Appalachian
STATE UNIVERSITY

Department of Economics Working Paper

Number 15-12 | October 2015

Behavioral effects of tax withholding on tax compliance: Implications for information initiatives

Christian Vossler
University of Tennessee

Michael McKee
Appalachian State University

Department of Economics
Appalachian State University
Boone, NC 28608
Phone: (828) 262-2148
Fax: (828) 262-6105
www.business.appstate.edu/economics

Behavioral effects of tax withholding on tax compliance: Implications for information initiatives[1]

Christian A. Vossler and Michael McKee

PRELIMINARY: PLEASE DO NOT CITE WITHOUT PERMISSION

[1] McKee is a professor in the Department of Economics, Appalachian State University, Boone, NC. Vossler is a professor in the Department of Economics, University of Tennessee, Knoxville, TN. This research was undertaken in partial fulfillment of IRS contract TIRNO-09-Z-00019. The views expressed are those of the authors and do not reflect the opinions of the IRS or of any researchers working within the IRS. We have, however, benefited from several discussions with Kim Bloomquist and Alan Plumley at the IRS. We thank Mick Jones and David Bruner for their work related to software development for the experiments. We further thank Jens Schubert and Nathan Murray for their research assistance. Earlier versions were presented at the ASSA meetings in 2014 and at several universities.

Behavioral effects of tax withholding on tax compliance: Implications for information initiatives

ABSTRACT

Using a framed field experiment with working adults and deliberate tax framing, this study reports on the effects of tax withholding on subsequent individual tax reporting behavior. We find interesting behavioral asymmetries related to tax withholding position, in particular that tax underreporting is increasing in the level of expected as well as unanticipated tax under-withholding, but is largely invariant to the level of tax over-withholding and unexpected decreases in liability. Two information initiatives we explore – group compliance information and information related to fiscal exchange – serve to affect tax reporting in part through its influence on withholding. A third information initiative – a service that resolves uncertainty over tax liability – decreases the level of evasion by over twice as much for those who have under- versus over-withheld. Using information from an extensive taxpayer questionnaire, we find several interesting associations between taxpayer characteristics and experimental tax reporting behavior.

JEL Classifications: H21, H26, C91, C92

Keywords: tax withholding; tax information services; social norms; tax reporting and enforcement; experimental methods; framed field experiment

1. Introduction

A central feature of many individual income tax systems is that there is tax withholding at the source of earnings, before salaries are paid. Tax withholding at source has been in effect in the U.S. since World War II, and generally felt to have led to increases in tax revenue. This revenue effect is in part due to the fact that withholding decreases tax underreporting opportunities. Nevertheless, the U.S. Internal Revenue Service (IRS) estimates that there is still a considerable amount of underreporting of the individual income tax. The gross underreporting "tax gap" was estimated to be $235 billion and $57 billion for the individual income tax and the self-employment tax, respectively, in tax year 2006 (Bloomquist et al., 2013). Using techniques of experimental economics, the focus in this study is on the interplay between tax withholding and tax reporting, with emphasis on the role of information initiatives (e.g. a service that resolves tax liability uncertainty). Our subject pool consists almost entirely of adults in the workforce rather than the more typical student population used in laboratory experiments, and further the experiment uses deliberate tax system framing.

Despite the prominence of income tax withholding, and importantly that taxpayers have some discretion over the amount of tax they have withheld, most work on tax compliance has abstracted away from the tax withholding decision. Focusing on empirical work, Clotfelter (1983), in one of the more comprehensive studies using field data (from the Taxpayer Compliance and Measurement Program, or TCMP) found that there is a positive correlation between withholding and tax reporting. Those who under-withhold are more likely to underreport. This result is confirmed by Chang and Schultz (1990) who also use TCMP data. Martinez-Vazquez, et al. (1992) use a laboratory experiment to investigate the withholding effect, but find little evidence to support a relationship between tax withholding position and

reporting. In the lab experiment of Schepanski and Shearer (1995), they assigned subjects to one of four situations: Expected Refund, Unexpected Refund, Expected (tax) Payment, and Unexpected (tax) Payment. Subjects in the Unexpected Payment setting were those most likely to be noncompliant. The main theoretical explanation the authors use to explain the observed effects of withholding position from their study and previous work is tied to prospect theory (Kahneman and Tversky, 1979). That is, if taxpayers focus on their current asset position, those in an over-withholding state would frame the reporting decision as a "gain" – and thus be less inclined to take a gamble by underreporting taxes – relative to those facing a tax payment who would instead frame the decision as a "loss".

There are some shortcomings that possibly limit the interpretation of results from previous studies. Previous field data studies analyze cross-section data and do not control for possible simultaneity between withholding and reporting. That is, one possibility is that taxpayers intending to underreport simply are those who choose to under-withhold. Further, with tax return data is it is not possible to disentangle the effects of being in an expected versus unexpected over or under-withholding position. Turning to the experimental work, the previous experiments have involved placing subjects into hypothetical choice scenarios. The authors justify their use of a hypothetical setting (reference is made to Grether and Plott, 1979 and to Lichtenstein and Slovic, 1973) but these references are to very much dated works. Without any financial incentives, empirical and theoretical evidence suggests that true preference revelation is unlikely in settings related to public good provision (Carson and Groves, 2007). Third, field and lab studies to date have focused on the binary outcome comply or not comply.

Our experimental setting introduces a tax withholding stage at the beginning of the tax year and a filing stage at the end of the tax year. In an induced-value setting subjects earn

income, elect a withholding level, and file a tax return by reporting tax liability – a continuous choice. Our experiments use real financial rewards with respect to tax withholding and tax reporting decisions. We allow the subjects to choose whether they will be in an (expected) refund or tax payment situation at the time of filing, but introduce tax liability shocks to gain insight on expected vis-à-vis unexpected withholding position. We introduce a withholding cost to reflect the deferred spending or foregone interest associated with pre-payment of taxes. We also introduce a penalty for under-withholding, to reflect a liquidity constraint. These are asymmetric – the withholding cost is less than the under-withholding penalty. This allows us to investigate competing motives for withholding decisions and the interaction between the withholding and tax reporting decisions. The experimental setting also introduces, as treatments: a taxpayer assistance service that resolves liability uncertainty; information on group tax compliance; and, a fiscal exchange.

Our focus in this study is on the effects of tax withholding and how this interacts with the individual taxpayer's innate attitude toward taxation and government (what are referred to as taxpayer segments) and with the provision of taxpayer information services. Attitudes are expected to be influenced by the fiscal exchange (the benefits the taxpayer perceives arising from taxes paid) and social norms (the tax reporting behavior of others). This set of interactions is of interest since tax withholding is an important part of the tax system and a taxpayer's attitude toward paying the full share of taxes is likely to be affected by the net tax position (additional taxes owed versus tax refund due) she finds herself in at the time taxes are to be filed.

Briefly, we find interesting asymmetries related to tax position, in particular that tax underreporting is increasing in the level of (possibly) expected tax under-withholding, but is largely invariant to the level of tax over-withholding. Although this confirms the basic result

regarding under-withholding and compliance found elsewhere in the literature, we have thus found that the magnitude of the under-withholding level is further of importance to tax compliance. Unexpected increases in tax liability serve to increase tax underreporting, whereas econometric evidence suggests either a null or a small, negative effect between an unexpected tax decrease and an increase in underreporting. Further, we find that better information on tax liability provided by an information "service" reduces tax under-reporting by more than twice as much for those in an under-withholding versus an over-withholding position.

Our data further show behavior that is consistent with other results reported in the literature continue to hold when tax withholding is included in the tax reporting setting: tax evasion decreases with enforcement effort, social norms of compliance, and perceived fairness of the fiscal exchange. Taxpayer experiences (from outside the lab) and characteristics are strongly tied to experiment behavior. Risk-averse individuals, women and older persons are less prone to underreporting. Participants with great underreporting opportunity in actuality underreport taxes to a greater extent in the experiment.

2. Theoretical Framework

The basic economic theory model of tax compliance (Allingham and Sandmo, 1972; Yitzhaki, 1974) characterizes a situation where a taxpayer faces a tax reporting "gamble" where she assesses the tradeoffs between the risks of penalty with the benefits of a lower tax payment. Here, we first briefly describe the theory model of Vossler and McKee (2013), and its implications for liability information services, which we used to inform the experimental design. Then, we extend this framework to consider the effects of social interactions, in particular social norms related to compliance and fairness at it related to the fiscal exchange. Finally, we provide

some intuitive discussion of how tax withholding is expected to affect tax compliance. We leave more formal modeling of the withholding-reporting nexus to future work.

The withholding decision is made prior to the filing decision, but it is conditional on the taxpayer's planned behavior at the filing stage (by backward induction) and so we will begin by examining the reporting decision. At the reporting stage a risk-neutral taxpayer chooses what to report on one or more "line items" on the tax form.[2] We assume that the taxpayer considers directly the tax liability associated with her line item reports which allows us to generally characterize the optimal decision regardless of whether the line item is associated with a credit, deduction, reported income, or otherwise. The audit probability is p, audits are completely random and independent of whether other persons are audited or the reported tax liability. Audits on tax returns perfectly reveal unpaid taxes separately for each line item on the tax form. In addition to being liable for unpaid taxes upon audit, there is a constant per-unit penalty $\beta > 0$ assessed on unpaid taxes.[3]

The actual tax liability on one or more line items is uncertain, and there may be an liability information service available to partially or fully resolve the uncertainty. Let x_l^0 denote the actual tax liability associated with line item l.[4] From the perspective of the taxpayer, tax liability is a random variable x_l with distribution function $F(x_l)$, which is assumed to have positive density $f(x_l)$ on the interval $[a_l, b_l]$. It is assumed that x_l^0 lies within the interval, i.e. the true tax liability is considered probable. Further, assume that there are institutional or other constraints on the range of amounts the taxpayer is allowed to enter, such that reports lie in the interval $[\underline{a}_l, \bar{b}_l]$, with $\underline{a}_l \leq a_l \leq b_l \leq \bar{b}_l$.

[2] To be clear, we use the term "line item" to denote any tax form entry that the taxpayer has discretion over what to report. For simplicity, we rule out simple mathematical errors that, to a large extent, are automatically discovered by the tax authority and not subject to penalty.
[3] Largely consistent with the literature, upon audit there is no refund or bonus associated with over-paid taxes.
[4] Note that the liability may be negative, such that taxpayers receive a refund.

For each line item on the tax form the taxpayer chooses a tax liability to report, denoted R_l. The optimal reporting problem is then one of choosing a vector of tax liabilities $\boldsymbol{R} = \{R_1, \dots, R_L\}$ in order to minimize expected costs:

[1] $$\min_{\boldsymbol{R}} \sum_l \left\{ R_l + p\left\{ (\beta+1) \int_{R_l}^{b_l} (x_l - R_l) f(x_l) dx_l \right\} \right\}.$$

The optimal reporting choice for a particular line item, R_l^*, is implicitly defined by

[2] $$1 = p(\beta+1) \int_{R_l^*}^{b_l} f(x_l) dx_l \qquad \forall l.$$

The interpretation is that the taxpayer minimizes cost by equating the marginal cost of taxes reported with the expected marginal cost of the audit. The first-order necessary conditions can instead be written as

[2'] $$F(R_l^*) = 1 - \frac{1}{p(\beta+1)} \qquad \forall l.$$

An interior solution exists for R_l^* on the interval $[a_l, b_l]$ if $\frac{1}{p(\beta+1)} < 1$. Otherwise, there is a corner solution $R_l^* = \underline{a}_l$, i.e. the taxpayer engages in maximum tax evasion.[5] It is possible in general for the optimal reported liability to be under, over or equal to the true liability. For instance, even if $E[x_l] = x_l^0$ (i.e. beliefs are unbiased) there is the potential value to over-report in expectation as it decreases the probability (and expected cost) of being found to have underreported.

When liability is certain, it is not possible to have over-reporting as optimal, as paying too much tax provides no benefit regardless of whether an audit occurs. Instead, under certainty, the solution is to fully comply when $\frac{1}{p(\beta+1)} < 1$, and to engage in maximum evasion when $\frac{1}{p(\beta+1)} > 1$. Thus, uncertainty in the former case – if anything – leads the taxpayer *away* from the

[5] If the line item is associated with a liability, for example, then this means reporting zero liability. However, if the line item is associated with a deduction or a credit, then this implies taking the maximum deduction or credit possible in which case x_l^* is as large and negative as possible.

truth.[6] In the latter case, uncertainty has no effect as the taxpayer will be at the corner solution of maximum evasion regardless.

2.1 The effect of liability information services on tax reporting

At the time of filing, the individual may face uncertainty regarding tax liability due to tax complexity, poor record keeping, or both. A liability information service provided by the tax agency can address this uncertainty and perhaps improve tax reporting accuracy. Vossler and McKee (2013) examine what they label a *helpful information service*. Letting $G(x_l)$ denote the distribution of tax liability after receiving the information service, a helpful service is one with $G(x_l) \leq F(x_l)$ for $x_l \leq x_l^0$ and $G(x_l) \geq F(x_l)$ for $x_l \geq x_l^0$, with strict inequality between distribution functions holding at least for some x_l. Although more general, this characterizes a service, such as the one we explore in the experiment, that is unbiased, i.e. $E[x_l|G] = x_l^0$; and reduces the uncertainty over x_l through, for example, reducing the variance [i.e. $Var(x_l|G) < Var(x_l|F)$]. As Vossler and McKee (2013) show, a helpful information motivates more truthful reporting when there is an interior solution to [1], or there is instead a corner solution of full evasion in the absence of the service.[7]

2.2 The Effect of Social Norms and Fairness of Fiscal Exchange (Public Good Provision)

Another strand of literature has focused on the role that social norms play in the tax compliance decision (Alm *et al.,* 1999; Torgler; 2002). This literature suggests that, in addition to the explicit expected costs imposed by the tax authority for non-compliance, there are

[6] This result is similar to that obtained by Beck and Jung (1989).
[7] To be clear, given that information services have a random outcome, this result is true "on average" rather than specifically for each taxpayer in each instance.

additional, implicit costs from deviating from a norm level of compliance.[8] Norms differs from fairness considerations in tax morale – the latter captures the taxpayer's perception of the tax burden, while the former addresses how others perceive the taxpayer's level of compliance. Hence, a taxpayer with high (low) morale employed in an industry with a norm of evasion (compliance) may be conflicted.[9] Let For simplicity, assume taxpayers incur a sanction, $\lambda > 0$, for deviating from the norm, αt_l,

[3] $$s(R_l) = \lambda(\alpha \bar{t}_l - R_l)^2$$

where $0 \leq \alpha \leq 1$.[10] These implicit costs can be the result of social sanctions, such as ostracism, resulting in a lost stream of future benefits from exchange with members of one's group, or emotional dismay, as a result of guilt or shame (Erard and Feinstein, 1994). The quadratic specification of these costs in equation assumes a deviation above or below the norm is penalized symmetrically.[11] In this situation, we can write the optimal reporting choice as

[4] $$R_l^* = \alpha \bar{t}_l + \frac{1}{2\lambda}\{p(1+\beta)\big(1 - F(R_l^*)\big) - 1\} \qquad \forall l.$$

Now the taxpayer equates reported taxes with that of the expected norm level of compliance, adjusted now for the discounted effect of the uncertain enforcement regime.

Next, we account for the public goods and services that are provided through collected taxes (Alm *et al.,* 1999). Such benefits (costs) create an incentive for increased compliance (evasion). Likewise, considerations such as for fairness (Fehr and Schmidt, 1999) are important. Assume each taxpayer holds a subjective perception, $\pi(R_l)$, regarding the resulting impact their

[8] The source of such a norm is beyond the scope of this paper. See Elster (1989) for a detailed discussion on the topic.
[9] Obviously, when the social norm is aligned with a taxpayer's morale, the incentives will reinforce each other.
[10] Hence, the norm is regarding the level of compliance not the level of taxes reported. Hence, if taxes differ by income, the norm implies reported taxes will differ by income.
[11] Alternatively, deviations above or below the norm could be penalized asymmetrically. For example, the social cost function could simply penalize deviations below the norm (Alm *et al.,* 1999), inducing higher compliance. However, one can imagine the norm working in the opposite direction. Hence, we employ a general specification of social norms that allows for a broad range of possibilities.

reported taxes will have on their own well-being, a tax morale. In general, these subjective perceptions can either be a benefit, or yet another cost, and may either be increasing or decreasing in reported taxes (i.e. $\pi_1 > 0$ or $\pi_1 < 0$ where π_1 is the first derivative of tax morale). For tractability, let tax morale take the form of a modified public good,

[5] $$\pi(R_l) = \gamma \sum_{i=1}^{N} R_{li} - \delta(R_{li} - \bar{R}_l)^2$$

where $i = 1, \ldots, N$ is an index of taxpayers and $\bar{R}_l$ is the average contribution to the public good. The parameter γ denotes the marginal per capita return from the provision of public goods and services, which may be perceived to be positive or negative (i.e. a public bad).[12] Typically, $\gamma > 1$ is necessary to induce voluntary provision of the public good (i.e. voluntary compliance). The presence of fairness considerations, however, complicates the decision considerably.[13] The quadratic term in the above equation implies those who pay more (less) than average, and perceive that to be unfair, $\delta > 0$, incur a psychological cost and are inclined to increase evasion (compliance) to compensate.[14]

Abstracting away from social norms related to compliance (i.e., let $\lambda = 0$), and assuming $\gamma > 0$ and $\delta > 0$, we can write the optimal reporting decision as

[6] $$R_l^* = \bar{R}_l + \frac{1}{2\delta}\{\gamma - 1 + p(1+\beta)(1 - F(R_l^*))\} \qquad \forall l.$$

Hence, the taxpayer will equate reported taxes with that of the average taxpayer, adjusted now for the discounted effect of both the public goods provision and the uncertain enforcement regime. The basic implication is that tax reporting is increasing in the average (or, with a fixed number of taxpayers, the total) taxes reported.

---

[12] For example, a criminal may perceive their tax dollars support law enforcement, which in turn, reduces the criminal's income. In this case, the objective is to minimize costs, which results in reduced compliance.

[13] Bordignon (1993) offers a slightly different approach to modeling such preferences. He assumes the fairness consideration enters the taxpayer's objective function as a constraint. This eliminates the possibility to tradeoff motives.

[14] The quadratic specification differs from that of Fehr and Schmidt (1999). It is assumed for simplicity when deriving the optimal tax compliance.

2.3 The Effects of Tax Withholding[15, 16]

Requiring taxpayers to have taxes withheld at the source alters the reporting decision as compared with the no withholding case. In making the withholding choice, the taxpayer must weigh the *ex ante* costs of paying taxes in advance of the tax reporting period against the costs associated with having to pay additional taxes (and perhaps interest and penalties) at the time of filing. The costs of withholding include the foregone consumption or investment opportunities during the year. The costs of being required to pay additional taxes at the time of tax filing can consist of the administratively imposed costs (interest and penalties) by the tax agency for under-withholding, liquidity costs, and the type of reference effect demonstrated by Kahneman and Tversky (1979). Facing uncertainty regarding true tax liability at the time of the withholding decision, the taxpayer must balance the certain costs of withholding against the expected costs of the under-withholding penalty as well as the reference effect costs posited in K&T. In effect, the withholding choice allows the taxpayer to either avoid or incur these expected costs. There is a cost of avoidance that is increasing in the magnitude of the avoidance. However, absent uncertainty over tax liabilities, taxpayers can choose whether to be in a tax refund or tax payment situation. Uncertainty at the withholding stage introduces the potential for unexpected tax payment situations to arise with the concurrent reference point effect. The uncertainty associated with being in a tax payment versus a tax refund situation is the mechanism that results in the

[15] The speculated interactions between tax withholding and tax reporting are inherently behavioral. In this section, we provide some economic intuition to help guide the empirical analysis.

[16] Some theoretical investigations, extending the basic framework of Allingham and Sandmo (1972) of the propensity to underreport taxes in the presence of withholding have been undertaken. Yaniv (1988) considers both employer and employee reactions. He notes that employers may evade by not fully remitting the amount they withhold from the workers' pay and employees may evade if the withholding is less than their full tax liability. The usual results hold. Propensity to underreport is lower as the audit probability and penalties increase. But this suggests that an effective enforcement mechanism is to reduce the taxpayer's incentive to under-withhold. When the reference point effects are incorporated, a policy of over-withholding becomes even more effective.

taxpayer facing the reference point effect and it is this effect that links the withholding and filing stages of the tax year. The decision models of K&T predict that the likelihood of underreporting at the time of filing will be driven by whether the taxpayer has under versus over-withheld. In the case of over-withholding the taxpayer on the positive side of the reference point effect and will be less willing to incur the risks associated with tax underreporting, holding enforcement effort fixed. The opposite will hold for the taxpayer in the under-withholding state. Here the taxpayer can also avoid the under-withholding penalty by underreporting taxes.[17]

Information services are not directly relevant to the withholding choice since at the beginning of the tax year (when the withholding decision is made) there is no means for the tax authority to provide relevant information. However, the availability of liability information services at the time of tax filing can have an indirect effect on the withholding choice via the effect on tax reporting. For a given withholding level the taxpayer can face three situations at filing time: withholding exceeds current legal tax liability and a refund is due; withholding falls short of legal tax liability and additional taxes plus an under-withholding penalty must be paid; withholding is exactly equal to legal tax liability. This last case is uninteresting, but nevertheless provides a baseline from which to compare behavior in the other cases.

Since taxes reported is a choice for the taxpayer (based on reported income and deductions claimed), the taxpayer can react to the tax situation she faces at the time of filing by adjusting income and deduction in relation to the tax withholding level. Since the liability information service allows the taxpayer to know her tax liability perfectly, it seems likely that the

[17] Since taxpayers will exhibit heterogeneous preferences, we expect that relative weights of the incentives to over or under-withhold will vary across our participants but that regularities will be identifiable when past audit results and taxpayer characteristics are controlled.

taxpayer in the under-withholding state would be more likely to purchase the information. This can lead to avoiding both the audit penalty and the under-withholding penalty.

On the other hand, a taxpayer in the refund position (taxes over-withheld) is getting a refund based on the reported income and deductions. The information may lead the taxpayer to report higher taxes and thus may be viewed as a bad. Thus we predict that those who have over-withheld will be less likely to expend resources to acquire information. If participants randomly fall in the tax payment or tax refund states, we would expect to see no aggregate effects from providing the liability information service.

3. Experimental Design

3.1 Overview

Relative to related tax compliance research, the decision setting here has been expanded to include the tax withholding decision coupled with the taxpayer information initiatives. The latter serves to reduce taxpayer uncertainty regarding tax liability at the time of filing. At the time of the tax withholding decision the taxpayer faces considerable uncertainty regarding income and deductions since this is the beginning of the tax year.[18] Finally, the experiment setting introduces social norms through information regarding the aggregate tax reporting behavior of the other participants and, as a treatment, the presence of a fiscal exchange in the form of a shared public good financed through tax collections.

As the theoretical framework presented in the previous section suggests, tax reporting behavior is potentially influenced by social norms and by the extent to which taxpayers perceive the "fiscal exchange" to be fair. To capture the effect of social norms and fiscal exchange, we

[18] To emphasize the uncertainty we do not allow for revisions to the withholding decision during the tax year. In the naturally occurring setting the individual may, of course, revise her W-4 to reflect changes in tax status arising from marriage, home purchase, or an added dependent. But there are transactions costs associated with such changes.

introduce feedback on the tax reporting behavior of others and we incorporate a public expenditure which benefits all participants equally. The latter is implemented as a "transfer payment" by which a fraction of the total taxes collected is returned to the participants; a public good.[19] Prior to the beginning of the period the participants choose an amount of taxes to have withheld (akin to completing a W-4). This amount is credited against taxes owed at the filing time. Since taxes are withheld from the beginning of the period, the true cost of a dollar withheld may be greater than one (e.g., forgone interest) or less than one (the reference effect of receiving a refund). At the time of the withholding decision the participant faces uncertainty regarding true income and the level of allowed deductions that may be claimed.

3.2 Experiment Setting Details[20]

There are a set of common characteristics of a decision round in the experiment. Participants earn income by performing a task, determine how much in taxes to have withheld, self-report their tax liability, and then face the possibility of audit and penalties for underreporting taxes. In the earnings task, participants are presented with a picture of either a jar of pennies, gumballs or jelly beans, and are asked to guess the number of elements in the jar. One-third of the participants with the closest estimate are placed in the high income group, the second third are placed in the middle income group and the next third in the lowest income group. The group assignment is in effect for a "series" of rounds, which we later describe. The earnings task rewards relative performance. In effect all compensation schemes have some component that is relative since there are time constraints facing even piece rate workers.

[19] In the instructions this is characterized as "roads, etc provided by government." The specific marginal per capita return is not revealed in the experiment.

[20] This section describes the experimental setting and design in detail. Sample subject computer screens and printed instructions are available in a Reviewer Appendix, available upon request.

At the beginning of a decision round the participants select the amount of taxes they wish to have withheld, from a discrete choice set of amounts, through a simplified W-4 form. Specifically, there are five withholding levels to choose from (tied to the number of "allowances" claimed). The withholding amounts span expected tax liability amounts as well as allow the participant to unambiguously over-withhold (by claiming zero allowances) or under-withhold (by claiming four allowances). Participants are provided with information on their income, standard and itemized deduction amounts. In particular, they are provided a range of possible income amounts, two possible standard deduction amounts, and a range of possible itemized deduction amounts. The true income, standard and itemized deduction amounts lie within the amounts provided. Overall, the information allows one to determine expected tax liability, albeit with considerable uncertainty. There is an added "withholding cost" equal to 10% of the amount withheld. This is intended to reflect factors such as discounting and opportunity cost. Completion of the form is time limited. Failure to complete the form in the time allowed results in the maximum level of tax withholding.[21]

After the withholding choice, and after a small delay, the participants are provided with their tax form. The tax form has the taxpayer report income and claim either a standard deduction (there are two possible amounts) or an itemized deduction (an amount of their choosing). Relative to the withholding stage, and to reflect that taxpayers are better informed after the tax year has ended, there is a *partial* resolution of tax liability uncertainty upon entry into the filing stage. At the time the tax form is to be completed some of this uncertainty will have been naturally resolved with the passage of time. For example, an individual will have received some information via her employer, bank statements, and so on that will partially

[21] That is, it is as if the participant was claiming zero exemptions on the W-4 form – the default value for anyone not submitting a W-4 to their employer.

resolve the income and deduction uncertainty faced when the level of tax withholding was selected. In the experiment, the range that contains their actual income and allowable itemized deduction is reduced by 25%, and further the true standard deduction amount is revealed. Final tax liability is the difference between earned income and deductions claimed, taxable income, multiplied by a tax rate of 50%.

Participants are free to alter their entries on the tax form up until they file or until the tax form times out. As they adjust their entries they can update their tax form by clicking on a "Do the Math" button. By clicking this button, the amount of reported taxes (i.e., taxable income multiplied by the tax rate) is calculated along with the corresponding tax payment or refund. The amount of withholding is automatically credited on the form. Thus, a tax payment is revealed if the player has reported tax liability in excess of their withholding; otherwise, the relevant tax refund amount (i.e., amount withheld minus reported taxes) is displayed. In the event that a tax payment is due, an under-withholding penalty is also reported, equal to 20% of the reported tax payment.

Following the tax reporting stage, there is an audit process that is completely random and is conducted independently for each participant. A graphic appears on the decision screen consisting of three balls in a box and the balls alternate colors (white and blue) and when the balls stop changing color the participant is audited (blue) or not (white).[22] If the player is selected for an audit, unpaid taxes (based on the actual income and allowable deduction amounts) are discovered and collected along with the penalty, which is equal to 300% of unpaid taxes. These audits occur with a known probability and are perfectly revealing.

[22] The audit process used in these experiments is completely random. While much of the IRS audit selection is based on endogenous rules, a purely random process avoids some of the complications that would arise from the use of relative reports. The use of the random process allows us to focus on the issues to be investigated in this series of experiments. Alm and McKee (2004) have examined the behavioral responses under endogenous processes and find the taxpayers attempt to coordinate reporting levels but that they have difficulty doing so.

Participants are informed that if they report more income than earned and/or claim less in deductions than allowed, they will not be refunded the taxes overpaid. That is, the audit process can never increase a participant's earnings. Allowing the form to time out without filing results in the automatic audit of the return and, since all entries are imputed to be zero, no deduction is claimed. Thus the participant faces the maximum tax liability and the penalty is based on this. We explicitly inform the participants that it is never in their interest to allow the form to time out. Some participants do allow the form to time out but this rarely occurs beyond the earliest rounds.

After the audit determination, participants are provided with a summary screen that reveals their actual income and deduction amounts, what they reported on the tax form, and a detailed breakdown of earnings from the round. Earnings are largely determined as the difference between actual income and taxes paid. The extra withholding cost (10% of withholding), any under-withholding penalty (20% of tax payment), and audit costs (unpaid taxes and audit penalty) are also subtracted from earnings as applicable.

3.3 Experiment Treatments and Sessions

We use a 2x2 between-subjects design where the main treatment variables are the presence/absence of a liability information service and the presence/absence of information on tax compliance and a partial redistribution of reported taxes (i.e., a public good). Implemented as a treatment variable, taxpayer liability information assistance may be offered prior to tax filing, in the tax reporting stage (treatments T3 and T4). To reflect the transaction costs associated with obtaining information there is a monetized cost for the service, equal to 50 lab dollars. There is still considerable uncertainty over tax liability in the tax filing stage, and upon purchase of the information service all uncertainty is resolved. In particular, the ranges of possible income and

itemized deduction amounts are reduced to the actual amounts upon receipt of the information service. To request the service participants simply click on an "Information" button. The information is always supplied and always correct.[23]

When the information on tax compliance and a public good is provided (T2 and T4), only the taxes voluntarily reported are used to finance the good.[24] Implicitly we are treating the penalties and unpaid taxes collected via the audit as the cost of the audit process. This assumption is made to emphasize the social norm aspect of the public good provided from the tax receipts. Penalties are not a part of a social norm of voluntary tax compliance in this setting. The public good multiplier is set quite low to reinforce that this is a tax reporting exercise not a public good provision game.[25] Specifically, 50% of taxes voluntarily paid are equally allocated to all group members.

A second element of the social norm is the extent to which the knowledge of others' tax reporting behavior affects our own. We provide end-of-round information on the actual taxes reported relative to the required taxes broken down by income class. The individual can compare her taxes reported relative to the average of her income group, as well as to reporting behavior of other income groups.

The audit rate is implemented as a within-subject treatment variable. The three audit rates used are 10%, 30% and 50%. Based on the theoretical framework presented earlier, in the absence of an information service and social norm/public good features, the audit rates are predicted to induce full evasion, partial evasion, and full compliance, respectively. Information

[23] Incomplete and/or incorrect services have been investigated elsewhere (see Alm *et al.,* 2010; and Vossler and McKee, 2013).

[24] In previous research we systematically turned on/off the tax compliance and public good elements in the design (see Vossler *et al*., 2012). As we found no interaction effects, in the current research we opted to simply turn on/off both features simultaneously.

[25] Previous work (Alm, Jackson, and McKee, 1992, 1993) placed greater emphasis on the fiscal exchange aspect of tax reporting behavior and report that the public good increases compliance in small group settings but the mechanics of the provision (the institutional setting) matter a great deal.

services, if requested, perfectly reveal information and as such should lead to full evasion in the 10% audit regime and full compliance otherwise. Parameters used for the experiments are reported in Table 1. All amounts are denominated in lab dollars.[26] The four treatments are presented in Table 2.

An experimental session consists of 18 paid rounds arranged into three blocks of six rounds each. At the beginning of each block the taxpayers (participants) earn income by completing a simple task. These earnings are in effect for the remainder of the block. For the duration of a series the audit probability is constant. All participants experience the three audit probabilities. To help control for order effects as well as increase the number of independent observations, there are two distinct taxpayer groups within each session. In the social norm/public good treatments, compliance information and the transfer payment are group-specific. In all sessions, the sequence of audit rates differs across groups.

There are 18 sessions in the experimental design, the distinguishing features of which are presented in Table 3. Given the interdependencies created by implementing the compliance/public good features, there are six sessions each of the two treatments with these features. For the remaining two treatments there are three sessions each. With three audit rates, there are six unique audit sequences. Each audit sequence is implemented once for the two treatments without the group interaction (i.e., no compliance and public good features), and replicated at each location for the remaining two treatments.

3.4 Participant Pools and Procedures

The experiments were conducted largely with employed adults from the Knoxville, TN and Boone, NC area. The experimental labs are located at the University of Tennessee and

[26] Lab dollars are converted to US dollars at the end of the session at the rate of 300 lab dollars to one US dollar.

Appalachian State University.[27] The labs both include two-dozen networked computers, a server, and software designed for this series of experiments. Recruiting at both sites was accomplished using the Online Recruiting System for Experimental Economics (ORSEE) developed by Greiner (2004). The participant databases were built using posters and email announcements to various community groups in each location. Registered persons were invited to specific sessions via email, and were permitted to participate in only one tax experiment session. Only participants recruited specifically for a session are allowed to participate, and no participant has prior experience in this specific experimental setting. Somewhat unique to this investigation is the use of a non-standard subject pool. Participants for this study were recruited from the general population living or working in the area near the two universities where the experiments were conducted. The resulting pool is much older than is typical, largely employed full-time (73%), and experienced in the completion of a W-4 and a tax return. Since the experimental setting is highly context specific (tax language is used throughout the instructions and the computer interface) this experience is likely transferred to the lab setting. Overall, there are 359 participants.[28] The number of participants in each session and the lab locations for the sessions are presented in Table 3.

The experiment was computerized, programmed and conducted with the experiment software z-Tree (Fischbacher, 2007). An experimental session proceeded in the following fashion. Each participant sits at a computer located in a cubicle, and is not allowed to communicate with other participants. An experiment moderator welcomes everyone for their

[27] Although the pools are intended to be comprised entirely of employed adults who are not full-time students, there is some inevitable leakage of students who are working part time into the subject pool given that the recruiting off campus is biased toward proximity to campuses to ensure greater participation in the experiments.
[28] In refining the instructions, procedures and experimental design, we also conducted three pilot sessions with full-time students. The last pilot session, aside from the student subject pool, is otherwise identical to Session 13. Data from this session, which does include some who are employed, can be made available upon request.

participation, explains that earnings are based on decisions in the experiment, decisions are anonymous, and that experimenter deception is not permitted. Then, the software is initialized and on-screen instructions first guide participants through a set of risk elicitation tasks modeled after Holt and Laury (2002), as amended by Bruner (2012). The experiment moderator answers any questions prior to decision making.

The instructions for the tax experiment are then conveyed by a set of printed instructions that are read allowed to ensure both common knowledge and that the participants at each site received exactly the same instructions (instructions are available as a Reviewer Appendix). The first practice round is conducted with the stage clocks (Withholding and Filing) turned off and with the experimenter directing the participants on the use of the interface. The second practice round (which also does not affect earnings) is conducted with the clocks running as in the paid rounds. Clarification questions are addressed at the end of the second training round. The participants are informed that all decisions are private; the experimenter is unable to observe the decisions, and the experimenter does not move about the room once the session starts to emphasize the fact that the experimenter is not observing the participants' compliance decisions. This reduces, to the extent possible, peer and experimenter effects that could affect the decisions of the participants and implements a double blind design in so far as in possible with the person running the experiments being the person who designed them. All actions that participants take are made on their computer station.

The experiment proceeds for 18 paid decision rounds, although the actual number of rounds is not pre-announced nor is the length of a series. After the final decision round, participants learn of their earnings from both the risk elicitation exercise and the tax experiment. Participants are then directed to complete both a demographic and taxpayer attitude debriefing

questionnaire. The demographic questionnaire elicits information on personal characteristics as well as tax filing experiences. The taxpayer attitude questionnaire is adapted from Kirchler and Wahl (2010). After the briefing is completed, participants are called up to the front of the room individually and paid their earnings in cash. Average earnings were approximately $80 for the session which lasted about two hours on average.

4. Results

Descriptive statistics for the experiment data used in the subsequent analysis are reported in Table 4. Overall, the participants represent a fairly diverse group. The average age is 38 and ranges from 18 to 68. In terms of employment experience, 73% classify themselves as employed full-time and 23% as part-time employed. Types of employment cover a wide breadth with the largest percentage (40%) being in the education area. Other heavily represented (identified) sectors include the food services sector (10%) and retail trade (7%). Average (personal) income was approximately $31,000 in the year 2012 with a large standard deviation indicating considerable variation across the participants. The overwhelming majority (89%) filed taxes last year and a very small number (15%) reported being listed as a dependent on another taxpayer's form. Thus, a large fraction of the pool had personal experience with tax filing. There is also a fair amount of variation in tax under-reporting opportunities as illustrated by the number of participants who self-reported having non-wage income (50%) and itemizing deductions (32%). The two largest filing status types are "single" (57%) and "married filing jointly" (26%). Of those identifying which form they used to file their 2012 return, approximately half of the pool reports using the standard 1040 form. Consistent with the population at large a vast majority of our participants claimed a tax refund last year (75%). About one third of the pool used a

professional tax preparer last year. A very small fraction of our pool utilized IRS taxpayer services last year (10%). While enhancing the external validity of our results, the diversity of this pool further allows for us to identify associations between experiment outcomes and taxpayer characteristics and experiences.

To analyze three outcomes of interest – tax withholding, demand for information services, and tax reporting – we estimate linear regressions. To investigate tax reporting we use the constructed outcome variable *Tax Underreported*, defined as the difference between one's actual and (expected) tax liability. This thus combines information from the individual income reporting and deduction decisions. Expected liability is based on the information set at the time of filing. As controls common across models we include the following. The variables *Compliance* and *Fairness* relate to variables identified in the theory section to capture the effects of our social norm and public good design elements, when these features are in effect. The dummy variables *Penalized* and *Not Penalized* are one-period lags of indicators for non-penalizing audits (i.e., individual was selected for audit but not penalized) and penalizing audits, respectively, to allow for behavioral responses to the audit process. *Earned income*, a dummy for whether the high standard deduction is allowed (*High Standard*), and the level of itemized deduction (*Itemized*) allow for differences in choices based on expected liability. *Audit Rate*, the probability of being randomly selected for audit, is included to control for enforcement effort.

Estimation is carried out using the experiment panel data of participant-specific outcomes. Variance-covariance estimators robust to heteroskedasticity and within-person serial correlation are used in all models.[29] To control for unobserved individual and time effects, we include both individual and round fixed effects. In models where we include additional variables

[29] For those less familiar with cluster-robust standard errors, note that this is a consistent estimator for the standard errors in the presence of an individual-level random effect.

that related to participant characteristics and tax experience, we necessarily omit the individual fixed effects (which would be perfectly collinear with these additional varaibles). Estimation results are presented in Tables 5 – 7. We first discuss models that exclude the participant control variables to focus on the main treatment effects.

4.1 Tax reporting

Three models related to tax underreporting are presented in Table 5. The amount of taxes under-withheld, over-withheld, and unanticipated changes in expected liability across stages (i.e., variables that measure the magnitude of the tax liability increase or decrease) capture nuances of one's tax position when filing. Further, the indicator variables *Info* × *Under-withheld* and *Info* × *Over-withheld* capture the effect of receiving information, allowing for differential effects based on whether one finds themselves in an under- or over-withholding position, respectively. We find important asymmetries related to tax position for all three sets of variables.

The coefficient on *Tax Under-withheld* in Model 1 suggests a strong and statistically significant relationship between (expected) tax position and underreporting. In particular, the equation suggests that participants under-report taxes by about 40 cents for every dollar under-withheld. In contrast there is, if anything, only a small effect between the amount of tax over-withheld and tax reporting. Turning to unanticipated shocks tied to the resolution of uncertainty across the withholding and reporting stages, there is also an important asymmetry. In particular, those who experience an unanticipated increase in liability underreport by about 24 cents for every dollar of unanticipated increase, where the effect is to decrease underreporting by 12 cents for every one-dollar in an unexpected tax liability decrease. Turning to information services, we find that those in an under-withholding state decrease under-withholding by almost 150 lab

dollars. To put this in perspective, in the absence of an information service participants on average underreport by almost 250 lab dollars when in an under-withholding state and so the information service decreases under-reporting rather substantially. The effect of the information service is less pronounced for those instead in an over-withholding state.

Increases in tax compliance and tax redistribution (as measured by *Compliance* and *Fairness*, respectively) reduce tax underreporting as suggested by theory. There is no significant average effect for those audited last round, regardless of whether they were found to be in violation or not. Similar to findings from related work (e.g. Vossler and McKee, 2013), we find that underreporting decreases with enforcement effort and expected liability.

4.2 Tax withholding and demand for liability information services

Table 6 presents linear probability models of liability information acquisition, using the sample of participants from T3 and T4 for which this service was available. Although the included explanatory variables are similar to those in the tax reporting models, there are some subtle differences. In particular, variables with labels that include the appendage "(2)" are constructed based upon the participant's information set upon entering the tax filing stage. This is different from the information set when filing only for those who obtain the information service.

About 40% of participants avail themselves of the liability information service when it is offered. Recall that there is a fee for this service and thus we have revealed willingness to pay for the information. Looking first at the effects of tax position at the time of filing, we find that those in an under-withholding position have a decreased demand for information. This effect is

reasonably small, with the probability of uptake declining by two percentage points. Liability shocks appear to have no effects on the information acquisition decision.

The largest identified drivers for information demand appear to be linked to enforcement. In particular, those audited but not penalized in the previous round have a 4 percentage-point higher uptake of the service. For every ten percentage-point increase in the audit probability, the probability of uptake increases by about three percentage points. This last result is intuitive as the value of information increases as the marginal cost of evasion increases. The demand for information is increasing in earned income, which is consistent with information being a normal "good".

Tax withholding in the experiment is a choice, and we have introduced financial incentives to motivate optimal withholding choices by the use of the add 10% withholding cost and 20% under-withholding penalty. Although these design choices were intended to reflect naturally-occurring considerations such as opportunity costs and liquidity constraints, these experiment parameters are not varied across participants. Further, there are no doubt other features related to dynamic behaviors and institutional realities that we abstract from to keep the decision setting tractable for participants.[30] Although we have considerable variation in tax withholding position, which gives us confidence in our results related to tax underreporting and taxpayer information services, our analysis of the drivers of withholding is relatively limited.

The results of an analysis of withholding behavior are presented in Table 7. We find that withholding is increasing in earned income (which is defined according to the information set in the withholding stage) in a manner that reflects that the financial incentives were strong: tax

[30] Research suggests that many taxpayers under-withhold due to a forced savings motive. In a recent paper, Jones (2012) provides evidence that "inertia" is also important, as taxpayers tend to be slow to respond to changes in their own tax status (e.g., the birth of a child) or changes in tax policy (e.g., the introduction of the Earned Income Tax Credit) and the result is over-withholding.

withholding increases about 50 cents for every 1 dollar increase in withholding, consonant with our 50% tax rate on taxable income. Further, withholding is increasing with the audit probability and tax fairness. There appears to be a statistically significant relationship between the participant's audit histories, but only in the case of a penalizing audit. Viewed in their entirety, these relationships suggest that withholding is partially driven by expected compliance but there are nevertheless drivers of the underreporting decision that are not considered in the withholding stage.

4.3 Simultaneity of outcome variables

It is certainly plausible that those intending to evade simply under-withhold more. On a similar note, those not interested in being compliant may be the same persons who do not purchase the liability information service when available. For the estimated relationships regarding tax under-reporting and information acquisition presented previously to be interpretable as causal, one must make the assumption that possible self-selection effects are adequately controlled for by included observables and fixed effects.

In our tax underreporting model, the variables *Tax Under-withheld, Tax Over-withheld, Info* × *Under-withheld* and *Info* × *Over-withheld* are possibly endogenous. To address this issue, we estimate an IV regression using Generalized Method of Moments (GMM) estimation. As is common for panel models with unit effects and some explanatory variables for which the strict exogeneity assumption may fail, the estimating equation is specified in first-differences. To facilitate estimation, we assume that one-period lags of the four variables based on the withholding and information acquisition choices can be considered sequentially exogenous (i.e. predetermined) in the sense that they are uncorrelated with current or future error terms. Under

this assumption, moment conditions based on second and higher-order lagged levels of these variables can be generated to facilitate estimation.[31]

Model 2 of Table 5 presents the GMM-IV estimates. Overall, the results are rather similar to those of Model 1. If anything, the withholding-effect asymmetries become more pronounced as the effect of tax over-withholding is even smaller than before and is no longer significant. The same is true concerned the effects of an unexpected tax decrease. The effects of tax under-withholding and positive liability shocks remain large in magnitude albeit similar to prior estimates, and statistically significant. Acquiring the liability information service continues to significantly decrease underreporting, with an effect over twice as large in the case where the taxpayer finds herself in a tax under-withholding situation. The magnitudes are slightly larger than in the non-IV model.

A similar GMM-IV estimation was undertaken for the model of liability information service uptake, and this likewise generated very similar results to the non-IV estimator.

4.4 Individual characteristics and taxpayer experience

Model 3 in Table 5 adds to the basic underreporting model several variables related to taxpayer characteristics and experience. The coefficient on the variable *Risk Averse*, which is an indicator for risk-averse individuals based on data from the risk elicitation exercise that preceded the tax compliance experiment, is negative and weakly significant suggesting that risk aversion is tied to less underreporting. Women tend to underreport less, which seems to be a universal empirical finding in the tax compliance literature, and the level of underreporting declines with age. Those who report being joint-filers exhibit lower underreporting and this is consistent with

[31] Our approach parallels the common Arellano-Bond estimator for dynamic panel models. The only difference is that the one-period lag of a subset of explanatory variables are treated as sequentially exogenous rather than the one-period lag of the dependent variable in the dynamic panel setting.

the fact both parties are liable if tax cheating is detected. Those who report higher levels of non-matched income engage in more underreporting. Those who itemized deductions in their 2012 tax return significantly underreport. Those who have to pay additional taxes on their last return underreported more in the experiment. These results suggest that our participant pool brings some "homegrown" tax filing experience to the lab and buttresses our argument that we have a framed field experiment. Inclusion of the demographic variables does not noticeably alter the effects of experiment settings, however, suggesting that random assignment into treatment maintains identification in the presence of participants' innate characteristics. The overall goodness of fit improves slightly when demographics are included.

Turning to the information acquisition and withholding models (Model 5 and Model 7), there are fewer links between these outcomes and participant characteristics and taxpayer experiences. On average, women are much less likely to purchase information (a 21 percentage-point difference), and those with a college degree are more likely to acquire information (a 14 percent difference). Risk aversion, being female, and being older increases withholding.

We summarize the main results as they relate to tax reporting, tax withholding and information interventions below.

Result 1. (Tax position matters.) Tax underreporting increases with the amount of tax under-withheld. Over-withholding has no little or no effect on underreporting.

Result 2. (Unanticipated shocks matter.) Tax underreporting increases (decreases) with the level of an unanticipated increase (decrease) in liability.

Result 3. (Information services matter.) Those who receive helpful tax liability information decrease under-reporting when in an under-withholding position. Otherwise, there is no effect on underreporting.

Result 4. (Social interactions matter.) Tax underreporting decreases as measures of compliance and fairness increase.

Result 5. (Evidence of external validity.) Tax underreporting is correlated with both participant demographics and tax reporting experiences. Of particular note is that those with greater underreporting opportunities in the field have greater tax underreporting in the experiment.

5. Discussion

Although tax withholding is a central component in the US income tax system there has been surprisingly little research conducted as to its effects on tax reporting by individuals. This study addresses the question of the effect of withholding behavior on tax reporting and examines the interaction between taxpayer information initiatives and tax withholding.

Our results related to tax underreporting support the notion that tax position – whether one finds themselves in an expected tax refund or payment scenario – is a significant behavioral driver. This is true for both planned withholding as well as for shocks that alter tax liability expectations after withholding but prior to filing. Both planned and unexpected under-withholding leads to significant and large increases in tax underreporting whereas planned or unexpected over-withholding has no discernable effect. These results are somewhat consistent with a Kahneman and Tversky (1979) reference point effect, which implies an asymmetric response to loss space versus gain space. In the tax under-withholding state, the taxpayer is required to pay additional taxes and penalties if the true level of liability is reported and this represents a loss from a reference point of having already paid taxes through withholding. Thus, a greater propensity to take the risk associated with the audit. In the tax over-withheld state the

taxpayer is strictly in gain space since a refund is due and so would be less willing to be exposed to the risk of an audit.

The behavioral asymmetries related to tax withholding position have important implications for information programs, whether provided by the tax agency or otherwise. One initiative we investigate in this study, an available service that resolves uncertainty over tax liability serves to reduce tax evasion in the experiment. Since the under- or over-withholding state is due at least in part to uncertain tax liabilities at the time the withholding decision is made, liability information services provided may alter the tax reporting decision. If the taxpayer is in the tax payment state (withholding is less than taxes owed), then she is relatively risk loving and hence less compliant. This should mean that a liability information service that resolves the uncertainty should have a relatively larger impact on reporting behavior. Consistent with this notion, we find that those in an under-withholding state reduce their evasion by more than twice as much as those in an over-withholding state.

As highlighted by the theoretical model of Vossler and McKee (2013), an unbiased liability information service that partially or fully resolves uncertainty leads to more truthful tax reporting behavior. If the service provided by the tax agency can be roughly characterized this way, targeting of the service to those less likely to report truthfully would, *a priori*, be more efficient than an information service that was not so targeted.

More broadly, our results related to the effects of information provision have potential policy application. Information programs or other inducements that motivate taxpayers to engage in over-withholding is likely to promote compliance. Reminding taxpayers of the angst regarding making a tax payment or the "insurance" afforded by planned savings may be effective in this regard. Further, since unexpected increases in tax liability have been demonstrated to increase

evasion, reminding taxpayers mid-year to make W-4 withholding changes in response to these unanticipated events may be fruitful. Interestingly, North Carolina recently introduced a modified withholding structure (a new "NC4") that emphasizes penalties for under-withholding and accentuates uncertainty at the time of the withholding decision. Presumably this will motivate some to move to an over-withholding position.

Our evidence suggests that evasion decreases with the reported compliance of others and with higher "fairness" or returns from the fiscal exchange. With respect to the former, one strategic approach would be to highlight a certain taxpayer group for which there would be evidence of a high compliance rate, such as an occupation class where most workers would have only matched income. Such targeted information would be more effective than, of course, highlighting overall compliance rates. With respect to the latter, our evidence is suggestive that expanding further signage and other mechanisms that emphasize "tax dollars at work" are likely to be effective at promoting tax morale. As we find that increased perceptions over fairness increase tax withholding, programs the increase tax morale have both direct and indirect effects on tax evasion.

Given the nature of our participant pool we have considerable diversity of tax filing experience, levels of income, sources of income, and age. Of course, targeting of the service to those less likely to report truthfully would, *a priori*, be more efficient than an information service that was not so targeted. The extensive debriefing questionnaire provides us with several potential explanatory variables related to tax underreporting. This provides suggestive evidence that information initiatives such as those describe above might be more productive for males, younger persons, and those with significant opportunities to underreport (e.g., those with unmatched income, deduction itemizers).

Further analysis of our data and more extensive use of interaction effects in our models may yield additional insights but this is a topic for further research. Additional work is also needed to address the question of why individuals under-withhold. Since there is an explicit penalty for doing so and since there is the potential loss effect from having to pay additional taxes, the persistence of under-withholding merits further analysis of our data. But this is a topic for future work.

6. References

Allingham, Michael G. and Agnar Sandmo. 1972. "Income Tax Evasion: A Theoretical Analysis." *Journal of Public Economics,* 1 (4): 323-338.

Alm, James, Todd L. Cherry, Michael Jones, and Michael McKee. 2010. "Taxpayer Information Assistance Services and Tax Compliance Behavior." *Journal of Economic Psychology,* 31: 577-586.

_____. 2012. "Social Programs as Positive Inducements for Tax Participation." *Journal of Economic Behavior and Organization*, 84(1): 85-96.

Alm, James, Betty R. Jackson, and Michael McKee. 2009. "Getting the Word Out: Enforcement Information Dissemination and Tax Compliance Behavior." *Journal of Public Economics* 93 (3-4): 392-402.

_____. 1993. "Fiscal Exchange, Collective Decision Institutions, and Tax Compliance." *Journal of Economic Behavior and Organization*, 22 (3): 285-303.

Alm, James and Michael McKee. 2004. "Tax Compliance as a Coordination Game." *Journal of Economic Behavior and Organization*, 54 (3): 297-312.

Alm, James, Gary H. McClelland and William D. Schulze. 1992. "Why Do People Pay Taxes?" *Journal of Public Economics*, 48 (1): 21 – 38.

_____. 1999. "Changing the Social Norm of Tax Compliance." *Kyklos*, 52 (2): 141 – 171.

Andreoni, James, Brian Erard, and Jonathan Feinstein. 1998. "Tax Compliance." *Journal of Economic Literature* 36 (2): 818-860.

Andreoni, James. 1995. "Cooperation in Public-Goods Experiments: Kindness or Confusion?" *American Economic Review*, 85 (4): 891 – 904.

Barro, Robert. 1974. "Are Government Bonds Net Wealth?" *Journal of Political Economy*, 82(6): 1095-1117.

Bordignon, Massimo. 1993. "A Fairness Approach to Income Tax Evasion." *Journal of Public Economics*, 52 (3): 345 – 362.

Bruner, David. 2012. "Multiple Switching Behaviour in Multiple Price Lists." *Applied Economics Letters*, 18(5): 417-420.

Buchanan, James and Richard Wagner. 1977. *Democracy in Deficit*. New York, Academic Press.

Carson, Richard, and Theodore Groves. 2007. "Incentive and Information Properties of Preference Questions." *Environmental & Resource Economics*, 37: 181-210.

Chang, Otto H. and Joseph J. Schultz Jr. 1990. The income tax withholding phenomenon: Evidence from TCMP data. *Journal of the American Taxation Association,* 12(1):88 – xx.

Chang, Otto H., Donald R. Nichols, and Joseph J. Schultz. 1987. "Taxpayer attitudes toward tax audit risk." *Journal of Economic Psychology*, 8(3):299 - 309, 1987.

Clotfelter, Charles T. 1983. "Tax evasion and tax rates: An analysis of individual returns." *The Review of Economics and Statistics*, 65(3): 363 - 373.

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

Cummings, Ronald, Jorge Martinez-Vazquez, Michael McKee, and Benno Torgler. 2009. "Tax Morale Affects Tax Compliance: Evidence from Surveys and an Artefactual Field Experiment." *Journal of Economic Behavior and Organization*, 70: 447-457.

Elster, Jon. 1989. "Social Norms in Economic Theory." *Journal of Economic Perspectives*, 3 (4): 99 – 117.

Fehr, Ernst and Klaus Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics*, 114(3): 817 – 868.

Feldman, Naomi E. Tax withholding and taxpayer behavior: Evidence from a 1992 presidential executive order. Venice Summer Institute, 2004.

Fischbacher, Urs. 2007. "z-Tree: Zurich Toolbox for Ready-made Economic Experiments." *Experimental Economics*, 10(2): 171-178.

Fortin, Bernard, Guy Lacroix, and Marie-Claire Villeval. 2007. "Tax evasion and social interactions." *Journal of Public Economics*, 91(8): 2089-2112.

Greiner, Ben, 2004. "The Online Recruitment System ORSEE 2.0 - A Guide for the Organization of Experiments in Economics," Working Paper Series in Economics 10, University of Cologne, Department of Economics.

Grether, David and Charles R. Plott. 1979. "Economic Theory of Choice and the Preference Reversal Phenomenon." *American Economic Review*, 69: 623-638.

Holt, Charles and Susan Laury. 2002. "Risk Aversion and Incentive Effects," *American Economic Review*, 92(5): 1644 – 1655.

Houser, Daniel and Robert Kurzban. 2002. "Revisiting Kindness and Confusion in Public Goods Experiments." *American Economic Review* 92 (4): 1062 – 1069.

Jones, Damon. 2010. "Inertia and Over-withholding: Explaining the prevalence of income tax refunds." Working Paper 15963, National Bureau of Economic Research.

Kahneman, Daniel and Amos Tversky. 1979. "Prospect Theory: An analysis of decision under risk." *Econometrica*, 47: 263-291.

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

Kirchler, Erich., & Wahl, Ingrid. (2010). Tax compliance inventory TAX-I: Designing an inventory for surveys of tax compliance. *Journal of Economic Psychology*, 31 (3), 331-346.

Lichtenstein, Sarah and Paul Slovic. 1973. "Response Induced Reversals of Preferences in Gambling: An extended replication in Las Vegas." *Journal of Experimental Psychology*, 101: 16-20.

Manksi, Charles F. 2000. "Economic Analysis of Social Interactions." *Journal of Economic Perspectives*, 14 (3): 115 – 136.

Martinez-Vazquez, Jorge, Gordon B. Harwood, and Ernest R. Larkins. 1992. Withholding position and income tax compliance: Some experimental evidence. *Public Finance Review*, 20(2):152 – 174.

Schepanski, A. and T. Shearer. 1995. "A prospect theory account of the income tax withholding phenomenon." *Organizational Behavior and Human Decision Processes*, 63(2):174-186.

Scotchmer, Suzanne. 1989. "Who Profits from Taxpayer Confusion?" *Economics Letters* 29 (1): 49 – 55.

Scotchmer, Suzanne and Joel Slemrod. 1989. "Randomness in Tax Enforcement." *Journal of Public Economics*, 38 (1): 17 – 32.

Shapiro, Matthew D. and Joel Slemrod. 1995. "Consumer response to the timing of income: Evidence from a change in tax withholding." Working Paper 4344, National Bureau of Economic Research.

Smith, Vernon L. and James M. Walker. 1993. "Monetary Rewards and Decision Cost in Experimental Economics." *Economic Inquiry*, 31 (2): 245 – 261.

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

Vossler, Christian and Michael McKee. 2013. "Efficient tax reporting: The effects of taxpayer information services." Working paper 13-25, Department of Economics, Appalachian State University.

Vossler, Christian, Michael McKee, David Bruner and Michael Jones. "Tax Reporting Behavior and Taxpayer Segmentation: Effects of Pre-populated Tax Returns." Final Report on TIRNO – 09-Z-00019 Task Order 003, May 2012.

Yaniv, Gideon. 1988. "Withholding and non-withheld tax evasion." *Journal of Public Economics*, 35(2):183-204.

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

Table 1: Experiment Parameters

Parameter / variable	Value(s)
Income (expected value, EV)	Low: 1250 Medium: 1750 High: 2250 Uncertainty range: +/- 500 in withholding stage
Standard Deduction	250 or 500
Itemized Deduction (EV)	Low: 250 High: 500 Uncertainty range: +/- 250 in withholding stage
Audit Probability	10%, 30%, or 50%
Penalty Rate	300% on unpaid taxes
Tax Rate	50% on taxable income
Under-withholding Penalty	20% of amount owed at tax filing
Withholding Cost	10% of amount withheld
Tax Filing Time	120 seconds
Withholding Time	35 seconds
Liability information service	If available, cost is 50 lab dollars to acquire

Table 2: Treatment Conditions

Treatment	Tax Withholding	Uncertain Income & Deductions	Taxpayer Assistance	Public Good & "Social Norm"	Sessions
T1	Yes	Yes	No	No	3
T2	Yes	Yes	No	Yes	6
T3	Yes	Yes	Yes	No	3
T4	Yes	Yes	Yes	Yes	6

Table 3: Experiment Schedule

Session	Treatment	Location	Group 1 Audit Sequence	Group 2 Audit Sequence	Participants
1	1 (Info=No; PG=No)	UT	10-30-50	30-50-10	23
2	1 (Info=No; PG=No)	UT	50-10-30	50-30-10	23
3	1 (Info=No; PG=No)	ASU	10-50-30	30-10-50	19
4*	2 (Info=No; PG=Yes)	UT	10-30-50	10-50-30	21
5	2 (Info=No; PG=Yes)	UT	30-10-50	30-50-10	20
6	2 (Info=No; PG=Yes)	UT	50-10-30	50-30-10	21
7	2 (Info=No; PG=Yes)	ASU	10-30-50	10-50-30	19
8	2 (Info=No; PG=Yes)	ASU	30-10-50	30-50-10	19
9	2 (Info=No; PG=Yes)	ASU	50-10-30	50-30-10	19
10	3 (Info=Yes; PG=No)	UT	10-50-30	30-10-50	22
11	3 (Info=Yes; PG=No)	ASU	10-30-50	30-50-10	17
12	3 (Info=Yes; PG=No)	ASU	50-10-30	50-30-10	17
13	4 (Info=Yes; PG=Yes)	UT	10-30-50	10-50-30	24
14	4 (Info=Yes; PG=Yes)	UT	30-10-50	30-50-10	22
15	4 (Info=Yes; PG=Yes)	UT	50-10-30	50-30-10	17
16	4 (Info=Yes; PG=Yes)	ASU	10-30-50	10-50-30	19
17	4 (Info=Yes; PG=Yes)	ASU	30-10-50	30-50-10	18
18	4 (Info=Yes; PG=Yes)	ASU	50-10-30	50-30-10	19

*Notes: Order of implementation was random and does not reflect the session number. *Due to time constraints, only 15 of 18 periods were completed.*

Table 4. Variable Description

Variable Name	Description	Mean	Std. Dev.
Tax Underreported	(Expected) taxes underreported (both line items)	126.076	261.057
Tax Under-withheld	'Expected Liability' minus 'Tax Withheld', if >0	101.932	169.612
Tax Over-withheld	'Tax Withheld' minus 'Expected Liability', if >0	214.450	230.414
Liability Increase	=1 if expected tax liability increases across stages	0.514	0.500
Liability Decrease	=1 if expected tax liability decreases across stages	0.324	0.468
Liability Info × Under-withheld	=1 if received info in under-withholding situation	0.076	0.266
Liability Info × Over-withheld	=1 if received info in over-withholding situation	0.131	0.338
Liability Info × Not Received	=1 if liability information service available, but not purchased	0.285	0.451
Compliance	(Lag) mean taxes paid minus mean taxes owed for income group; =0 if 'Social Interactions'=0	-72.873	157.781
Fairness	(Lag) mean taxes paid (common group-level); =0 if 'Social Interactions'=0	377.816	286.279
Penalized	=1 if audited and penalized in previous round	0.195	0.396
Not Penalized	=1 if audited and not penalized in previous round	0.105	0.306
Earned Income	(Expected) earned income when filing	1776.901	432.290
High Standard	=1 if 500 Standard deduction allowed	0.504	0.500
Itemized	(Expected) allowable itemized deduction	365.013	142.455
Audit Rate	audit probability; .1, .3 or .5	0.299	0.164
Social Interactions	=1 if info on compliance displayed & taxes partially reallocated	0.660	0.474
Round	Round in session, 1 to 18	9.426	5.159
Risk Averse	=1 if selected sure bet in 70%, 80% or 90% lottery	0.386	0.487
Employed Full-time	=1 if participant employed full time	0.724	0.447
Female	=1 if participant is female	0.575	0.494
Age	Participant's age, in years	38.181	13.193
College Degree	=1 if participant has college degree	0.406	0.491
Total Income	Participant's 2012 (individual) income, in $1000s	30.607	24.656
Unmatched Income	Percentage of total income that is unmatched	8.760	19.506

Asked for Advice	=1 if participant used tax advice from a non-tax professional	0.221	0.415
Used Prep Service	=1 if participant used a professional tax preparation service to file 2012 return	0.301	0.459
Used Prep Software	=1 if participant used tax software (e.g. TuboTax) when preparing his/her 2012 return	0.453	0.498
Payment 2012	=1 if participant paid taxes upon filing for 2012	0.126	0.332
Filed Jointly	=1 if married filing jointly on 2012 return	0.265	0.442
Itemized Deductions	measure of evasion opportunity; =1 taxpayer itemized for 2012 tax return	0.243	0.429
Percent Tax Paid	Percentage of (expected) taxes paid	82.852	38.473
Tax Withheld	Amount withheld in withholding stage	805.477	397.575
Liability Service	=1 if liability information service available	0.492	0.500
Purchased Service	=1 if liability service purchased	0.208	0.406
High Itemized	=1 if participant faces high itemized deduction range in withholding stage	0.495	0.500
Expected Liability	(Expected) income minus (expected) deductions, multiplied by the tax rate of 50%	692.318	225.07

Table 5. Tax Underreporting Models: Estimation Results

Dependent Variable: Overall tax underreporting, in lab dollars (*Tax Underreported*)			
	Model 1	**Model 2**	**Model 3**
Tax Under-withheld	0.40** (0.04)	0.45** (0.07)	0.52** (0.05)
Tax Over-withheld	−0.06** (0.02)	−0.01 (0.06)	0.03 (0.03)
Unexpected Tax Increase	0.24** (0.06)	0.23** (0.12)	0.21** (0.07)
Unexpected Tax Decrease	−0.12** (0.05)	0.05 (0.13)	−0.17** (0.06)
Liability Info × Under-withheld	−146.56** (24.84)	−174.21** (43.26)	−69.84** (27.39)
Liability Info × Over-withheld	−65.16** (19.87)	−83.75** (31.98)	−13.57 (17.40)
Liability Info × Not Received			14.62 (19.07)
Compliance	−0.09** (0.02)	−0.09* (0.05)	−0.12** (0.03)
Fairness	−0.16** (0.05)	−0.15* (0.08)	−0.36** (0.06)
Penalized	3.72 (6.70)	−27.26 (18.17)	17.42** (8.10)
Not Penalized	−10.66 (7.85)	24.66 (21.80)	−64.67** (9.91)
Earned Income	0.08** (0.01)	0.07** (0.01)	0.07** (0.01)
High Standard	−24.22** (5.38)	−14.23** (13.89)	−25.92** (6.91)
Itemized	−0.08** (0.02)	−0.03 (0.04)	−0.06** (0.02)
Audit Rate	−179.56** (24.81)	−211.12** (64.20)	−143.32** (26.43)
Social Interactions			181.10** (39.44)
Risk Averse			−30.82** (15.12)
Employed Full-time			28.54 (20.51)
Female			−44.75** (17.02)
Age			−2.06** (0.88)
College Degree			12.18 (20.18)
Total Income			−0.14 (0.51)
Unmatched Income			0.73** (0.35)
Asked for Advice			−22.93 (20.83)
Used Prep Service			−28.72 (23.86)
Used Prep Software			−3.60 (22.21)
Payment 2012			41.24** (21.48)
Filed Jointly			−36.79** (20.06)
Itemized Deductions			35.37* (19.60)
Constant			124.57** (47.57)
Round Fixed Effects?	Yes	Yes	Yes
Individual Fixed Effects?	Yes	Yes	No
IV estimation?	No	Yes	No
Number of Observations	6378	6001	6180
R^2	0.597		0.293

*Notes: * and ** denote estimates that are statistically different from zero at the 10% and 5% significance levels, respectively. Standard errors (parentheses) are clustered at participant-level.*

Table 6. Information acquisition model (Treatments 3 and 4 only)

Dependent Variable: =1 if liability service purchased; =0 otherwise (*Purchased Service*)		
	Model 4	**Model 5**
Tax Under-withheld (2) (in 100s)	−0.0237** (0.0084)	−0.0421** (0.0121)
Tax Over-withheld (2) (in 100s)	−0.0042 (0.0043)	−0.0113 (0.0109)
Unexpected Tax Increase (2) (in 100s)	−0.0111 (0.0162)	−0.0497** (0.0236)
Unexpected Tax Decrease (2) (in 100s)	0.0058 (0.0180)	−0.0080 (0.0253)
Compliance (in 100s)	0.0024 (0.0053)	−0.0118 (0.0089)
Fairness (in 100s)	0.0095 (0.0090)	−0.0046 (0.0198)
Penalized	−0.0187 (0.0184)	−0.0802** (0.0276)
Not Penalized	0.0413** (0.0217)	0.1494** (0.0341)
Earned Income (2) (in 100s)	0.0055** (0.0023)	0.0061* (0.0034)
High Standard	−0.0274* (0.0141)	−0.0528** (0.0209)
Itemized (2) (in 100s)	−0.0004 (0.0046)	−0.0074 (0.0060)
Audit Rate	0.2759** (0.0592)	0.2635** (0.0635)
Social Interactions		−0.0023 (0.1218)
Risk Averse		0.0156 (0.0606)
Employed Full-time		−0.0331 (0.0766)
Female		−0.2063** (0.0608)
Age		−0.0030 (0.0030)
College Degree		0.1359** (0.0671)
Total Income		−0.0001 (0.0020)
Unmatched Income		0.0019 (0.0014)
Asked for Advice		0.0081 (0.0745)
Used Prep Service		0.0560 (0.0837)
Used Prep Software		0.0337 (0.0698)
Payment 2012		−0.0497 (0.0801)
Filed Jointly		−0.0644 (0.0811)
Itemized Deductions		−0.0201 (0.0645)
Constant	0.4644** (0.0548)	0.5825** (0.1350)
Round Fixed Effects?	Yes	Yes
Individual Fixed Effects?	Yes	No
Number of Observations	3144	3000
R^2	0.629	0.123

*Notes: * and ** denote estimates that are statistically different from zero at the 10% and 5% significance levels, respectively. Standard errors (parentheses) are clustered at participant-level.*

Table 7. Tax Withholding Behavior

Dependent Variable: Amount of tax withheld (*Tax Withheld*)		
	Model 6	**Model 7**
Liability Service		−16.83 (24.37)
Compliance	0.03 (0.03)	0.05 (0.03)
Fairness	0.14** (0.06)	0.26** (0.09)
Penalized	22.38** (10.15)	2.55 (12.13)
Not Penalized	3.67 (10.57)	14.55 (10.57)
Earned Income (WH)	0.49** (0.01)	0.51** (0.02)
High Itemized	5.76 (6.49)	4.12 (7.86)
Audit Rate	151.03** (31.02)	132.77** (31.94)
Social Interactions		−78.52 (57.66)
Risk Averse		67.51** (24.39)
Employed Full-time		−7.43 (30.48)
Female		96.33** (25.43)
Age		5.82** (1.37)
College Degree		−33.21 (28.91)
Total Income		−0.97 (0.60)
Unmatched Income		−0.81 (0.65)
Asked for Advice		−9.34 (31.76)
Used Prep Service		−23.02 (35.40)
Used Prep Software		−34.85 (29.86)
Payment 2012		6.03 (30.20)
Filed Jointly		−49.33 (31.77)
Itemized Deductions		−4.63 (28.29)
Constant	−628.75** (23.63)	−363.61** (68.70)
Round Fixed Effects?	Yes	Yes
Individual Fixed Effects?	Yes	No
Number of Observations	6399	6201
R^2	0.643	0.341

*Notes: * and ** denote estimates that are statistically different from zero at the 10% and 5% significance levels, respectively. Standard errors (parentheses) are clustered at participant-level.*