private toss of a die. Subjects can report truthfully or lie – there is no way that the experimenter can verify the result of the subject's die toss.

The authors of this game (and subsequently many others) find that a significant percentage of subjects lie – and this is the case in our experiment. The extent of lying is inferred by the extent to which the frequency distribution of the die outcomes (1 through 6) deviates from a uniform distribution.

Figure 3 presents the distribution of reported results for High and Low Performance types. The High Performance types are not shy about lying. About 70 percent of high ability subjects report having tossed a six. The Low Performance types are somewhat more shy about lying – although clearly they lie. About 40 percent report rolling a six – another 30 percent of low ability types report a five.[12] This experiment provides subjects with an anonymous, uncomplicated and non-strategic setting in which they have an opportunity to cheat. The high level of "extreme" cheating by our High Performance types in the die-rolling experiment provides further support for our conjecture about ability and cheating.

[12]This suggests that over 50 percent of the High Performance types lied about rolling a six compared to around 22 percent of the Low Performance types. Its also the case that about 15 percent of the Low Performance types lied about rolling a five. On average about 40 percent of the subjects lied about their die outcome which is somewhat higher than the 0.216 estimated in the Abeler, Nosenzo and Raymond (2016) meta-study of the die game.

Figure 3: Die Toss: Reported Results for Low and High Ability Types



Performance type and wealth are correlated in the tax compliance game. Observing more cheating by High Performance types in our tax compliance experiments could simply reflect resistance to compliance as the size of that contribution gets bigger – suggesting $\beta = 0$ and $\alpha \neq 0$ in the cheating calculus. We address this with the multivariate analysis in the next section. But our results concerning generalized cheating behavior also speak to this issue. As we demonstrated above, when our High Performance types unexpectedly perform poorly – and hence earn less money – we do not see a change in their cheating behavior. This suggests that their "cheating reflex" is not simply responding to the magnitude of their tax obligation.

A second insight from the analysis concerning generalized cheating behavior is that this "cheating reflex" by High Performance types is not confined to the tax compliance game. We observe the cheating behavior of High Performance types in other games where the payoffs from cheating are entirely unrelated to their wealth from the real effort tasks. High Performance types give less money in a conventional Dictator Game and also are more likely to cheat in a classic die game in which they privately report the results of tossing a die. Their cheating behavior in these two games, unrelated to the RET, cannot be confounded with a wealth effect. Hence we are confident that high levels of cheating behavior by High Performance types in the tax compliance games indicates that higher ability is negative correlated with intrinsic payoffs from tax compliance.

4.3 Multivariate Analysis

Cheating is measured in two fashions in this experiment: 1) simply whether or not the subject correctly reported her income from the real effort task; and 2) the percent of a subject's actual income from the real effort task that was not reported. Here we estimate logit models with a dichotomous dependent variable coded zero for subjects who reported their actual winnings and coded one for those who reported amounts that deviated from their actual winnings. We report clustered standard errors to reflect that fact that each subject plays the tax compliance game ten times. [13]

Table 4 reports logit regression results for the dichotomous measure of cheating. The first results column in Table 4 presents the estimated coefficients for all 2,840 decisions taken by 284 subjects in the 14 lab sessions in which participants were randomly matched to fixed partners. Included in the model is # of Additions (number of correct additions) which is our measure of performance and Compliance Cost (Tax Rate X Actual Earnings). The significant coefficient on performance suggests that across all treatment contexts High Performance types are much more likely to cheat than Low Performance types. Compliance Cost, as was suggested by Table 2, matters less than anticipated – the coefficient is small and statistically insignificant. Cheating is correlated with one's performance type rather than with the actual cost of compliance.

In the second column of Table 4 we report results for a fully-specified logit regression model that includes dummy variables for each of the treatment sessions and treatments interacted with the performance variable. Of particular interest are these interaction terms – two of them are reasonably large. For those randomly assigned to low wages in the status treatment, performance seems to matter disproportionately more than in the other treatments. And there is evidence here

[13]Table 5 in Appendix 1 replicates the analysis in Table 4 with a dependent variable that measures the percent of a subject's earnings that were not reported. The results essentially confirm the findings from Table 4.

that the Redistribution treatment significantly moderated the correlation between performance (# of Additions) and cheating.

The subsequent columns of Table 4 present logit regression results separately for each of the treatment sessions. For the Baseline treatment there is a positive relationship between performance and cheating (although the estimate is imprecisely estimated) and cheating increases with the Compliance Cost.

In the standalone Status Model the correlation between performance (# of Additions) and cheating is positive and precisely estimated. For those with low status in these sessions, there is a strong correlation between performance and cheating – performing well in the "unfair" treatment results in particularly high levels of cheating. The coefficient on the High Salary X Additions interaction term suggests that the correlation between performance and cheating for those assigned to the high status condition, while still positive, is moderated relative to the low status subjects. Cost of compliance is statistically insignificant in this model.

Table 4: Logit model of cheating regressed on performance

	Full	Full	Baseline	Status	Shock	Redistribute	Non-Fixed
# of Additions	0.181	0.144	0.056	0.769	0.164	0.111	0.157
	(0.040)	(0.072)	(0.077)	(0.187)	(0.087)	(0.074)	(0.077)
Compliance	0.0005	0.001	0.005	−0.003	0.004	−0.001	−0.001
Cost	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
High Status		−1.087					
		(1.207)					
Low Status		−4.349					
		(1.833)					
No Shock		−0.594					
		(1.242)					
Receive Shock		−1.149					
		(1.311)					
Redistribute		1.330					
		(1.039)					
# of Additions X		0.089					
High Status		(0.097)					
# of Additions X		0.552					
Low Status		(0.187)					
# of Additions X		0.064					
No Shock		(0.103)					
# of Additions X		0.117					
Receive Shock		(0.109)					
# of Additions X		−0.090					
Redistribute		(0.092)					
High Status				3.356			
				(2.054)			
# of Additions X				−0.401			
High Status				(0.203)			
Receive Shock					−1.233		
					(1.222)		
# of Additions X					0.057		
Receive Shock					(0.083)		
Constant	−0.556	−0.540	−0.757	−4.894	−1.279	0.794	−0.269
	(0.426)	(0.683)	(0.623)	(1.799)	(1.077)	(0.739)	(0.648)
AIC	2354.64	2281.27	647.1	435.47	442.26	693.22	932.68
Observations	2,840	2,840	720	720	560	840	880

For the Shock version of the tax compliance experiment, the Performance coefficient is positive and statistically significant. The Receive Shock X Additions interaction term is positive but the magnitude of the interaction coefficient is relatively small and it is not precisely estimated. The "Redistribute" Model in Table 4 presents the impact of a treatment in which each group's tax revenues are distributed unequally with a higher percentage going to the two subjects that performed most poorly. While the coefficients on # of Additions and Compliance Cost remain positive they are not statistically significant.

Learning. We conjecture that individuals know their Performance type – this is not something that they learn as part of the game. There are a number of features of the experimental design and multivariate estimation that allow us to assess the extent to which this is in fact the case. First, with respect to the experimental design. Subjects are never informed about the distribution of performance types or of the RET income distribution. Hence, the robust correlation we observe between performance and cheating suggests that subjects are very aware of their "Performance type". Any notion that cheating is emboldened simply by being rich, (i.e., by one's wealth relative to others in the population) or by having won a competition (in which there are identifiable losers), is clearly challenged by these experimental results. Subjects are never explicitly provided with information regarding their relative "wealth" or "success" in the repeated game.

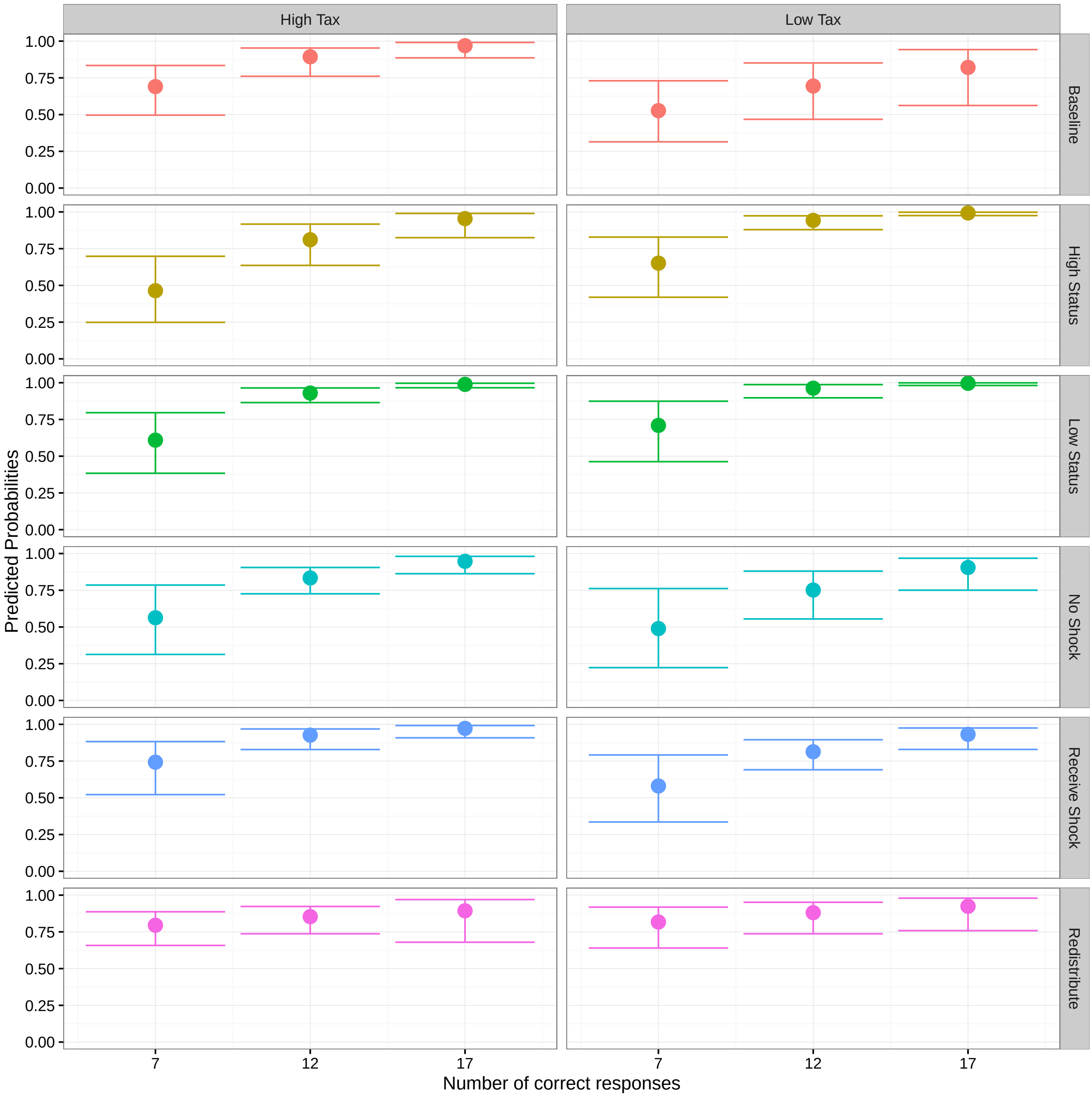
In spite of never being told about their relative performance, subjects might learn this over the repeated play of the same game. Subjects play the tax compliance game over 10 periods which might result in some learning about one's performance levels; in particular, subjects observe how much they earn and also their revenues from tax redistribution. But learning unlikely accounts for the Performance effect we observe on cheating. To control for learning, we estimated all of the models in Table 4 for only the first period of each of the different treatment sessions. We present the results in Table 6 in the Online Appendix. For the most part we find that Performance remains strongly, and positively, correlated with cheating. Its unlikely that our results in Table 4 are an artefact of learning.

As we pointed out earlier, the structure of these tax compliance games resemble a standard public goods game. Recall that subjects are randomly matched to partners and they keep these partners over the 10 iterations of the game. As a result, strategic considerations could lead subjects, particularly the Low Performers, to comply with prevailing tax rates in order to encourage others, in particular High Performers, to cooperate, i.e., report their actual earnings, in subsequent rounds of the game. There is also good evidence to suggest that in these public goods games, subjects become particularly uncooperative as they play the final rounds of the game. One could imagine that our Performance type effects could be confounded with strategic behaviour by subjects playing our repeated tax compliance game. These concerns primarily result from our decision to have a fixed partner matching design. Accordingly, we implement a series of sessions in which subjects were randomly re-matched to a new group after each iteration of the tax compliance game. We present, in the last column, results for the these Non-fixed matching sessions (i.e., the Non-Fixed model). Again consistent with the results in Table 4, the Performance coefficient is positive and precisely estimated while the Cost of Compliance coefficient is not significant. There is no evidence that our results are an artefact of strategic reasoning and learning associated with the fixed partner matching design.

Predicted Rates of Cheating The estimated coefficients in Table 4 constitute strong support for our conjecture that High Performance types will cheat more than Low Performance types. Surprisingly, Compliance Cost had a much weaker affect on the decision to cheat. And for the most part there was not much contextual variation in this performance effect on cheating. Figure 4 summarizes the model estimates from Figure 4 by presenting the predicted probabilities of cheating for particular performance levels and for the 10% and 30% tax regimes. We generate predicted probabilities for three performance levels for which we have large numbers of observed values in the actual data: 7, 12, and 17 correct additions.

Figure 4 confirms our conclusions from the multivariate discussion. First, cost of compliance does not appear to have a particularly strong affect on cheat – the differences between the 10% and 30% tax regimes are not large. On the other hand performance clearly matters – we see a reasonably strong positive correlation between performance and the probability of cheating. And in most cases, predicted cheating at the 17 correct addition level is significantly higher than predicted cheating for those correctly adding 7 pairs of two-digit numbers. The one exception is the Redistribution treatment in which there is no discernable difference between High and Low Performance types. But this appears to result from a rise in cheating by Low Performance types under redistribution rather than a drop in cheating by the High Performers.

Figure 4: Predicted Probability of Cheating for Treatments: 10% & 30% Tax



5 Discussion

The goal of this essay is to understand why we cheat. Tax compliance experiments are the means by which we identify the causal mechanisms associated with cheating. Our results are consistent with recent findings indicating heterogeneity in lying and cheating (Gneezy, Rockenback and Serra-Garcia 2013, Cappelen, Sorensen and Tungodden 2013, Hurkens and Kartik 2009). Subjects in our experiments who perform better on real effort tasks cheat more in this tax compliance game. This supports our central claim, Conjecture 1, that ability triggers cheating.

The cost of complying with taxation should affect cheating – our Conjecture 2. We observe cheating by Low and High Performance types in different tax treatments and with different earnings. Surprisingly, there is only weak evidence that cheating rises when the price of tax compliance increases. Performance type is a much better predictor of the likelihood of cheating.

Conjecture 4 suggests that High Performers cheat less in contexts where they perceive luck or status to play an important role in determining their success. When our subjects' earnings are associated with luck, there is no signifiant moderation in the cheating gap between high and low performance types. For those assigned to the high wage treatment, the correlation between Performance and tax cheating is essentially the same as in the treatment in which all wages were the same. For those assigned to the low wage treatment we actually see a significant increase in the positive correlation between Performance and cheating. In another treatment, half the subjects were randomly assigned to receive a lucky bonus after they competed their real effort task. The correlation between Performance and cheating remained positive and significant for both those receiving and not receiving the random bonus payment.

Context was not irrelevant in our experimental sessions. Implementing a more redistributive tax regime eliminated the difference in cheating between High and Low Performance types – our Conjecture 5. But this results because Low Performers cheat more rather than because High Performance types cheat less!

Conjecture 6 suggests that this predilection for cheating by High Performance types is not confined to tax compliance but rather reflects a general behavioral trait. Results from choices made in other games by the same subject confirm this is the case: High Performance types give less money in a conventional Dictator Game and also are more likely to cheat in a classic die game in which they privately report the results of tossing a die.

Implications Cheating is costly. Annual bribes in the world are estimated to exceed \$1 Trillion and these negatively affect provision of public services and increase inequality (International Monetary Fund 2016). There is growing evidence of the economic costs associated with cheating – for example, Balafoutas et al. (2015) on the economic efficiency costs associated with tax evasion in credence goods. It comes as no surprise that with globalization and technological advances in financial transactions, the rich availed themselves of tax havens and banking services that allow them to avoid paying taxes. A growing portion of the wealth of the richest Americans is held in off-shore banking havens (Zucman 2015).

Reducing cheating likely improves the functioning of markets and of our political institutions. The public and private sectors invest considerable resources aimed at reducing cheating. Achieving this goal, though, requires a better understanding of why, and under what circumstances, individuals cheat.

Our findings suggest why poorly-specified models of the decision to cheat can generate unintended, if not contradictory, policy outcomes. Most accounts of cheating ignore its heterogeneity in the population. We find though that High Performance types have much higher proclivities for cheating than do Low Performers. Policies designed to reduce cheating, but that ignore this heterogeneity, could have no effect, or possibly result in perverse outcomes. In fact, maybe there are circumstances in which we want to facilitate cheating!

Efforts to increase tax compliance are a case in point. A popular recent strategy, building on experimental evidence, is to promote timely tax compliance with framing strategies that appeal to intrinsic motivations or reputational concerns (Blumenthal, Christian and Slemrod 2001, Fellner, Sausgruber and Traxler 2013, Castro and Scartascini 2015). Many of these efforts (and the findings on which they are based) ignore the Low/High Performance heterogeneity in the taxpayer population. Our results though suggest that ignoring heterogeneity in the case of these policies might generate higher tax revenues but at a cost – higher levels of post-tax income inequality. Its not simply that High Performance types (who are often rich) cheat more at their taxes. But its also the case that Low Performance types, typically the poor, respond more positively to these framing appeals than the High Performance types, i.e, the rich.[14]

These results also concern the design of employment contracts. Our results are consistent with Gill, Prowse and Vlassopoulos (2013) who find that subjects who are more productive in their real effort tasks also cheat more. It is true that

[14]The asymmetry we find in our experiments resembles the heterogeneous effect that Dwenger et al. (2015) report for their "compliance" rewards treatments that consisted of either a social or private recognition for compliance with the German local church tax. They find that it has a positive effect on subjects who are baseline "donor" types (those who overpay in pre-treatment) and a negative effect on subjects who are baseline "evader" types (those who underpay in pre-treatment).

employee theft represents a serious cost to the economy. On the other hand, its not clear that simply adopting policies that eliminate the possibility of cheating would be optimal. In a competitive labor market, firms who want to attract High Performance types may need to create, or at least tolerate, opportunities for cheating. Employment contracts that make it impossible, or extremely costly, to cheat might not be in the best interests of competitive firms.

A related issue concerns the reduction of corrupt practices in the public sector. Evidence of corruption in the public sector is voluminous. An influential literature inspired by Becker and Stigler (1974), in particular, identifies raising wages as one strategy for reducing corruption in developing countries where public sector salaries are very low (Akerloff and Yellen 1990). These insights have had a significant impact on policies designed to reduce public sector corruption (Olken and Pande 2012). Our strong correlation between performance and cheating proclivities suggests that simply raising wages might result, at least in the short term, in perverse outcomes. The higher salaries (along with the opportunities to cheat) could entice High Performance types into public sector employment which could raise the level of corrupt activity. This might explain why some efforts at reducing public sector corruption by increasing wages in low-income developing countries have disappointed and, in some instances, had the opposite outcome.[15]

Cheating results in significant economic costs both in the public and private sectors. Cheating proclivities are strongly correlated with ability or performance. This has consequences for how we model cheating behavior and for policies designed to reduce its occurrence in the public and private sectors.

[15]A case in point is the experience the Ghana government had with their civil service salary reform. Raising the salaries of enforcement officers resulted in higher levels of transport-related bribes (Foltz and Opoku-Agyemang 2016)!

References

Abeler, Johannes, Daniele Nosenzo and Collin Raymond. 2016. "Preferences for truth-telling." Manuscript.

Akerloff, G. and J. Yellen. 1990. "The Fair Wage Hypothesis and Unemployment." *Quarterly Journal of Economics* 105(2):255–283.

Alesina, Alberto and George-Marios Angeletos. 2005. "Fairness and Redistribution." *American Economic Review* 95(4):960–980.

Allingham, M.G. and A. Sandmo. 1972. "Income Tax Evasion: A Theoretical Analysis." *Journal of Public Economics* 1:323–338.

Alm, James, Betty R. Jackson and Michael McKee. 1992. "Estimating the Determinants of Taxpayer Compliance with Experimental Data." *Economic Development and Cultural Change* 39(4):107–14.

Alm, James, Kim M. Bloomquist and Michael Mckee. 2015. "On the External Validity of Laboratory Tax Compliance Experiments." *Economic Inquiry* 53(2).

Alm, James and Michael McKee. 2006. "Audit Certainty, Audit Productivity, and Taxpayer Compliance." *National Tax Journal* 59(4):801 – 816.
URL: *http://search.ebscohost.com/login.aspx?direct=truedb=bthAN=23639630site=ehost-live*

Anderhub, V., S. Giese, W. Guth, A. Hoffmann and T. Otto. 2001. "Tax evasion with earned income - an experimental study." *FinanzArchiv* 58:188–206.

Andreoni, James. 1995. "Cooperation in Public Goods Experiments: Kindness or Confusion." *American Economic Review* 85:891–904.

Andreoni, James and John Miller. 2002. "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism." *Econometrica* 70(2):737–753.

Ariely, Dan. 2012. *The Honest Truth about Dishonesty: How We Lie to Everyone – Especially Ourselves.* Harper Collins.

Atkinson, Anthony B. 2015. *Inequality: What Can be Done?* Harvard University Press.

Balafoutas, Loukas, Adrian Beck, Rudolf Kerschbamer and Matthias Sutter. 2015. "The hidden costs of tax evasion.: Collaborative tax evasion in markets for expert services." *Journal of Public Economics* 129:14 – 25.
URL: *http://www.sciencedirect.com/science/article/pii/S0047272715001176*

Becker, Gary S and George J. Stigler. 1974. "Law Enforcement, Malfeasance, and Compensation of Enforcers." *The Journal of Legal Studies* 3(1):1–18.

Becker, Winfried, Heinz-Jurgen Buchner and Simon Sleeking. 1987. "The impact of public transfer expenditures on tax evasion: An experimental approach." *Journal of Public Economics* 34(2):243–252.

Benabou, Roland and Jean Tirole. 2011. "Laws and Norms." NBER Working Paper 17579.

Besley, Timothy, Anders Jensen and Torsten Persson. 2015. "Norms, Enforcement, and Tax Evasion." Centre for Economic Policy Research Discussion Paper 7592.

Blanco, C., J. Grant, N. M. Petry, H. B. Simpson, A. Alegria, S.-M. Liu and D. Hasin. 2008. "Prevalence and Correlates of Shoplifting in the United States:

Results From the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC)." *The American Journal of Psychiatry* 165(7):905–913.

Blumenthal, Marsha, Charles W. Christian and Joel Slemrod. 2001. "Do Normative Appeals Affect Tax Compliance? Evidence from a Controlled Experiment in Minnesota." *National Tax Journal* 54(1):125–138.

Bradly, J.C. 1987. "Income tax evasion and the tax schedule: Some experimental results." *Public Finance* 42(3):357–383.

Cappelen, Alexander W., Erik O. Sorensen and Bertil Tungodden. 2013. "Why do we lie?" *Journal of Economic Behavior Organization* 93:258–265.

Castro, Lucio and Carlos Scartascini. 2015. "Tax Compliance and Enforcement in the Pampas: Evidence from a Field Experiment." *Journal of Economic Behavior Organization* 116:65 – 82.

URL: *http://www.sciencedirect.com/science/article/pii/S0167268115001067*

Chan, K.S., S. Mestelman, R. Moir and R.A. Muller. 1996. "The voluntary provision of public goods under varying income distribution." *Canadian Journal of Economics* 29:54–69.

Chan, K.S., S. Mestelman, R. Moir and R.A. Muller. 1999. "Heterogeneity and the voluntary provision of public goods." *Experimental Economics* 2:5–30.

Cherry, Todd L., L.P. Frykblom and J.F. Shogren. 2002. "Hardness the Dictator." *American Economic Review* 92:1218–1221.

Cherry, Todd L., Stephan Kroll and Jason Shogren. 2005. "The Impact of Endowment Heterogeneity and Origin on Public Good Contributions: Evidence from the Lab." *Journal of Economic Behavior and Organisation* 57:357–365.

Cowell, Frank. 1990. *Cheating the Government: The Economics of Evasion.* MIT University Press.

Dubois, David, Derek D. Rucker and Adam D. Galinsky. 2015. "Social Class, Power, and Selfishness: When and Why Upper and Lower Class Individuals Behave Unethically." *Journal of Personality and Social Psychology* 108(3):436–449.

Durham, Yvonne, Tracy S. Manly and Christina Ritsema. 2014. "The Effects of Income Source, Context, and Income Level on Tax Compliance Decisions in a Dynamic Experiment." *Journal of Economic Psychology* 40:220–233.

Dwenger, Nadja, Henrik Kleven, Imran Rasul and Johannes Rincke. 2015. "Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany." MPI for Tax Law and Public Finance.

Engel, Christoph. 2011. "Dictator Games: A Meta Study." *Experimental Economics* 14:583–610.

Falkinger, Josef. 1995. "Tax Evasion, Consumption of Public Goods, and Fairness." *Journal of Economic Psychology* 16(1):63–74.

Fellner, Gerlinde, Rupert Sausgruber and Christian Traxler. 2013. "Testing Enforcement Strategies in the Field: Threat, Moral Appeal and Social Information." *Journal of the European Economic Association* 11(3):634–660.
URL: *http:https://dx.doi.org/10.1111/jeea.12013*

Fischbacher, U. 2007. "z-Tree: Zurich Toolbox for Ready-made Economic Experiments." *Experimental Economics* 10:171–78.

Fischbacher, Urs and Franziska Follmi-Heusi. 2013. "Lies in Disguise: An Experimental Study of Cheating." *Journal of the European Economic Association* 11(3):525–547.
URL: *http://dx.doi.org/10.1111/jeea.12014*

Fischbacher, Urs, Simon Gchter and Ernst Fehr. 2001. "Are people conditionally cooperative? Evidence from a public goods experiment." *Economics Letters* 71(3):397 – 404.
URL: *http://www.sciencedirect.com/science/article/pii/S0165176501003949*

Foltz, Jeremy D. and Kweku A. Opoku-Agyemang. 2016. "Do Higher Salaries Lower Petty Corruption? A Policy Experiment on West Africa's Highways." Manuscript.

Francis, G. 2012. "Evidence that Publication Bias Contaminated Studies Relating Social Class and Unethical Behaviour." *Proceedings of the National Academy of Science* 109.

Frank, Robert H. 2016. *Success and Luck*. Princeton and Oxford: Princeton University Press.

Gibson, Rajna, Carmen Tanner and Alexander F. Wagner. 2013. "Preferences for Truthfulness: Heterogeneity among and within Individuals." *American Economic Review* 103(1):532–548.

Gill, David, Victoria Prowse and Michael Vlassopoulos. 2013. "Cheating in the workplace: An experimental study of the impact of bonuses and productivity."

Journal of Economic Behavior Organization 96:120 – 134.
URL: *http://www.sciencedirect.com/science/article/pii/S0167268113002436*

Gneezy, Uri, Bettina Rockenback and Marta Serra-Garcia. 2013. "Measuring Lying Aversion." *Journal of Economic Behavior Organization* 93:293–300.

Gneezy, Uri, Stephan Meier and Pedro Rey-Biel. 2011. "When and Why Incentives (Don't) Work to Modify Behavior." *Journal of Economic Perspective* 25(4):1–21.

Heinemann, Friedrich and MartinG. Kocher. 2013. "Tax compliance under tax regime changes." *International Tax and Public Finance* 20(2):225–246.
URL: *http://dx.doi.org/10.1007/s10797-012-9222-3*

Hoffman, E., K. McCabe and V. Smith. 1996. "Social Distance and Other-Regarding Behavior in Dicatator Games." *American Economic Review* 86:653–60.

Holt, Charles A. and Susan K. Laury. 2002. "Risk Aversion and Incentive Effects." *American Economic Review* 92:1644–1655.

Hurkens, Sjaak and Navin Kartik. 2009. "Would I Lie to You? On Social Preferences and Lying Aversion." *Experimental Economics* 12:180–192.

International Monetary Fund. 2016. Corruption: Costs and Mitigating Strategies. Technical Report SDN/16/05 International Monetary Fund.

Kajackaite, Agne and Uri Gneezy. 2015. "Lying Costs and Incentives." Mimeo.

Levitt, Stephen and John List. 2007. "What Do Laboratory Experiments Measuring Social Preferences Reveal about the Real World." *The Journal of Economic Perspectives* 21(7):153–174.

Luttmer, E. and M. Singhal. 2014. "Tax Morale." *Journal of Economic Perspectives* (1):149–168.

Major, Brenda. 1994. From Social Inequality to Personal Entitlement: the Role of Social Comparisons, Legitimacy Appraisals, and Group Membership. Vol. 26 of *Advances in Experimental Social Psychology* Academic Press pp. 293 – 355.
URL: *http://www.sciencedirect.com/science/article/pii/S0065260108601562*

Major, Brenda and Maria Testa. 1989. "Social comparison processes and judgments of entitlement and satisfaction." *Journal of Experimental Social Psychology* 25(2):101 – 120.
URL: *http://www.sciencedirect.com/science/article/pii/0022103189900073*

Olken, Benjamin A. and Rohini Pande. 2012. "Corruption in Developing Countries." *Annual Review of Economics* 4(1):479–509.
URL: *http://dx.doi.org/10.1146/annurev-economics-080511-110917*

Piff, Paul K., Daniel M. Stancato, Stphane Ct, Rodolfo Mendoza-Denton and Dacher Keltner. 2012. "Higher social class predicts increased unethical behavior." *Proceedings of the National Academy of Science of the United States of America* .

Piketty, Thomas. 2014. *Capital in the Twenty-First Century*. The Belknap Press of Harvard University Press.

Posner, Eric A. 2000. "Law and Social Norms: The Case of Tax Compliance." *Virginia Law Review* 86:1781–1819.

Schechter, Laura. 2007. "Traditional trust measurement and the risk confound: An experiment in rural Paraguay." *Journal of Economic Behavior Organization* 62(2):272 – 292.
URL: *http://www.sciencedirect.com/science/article/pii/S0167268105002507*

Schurr, Amos and Ilana Ritov. 2016. "Winning a competition predicts dishonest behavior." *Proceedings of the National Academy of Sciences* .
URL: *http://www.pnas.org/content/early/2016/01/25/1515102113.abstract*

Slemrod, Joel. 2007. "Cheating Ourselves: The Economics of Tax Evasion." *The Journal of Economic Perspectives* 21:25–48.

Sonnemans, Joep, Arthur Schram and Theo Offerman. 1999. "Strategic behavior in public good games: when partners drift apart." *Economics Letters* 62(1):35 – 41.
URL: *http://www.sciencedirect.com/science/article/pii/S0165176598002031*

Spicer, Michael W. and Lee A. Becker. 1980. "Fiscal Inequity and Tax Evasion: An Experimental Approach." *National Tax Journal* 33(2):171–75.

Swanner, Jessica K. and Denise Beike. 2015. "Throwing You Under the Bus: High Power People Knowingly Harm Others When Offered Small Incentives." *Basic and Applied Social Psychology* 37(5):294–302.
URL: *http://www.tandfonline.com/doi/abs/10.1080/01973533.2015.1081851*

Teyssier, Sabrina. 2012. "Inequity and Risk Aversion in Sequential Public Good Games." *Public Choice* 151:91–119.

Trivedi, Viswanath Umashanker and Janne O.Y.Chung. 2006. "The Impact of Compensation Level and Context on Income Reporting Behavior in the Laboratory." *Behavioral Research in Accounting* 18(1):167–183.

Vincent, Lynne and Maryam Kouchaki. 2015. "Creative, Rare, Entitled, and Dishonest: How Commonality of Creativity in One's Group Decreases an Individual's Entitlement and Dishonesty." *Academy of Management Journal* .
URL: *http://amj.aom.org/content/early/2015/09/03/amj.2014.1109.abstract*

Zucman, Gabriel. 2015. *The Hidden Wealth of Nations: The Scourge of Tax Havens.* University of Chicago Press.