Florida Tax Review

Volume 22 Article 6

Rewarding Honest Taxpayers: An Experimental Assessment

Emily A. Satterthwaite
University of Toronto

FLORIDA TAX REVIEW

Volume 22 *2018* *Number 1*

REWARDING HONEST TAXPAYERS: AN EXPERIMENTAL ASSESSMENT

by

*Emily A. Satterthwaite**

ABSTRACT

Shrinking budgetary allocations for tax enforcement at the U.S. federal level have placed an unprecedented premium on low cost policies that promote voluntary tax compliance. In other jurisdictions, tax administrators have experimented with rewarding taxpayers for voluntarily complying with tax laws, but there has been an absence of reward-focused policy experimentation in the United States. To explore the efficacy of rewards among U.S. taxpayer populations, a multi-period online tax reporting experiment was conducted featuring a simple reward intervention: a token monetary amount pre-announced and provided to participants who were audited and found to have fully complied. The reward failed to increase average post-audit compliance levels as compared to the no-reward control condition, regardless of whether random audits or non-random (i.e., conditional on past detected evasion) audits were used. However, the reward treatment condition in combination with random audits was strikingly effective with respect to an alternative measure of tax compliance: "consistent

* Assistant Professor, University of Toronto Faculty of Law. Thank you to Leandra Lederman, Zachary Liscow, Rebecca Morrow, and participants at the 2017 Junior Tax Scholars' Workshop for valuable comments. Support from the University of Toronto Faculty of Law's Faculty Research Stipend and the Connaught Fund's New Researcher Award is gratefully acknowledged. Joffre Brunet, Alex Teoh, and Michael Hassar provided outstanding research assistance. All errors are mine.

compliance," or the outcome in which a participant voluntarily reports all of her income in each and every period of the experiment. When used in conjunction with random audits, the reward treatment caused consistent compliance to rise by 89% as compared to the no-reward control condition (statistically significant at the 5% level). These results suggest that pairing token monetary rewards with random audits may help maintain taxpayers' commitments to voluntary compliance over time. Such findings may justify conducting field experiments to better understand the effects of reward programs on real-world taxpayer populations.

Introduction 202
I. Related Literature 213
A. Rewards Generally 213
B. Tax Contexts 214
1. Lotteries 215
a. Laboratory Experiments 215
b. Field Experiments 217
2. Monetary Rewards 218
II. Description of Experiment 220
A. Experimental Design 220
1. Eligibility 220
2. Set-Up 222
a. Audit Rule Conditions 222
b. Reward Rule Conditions 224
c. Other Parameters: Instruction Screens 224
B. Rational Expected Utility Model Predictions 226
III. Results 228
A. Summary of Data 228
B. Effect of Audits and Rewards on Average Compliance 232
C. Effect of Audits and Rewards on Perfect Post-Audit Compliance 232
D. Effect of Audits and Rewards on Consistent Compliers 235
IV. Conclusion 237
Appendix 238

Introduction

In recent years, U.S. federal tax administration and enforcement resources have been drastically curtailed.[1] At the same time, complex and hastily adopted new tax legislation,[2] coupled with rapidly changing income-earning patterns of the gig economy,[3] have put the enforcement capacities of the Internal Revenue Service ("IRS") under increasing strain.[4] To respond to these developments and to deter evasion among taxpayers who may be emboldened by news reports of an agency stretched thin, interest in developing innovative policy interventions that can bolster voluntary tax compliance at low cost has grown. For example, a series of U.S-based field experiments that have focused on the design details and efficacy of audit threat letters, which are inexpensive

1. Since 2010, the IRS's budget has been slashed by $900 million, resulting in 21,000 fewer employees. *See* Bryan Naylor, *Now That the GOP Tax Bill Is Approved, the IRS Gets Busy*, NPR (Dec. 21, 2017), https://www.npr.org/2017/12/21/572326998/now-that-the-gop-tax-bill-is-approved-the-irs-gets-busy.

2. *See* 1 Nat'l Taxpayer Advocate, 2017 Annual Report to Congress vii–viii ("In recent weeks, there has been considerable discussion about how the IRS has been beaten down by continuing funding cuts and about concerns the agency is stretched so thin it will not be able to properly implement tax reform. . . . *The IRS absolutely needs more funding.* It cannot answer the phone calls it currently receives, much less the phone calls it can expect to receive in light of tax reform, without adequate funding.").

3. *See* Kathleen DeLaney Thomas, *Taxing the Gig Economy*, 166 U. Pa. L. Rev. 1415, 1434–35 (2018) ("Underreporting business receipts, which once was the primary source of tax evasion among small business owners, is thus becoming virtually impossible as third-party information reporting expands and the use of cash declines. . . . [M]any small [gig economy] business owners offset increases to their reported receipts by simply increasing their reported business deductions, resulting in little change to net income. The overstatement of deductions—previously a much smaller problem than understated receipts—has essentially become the new tax enforcement challenge for the IRS in the wake of recent technological advancements.").

4. *See generally* Brandon Debot et al., *Trump Budget Continues Multi-Year Assault on IRS Funding Despite Mnuchin's Call for More Resources*, Ctr. on Budget & Pol'y Priorities (Mar. 16, 2017), https://www.cbpp.org/sites/default/files/atoms/files/3-14-17bud.pdf.

to deploy and simple to adjust in response to research findings, address this demand for cost-effective new enforcement tools.[5]

This laboratory-style study was conceived of and designed in a similar spirit. It draws conceptual inspiration from proposals by academics to compensate taxpayers for the cost of being subjected to random audits[6] and empirical inspiration from recent policy initiatives outside the United States, where taxing authorities may face similar, or more severe, constraints. Hilke Brockmann and co-authors stated in 2016:

---

5. *See, e.g.*, Marsha Blumenthal et al., *Do Normative Appeals Affect Tax Compliance? Evidence from a Controlled Experiment in Minnesota*, 54 Nat'l Tax J. 125, 128–32 (2001); Ben S. Meiselman, *Ghostbusting in Detroit: Evidence on Nonfilers from a Controlled Field Experiment*, 158 J. Pub. Econ. 180 (2018); Joel Slemrod et al., *Taxpayer Response to an Increased Probability of Audit: Evidence from a Controlled Experiment in Minnesota*, 79 J. Pub. Econ. 455, 461 (2001); Ricardo Perez-Truglia & Ugo Troiano, *Shaming Tax Delinquents* (Nat'l Bureau Econ. Research, Working Paper No. 21264, 2015 & rev. 2018), https://www.nber.org/papers/w21264.pdf.

For research in the value-added tax context, see Dina Pomeranz, *No Taxation Without Information: Deterrence and Self-Enforcement in the Value Added Tax*, 105 Am. Econ. Rev. 2539, 2539–40 (2015) (studying the efficacy of audit threat letters in the context of a value-added tax); Francesco Drago et al., *Compliance Behavior in Networks: Evidence from a Field Experiment* (Inst. of Labor Econ., Discussion Paper No. 9443, 2015), http://ftp.iza.org/dp9443.pdf (studying the effect of audit letters on compliance with the Austrian TV tax). For recent surveys of the audit threat field experiment findings, see Michael Hallsworth, *The Use of Field Experiments to Increase Tax Compliance*, 30 Oxford Rev. Econ. Pol'y 658, 661–4 (2014); Joel Slemrod, *Tax Compliance and Enforcement: New Research and Its Policy Implications* (Mich. Ross Sch. of Bus., Working Paper No. 1302, 2016), https://deepblue.lib.umich.edu/bitstream/handle/2027.42/117359/1302_Slemrod.pdf?sequence=1&isAllowed=y.

6. *See* Joseph Bankman, *Tax Enforcement: Tax Shelters, The Cash Economy, and Compliance Costs*, 31 Ohio N.U. L. Rev. 1 (2005) (proposing compensating taxpayers for costs of random audits); Sarah B. Lawsky, *Fairly Random: On Compensating Audited Taxpayers*, 41 Conn. L. Rev. 161 (2008) (analyzing the merits of providing compensation in the context of random audits; discussing other proposals by policymakers and tax policy experts to compensate taxpayers for random audit costs); *see also* Joshua D. Rosenberg, *The Psychology of Taxes: Why They Drive Us Crazy, and How We Can Make Them Sane*, 16 Va. Tax Rev. 155 (1996) (addressing the use of tax holidays as rewards for taxpayers who disclose difficult-to-detect avoidance transactions).

"Pressed by high spending requirements and high political obstacles to tax increases, some governments have recently experimented with recovering some of these losses through a new *rewards approach* to tax compliance. The idea is to curb tax evasion by providing positive rewards for individual tax compliance."[7]

Prizes, lotteries, tax holidays, and other incentives have been the subject of field experiments conducted by governments in a number of countries, including Argentina,[8] Germany,[9] and Uruguay.[10] In the context of value-added taxes ("VATs"), evidence is emerging that receipt-based lotteries[11] have been effective in incentivizing consumers to ask

---

7. Hilke Brockmann et al., *Happy Taxation: Increasing Compliance Through Positive Rewards?*, 36 J. Pub. Pol'y 381, 382 (2016).

8. *See* Paul E. Carrillo et al., *Do Rewards Work?: Evidence from the Randomization of Public Works* (Inter-Am. Dev. Bank, Working Paper No. 794, 2017), https://publications.iadb.org/handle/11319/8246.

9. *See* Nadja Dwenger et al., *Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany*, Am. Econ. J.: Econ. Pol'y, Aug. 2016, at 203, 206.

10. *See* Thad Dunning et al., *Positive vs. Negative Incentives for Compliance: Evaluating a Randomized Tax Holiday in Uruguay* (May 8, 2015) (unpublished manuscript), https://ssrn.com/abstract=2650105 (using a field experiment of various taxpayer informational interventions to study "the effects of a randomized lottery in Montevideo, Uruguay, in which the municipal government raffles tax holidays to [randomly selected] good taxpayers who are current on past payments. . . . across four kinds of taxes (property, vehicle, sewage, and head) . . . reward[ing] them with a year free of tax payments"; noting that this type of tax-holiday lottery is "increasingly popular among municipal governments in the developing world").

11. *See* Marco Fabbri, *Shaping Tax Norms Through Lotteries*, 44 Int'l Rev. L. & Econ. 8, 8 (2015) ("[A] number of countries enacted policies to engage customers in contrasting [sic] VAT and RST evasion. . . . The government institutes a lottery and announces a prize. A serial number is printed on all sales receipts and the individual owning the receipt with the number corresponding to the lottery extraction is entitled to claim the prize. . . . The importance of the sales receipt for tax compliance comes from the fact that, in many countries, it represents the proof of the existence of a monetary transaction. . . . Once the invoice has been remitted, it becomes difficult for business owners to hide information regarding the business volume and taxable income.").

for VAT receipts and sellers to correctly report their sales for VAT purposes.[12]

However, research on rewarding taxpayers for voluntary compliance is still in its early stages.[13] As Leandra Lederman states in a recent paper that reviews empirical studies of tax enforcement and voluntary compliance: "There are few studies of what effects rewards would have in the context of tax compliance. . . ."[14] And, despite suggestions that experimenting with reward programs may be on the IRS's radar, no U.S.-based field experiments involving rewards appear to have been announced to date.[15]

This Article uses low-cost experimental methods to contribute to the literature on rewarding taxpayers for voluntarily complying with their tax reporting obligations. Specifically, it evaluates the effectiveness of providing honest taxpayers with nominal rewards following an audit in an online experimental setting. Might such reward programs represent an underappreciated enforcement tool and, as such, merit field experimentation?

---

12. *See* Jonas Fooken et al., *Improving VAT Compliance—Random Awards for Tax Compliance* 3–4 (Eur. Comm'n Tax'n Papers, Working Paper No. 51-2014, 2014), https://ec.europa.eu/taxation_customs/sites/taxation/files/resources/documents/taxation/gen_info/economic_analysis/tax_papers/taxation_paper_51.pdf.

13. *See* Benno Torgler, Tax Compliance and Tax Morale: A Theoretical and Empirical Analysis 19 (2007) ("[W]hile most studies focus on *punishment*, experiments have started to analyse the effects of *rewards* on tax compliance. . . . However, . . . the analysis of positive rewards . . . is still in its infancy. . . .").

14. Leandra Lederman, *Does Enforcement Reduce Voluntary Tax Compliance?*, 2018 BYU L. Rev. 627, 636 (2018) (concluding the sentence with "which presumably is not a 'high-interest' task"). Lederman's reference is to meta-analyses she discusses earlier in the article, which "reached different results with respect to whether a tangible reward for doing an interesting activity undermined or increased motivation." *Id.* at 633 (noting that "they generally found that a tangible, expected reward given simply for performing a particular task"; describing one meta-analysis as "finding small but statistically significant negative effect both for high-interest tasks and for tangible rewards overall").

15. *See* Carrillo et al., *supra* note 8, at 3 n.3 ("The IRS . . . is evaluating the introduction of rewards as part of their programs according to conversations held with IRS personnel in charge of the process.").

From a theoretical perspective, two related but separate debates in the vast academic literature on tax compliance animate this objective.[16] These debates feature the same two contestants in rotating roles: the carrot and the stick.[17] In the particular context of voluntary tax compliance, the carrot represents a subsidy: a reward, prize, bonus, chance to win a lottery, or other positive inducement for honest reporting of income. The stick represents a sanction. The most common forms of sanctions are audits and penalties, where penalties are assessed as a percentage of unreported income detected upon audit.

---

16. For a recent summary of this literature that specifically addresses the role of rewards, see generally Lederman, *supra* note 14.

17. *See* Bruno S. Frey & Lars P. Feld, *Deterrence and Morale in Taxation: An Empirical Analysis* 7 (CESifo Working Paper No. 760, 2002), http://www.cesifo-group.de/DocDL/760.pdf (noting role for both carrots and sticks). Notably, there is a large and fascinating academic and popular literature on the use of competing price instruments (sticks versus carrots) outside the context of tax compliance. *See, e.g.*, Ian Ayres, Carrots and Sticks: Unlock the Power of Incentives to Get Things Done 45–70 (2010). Closer to the tax context is a growing literature on the choice between subsidies versus taxes to address behavior that carries externalities or spillovers from the behavior that affect people other than the actor. Brian Galle, *The Tragedy of the Carrots: Economics & Politics in the Choice of Price Instruments*, 64 Stan. L. Rev. 797, 801, 825, 828, 831 (2012) [hereinafter Galle, *Carrots*). Within this literature, one of the contributions of Galle's 2012 paper in particular is to extend "the more sophisticated theoretical apparatus of the negative externalities literature to positive externalities . . . [through a consideration] . . . of whether penalties for failure to produce positive externalities would be as effective as, or better than, a subsidy." *Id.* at 832. The application of this analysis to the case of voluntary compliance, which has been shown to feature positive spillovers on others' choice to comply, is ripe for consideration. It may be particularly important in light of evidence that positive-externality tax compliance may systematically depart from the predictions of a rational-choice model (as in the case of post-random audit, bomb crater effects); this raises the related issue of how salient a reward for tax compliance should be. *See* Brian Galle, *Carrots, Sticks, and Salience*, 67 Tax L. Rev. 53, 54–55 (2013) (addressing the choice of carrots and sticks in settings in which "market actors are sometimes imperfectly rational"; arguing that "it is possible that salience can improve carrots by such an extent that they might actually become a viable policy option, even though under standard price theory sticks are almost always the better choice"). I pursue none of these issues in this Article, but I flag them as of great interest for future work.

In the first debate, the carrot and the stick are potential substitutes that may have differing capacities to promote tax compliance.[18] Although there is a body of research in psychology that has emphasized the efficacy of carrots in inducing desired behavioral outcomes,[19] the standard expected utility model of income tax compliance and most empirical research on its predictions features sticks.[20] The model conceptualizes individual income tax compliance as a choice between a risk-free (reported, after-tax income) asset and a risky (unreported, untaxed income) asset.[21] An increase in the probability of audit, or in the penalty for detected evasion, causes taxpayers to shift towards the risk-free asset of reported income.[22] This means that—all else equal—more sticks (in the case of audits), or bigger sticks (in the case of penalties), should increase tax compliance.[23] According to this conceptual approach, carrots in the form of rewards for compliant taxpayers can have the same effect: they change the price of compliance relative to evasion.[24]

In the second debate, the carrot and the stick are lumped together as a single potential villain. At issue is the efficacy of using prices in

---

18. *See* Josef Falkinger & Herbert Walther, *Rewards Versus Penalties: On a New Policy Against Tax Evasion*, 19 Pub. Fin. Q. 67, 69 (1991) (providing a theoretical analysis of rewards).

19. *See* Lars P. Feld & Bruno S. Frey, *Tax Compliance as the Result of a Psychological Tax Contract: The Role of Incentives and Responsive Regulation*, 29 L. & Pol'y 102, 110 (2007) (noting, "it appears to be common knowledge among psychologists that rewards lead to better outcomes than punishment").

20. *See* Lederman, *supra* note 14, at 631 (observing that "the tax-enforcement context typically focuses on fines rather than rewards").

21. *See* Michael G. Allingham & Agnar Sandmo, *Income Tax Evasion: A Theoretical Analysis*, 3 J. Pub. Econ. 323, 324 (1972) (the baseline model's exogenous parameters are actual income (which "is known by the taxpayer but not by the government's tax collector"), rate of tax on declared income, probability that the taxpayer will be audited, and a penalty rate (which is constrained to be higher than the tax rate) on undeclared income detected in an audit; the taxpayer's decision variable is declared, or reported, income); T.N. Srinivasan, *Tax Evasion: A Model*, 2 J. Pub. Econ. 339, 340 (1973).

22. *See* Allingham & Sandmo, *supra* note 21, at 325–26 (showing that the comparative statics of voluntary compliance are unambiguously positive for the following parameters: audit rates, penalties/fines upon detection, and, where included, detection rates upon audit).

23. *Id.*

24. *See* Galle, *Carrots*, *supra* note 17, at 801.

the form of carrots *or* sticks, as opposed to relying on the "intrinsic motivations" of individuals to pay their taxes honestly.[25] In this debate, the use of either a carrot or a stick risks crowding out taxpayers' commitments to voluntary tax compliance.[26] The crowd-out problem rears its head when a sanction or subsidy chills individuals' willingness to voluntarily comply, thereby unintentionally increasing rather than decreasing tax evasion.[27] Crowding-out might occur in a variety of circumstances, including when a specific individual decreases her compliance after directly experiencing an audit.[28] Or it might occur in a more

---

25. *See* Bruno Frey, *Crowding Out and Crowding In of Intrinsic Preferences*, *in* Reflexive Governance for Global Public Goods 78 (Eric Brousseau et al. eds., 2012); *see also* Feld & Frey, *supra* note 19, at 106 (describing a theory of a "psychological tax contract, [in which] punishment still plays a role in order to provide deterrence. But the satisfaction of taxpayers with what they get from the other contract party, that is, the government, mainly influences their tax morale."); *id.* at 105 ("The idea of intrinsic motivation is largely attached to psychology. A group of cognitive social psychologists have identified that, under particular conditions, monetary (external) rewards undermine intrinsic motivation. Giving of rewards for undertaking an activity has indirect negative consequences as rewards lead to the expectation of future rewards such that desired behavior is undertaken only if rewards are provided." (footnote omitted)).

26. *See* Lederman, *supra* note 14, at 629 (observing that some scholars "argue that enforcement ultimately will have the perverse effect of *reducing* voluntary tax compliance").

27. *See* Feld & Frey, *supra* note 19, at 105–6 ("External interventions undermine intrinsic motivation when they are perceived to be intrusive by the individuals concerned ('crowding-out effect')"; further, "[t]axpayers' reward from that contract must be understood in a broad sense going beyond pure exchanges of goods and services for the payment of a tax price. . . . A genuine reward is therefore obtained only if taxpayers as citizens have an inclusive, respectful relationship with the community. . . . As deterrence and tax morale interact, it would be counterproductive solely to rely on punishment or monetary (non-authentic) rewards, because tax morale can be undermined. A dynamic relationship results, in which deterrence, monetary rewards, fiscal exchange, but also decision-making procedures and the treatment of taxpayers play a role."); Daniel M. Kahan, *The Logic of Reciprocity: Trust, Collective Action, and Law*, 102 Mich. L. Rev. 71, 83 (2003).

28. *See* Lederman, *supra* note 14, at 690 (distinguishing between individual-level and more generalized deterrence of voluntary compliance

generalized way, when the use of sanctions or subsidies results in decreased average compliance rates within a particular community or across a broader population.[29]

Evidence of the former circumstance—individual-level crowding-out—has been observed in laboratory settings[30] and has been investigated in several field experiments.[31] Some of these studies have documented a "bomb crater effect" following audits, which is a particular form of crowding-out that is specific to the taxpayer who has been audited.[32] It refers to a taxpayer's reduced compliance (relative to her pre-audit compliance level) after experiencing an audit.[33] As explained

---

as a result of audits: "Note that the focus here [in a particular study] is on *specific deterrence*—that is, the effect on the audited taxpayer. The regime-level effect of audit rates is the focus of many studies, and many studies do not separate out the effects on those actually experiencing an enforcement action.").

29. *See generally id*.

30. *Id*. at 673 *et seq.* (describing laboratory experiments relating to post-audit compliance and investigating the bomb crater effect).

31. *Id*. at 683 (describing field experiments relating to post-audit compliance, noting that, in contrast to "[t]he laboratory experiments . . . [, which] generally focus on the effects of audits on all audited participants. By contrast, field experiments have raised the prospect of a post-audit decline in compliance only of the subgroup of individual taxpayers found compliant on audit").

32. *See, e.g.*, Luigi Mittone, *Dynamic Behaviour in Tax Evasion: An Experimental Approach*, 35 J. Socio-Econ. 813, 823–24 (2006) ("Even if the trends are highly unstable and apparently follow some sort of random walk, there is a sort of constancy in the rounds immediately after a fiscal audit, which is almost always followed by a systematic increase in tax evasion. This increase generally has its lowest peak in correspondence to the round immediately after the fiscal audit, and sometimes lasts for more than one round. This shall be called the 'bomb crater effect' . . . the subjects decide to evade immediately after a fiscal audit because they believe that it cannot happen twice in the same place (time).").

33. *See* Barbara Kastlunger et al., *Sequences of Audits, Tax Compliance, and Taxpaying Strategies*, 30 J. Econ. Psychol. 405, 417 (2009) ("It is shown that the effectiveness of audits and fines—suggested by the standard economic model as the most relevant determinants deterring from tax evasion—cannot completely be confirmed. Especially, the finding of the bomb crater effect, shows that, rather than increasing or strengthening compliance, audits can lead taxpayers to develop strategies to 'escape' and thus have the

by Francesco Guala and Luigi Mittone, the name evokes the following scenario:

> They say that troops under heavy enemy fire hide in the craters of recent explosions, for they believe it is highly unlikely that two bombs will fall exactly in the same spot within a short time period. Something similar seems to happen in the tax experiments: immediately after each audit, tax payments fall sharply (i.e. evasion increases).[34]

A recent experimental study found a bomb crater effect that was sensitive to the audit rule used by the tax authority: it appeared in response to random audits, but it did not appear in response to audits that were conditioned on prior detected evasion ("endogenous audits") and announced as such to taxpayers.[35] The bomb crater studies raise the question of whether measures used in combination with random audits might ameliorate their unintended post-audit consequences. Given the well-documented centrality of random audits to effective tax enforcement,[36]

---

opposite than expected effect."); Boris Maciejovsky et al., *Misperception of Chance and Loss Repair: On the Dynamics of Tax Compliance*, 28 J. Econ. Psychol. 678, 684–5 (2007) (finding post-audit compliance bomb craters in two audit experiments).

34. Francesco Guala & Luigi Mittone, *Experiments in Economics: External Validity and the Robustness of Phenomena*, 12 J. Econ. Methodology 495, 505 (2005).

35. *See* Emily Satterthwaite, *Can Audits Encourage Tax Evasion?*, 20 Fla. Tax Rev. 1, 58–59 (2016) ("The summary table [noting that for random audits, post-audit compliance dropped by 8%, significant at 1%; for endogenous audits, post-audit compliance increased by 3%, significant at 1%] highlights not only the magnitude of the bomb crater and deterrent effects but also that these effects are statistically significant—they matter even after accounting for the variance in the compliance rates that was observed in Figure 1 through Figure 3.").

36. *See* Lawsky, *supra* note 6, at 167–68 ("Although relatively few taxpayers are selected for random audits, random audits are key to IRS enforcement. The importance of these audits lies in their very randomness, because it is this randomness that permits the IRS to gather accurate information. . . . In

such measures could represent important enhancements to self-assessment tax systems.[37]

Taxpayer reward programs are natural candidates. The most recent and comprehensive empirical study (to the best of this author's knowledge) of the effectiveness of taxpayer reward programs found that their success is highly sensitive to their design.[38] As a result, the treatment experimental conditions explored in this Article were structured to test the effectiveness of the simplest and most easily-executed form of taxpayer compliance rewards: token monetary transfers in consideration for full compliance.[39]

In particular, this Article experimentally investigates the interaction between audits, rewards, and voluntary compliance over time in a U.S. taxpaying context.[40] It reports the results of a multi-period online tax reporting survey administered in May 2017 to approximately 400 U.S. participants on the Amazon Mechanical Turk ("MTurk") task completion

---

short, while relatively few taxpayers are picked for audit randomly, random audits provide information that is essential to tax enforcement.").

37. Cementing full and honest compliance as a strong norm among taxpayers has obvious importance to tax agencies. *See* Leandra Lederman, *The Interplay Between Norms and Enforcement in Tax Compliance*, 64 Ohio St. L.J. 1453, 1465 (2003); *see also* Jon S. Davis et al., *Social Behaviors, Enforcement, and Tax Compliance Dynamics*, 78 Acct. Rev. 39 (2003) (modeling the conditional cooperator approach incorporating norms of others).

38. *See* Carrillo et al., *supra* note 8, at 3, 28–29.

39. *See* Lawsky, *supra* note 6, at 207 (proposing a scheme of "nominal compensation" and noting that "[i]f demoralization costs can be eliminated or reduced by nominal compensation, some amount of compensation for random audits might be particularly desirable").

40. *See* Satterthwaite, *supra* note 35, at 1–3 (assessing the audit-rule sensitivity of the bomb crater effect). With respect to its U.S. focus, as discussed *infra* at Part II.A.1, the experiment carefully restricted the eligibility of participants to U.S. residents and sought in its framing and presentation to simulate the U.S. taxpaying environment. This is because few dynamic tax compliance experiments have been conducted with U.S. resident participants or in an experimental context designed to resemble the U.S. income tax system. An important exception is Alm and coauthors' 1992 experiment, discussed in Part I.B.2, *infra*.

platform.[41] The survey evaluated the performance of rewards in combination with both random and endogenous audits.[42]

Analysis of the data showed that the reward treatment was ineffective in improving average post-audit compliance rates under either audit rule. Rewards also failed to increase the likelihood that audits—of either type—were followed by perfect post-audit compliance in the period immediately following the audit. However, the reward treatment's performance stood out with regard to a specific metric relating to the durability of taxpayers' voluntary compliance over time: when used in combination with random audits, the reward increased the likelihood that participants in the experiment would fully report all of their income in each of the 60 periods of the experiment ("consistent compliance") by 89% as compared to the no-reward condition (statistically significant at the 5% level).

Average compliance rates exhibit deterioration over time in most experimental dynamic audit studies; they slope downward over time.[43] Therefore, one interpretation of the result regarding consistent compliance is that using rewards in combination with random audits may prevent the progressive *erosion* of taxpayers' commitments to voluntary compliance. Overall, the results suggest that reward programs featuring nominal monetary amounts may have some role to play in combination with random audits. Such findings may be useful in assessing whether reward-related field experiments involving real-world taxpayer populations are warranted.

The Article proceeds as follows. Part I reviews the literature on how rewards affect behavior in the tax context, particularly in connection with existing research on how audit rules may affect the dynamics of post-audit tax compliance. Part II describes the design of the online survey experiment and demonstrates that the prediction of the rational expected utility model is full evasion in all experimental conditions. Part

---

41. While the flexible MTurk platform is increasingly common in social science research, it has been used only rarely in the tax compliance context. *See* Satterthwaite, *supra* note 35, at 4, 9 (using the MTurk platform, and noting some of its advantages and disadvantages).

42. *See infra* at Part II.

43. *See* Mittone, *supra* note 32, at 822–25; Satterthwaite, *supra* note 35, at 53–55 (showing downward trend of average compliance over time for random audits).

III summarizes the data and reports the results of the experiment. Part IV briefly concludes.

I. Related Literature

A. Rewards Generally

The study of rewards' effectiveness in producing a desired behavior occupies a substantial literature. Numerous scholars have noted that while rewards may increase the measurable returns to performing a specific task, they may also decrease an individual's intrinsic motivation to perform that task.[44] The non-tax literature on rewards has examined their impact in a number of contexts, including blood donations[45] and "inspection games."[46] Summarizing some of this literature, Uri Gneezy

---

44. *See* Bruno S. Frey & Reto Jegen, *Motivation Crowding Theory*, 15 J. Econ. Surveys 589, 589 (2001); Bruno S. Frey, *On the Relationship Between Intrinsic and Extrinsic Work Motivation*, 15 Int'l J. Indus. Org. 427, 432 (1997) ("The more a *reward is contingent on the performance* desired by the principal, the more strongly the locus of control is shifted from intrinsic to extrinsic incentives, and the more is intrinsic motivation crowded out.").

45. Nicola Lacetera et al., *Will There Be Blood? Incentives and Displacement Effects in Pro-Social Behavior*, Am. Econ. J.: Econ. Pol'y, Feb. 2012, at 186, 188–90 (finding that paying for blood donations did not increase the quantity of blood donated); Lederman, *supra* note 14 (studying the effect of paying for contributions of blood as opposed to a policy of voluntary blood donations; presenting evidence that "paying for blood donations decreases the donations' quality because payment creates incentives for those with infectious diseases or drug addictions to conceal those conditions" (discussing Richard M. Titmuss, The Gift Relationship 245–46 (1970)). The example of blood donation is often cited as evidence that price mechanisms to encourage behavior do not always work, and for the proposition that the motivation behind the behavior must be carefully understood. Lacetera et al., *supra*, at 220–21.

46. *See* Daniele Nosenzo et al., *Encouraging Compliance: Bonuses Versus Fines in Inspection Games*, 30 J.L. Econ. & Org. 623, 631–33 (2013) (giving subjects the role of employer or employee in which the employee had the option to work or shirk in each period while the employer had the option of inspecting or not inspecting. The study separated the subjects into three groups: a control group, a group that received bonuses for compliance if inspected, and a group that received fines for non-compliance if inspected. The fine treatment

and coauthors state, "[m]onetary incentives [can] have two kinds of effects: the standard direct price effect, which makes the incentivized behavior more attractive, and an indirect psychological effect."[47] Their review suggests that providing rewards can introduce distrust to a relationship, making compliance less likely.[48] For greater specificity, however, this Part summarizes the literature on rewards that has arisen in the particular context of tax compliance.

B. Tax Contexts

Lederman concludes in her recent analysis of the interplay between enforcement and subsequent voluntary tax compliance that sanctions typically increase compliance consistent with the rational expected utility model.[49] She cites evidence indicating that sanctions may fall short of their goals in the immediate post-audit periods (e.g., bomb crater effects). She discusses studies that show that audits may induce lower post-audit compliance among self-employed or business taxpayers who are found compliant on audit. She argues that such a "result is consistent with the deterrence model: a positive outcome [of an audit, where no unreported income is found] is not much of a deterrent."[50]

It is plausible that the use of rewards will generate similar results. However, as Lederman emphasizes, this question has received comparatively little treatment.[51] Paul Carrillo and coauthors note in a

---

was effective in reducing shirking compared to the no-fine treatment, but the effect of the bonus on shirking was statistically insignificant).

47. Uri Gneezy et al., *When and Why Incentives (Don't) Work to Modify Behavior*, J. Econ. Persp., Fall 2011, at 191, 192 (noting that this psychological effect can sometimes have the opposite effect that the monetary incentive was intended to have).

48. *Id.* at 199.

49. Lederman, *supra* note 14, at 699 ("The notion that deterrence, perversely, will reduce tax compliance, generally is not consistent with the ample empirical evidence in the United States and elsewhere.").

50. *See id.* at 698 (generally showing that studies find that audits and audit threats increase tax compliance; increased penalties generally have little effect if not paired with increased audits; and only for taxpayers who themselves are audited and found compliant have some studies found that subsequent tax payments decrease).

51. *Id.* at 631; *see also* Brockmann et al., *supra* note 7, at 386 ("Although tax practitioners show a keen interest in positive rewards, the

recent paper, "While studying the effect of rewards has become conspicuous in the behavioral economics literature, it has been mostly absent from the burgeoning experimental empirical literature on tax compliance."[52] This section describes the existing studies.

1. Lotteries

Other than in the specific context of a VAT, the use of prizes in the form of lotteries to reinforce voluntary compliance has been found to increase tax compliance in laboratory and field experiments.[53] This literature suggests that lotteries with high-value prizes offered with relatively low probabilities of success are effective "because individuals tend to overweigh the low probability of winning the jackpot, and hence generally judge the attractiveness of lotteries by the size of the jackpot and not by the probability of winning it."[54] There are both laboratory and field experiments evaluating lotteries.[55]

a. Laboratory Experiments

Two recent laboratory studies examine the effectiveness of lottery-based rewards. In the first, Cecile Bazart and Michael Pickhardt experimentally test, using student subjects in Germany and France, the impact of

---

empirical tax compliance literature has largely ignored them so far." (citation omitted)); Barbara Kastlunger et al., *What Goes Around Comes Around? Experimental Evidence of the Effects of Rewards on Tax Compliance*, 39 Pub. Fin. Rev. 150, 151 (2011) ("Although the call for introducing positive incentives into tax policy has been made quite often, empirical evidence of the effects of rewards on compliance and tax evasion strategies is scarce." (citations omitted)).

52. *See* Carrillo et al., *supra* note 8, at 2.

53. *See* Brockmann et al., *supra* note 7, at 386.

54. *Id.* at 387 (citation omitted).

55. The literature review in this Part focuses on studies outside the value-added tax context, as those are not relevant to the current U.S. tax landscape. However, as noted by Bazart and Pickhardt, reward programs involving receipt-based value-added tax lotteries have largely found that such lotteries increase compliance. These programs incentivize accurate use of VAT invoices by supplier businesses by allowing consumers to submit purchase receipts as lottery tickets. Cecile Bazart & Michael Pickhardt, *Fighting Income Tax Evasion with Positive Rewards*, 39 Pub. Fin. Rev. 124, 126 (2011).

lottery winnings on compliant taxpayers. In their experiment, the probability of an audit was endogenously determined by the choice of the participant to contribute above a certain threshold.[56] They found that participants in the two countries behaved similarly and that positive rewards in the form of individual lottery winnings for fully compliant taxpayers have a strong positive impact on tax compliance, particularly for male taxpayers.[57]

In the second study, Hilke Brockmann and coauthors measured the impact of providing a reward to honest compliers in the form of a lottery ticket (with a 1 in 800,000 chance of winning 10,000 euros) or a "donation" (the ability to choose where tax revenues were directed).[58] They found that the reward decreased rather than increased tax compliance, although the effect was not statistically significant.[59] However, there was a substantial difference between male and female participants' reaction to the rewards.[60] The researchers suggested that providing a reward introduced the idea that paying taxes was a voluntary behavior and therefore weakened participants' sense of civic responsibility and their commitment to the normative obligation to pay taxes.[61]

---

56. *Id.* at 131–32 (additionally, tax revenues were returned in the form of public goods).

57. *Id.* at 145.

58. Brockmann et al., *supra* note 7, at 386, 399 ("Although tax lotteries seem to be popular, their revenue and compliance effects are unclear"; "positive rewards may, under certain conditions, crowd-out threat-based or norm-based motivations for tax compliance," showing that excessive positive rewards can deter tax compliance as they de-emphasize the threat of audits and fines and make tax compliance appear less obligatory. (citations omitted)); *id.* at 390–92 (assigning participants randomly to either a baseline, a donation, or a lucky (lottery) group; all participants had the same probability of being audited and the same fines for non-compliance.).

59. *Id.* at 392–93 (finding a result contrary to the researchers' initial hypothesis).

60. *Id.* at 393–97 (finding that the rewards caused a substantial increase in tax evasion among the male participants whereas a slight reduction in tax evasion was observed amongst female participants; theorizing that the reward treatments made participants much more aware of the risks of non-compliance and consider factors such as the probabilities of audits; this resulted in more risk-seeking behavior (non-compliance) being observed in male participants rather than female participants).

61. *Id.* at 397.

b. Field Experiments

A recent study by Nadja Dwenger and coauthors uses a "church tax" in Bavaria, Germany, to examine taxpayers' motivations for tax compliance in a system with limited enforcement.[62] The treatment group received rewards for voluntary compliance in the form of a lottery ticket (with a low probability of success); the lottery involved public recognition of honest compliance.[63] The researchers hypothesized that, because low-probability lottery rewards have such minimal impact on the expected payoff of individuals (i.e., their extrinsic motivation), any observed effect could be attributable to the effect of the reward in increasing individuals' intrinsic motivations to comply.[64] They found that those who were evading taxes prior to the introduction of the reward system were more likely to evade taxes after the system was implemented.[65] Similar to Brockmann and coauthors, the authors speculated the reward highlighted the voluntary nature of the tax system.[66] For those who were already compliant, the introduction of the reward had no effect on compliance (e.g., it did not cause a crowding-out effect), but it did increase the likelihood that a compliant taxpayer would voluntarily *increase* their payment to the church in the form of a donation.[67]

In addition, a study on the timing and rollout of lottery-based rewards found that rewards are more effective when they are socially salient.[68] Paul Carrillo and coauthors measured the impact of a reward program in Santa Fe, Argentina, in which individuals who paid their

62. *See* Dwenger et al., *supra* note 9, at 207 (describing the regime as limited enforcement because individuals self-report income information and, although the church has the ability to verify this information, it has never done so in the past).

63. *Id.* at 213 (noting that the chance of winning the lottery was 1/1000).

64. *Id.* (suggesting that intrinsic motivation could be affected in this context if the rewards would change an individual's perception about the voluntary nature of them contributing to a public good).

65. *Id.* at 227 (finding that the probability of evasion increased by 1.27%).

66. *Id.* at 206.

67. *Id.* at 227 (suggesting that it is important to understand whether an individual is intrinsically or extrinsically motivated before the introduction of a reward to better predict its effect).

68. *See* Carrillo et al., *supra* note 8, at 1.

municipal property taxes were entered into a lottery to win a sidewalk abutting their property.[69] They found a positive and persistent effect of winning the lottery on winners' future tax compliance.[70] With respect to those who did not win the lottery, they hypothesized that voluntary compliance would increase because the sidewalks highlight the reciprocity benefits of voluntary compliance (i.e., by emphasizing that taxpayers receive benefits from taxes).[71] In addition, or alternatively, sidewalk construction may foster pro-compliance spillovers through peer effects by raising the profile of voluntary compliers in one's neighborhood.[72] Indeed, they found that the neighbors of lottery winners who were not eligible to participate in the lottery (i.e., non-compliant taxpayers) were more likely to pay their taxes after the lottery win took place, suggesting the power of peer effects in well-structured reward systems, particularly at the municipal level. [73]

2. Monetary Rewards

A 2011 paper by Barbara Kastlunger and coauthors reported the effects on compliance of offering rewards to honest taxpayers (at two reward levels, as compared to a control condition of no rewards) in a laboratory experiment.[74] The paper hypothesized that "by rewarding compliance, the difference between the two options [rewards vs. penalties]

69. *Id.* at 10.

70. *Id.* at 15 (speculating that this was due to taxpayers seeing the benefits they received from their tax payments, increasing reciprocity for that individual).

71. *Id.* at 19 (leveraging these spillover effects may be a key way to increase the effectiveness of a reward program).

72. *Id.*

73. *Id.* at 22 ("These are relatively large, statistically significant, and robust effects. Again, it is almost unheard of reward programs having positive spillover effects on people other than the winners of the rewards.").

74. *See* Kastlunger et al., *supra* note 51, at 153–56 (providing for 60 periods in the experiment; all participants were informed that there was a 15% chance of being audited and the tax rate was 20%. Participants decided how much tax to evade in each period aiming to maximize their wealth. Participants who were found to have complied with an audit either received nothing (control group) or received 200 or 400 credits (treatment groups)).

decreases and tax compliance gains in attraction."[75] The authors found that "all-or-nothing" behavior increased when rewards for honest compliance were provided, meaning that participants either fully complied or reported no income.[76] The experiment also examined the behavior of participants after they had been audited and found a higher rate of compliance in subsequent periods among participants who received rewards than those in the control group who did not.[77] However, revenue collected overall was not affected by the use or the size of that reward.[78] This Article's results are largely consistent with the findings of Kastlunger et al.

Another study used experimental methods to test non-lottery-based monetary rewards.[79] James Alm and coauthors recruited U.S. resident university students to participate in an experiment studying the effects of four treatments: (1) a small, certain (conditional on full compliance) monetary reward; (2) a bigger but uncertain monetary reward in the form of a lottery ticket; (3) a group-wide reward in the form of a public good; and (4) a reward in the form of reduced audit probability in subsequent periods.[80] Voluntary compliance was higher in all the reward treatments as compared to the no-reward control condition, with the highest levels achieved in the conditions with the fixed certain reward

75. *Id.* at 154.

76. *Id.* at 162 (interpreting this finding to mean that when a reward is present taxpayers will either try to earn the reward or aim to receive additional income from tax evasion and forfeit the reward).

77. *Id.* at 162–63 (finding this result to be in line with the literature on operant conditioning).

78. *Id.* at 162 (noting that the impact of rewards on tax revenue has been inconclusive).

79. A 2003 experiment conducted by Benno Torgler with professional participants from Costa Rica is not included here. Although it tested the efficacy of a certain monetary rewards, it was "one-shot" (i.e., did not look at behavior over multiple reporting periods). It found that the compliance rate in the reward treatment (noting here the small sample size—only 13 participants in the treatment) was 100%. Benno Torgler, *Beyond Punishment: A Tax Compliance Experiment with Taxpayers in Costa Rica*, 18 Revista de Análisis Económico, No. 1, 2003, at 27, 50–56.

80. James Alm et al., *Deterrence and Beyond: Toward a Kinder, Gentler IRS*, *in* Why People Pay Taxes: Tax Compliance and Enforcement 311, 313–19 (Joel Slemrod ed., 1992).

and the lottery ticket, respectively,[81] notwithstanding that these conditions had the same expected value.[82]

There have also been experiments outside the context of rewarding individual-level compliance in response to audits.[83] For example, one of these studies explored the efficacy of various treatments that rewarded taxpayers by returning or redistributing all or a portion of the revenue collected.[84]

II. Description of Experiment

In May 2017, 404 participants on Amazon's MTurk portal completed a multi-period online tax reporting survey.[85]

A. Experimental Design

1. Eligibility

Following the best practices recommended by Kuziemko and coauthors for screening study participants in the MTurk environment, six different checks were embedded in the experiment's design to ensure that

---

81. *Id.* at 321–22 (noting that the increase in compliance was primarily due to changes in the frequency of extreme tax compliance behavior in which participants evaded all of the tax or complied fully).

82. *Id.* at 323; *see also* Julie H. Collins, *Commentary*, *in* Why People Pay Taxes, *supra* note 80, at 330, 330–32 (discussing the Alm et al. experiment and noting that not all five conditions (one control and four treatment) had the same expected value, so interpretation of relative levels of tax compliance is not straightforward).

83. *See* Marina Bornman & E M (Lilla) Stack, *Specific Rewards for Tax Compliance: Responses of Small Business Owners in Ekurhuleni, South Africa*, 13 eJ. Tax Res. 799 (2015).

84. *See* Martin Fochmann & Eike B. Kroll, *The Effects of Rewards on Tax Compliance Decisions*, 52 J. Econ. Psychol. 38, 38–40 (2016).

85. MTurk has plusses and minuses with respect to experimental surveys of the sort deployed here. On the plus side, the participant pool may be particularly appropriate because all MTurk workers receive independent contractor income from Amazon. While this is likely to be third-party reported, they are faced with a greater range of evasion possibilities because of the scope for claiming expenses as deductions from gross income. *See* Satterthwaite, *supra* note 35, at Appendix A.

(1) workers' participation would be executed in a careful and serious manner; (2) only U.S. resident individuals would be permitted to complete the survey; and (3) such individuals would be permitted to participate only once.[86]

First, eligible participants were limited to MTurk workers who have minimum MTurk performance ratings. For this study, it was possible to require that workers have a past MTurk completion rate of more than 95% on prior tasks. This requirement was adopted to increase the likelihood that the worker would be an individual who would take the task seriously and would continue with it to completion, if not to earn the wage then to maintain her high MTurk task completion rating, which preserves her eligibility for future tasks.

Second, participants were required to take a quiz after being introduced to the set-up (see below) but before progressing to Period 1. The quiz was structured as a multiple choice test, and passing and continuing with the task required correct answers to all the questions (which covered all of the key parameters of the experiment: tax rate, audit rate, penalty rate, number of periods, and confidentiality of information).[87] This check was designed to serve two purposes: first, to ensure that participants comprehended the setup of the survey and, second, to alleviate concerns about robots, or "bots."[88]

Third, as a supplemental layer of security against bots, workers were required to manually type in a code to a field on the last page of the task before they were able to claim their reward.[89]

Fourth, to make sure only U.S. residents were eligible to complete the task and serve as participants in the experiment, only workers with U.S.-based internet protocol addresses were permitted to participate.[90] Further, Amazon's functionality was used to directly select

---

86. *See* Ilyana Kuziemko et al., *How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments*, 105 Am. Econ. Rev. 1478, 1478–98 (2015).

87. Participants who had errors could re-take the quiz before proceeding.

88. A bot is a hacker or hacker-created algorithm that can breach the MTurk portal and exploit it by clicking through to completion and illegitimately claiming a wage

89. In this task, the worker was required to type the word "banana" before their results could be approved for payment.

90. Due to the widespread use of virtual private networks, however, this last provision would not have been sufficient on its own.

U.S.-based workers. This means that only those workers who submitted directly to Amazon (as is required when signing up for an MTurk account) a U.S. address and social security number, and who had those sources of identification verified by Amazon, were eligible. Last, the task included a special screen at the beginning of the survey that required the worker to certify that he or she is a resident of the United States.

Fifth, to prevent participants from skipping steps by advancing without completing all fields, popup windows were used as "progress blockers" when a worker attempted to advance prematurely.

Finally, to address the risk that a given MTurk worker would complete the survey more than once, thereby biasing the results, Amazon's "once only" functionality was enlisted to ensure that workers who had accepted the task in the past were blocked from accepting it again.

2. Set-Up

The experiment had two separate dimensions. Each was randomly and independently assigned to participants.[91]

a. Audit Rule Conditions

The first dimension varied the audit rule by assigning one of two audit rule conditions intended to correlate approximately with those used by tax agencies: random audits versus non-random targeted (endogenous) audits.[92] The control audit rule used an independent and identically distributed random variable with a 10% uniform distribution to select participants for audit.

The treatment (endogenous) audit rule also used random audits as the baseline selection method but, in the event that such an audit revealed a participant to have unreported income, the participant was re-audited in the immediately succeeding period. Otherwise, the participant was simply returned to the pool that was subject to random audits. However, to constrain the average audit rate to be equal across

91. The experiment added a second layer of experimental conditions (reward-no reward) experimental conditions to the setup (random audit-endogenous audit) used in a prior paper on more general patterns of post-audit compliance. *See* Satterthwaite, *supra* note 35, at 47–49.

92. The two audit rules were nearly identical to that used in Satterthwaite, *id.* at 47–48.

the two experimental conditions (at approximately 10%), the endogenous audit rule also included a cap on the total number of repeat audits experienced by participants in the endogenous audit treatment: a participant could not be audited in the endogenous treatment condition more than six times (out of 60 periods), to artificially create an audit rate that was approximately equal to that of the control audit condition.

There were two reasons for this cap on the total number of repeat audits experienced by participants in the endogenous audit condition. First, as suggested above, constraining the parameter of the audit rate across experimental conditions is important for causal inference. Interpreting compliance results across experimental conditions (under a random audit rule and an endogenous audit rule) is difficult when one of the parameters that typically influences compliance varies along with the audit rule.[93] Holding the audit rate constant across experimental conditions allows the researcher to isolate the effects of the treatment condition. Any other approach would require disentangling the effect of the *type* of audit treatment from the effect of an increase in the average *rate* of audit.

Second, the cap evokes a "fixed auditor resource" approach to distributing audits across the participant population. Rather than simply drawing audits according to a random variable, as in the control audit condition, the endogenous audit rule follows the approach of Collins and Plumlee, whose study compared joint tax reporting and work-effort decisions under differing tax rates and audit scheme conditions.[94] To ensure that the audit probability remained fixed across the conditions, they held the absolute number of audits fixed at two audits for every ten taxpayers. They described this as "somewhat analogous to the situation in . . . [which] audit resources are fixed by a binding budget constraint."[95]

Participants assigned to the treatment (endogenous) audit condition were exposed to the following additional language as part of the instruction screens (in addition to that described below): "If you are

---

93. *Id.* at 49 ("Simply performing repeat audits on taxpayers found to be noncompliant implies that the overall audit rate would be much higher than if a random audit rule at a fixed percentage were followed.").

94. Julie H. Collins & R. David Plumlee, *The Taxpayer's Labor and Reporting Decision: The Effect of Audit Schemes*, 66 Acct. Rev. 559, 561 (1991) ("fix[ing] at a constant level the total amount of auditing that the taxing authority can employ in each [audit] scheme [random, cutoff, or conditional]").

95. *Id.* at 561.

audited and have unreported income, you will be flagged for repeat audits in the future. This means you will be more likely to be selected for a future audit than if you have reported all income honestly."[96]

b. Reward Rule Conditions

The second dimension of the experiment had two reward experimental conditions, which were also randomly assigned to participants at the outset of the survey. The control rule was simply not giving a reward under any circumstances, and the treatment rule provided a token reward that was transferred with certainty (i.e., not via a lottery) to participants who were audited and found to have fully reported their income for that period. There was no cap on the number of times a reward could be received following a "no change" audit.

Participants assigned to the treatment reward condition were exposed to the following additional language as part of the instruction screens (in addition to that described below): "If you are chosen for audit and have honestly reported all your income, you will receive an immediate reward of **$10** after the audit. This reward will be added to your accumulated net income. You are eligible to earn the honest compliance reward each time you are audited" (boldface type in MTurk experiment).[97]

c. Other Parameters: Instruction Screens

To ensure valid causal inference, the key experimental parameters were held constant across all experimental conditions: number of tax periods

---

96. The second sentence was added to clarify the "flagging" language from Satterthwaite, *supra* note 35, at 47.

97. The presence of a reward increases the expected value of compliance, which arguably complicates the causal interpretation of my results. However, the reward was set such that it had a vanishingly small economic value to participants in expectation. A fully compliant participant who was audited the average number of times (six) would receive $60 in rewards within the experiment. This payment would increase her accumulated net income by about one-tenth of 1% (average accumulated net income was $44,733). Because the performance-based bonus on MTurk was set at 0.001% of participants' accumulated net income over all 60 periods, the upper bound monetary value to the MTurk worker is $0.0006. *See* Part II.B, *infra*.

(60); average audit rate (10%[98]); tax rate (30%[99]); penalty rate (100% of tax on unreported income[100]); average income subject to tax (randomly selected in the $800 to $1200 range; specified to be non-third-party-reported[101]); and the performance-based incentive structure of the experiment (participants earned a "bonus" in addition to the fixed $3.00 reward for completing the task that was based on the aggregate

98. From instruction screens of survey: "Even if you report the full amount of your income, you may be selected for audit. Audits occur with an average probability of 10 percent. This means that the audit rate, averaged over all taxpayers, is 10 percent. Audits examine only your reporting in the immediately prior tax period. If you do not report all of your income and you are audited, any income you did not report ('unreported income') will be detected with certainty. If you are audited, there will be a short waiting period (between 5 and 30 seconds) that simulates the time and hassle required for an audit."

99. From instruction screens of survey: "Your tax rate is 30 percent. This means that, for $100 of income reported, $30 is payable as tax and $70 is your 'after-tax' amount."

100. From instruction screens of survey: "Any unreported income will be taxed at 30 percent. You will also owe a penalty for evasion at a rate of 100 percent of taxes on unreported income. Example: if you are audited and you have $100 of unreported income, you will owe $30 in taxes and a $30 penalty for evasion, for a total payment to the government of $60."

101. From instruction screens of survey:

> In each tax period, you will be given an amount of earned income ("income"). Think of this income as: Being earned from self-employment activities, such as consulting services or other work as an independent contractor [and] [b]eing additional, or extra, income over and above what you may have from wages, investments or other regular sources that are automatically reported by the payer to the government. This amount will range between $800 and $1200 (in increments of $10). . . . You are required by law to report the income you earn each tax period on your tax return. However, none of your income in this HIT [a "human intelligence task," or what projects are called on MTurk] is subject to third-party reporting to the government. This means that the payer of this income is not required to inform the government that you were paid the income, but they can do so voluntarily.

after-tax-and-penalty net income earned over all periods[102]). The bonus component was tracked by a "reward calculator," which appeared as part of a summary page following each period of the experiment and listed the accumulated net income of the participant.

B. Rational Expected Utility Model Predictions

The rational expected utility model of tax compliance provides a benchmark for assessing the tax reporting behavior observed in experimental situations.[103] Under this model, a participant is assumed to make choices that maximize her "expected utility." "Utility" is a catch-all term for the benefits derived by an individual from a given choice. In this experimental context, it is assumed that utility is equivalent to the monetary compensation to the MTurk participant for taking the survey. Arguably, this assumption is more realistic than in a real-world environment, where non-monetary preferences may significantly influence behavior. (For instance, a person may harbor a sense of civic duty relating to the payment of taxes, and so may derive pleasure from paying taxes in a fashion not captured by monetary compensation outcomes.)

In addition, under uncertain conditions, a rational actor is assumed to maximize the "expected value" of her utility (sometimes referred to as "expected utility"). This expected value is equal to the sum of: the utility of each possible outcome multiplied by the probability of that outcome occurring. This sum yields the average utility across outcomes.

---

102. From instruction screens of experiment:

> Your earnings from completing this HIT have two components: a fixed component as listed on MTurk [and] a bonus component based on your performance in the HIT. This is measured by your "net income". Your "net income" means your income for each period, less taxes and penalties. Your bonus will be calculated at the end of the HIT as a fixed percentage of your net income totaled across all periods (your "accumulated net income"). The higher your accumulated net income, the higher your bonus will be at the end of the HIT.

103. This is a summary of the standard model offered in Allingham & Sandmo, *supra* note 21, at 324–26.

In this experiment, uncertainty arises in connection with the likelihood of being audited. The only variable under a participant's control is her compliance rate (i.e., the proportion of her income that she chooses to report). The Appendix shows mathematically that the optimal solution to the expected utility problem is determined by the probability of being audited.[104] If there is a greater than 50% chance of being audited, reporting more of one's income increases expected utility. Thus, it is optimal to report 100% of one's income in this scenario. If the chance of being audited is below 50%, however, increasing one's reporting rate *decreases* expected utility, so it becomes optimal to report no income. Finally, if the probability is exactly 50%, a rational actor is indifferent as to the percentage of their income that they report. This yields four predictions corresponding to each of the four combinations of experimental conditions (random audit-no reward, random audit-reward, endogenous audit-no reward, and endogenous audit-reward).

> 1. In the random audit condition with no reward, it is never optimal to report (any amount of) one's income.

Given the tax rate and penalty for being caught cheating in the random audit condition, a rational actor would not report any income if the probability of being audited is below 50%. Since the audit rate is known to be 10%, the rational expected utility model predicts that participants will report zero income in each period.

> 2. On the other hand, the endogenous audit condition with no reward creates uncertainty that can lead to an optimal choice of reporting all of one's income.

Since the mechanism of the endogenous audit rule is not disclosed to participants, they cannot precisely know the risk of audit after they have been "flagged." Thus, they must rely on their perception of the risk of audit in making their reporting decisions. If that perceived risk is sufficiently high (e.g., above 50%), they will choose to report all of their income.

> 3. The reward should have no effect in the random audit condition.

---

104. *See* Appendix, *infra* (calculating and explaining the predictions of the rational expected utility model in the context of the experimental setup).

The effect of the reward is proportional to income in a given period. Since it is so small, even at the lowest possible income of $800, it should have no impact on the reporting decision. In the no-reward condition, an expected utility-maximizer would report all her income only if the probability of audit was above 50%. In the reward condition, that threshold is lowered—at the most—to about 49%. Because the audit probability remains constant at 10%, the optimal strategy (of reporting no income) remains the same.

> 4. The reward should have little to no effect in the endogenous audit condition.

The reward should not affect predicted behavior in the endogenous audit condition. It is highly unlikely that a participant in the endogenous audit condition would estimate the likelihood of being audited with such precision that the effect of the reward (at most 1%) would affect behavior.

III. Results

A. Summary of Data

Four hundred and four participants completed the survey. One was dropped from the sample for failing to correctly answer the screening questions that tested participants' understanding of the instructions. Another two participants were removed for reporting income higher than the amount assigned to them in the first period.[105] Thus, 401 subjects remained in the sample for analysis (Tables 1 and 2).

Figure 1 shows that compliance rates in each of the four combinations of experimental conditions bear little correspondence with the predictions of the standard expected utility model, which predicts a strategy of complete evasion. Even in the "pure" control (random audits, no reward), average compliance exceeds 60% in every period.

Figure 2 summarizes how compliance changes over time in the experiment. All four experimental conditions have downward sloping

105. For all three observations, an error occurred that was likely related to the browser being used. When used with the (recommended and required) Chrome browser, the experiment included restrictions to prevent both of these situations.

Table 1: Assignment to Experimental Conditions

	No Reward	Reward	Total
Random audits	94 (23.4%)	105 (26.2%)	199 (49.6%)
Endog. audits	102 (25.4%)	100 (24.9%)	202 (50.4%)
Total	196 (48.9%)	205 (51.1%)	401 (100%)

Table 2: Breakdown of Compliance by Experimental Condition (95% normal-based confidence intervals in brackets)

	Random audits		Endogenous audits	
	No reward	Reward	No reward	Reward
Compliance	0.71 [0.64, 0.77]	0.77 [0.71, 0.83]	0.84 [0.79, 0.89]	0.88 [0.84, 0.92]
Percentage of periods with full compliance	56.6% [48.8, 64.3]	67.9% [60.8, 75.0]	71.5% [65.3, 77.8]	80.1% [75.0, 85.3]

Figure 1:

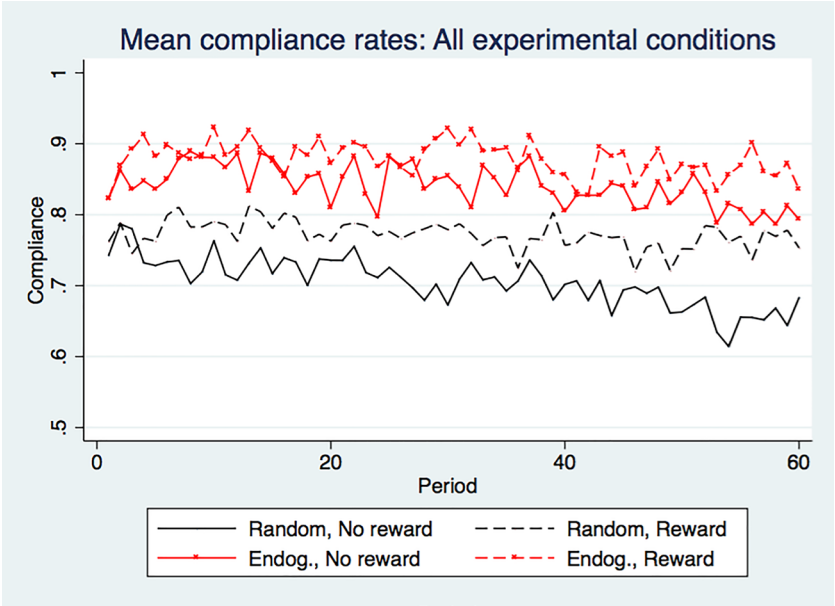
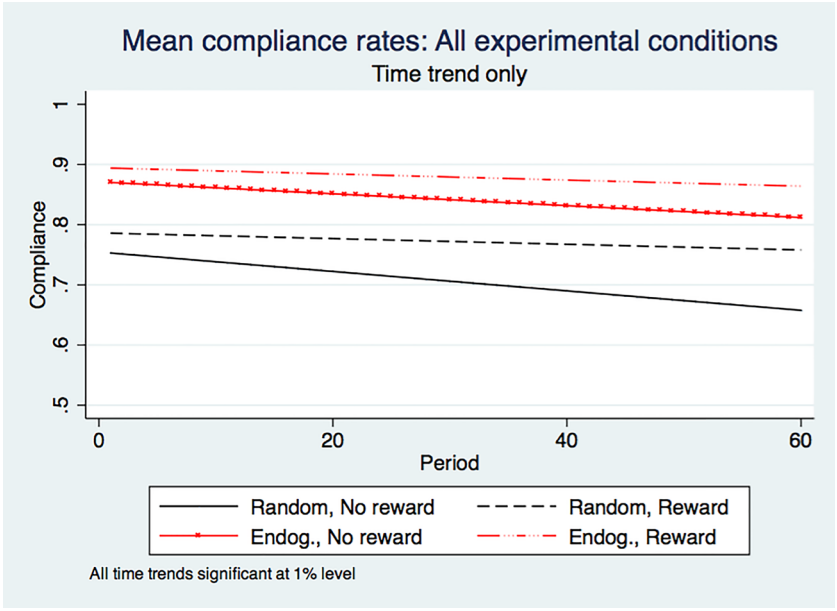


Figure 2:



compliance time trends (and, in each condition, the time trend is significant at the 1% level). This is consistent with a pattern of participants "learning" over time, in the sense that whatever they are learning helps them converge to the rational (expected utility) response of zero reporting. Additionally, in the endogenous condition, a participant may exhaust the capped number of audits if she is audited and found to be noncompliant more than six times. Audits will, in this instance, tail off and this naturally can be expected to reduce compliance.

By contrast, in the random audit condition, participants have the opportunity to "learn by doing" based on their experience with evading: as the accumulated net income from evading surpasses the probabilistic downside of being penalized and assessed for back taxes upon audit, they will evade more.

Figure 3 summarizes the audit rule results. Consistent with prior studies on the bomb crater effect, Figure 3 shows that the audit rule does have a significant effect. There is a gap between the confidence regions that persists over time (and is statistically significant). Here, endogenous audits foster higher average compliance than random audits. It is notable that the higher compliance rate for endogenous audits is downward-sloping but fairly linear: that is, it does not drop off even though in many

Figure 3:

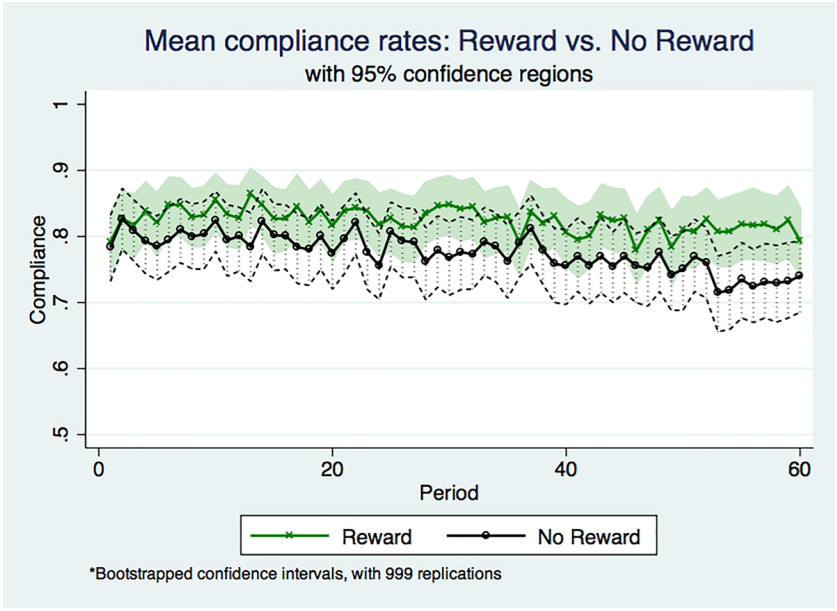
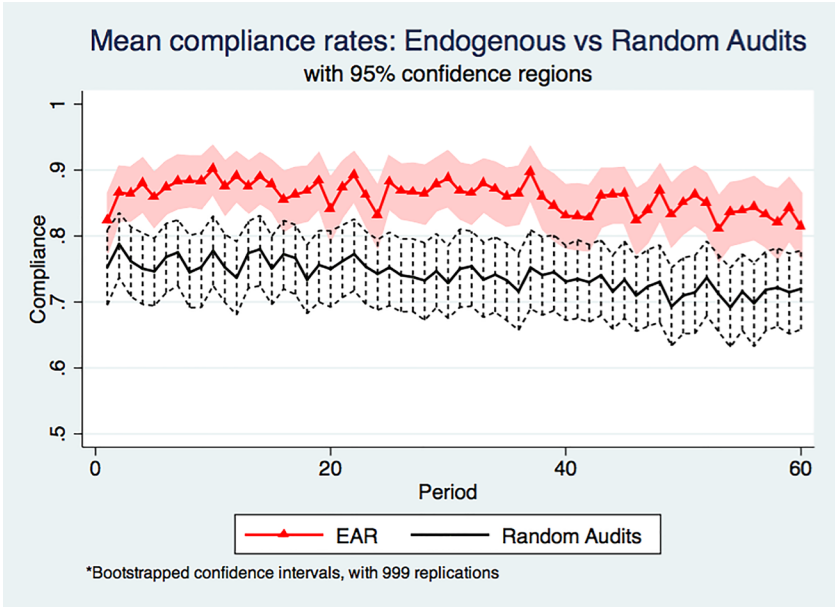


Figure 4:



cases the upper limit on audits may have been reached (i.e., at the cap of six). One interpretation of this trend might be that the experience of being repeat-audited under the endogenous audit rule fosters durable compliance behavior.

Figure 4 summarizes the reward rule results. It separates out average compliance by reward condition and shows that there is significant overlap between the confidence regions for the reward versus the no-reward condition. As confirmed in the regressions below, this indicates rewards are not effective in increasing average compliance.

B. Effect of Audits and Rewards on Average Compliance

Below, the impact of the audit and reward experimental conditions on compliance is reported using a linear regression model. The unit of analysis is the participant-period.

Table 3 indicates that post-audit decreases in compliance—the bomb crater effect—occurs in response to random but not endogenous audits. (This replicates on a new sample of participants the result of Satterthwaite (2016)). Specifically, the coefficients on β_1 (random audits—no rewards) are negative and significant at 1%; the coefficient on β_2 (endogenous audits—no rewards) are positive and significant at 1%. However, the reward treatment has no effect on overall compliance or post-audit compliance under either audit rule.

To interpret Table 4, note that the effect of an audit on compliance in the post-audit period is equal to the sum of the coefficient for "Post-Audit" and the coefficient of all applicable interactions. The table shows that random audits cause a significant (at the 1% level) post-audit bomb crater in compliance of approximately 8%, and endogenous audits cause a significant (again at the 1% level) *increase* in post-audit compliance of about 3% (each as compared to the non-post-audit baseline for random and endogenous audits, respectively). The reward condition has an insignificant effect on compliance, regardless of the audit rule.

C. Effect of Audits and Rewards on Perfect Post-Audit Compliance

To discern the effect of audits on the decision to fully comply post-audit (honest compliers), the compliance measure for each participant-period was transformed into a binary variable that takes the value "1" when post-audit voluntary compliance is 100% for that participant-period and

Table 3: Effect of Audits on Next-Period Compliance (outcome variable: compliance rate)

	(1) Experimental condition + interactions	(2) Including time trends
Constant	0.713***	0.762***
	(0.0349)	(0.0322)
Endogenous	0.126**	0.106**
	(0.0434)	(0.0404)
Reward	0.0684	0.0328
	(0.0469)	(0.0443)
Endog. × reward	–0.0289	–0.00737
	(0.0581)	(0.0549)
β_1: Post-audit[A]	–0.0776**	–0.0768**
	(0.0250)	(0.0250)
β_2: Endog × post-audit	0.115**	0.110**
	(0.0350)	(0.0346)
β_3: Reward × post-audit	–0.0115	–0.0116
	(0.0347)	(0.0347)
β_4: Endog. × reward × post-audit	–0.0130	–0.0112
	(0.0484)	(0.0479)
Time trends by experimental cond.	No	Yes
	24060	24060

Standard errors in parentheses.

[A]Post-audit indicates period following a random audit, i.e., is equal to zero following an endogenous audit.

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Table 4: Marginal Effects of Audits on Compliance (using Model 2 of Table 3, which includes time trends)

	Marginal effect of audit
Random audits, no reward	$\beta_1 = -0.0768$**
Random audits, reward	$\beta_1 + \beta_3 = -0.0894$
Endog. audits, no reward	$\beta_1 + \beta_2 = +0.0322$**
Endog. audits, reward	$\beta_1 + \beta_2 + \beta_3 + \beta_4 = +0.0094$

zero otherwise. A logistic regression was then used to determine how being audited in the previous period affects the likelihood of perfect post-audit compliance.

Table 5 indicates that the effect of random audits on the likelihood of perfect post-audit compliance is negative and significant (at 5%). By contrast, an endogenous audit increases the likelihood that a participant will choose perfect post-audit compliance (also significant at 5%). The reward does not have a significant impact on the likelihood of perfect compliance, under either audit rule. Table 6 summarizes the odds ratios that can be derived from Table 5's (specification (2)) logistic regression.

Table 5: Logistic Regression of Binary Transformation of Compliance (outcome variable: perfect compliance)

	(1) Experimental condition + interactions	(2) w/ time trends
Endogenous	1.852**	1.659*
	(0.422)	(0.388)
Reward	1.659*	1.417
	(0.386)	(0.343)
Endog. × reward	0.968	1.108
	(0.318)	(0.380)
Post-audit[A]	0.774*	0.775*
	(0.0881)	(0.0883)
Endog × post-audit	1.652*	1.654*
	(0.336)	(0.334)
Reward × post-audit	0.851	0.847
	(0.137)	(0.137)
Endog. × reward × post	1.164	1.174
	(0.346)	(0.348)
Time trends by experimental cond.	No	Yes
N	24060	24060

Exponentiated coefficients; Standard errors in parentheses.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Table 6: Odds of Perfect Compliance Following an Audit, by Experimental Condition (from Table 5, specification (2))

	Odds ratio
Random audits, no reward	0.774*
Random audits, reward	0.657
Endog. audits, no reward	1.282*
Endog. audits, reward	1.276

D. Effect of Audits and Rewards on Consistent Compliers

So far, the results indicate that the reward condition does not meaningfully affect post-audit compliance, either positively or negatively. However, the preceding analysis scrutinizes compliance decisions on a period-by-period basis for each participant (i.e., using the participant-period as the unit of analysis). This obscures the overall strategies pursued by participants across periods and leaves open the question of whether rewards, in combination with either audit rule, impact the likelihood that individuals will choose to report 100% of their income in each and every period of the experiment (e.g., behave as "consistent compliers" or engage in "consistent compliance").

One hundred twenty of the 401 subjects in the experiment (29.9%) behaved as consistent compliers. The distribution of consistent compliers across experimental groups is shown in Table 7. The difference in the proportion of consistent compliers between the reward and no-reward conditions in the presence of random audits, but not endogenous audits, is significant at the 5% level (denoted by *).

Table 7: Percentages of Consistent Compliers

	No reward	Reward	All	Difference	P-value (1-sided Fisher's exact test)
Random audits	22.3%	35.2%	29.1%	+12.9%*	0.032
Endog. audits	28.4%	33.0%	30.7%	+4.6%	0.291
All	25.5%	34.1%	29.9%	+8.6%*	0.037

To further investigate the possible link between the reward treatment and consistent compliance, a logistic regression (Table 8) assessed whether the reward was effective in increasing the likelihood of consistent compliance in conjunction with either or both audit rules. The first specification is the basic regression; the second includes participant gender, race, and age covariates. The magnitude of the estimate of the effect of rewarding compliance following a random audit (the first line of results) increases slightly from 1.89 to 1.96 when demographic controls are added. Its statistical significance remains the same (5%).

The results in Table 8 show that the reward was effective (at the 5% significance level) in increasing the likelihood of consistent compliance when a random audit rule is used. The reward had no significant effect on consistent compliance when the endogenous audit rule was used.

Table 9 quantifies the results described in Table 8 using odds ratios. It shows that the reward condition increased consistent compliance by 89% relative to the no-reward condition (significant at 5%).

Table 8: Logistic Regression: Experimental Conditions on Consistent Compliance

	(1) Allhonest	(2) Allhonest
Allhonest		
Reward	1.891*	1.961*
	(0.607)	(0.654)
Endog.	1.381	1.330
	(0.457)	(0.459)
Endog. × reward	0.655	0.757
	(0.291)	(0.352)
Female		2.365***
		(0.554)
Race		Yes
Age		1.032**
		(0.0112)
N	401	401

Standard errors in parentheses.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Table 9: Odds Ratios of Consistent Compliance by Experimental Condition

	Odds ratio
Random audits, no reward	(Control) = 1
Random audits, reward	Control × reward = 1.891*
Endog. audits, no reward	Endogenous = 1.381
Endog. audits, reward	Endog. × reward = 1.711

Rewards also increased the likelihood of consistent compliance under the endogenous audit treatment, but the effect did not achieve statistical significance.

IV. Conclusion

This Article addresses the question of whether participants' propensities to comply with an income tax over multiple periods is sensitive to the use of rewards. An experimental condition in which a (pre-announced) reward was provided following a random audit that yielded a "no change" result failed to increase average post-audit compliance levels as compared to the no-reward control condition, regardless of the audit rule used. However, the reward condition in combination with a random audit rule was effective with respect to an alternative measure of tax compliance: "consistent compliance," or the outcome in which a participant voluntarily reports all of her income in each and every period of the experiment. When used with random audits, the reward treatment caused consistent full compliance to rise by 89% as compared to the no-reward condition (statistically significant at 5%). This result suggests that pairing token monetary rewards with random audits may hold previously unrecognized potential to maintain taxpayers' commitments to voluntary compliance over time and, to this extent, field experimentation to explore the robustness of these effects under real-life taxpaying circumstances may be warranted.

Appendix: Predictions of the Rational Expected Utility Model

Let:

U_i : Utility of participant *i* for the experiment
U_{it} : Utility of participant *i* in period *t*
Y_{it} : Income of participant *i* in period *t*
Y_{it}^{NET} : After-tax income of paticipant *i* in period *t*
p : Probability of audit
r_{it} : Compliance rate; the proportion of income participant *i* reports in period *t*

The experiment is designed so that participants' rewards for their performance are measured by their accumulated net incomes. Under the assumption that participants' utilities from the experiment are equivalent to the sum of their net incomes across all 60 periods. Utilities can be expressed as:

$$U_i = \sum_{t=1}^{60} U_{it} = \sum_{t=1}^{60} Y_{it}^{NET}$$

Optimizing under uncertainty and finding the indifference point

Under uncertainty, a rational actor will maximize the expected value of her utility. Expected utility is equal to the sum of: the utility of each possible outcome multiplied by the probability of that outcome occurring. In this experiment, the uncertainty arises in connection with the likelihood of being audited (p), so the expected utility of a participant can be expressed as:

$$E[U_i] = E\left[\sum_{t=1}^{60} Y_{it}^{NET}\right] = \sum_{t=1}^{60} E[Y_{it}^{NET}] \quad (1)$$

Where,

$$E\left[Y_{it}^{NET}\right] = (1-p)Y_{it}^{NO\,AUDIT} + pY_{it}^{AUDIT} \quad (2)$$

Maximizing the expected utility in each period results in maximizing utility overall. Thus, the single-period decision can be generalized to a multi-period setting.

In this Article's survey experiment, the tax rate is 30%, and the penalty for unreported income is 100% of tax owing.[106] Thus, there is an effective tax rate of 60% on unreported income and we can express (2) for a given value of income as:

$$E\left[Y_{it}^{NET} \mid Y_{it} = Y\right] = (1-p)\left[(1-r_{it})Y + 0.7r_{it}Y\right] + p\left[0.4(1-r_{it})Y + 0.7r_{it}Y\right] \quad (3)$$

Simplifying and combining like terms:

$$E\left[Y_{it}^{NET} \mid Y_{it} = Y\right] = \left[1 - 0.6p + (0.6p - 0.3)r_{it}\right]Y \quad (4)$$

The only variable under a participant's control is her compliance rate (i.e., the proportion of her income that she chooses to report, or r_{it}). Taking the derivative of the expected utility function (1) with respect to r_{it} shows how the expected utility responds to changes in the compliance rate.

$$\frac{\partial E[U_{it} \mid Y_{it} = Y]}{\partial r_{it}} = (0.6p - 0.3)Y \quad (5)$$

Setting the derivative equal to zero, this function shows that the optimal solution is determined by the probability of being audited (p). Specifically,

$$\begin{aligned} &\text{If } p > 0.5 \rightarrow r_{it}^{*} = 1 \\ &\text{If } p < 0.5 \rightarrow r_{it}^{*} = 0 \\ &\text{If } p = 0.5 \rightarrow r_{it}^{*} = x \ \forall \ x \ \epsilon [0,1] \end{aligned} \quad (5A)$$

106. The instructions read: "participants must pay all taxes on unreported income plus a penalty equal to 100% of taxes on unreported income."

If there is a greater than 50% chance of being audited, increasing compliance increases expected utility. Thus, it is optimal for a participant to report 100% of her income. If the chance of being audited is below 50%, as it is in both experimental conditions, increasing one's compliance rate *decreases* expected utility, so it becomes optimal to report no income. Finally, if the probability is exactly 50%, a rational actor will be indifferent as to what percentage of their income they report (they might roll a dice to make a compliance decision).

Informally generalizing the indifference point

The above result can be generalized. Because the indifference point was calculated by setting the derivative equal to zero, and the effective tax rate on unreported income ("penalty rate") is 60%:

$$p^{Indiff} = \frac{0.3}{0.6} = \frac{Tax\ Rate}{Penalty\ Rate} \tag{6A}$$

Multiplying both the numerator and denominator by income Y yields:

$$\begin{aligned} p^{Indiff} &= \frac{Tax\ Rate}{Penalty\ Rate} X \frac{Y}{Y} \\ &= \frac{Taxes\ paid\ when\ reporting\ all\ income}{Penalty\ for\ reporting\ no\ income} \end{aligned} \tag{6B}$$

The effect of the Endogenous Audit Rule

The experimental design implies that the optimal strategy for participants assigned to the endogenous audit rule cannot be determined analytically because it depends on the subjective information possessed by the participants. Participants purposefully are given limited information: they are told that if they are caught underreporting their income, they "will be 'flagged' for audit in subsequent periods," yet it is specified that the "average audit rate is 10 percent." Participants are not told the actual mechanisms by which the endogenous audit rule ("EAR") works, namely that cheaters will be audited again in the following period, but that the number of audits is capped at six. Thus, the optimal strategy depends on their *perceived* risk of audit, which is in turn dependent on figuring out the rules that govern the EAR.

However, some insight can be gained into the rational expected utility-maximizing behavior of the participants under the EAR. Assume that participants are maximizing their income in each period, and thus the only effect of the EAR is through the perceived risk of audit. Put differently, assume that participants make their reporting decision in each period independently, aside from the fact that they use the results of previous periods to infer the probability that they will be audited in the current period. This assumption can be illustrated with an example of what it is ruling out. We are assuming that a participant will *not* consider future periods when deciding whether to cheat in the current period—e.g., that participants do not systematically change their strategies in response to the EAR. (An example of such a change would be to maximize current net income *and* minimize the future risk of being audited.) Under that assumption, the equation and results in (5) and (5A) can be generalized to

$$\frac{\partial E\left[U_{it} \mid Y_{it} = Y\right]}{\partial r_{it}} = (0.6\tilde{p}_{it} - 0.3)Y \tag{7}$$

$$\begin{aligned} &\text{If } \tilde{p}_{it} > 0.5 \rightarrow r_{it}^{*} = 1 \\ &\text{If } \tilde{p}_{it} < 0.5 \rightarrow r_{it}^{*} = 0 \\ &\text{If } \tilde{p}_{it} = 0.5 \rightarrow r_{it}^{*} = x \ \forall \ x \ \epsilon[0,1] \end{aligned} \tag{7A}$$

Where $\tilde{p}_{it}$ is the risk of audit as perceived by participant *i* in period *t*.

Another channel, using the indifference point

The above informal generalization of the indifference point could offer another channel through which participants' perceptions could affect their reporting decision. As we found,

$$p^{Indiff} = \frac{Tax\ Rate}{Penalty\ Rate} = \frac{Taxes\ paid\ when\ reporting\ all\ income}{Penalty\ for\ reporting\ no\ income}$$

Increasing the penalty for cheating decreases the indifference point. Thus, if being "flagged" in the EAR is interpreted as simply increasing the cost of cheating (conditional on being caught), it would lower the point at which the participant decides to report all of their income.

The effect of the reward

The reward slightly changes the expected utility of reporting 100% of one's income. Expressing this in terms of equation (3) with $r_{it} = 1$,

$$E\left[Y_{it}^{NET} \mid Y_{it} = Y, r_{it} = 1\right] = 0.7Y + 10p \tag{8}$$

Naturally, this will not change the solution for values of p above 0.5, where it was already optimal to report 100%. It will simply lower the point at which it becomes optimal to switch from reporting nothing to reporting 100%. To determine that value, we can solve for the indifference point (the value of p where a rational actor is indifferent across all compliance rates). Using equations (4) with $r_{it} = 0$, and (8):

$$E\left[Y_{it}^{NET} \mid Y_{it} = Y, r_{it} = 0\right] = E\left[Y_{it}^{NET} \mid Y_{it} = Y, r_{it} = 1\right]$$

$$(1 - 0.6p)Y = 0.7Y + 10p$$

$$p^{Indiff} = \frac{0.3}{0.6 + \frac{10}{Y}}$$

The result aligns with the intuition developed above: as a participant's income increases, the effect of the reward decreases, and the indifference point moves towards 0.3/0.6 = 0.5, which is the indifference point without the reward. In our experiment, income is a random value between 800 and 1200, with a mean of 1000. Using these values, we can illustrate the (negligible) effect on the indifference point:

$$\begin{aligned} &If\ Y = 800, \quad && p^{Indiff} = 0.4898 \\ &If\ Y = 1000, \quad && p^{Indiff} = 0.4918 \\ &If\ Y = 1200, \quad && p^{Indiff} = 0.4932 \end{aligned}$$

Note that this result is the same in the endogenous audit condition. However, there it would be expressed in terms of perceived probability.