

ANDREW YOUNG SCHOOL
OF POLICY STUDIES

Audit Certainty, Audit Productivity, and Taxpayer Compliance

James Alm and Michael McKee

Summary

Strategies for dealing with evasion include such standard policies as stricter enforcement (e.g., increased audit rates, more extensive audits, larger penalties). However, the exact responses of taxpayers to these enforcement measures are quite difficult to measure with existing field data, and so are not known precisely.

In this paper we use experimental methods to examine how individuals respond in their compliance decisions to a “certain” probability of audit and to information concerning the “productivity” of an audit. Our experimental design captures the essential features of the voluntary income reporting and tax assessment system used in many countries. Human subjects in a controlled laboratory environment earn income through their performance in a task, with their actual income determined by their relative performance in this task. The subjects must then decide how much of this income to report to a tax agency. Taxes are paid on reported income, and no taxes are paid on unreported income. However, unreported income may be discovered via a random audit, and the subject must then pay the owed taxes plus a fine based on the unpaid taxes. This reporting, audit, and penalty process is repeated for a given number of rounds that each represent a tax period, and is replicated with different sets of subjects. At the completion of the experiment, each subject is paid earnings equal to the laboratory market earnings converted to U.S. dollars at a pre-announced conversion rate.

Our main policy question is the compliance effect of audits. Our experimental design specifically addresses this question by varying the prior information concerning audit probabilities and by varying the productivity of the audit. We inform some individuals that their return will be audited with certainty prior to their making their compliance decision, while other individuals receive information that they will *not* be audited. We also inform individuals of the productivity of the audit, by announcing how much unreported income will be discovered via the audit. Our design therefore allows us to observe the behavioral responses of those who are told that they will be audited and of those who know that they will not be audited, when the productivity of the audit itself varies.

We find that the announcement of audits increases the compliance rate of those who are told that they will be audited. However, the compliance rate of those who know, by elimination, that they will not be audited falls, and the net effect is that overall compliance falls. We also find that subjects respond to increased audit effectiveness only when they are expecting to be audited; that is, more productive audit procedures alone are insufficient inducements to comply unless accompanied by a heightened expectation of the audit itself.

Audit Certainty, Audit Productivity, and Taxpayer Compliance

James Alm and Michael McKee

Abstract

Strategies for dealing with evasion include such standard policies as stricter enforcement (e.g., increased audit rates, more extensive audits, larger penalties). However, the exact responses of taxpayers to these enforcement measures are quite difficult to measure with existing field data, and so are not known precisely. In this paper we use experimental methods to examine how individuals respond in their compliance decisions to a “certain” probability of audit and to information concerning the “productivity” of an audit. Our design informs some individuals that their return will be audited with certainty prior to making their compliance decision, while other individuals receive information that they will not be audited; we also inform individuals of the productivity of the audit by stating how much unreported income will be discovered via the audit. We find that the announcement of audits increases the compliance rate of those who are told that they will be audited. However, the compliance rate of those who know that they will not be audited falls, and the net effect is that overall compliance falls.

Key Words: Tax evasion, experimental methods

JEL Codes: H2, H26, C91

Audit Certainty, Audit Productivity, and Taxpayer Compliance

James Alm
Professor and Chair
Department of Economics
Andrew Young School of Policy Studies
Georgia State University
P.O. Box 3992
Atlanta, GA 30302-3992
Phone: 404 651 0420
Fax: 404 651 4985
Email: jalm@gsu.edu

Michael McKee
Professor and Arthur Child Chair in Defence Economics
Department of Economics
University of Calgary
Calgary, Alberta
Canada T2N 1N4
Phone: 403 220 5291
Fax: 403 282 5262
Email: mmckee@ucalgary.ca

INTRODUCTION

Evasion is a pervasive problem in all tax systems. For example, the “tax gap” in the United States individual income tax – or the amount of income taxes not collected due to noncompliance – is currently estimated by the Internal Revenue Service (IRS) at roughly \$300 billion, and it is widely suspected that similar, even larger tax gaps exist in other countries. Strategies for dealing with evasion include such standard policies as stricter enforcement (e.g., increased audit rates, more extensive audits, larger penalties). However, the exact responses of taxpayers to these enforcement measures are quite difficult to measure, and so are not known precisely. In this paper we use experimental methods to examine how individuals respond in their compliance decisions to a “certain” probability of audit and to information concerning the “productivity” of an audit. We find that the announcement of audits increases the compliance rate of those who are told that they will be audited. However, the compliance rate of those who know that they will not be audited falls, and the net effect is that overall compliance falls.

A fundamental difficulty in any attempt to estimate the response of individuals to an increased audit rate is the lack of reliable information on taxpayer compliance. After all, individuals have strong incentives to conceal their cheating. Several methods have been employed in an attempt to overcome this difficulty.

The best available source of information on compliance in the U.S. is the Taxpayer Compliance Measurement Program (TCMP) of the IRS. From 1965 to 1988, the IRS conducted detailed line-by-line audits of a stratified random sample of roughly 50,000 individual tax returns on a 3-year cycle. These audits yielded an IRS estimate of the taxpayer's “true” income, which allowed measures of individual and aggregate income tax evasion to be calculated. Such estimates of noncompliance are probably the most accurate that are available. Even so, however,

TCMP data have some serious and well-recognized deficiencies: the audits do not detect all underreported income, nonfilers are not often captured, honest errors are not identified, and final audit adjustments are not included. Importantly, the individual-level data have not been used by researchers to estimate the impact of a higher audit rate on taxpayer compliance, and they are also now quite dated.¹

One creative attempt to estimate the impact of higher audit rates on compliance behavior is the work of Dubin, Graetz, and Wilde (1990), who used U.S. state-level reporting data for the years 1977 to 1986 to estimate the effects of change in state audit rates on reported income. They concluded that the steady decline in audit rates over this decade led to a significant decline in income tax collections. More recent work by Dubin (2004) gives a similar answer. Unfortunately, however, these studies do not have direct measures of noncompliance. The studies must also contend with various econometric issues (e.g., the endogeneity of the audit rates), and they are unable to control for all variables that affect taxpayer reporting decisions, including changes in the tax laws and in economic conditions.

In an effort to test more directly the deterrence effects of audits, Slemrod, Blumenthal, and Christian (2001) worked with the State of Minnesota to conduct a controlled field experiment. A stratified random sample of roughly 2000 Minnesota taxpayers was selected, and these taxpayers were told via a letter sent to them by the Minnesota Department of Revenue in January 1994 that the tax returns they were about to file would be “closely examined”. The intent of the experiment was to see whether individuals informed about an increase in the probability of audit prior to filing a tax return would in fact increase their compliance. Slemrod, Blumenthal, and Christian (2001) found that low- and middle-income taxpayers responded as predicted, by increasing their reported levels of income; surprisingly and perversely, they also

¹ See Feinstein (1991) for an econometric study of noncompliance that uses TCMP data.

found that high-income taxpayers actually reduced their reports, even in the face of a higher probability of audit. Like Dubin, Graetz, and Wilde (1990) and Dubin (2004), Slemrod, Blumenthal, and Christian (2001) did not have actual information on individual compliance, but instead had to focus on investigating the income reporting by the taxpayers, an indirect measure of compliance. Since the experiment employed a between-subjects design, they also were unable to provide any estimate of the actual response of individuals to the “certain” audit.²

Some studies have utilized survey data in an effort to estimate the impact of audit rates on compliance. For example, Kinsey (1992) and Sheffrin and Triest (1992) examined individual survey data, and found that compliance increases with a greater (perceived) probability of audit. While surveys can generate useful information, their accuracy is uncertain. Individuals may not remember their reporting decisions, they may not respond truthfully or at all, and the respondents may not be representative of all taxpayers. Surveys are also unable to control for many relevant determinants of noncompliance. Finally, surveys cannot determine the direction of causality between evasion and its determinants; that is, statements about, say, the audit rates and compliance may result from an *ex post* rationalization for noncompliance rather than be the *ex ante* cause of noncompliance.³

Despite the many insights of this literature, therefore, the effect of higher audit rates on compliance remains unanswered. Even the work of Slemrod, Blumenthal, and Christian (2001) is unable to answer fully this question. Aside from their lack of access to actual compliance

² With a between-subjects design (e.g., some taxpayers are in the treatment setting and others are in a control group) we can be less certain that observed differences in behavior are due to the experimental treatment. Instead, behavior differences may be due to unobserved subject characteristics.

³ Recurring themes in the use of survey data to investigate recalled behavior are recollection bias and reluctance to report on past illegal behavior. Elffers (1991) reports the results of surveys eliciting reported illegal (delinquent) behavior and cites several examples of studies in which respondents were guilty of dynamic inconsistency in their responses. A recent study of recollection bias in the case of terrorist attacks by Viscusi and Zeckhauser (2005) finds that individuals reported they were aware of the possibility of terrorist attacks prior to September 11, 2001 at levels that are inconsistent with policy actions before and after this event. They apply the term “recollection bias” to this behavior.

information, they cannot address the important behavioral question concerning the effect on those not receiving the informational letter from the revenue department. Perhaps the only effect is indirect, but, if a particular taxpayer receives notification informing her that she will be audited and a second taxpayer receives no letter, is the second person to assume that she will not be audited? If so, the response of the non-recipients must be taken into account. Since the Minnesota experiment did not inform the non-recipients of their status, conjectures about their own audit prospects are open questions. Finally, one cannot infer how the recipients of the letters interpreted the phrase “closely examined” in the letter.

Such difficulties with field data have led other researchers to utilize laboratory experiments to investigate the responses of human subjects to changes in enforcement parameters (Friedland, Maital, and Rutenberg 1978; Becker, Buchner, and Sleeking 1987; Webley, Robben, Elffers, and Hessing 1991; Alm, Jackson, and McKee 1992, 1993; Alm, McClelland, and Schulze 1992). A typical result from Alm, Jackson, and McKee (1992) is that an increase of 10 percent in the audit rate will increase compliance by 2 percent. Alm, McClelland, and Schulze (1992) find that this impact appears to be small and nonlinear, so that the deterrent effect of a higher audit rate eventually diminishes.

In this paper we utilize laboratory experiments to examine the response of individuals to audit rates and productivity. In the lab we are able to generate information on how human subjects make their actual compliance choices and how these choices vary with a change in the probability and the productivity of an audit. Our design informs some individuals that their return will be audited with certainty prior to their making their compliance decision, while other individuals receive information that they will *not* be audited. We also inform individuals of the “productivity” of the audit, as a means of incorporating the “closely examined” statement, by

announcing how much unreported income will be discovered via the audit. Our design therefore allows us to quantify the “closely examined” language of the Minnesota experiment, and it also allows us to observe the behavioral responses of those who are told that they will not be audited. We find that the announcement of audits increases the compliance rate of those who are told that they will be audited. However, the compliance rate of those who know, by elimination, that they will not be audited falls, and the net effect is that overall compliance falls. We also find that subjects respond to increased audit effectiveness only when they are expecting to be audited; that is, more productive audit procedures alone are insufficient inducements to comply, unless accompanied by a heightened expectation of the audit itself.

THEORY

The economic model of income tax evasion (Allingham and Sandmo 1972) is based on the economics-of-crime approach pioneered by Becker (1968). This model focuses on the income reporting behavior of taxpayers, and ignores other forms of evasion such as non-payment, excessive reporting of deductions, and non-filing. In its simplest form, an individual is assumed to receive an income I and must choose how much of this income to declare to the tax authorities. The individual pays taxes at rate t on every dollar D of income that is declared, while no taxes are paid on underreported income. However, the individual may be audited with a fixed, random probability p ; if audited, then all underreported income is discovered, and the individual must pay a penalty at rate f on each dollar that he or she underreported. The individual's income I_C if caught underreporting equals $I_C = I - tD - f[t(I - D)]$, while if underreporting is not caught income I_N is $I_N = I - tD$.

An individual is assumed to choose declared income to maximize the expected utility $E U(I)$ of the evasion gamble, or

$$E U(I) = p U(I_C) + (1-p) U(I_N), \quad (1)$$

where E is the expectation operator and utility $U(I)$ is a function only of income. This optimization generates a standard first-order condition for an interior solution; given concavity of the utility function, the second-order condition is satisfied.⁴

Comparative statics results are easily derived. It is straightforward to show that an increase in the probability of detection p and the penalty rate f unambiguously increase declared income. An increase in income has an ambiguous effect on declared income, which depends upon the individual's attitude toward risk. If the individual exhibits decreasing absolute risk aversion, then higher incomes (and wealth) are associated with lower levels of compliance.

The standard model has been modified in a number of ways.⁵ Even so, a standard result is that, given actual audit and fine rates, most people should rationally choose to cheat. Introducing risk-aversion certainly affects the predictions, but even so the basic expected utility model of tax evasion is generally unable to explain observed compliance rates.

A more fruitful modification is the incorporation of non-expected utility behavior into the analysis (e.g., individuals exhibiting loss aversion or more extreme forms of risk aversion, such as rank dependent expected utility), which generates predictions of compliance more consistent with observed behavior (Bernasconi 1998). For example, for individuals described by a rank

⁴ The first- and second-order conditions are, respectively (where each prime denotes a derivative),

$$\partial E U(I)/\partial D = p t(f-1)U'(I_C) - (1-p)tU'(I_N) = 0$$

$$\partial^2 E U(I)/\partial D^2 = p[t(f-1)]^2 U''(I_C) + (1-p)t^2 U''(I_N) < 0.$$

⁵ See Cowell (1990), Andreoni, Erard, and Feinstein (1998), and Alm (1999) for a discussion of the standard evasion model and its many variants.

dependent expected utility model, we can modify the basic maximization problem of equation (1) to one in which an individual maximizes

$$E U(I) = g p U(I_C) + (1-g) (1-p) U(I_N), \quad (2)$$

where g serves to overweight the probability of the “bad” outcome (or detection and punishment). This alternative approach also helps illuminate the roles of information dissemination by the tax authority. Official information provided by the tax authority that describes audits and their effectiveness should increase the weighted probability of an audit. In contrast, if the information highlights the absence of audits or their lack of effectiveness, we would expect that the information would lower the weighted probability of an audit.

Providing taxpayers with the knowledge that their return will be audited implies that the probability of the audit is now 1.0, in the limit. Rational individuals will respond to this information with perfect compliance. Relaxing the assumption that the audits are 100 percent successful in finding evaded taxes implies that the effective probability of the audit is revised downward. If individuals are able to compute correctly compound lotteries, the compliance effect of changing the audit probability is the same as the effect of an equivalent change in audit productivity. The experiments reported in this paper investigate the effects of informing taxpayers, prior to filing, that their returns will be audited, and also examine the effects of varying levels of audit productivity. We return to the issue of the equivalence of increasing audit probability and audit productivity in the concluding section of the paper.

EXPERIMENTAL DESIGN

The experimental design captures the essential features of the voluntary income reporting and tax assessment system used in many countries.⁶ Human subjects in a controlled laboratory environment earn income through their performance in a task, with their actual income determined by their relative performance in this task. The subjects must then decide how much of this income to report to a tax agency. Taxes are paid on reported income, and no taxes are paid on unreported income. However, unreported income may be discovered via a random audit, and the subject must then pay the owed taxes plus a fine based on the unpaid taxes.⁷ This reporting, audit, and penalty process is repeated for a given number of rounds that each represent a tax period, and is replicated with different sets of subjects. At the completion of the experiment, each subject is paid earnings equal to the laboratory market earnings converted to U.S. dollars at a pre-announced conversion rate.

Our main policy question is the compliance effect of audits. Our experimental design specifically addresses this question by varying the prior information concerning audit probabilities and by varying the productivity of the audit.

Since these are experiments designed to inform policy makers, they must satisfy Smith's (1982) precept of "parallelism". Parallelism is satisfied when the experimental setting captures the essential elements of the decision problem faced in the naturally occurring setting. It is not necessary (nor is it desirable) that the experiment setting implement all of the complexity of the naturally occurring setting (Plott 1987). As implemented, our experimental design follows the main elements of Alm, Jackson, and McKee (1992, 1993) and Alm, McClelland, and Schulze

⁶ The full set of experimental instructions is available upon request.

⁷ It may be argued that current audit practice implements purely endogenous audits, since a taxpayer either elicits an audit or not depending on his or her "score" in an audit rule. However, whether a taxpayer is actually audited depends on the score and on the audit budget of the tax authority. Since the taxpayer cannot know this latter item with certainty, there remains a random component to the audit process.

(1992), but incorporates some additional features to improve parallelism with taxpayers' decision-making in the naturally occurring world. For example, in the current design, subjects earn income by performing a task (rather than receiving an endowment of income), they disclose income via a tax form, and they face an audit process similar to that in the naturally occurring setting.⁸ This sequence is closer to the naturally occurring world in which individuals earn income and then must report this income. Further elements of parallelism are also generated. The periods are timed to implement the requirement that taxes be filed on time, and the penalties are set at levels that are broadly consistent with those imposed by the IRS (e.g., the penalty for non-filing is 10 percent plus the penalty for reporting zero). A non-filer is automatically audited. The experiments utilize tax language in the instructions, as well as in the computer interface used to present information and to elicit income reporting behavior. While the stakes are small, the decision setting is also simplified relative to that of the natural setting.

Subjects are recruited from undergraduate introductory classes in economics and business.^{9 10} Upon arrival at the lab, the subjects are organized into groups of eight persons with two groups in each session. The subjects do not know who is in their group, only the number in their group; they also know that there are two groups in the session. Basic instructions are

⁸ The earnings task is admittedly simple, but it generates significant heterogeneity of income. It also contributes to the notion that a higher income is (at least somewhat) “deserved”.

⁹ Recruiting was conducted through announcements in various classes and a sign up via a web page in which the subjects posted their contact information and the time blocks of their availability. Subjects were permitted to participate in only one tax experiment, although other experimental projects were ongoing at the time and many participated in other types of experiments. We actively discourage “snowball” sampling, in which recruited subjects bring additional subjects to a session. When we recruit subjects, we do not reveal the exact nature of the experiment.

¹⁰ Although our subjects are recruited from classes offered by the department of economics and the college of business, the subjects are more heterogeneous than this might imply. At the introductory level, the courses are taken by a wide set of students who will ultimately major in a variety of fields including sociology, engineering, and science. More importantly, the decision setting for our experimental treatments is purely individual. To the extent that others have found “subject effects” in experiments, these have tended to be in settings in which other-regarding behavior (e.g., altruism) plays a role. In our present setting, no such factors enter. The basic precepts of experimental economics (Smith 1982) include salience and reward dominance, such that subject effects are dominated by the rewards to the decision being studied. Based on our subjects' interest in the level of the cash received, we feel we have established reward dominance.

provided via hardcopy, while the main instructions are provided via a series of computer screens and practice rounds. Subjects are not allowed to communicate with one another during the session. They are told that the experiment will last an unknown number of periods; in actual practice the number of sessions is predetermined, and the sessions last for 30 real rounds. After the practice rounds are completed, any final procedural questions are answered. The full experiment then begins. Sessions last approximately 90 minutes, and each subject's earnings range from \$19 to \$37, depending upon his or her performance during the experiment. Subjects are told that payments will be made in private at the end of the session, that all responses are anonymous, and that the only record of participation denoting their name is the receipt signed when receiving their payments.

During the experiment each round (or period) proceeds as follows. The earnings task requires the subjects to sort the digits 1 through 9 into the correct order from a randomized order presented in a 3 by 3 matrix. They do this by pointing the computer mouse at the numbers and "clicking" on the numbers in the correct sequence. On the computer screen a 3 by 3 matrix with the digits in random order appears on the right side of the screen, and, as the numbers are "clicked", they appear in a 3 by 3 matrix on the left side of the screen.¹¹ A counter on the screen shows the elapsed time from when the first number is "clicked", and, when all nine have been ordered, the subject clicks the "Continue" button to transmit this time to the server. Actual earned income is determined by the relative speed of performance, with the fastest performer receiving the highest income (100 lab dollars) and the slowest performer receiving the lowest income (60 lab dollars).

Once all subjects have completed the income task, they are informed via the computer of their income for the round and presented with a screen that resembles a tax form in which they

¹¹ All computer screens used in this research are available upon request.

report their income. This screen informs the subjects of the tax policy information in effect for the session. Subjects must decide how much of their income to report. In all treatments they are informed of the current tax rate and the penalty rate applied to unreported income. As noted above, these experiments present the instructions and computer interface using tax language. In keeping with the central objective of this investigation, certain parameters (e.g., the tax rate and the penalty rate) are fixed throughout the experiments so that we may focus on the effect of information concerning audits and audit results. All audits investigate only the current period disclosure.

The experimental design implements four basic treatments, as shown in Table 1. There are four audit rates (0.05, 0.10, 0.30, and 0.40), and these are applied in each of the audit announcement treatments and the audit productivity treatments.¹² In each session the subjects experience two audit probabilities, and these are in place for 15 rounds each. Two sessions are run for treatments T2, T3, and T4; in one session for each the audit rates are 0.05 and 0.3, and in the second session the audit rates are 0.1 and 0.4. A third session is conducted for the baseline treatment T1, where the audit rates are 0.05 and 0.3. The tax rate is set at 0.35 throughout the experiments, and the fine rate is set at 150 percent. There is no public good financed by the tax payments. The currency used in the experiment is called “lab dollars”, and subjects are told that all lab dollars earned during the experiment will be redeemed for cash at the end of the experiment at a fixed conversion rate of 90 lab dollars per 1 U.S. dollar.

The process of determining whether an individual is audited is generated by a draw from an animated bingo cage that appears on each subject’s computer screen, a representation that

¹² These rates are higher than the oft-reported IRS audit rate. However, this reported rate usually refers to full audits and the IRS, in fact, conducts a wide range of audits including line matching and requests for information, which are much more frequent.

looks very similar to the mechanisms used in many state lotteries.¹³ In the computerized representation of a bingo cage (or bucket), there are 20 (white and blue) balls; a white ball signifies “no audit”, and a blue one denotes an audit. On the screen, the balls bounce around briefly, and then a door opens at the top of the bucket and one of the balls exits the bucket through this door. In sessions in which the audit selection is pre-announced (treatments T3 and T4), subjects are presented with the draw of the ball from the bucket *before* the subjects are required to file their tax report, as in the case of the Minnesota taxpayers receiving the letter from the revenue department. In the sessions in which the audit selection occurs *after* the tax report is filed (treatments T1 and T2), the bingo cage appears on the screen after the tax report has been filed. After the audit process has been completed, the subjects are shown a screen that provides the earnings and audit outcome summary for the round. In all treatments the round ends with a screen that reports the results of each completed round (earned income, taxes paid, audit occurrence, fine paid, final income) to the individual subjects.

In the event of an audit, subjects are told that the audit will discover a certain percentage of unreported income. These audit productivity rates are 100 percent (treatments T1 and T3) and 65 percent (T2 and T4); that is, when the audit productivity rate is 100 percent, all unreported income is discovered with an audit, while only 65 percent of unreported income is discovered when the audit productivity rate is 65 percent.

The experimental platform consists of 16 notebook computers, a server, and software designed for this series of experimentation. Sessions were conducted on the University of Tennessee at Knoxville. Methods adhere to all guidelines concerning the ethical treatment of

¹³ This audit selection approach is similar to that used in some previous evasion studies (Cummings, Martinez-Vazquez, and McKee 2005), but differs from Alm, Jackson, and McKee (1992, 1993) and Alm, McClelland, and Schulze (1992) in which a mechanical bingo cage was used.

human subjects. Instructions are available upon request. A total of 144 subjects participated, and the number participating in each treatment is shown in Table 1.

BEHAVIORAL HYPOTHESES AND METHODS

There are several basic behavioral hypotheses that are typically investigated in compliance studies, including the effects of economic variables like income and wealth on compliance. The focus of our current research is the effect of information about specific audit prospects and the productivity of these audits. In the field, individuals have varying levels of information regarding the objective probability of an audit. The tax authority may formally announce an audit rate, there may be less precise information reported in the press, or the audit rate may be a complete secret. The available information allows individuals to form priors regarding the probability that they will be audited. This expected audit probability is predicted to influence behavior directly. The exact influence can be investigated by varying the information provided to individuals over time. In the lab, it is straightforward to vary the information because we can tell individuals their exact audit prospects in each period by announcing the probability of audit. Further, we can inform the individual as to whether he or she is to be selected for an audit *before* the tax return is filed or we can wait until *after* the return is filed to select for audit, and we can also tell subjects the productivity of the audit.

Expected utility maximizers will increase compliance in response to being told that they will be audited with certainty. Similarly, individuals with rank dependent expected utility or loss aversion preferences will likely focus on the “bad” outcome (e.g., being audited), and so information regarding their specific audit prospects should lead them to update their priors to

increase their subjective audit risk. Information concerning the productivity of the audit should have less impact.

Specifically, we construct these hypotheses for responses to official information announcements:

H1: Individuals who are told that they will be audited prior to their filing decision will comply perfectly.

H2: Individuals who are told that they will not be audited will record very low (zero) levels of compliance.

To the extent that paying taxes is viewed as a social contract (Alm, Jackson and McKee 1993; Frey 1992, 1997), lowering the productivity of audits may actually increase compliance. In essence, the less intrusive audits denote greater trust and engender higher “intrinsic motivation” to pay taxes, and this will be reciprocated through higher levels of compliance. On the other hand, rational self-interest would suggest that less productive audits will be viewed as an opportunity for increased evasion and compliance will fall. Further, even compliant individuals may lower their compliance when audit productivities fall as these persons may view the system as unfair: those who do cheat will be less likely to pay their full share. As noted above, for individuals exhibiting rank dependent expected utility, the effects of changes in the productivity of the audit should have little impact unless these changes are extreme. On balance, then, we have the following additional hypothesis:

H3: A lower audit productivity rate will decrease compliance.

We examine these hypotheses in several ways. Aside from analysis of simple descriptive statistics, our primary approach consists of a generalized least squares (GLS) regression model that explains income tax compliance as a function of the various fiscal parameters (plus income and “wealth”, or accumulated income). The model includes subject-specific effects to control

for individual characteristics, and allows for heteroskedastic errors across individuals.¹⁴ The experimental data constitute a panel with 144 subjects and 30 decision rounds. Each subject makes one decision in a round, or the amount of income to declare. The variables that affect this decision are the experimental treatments, the results from previous rounds, and certain subject characteristics. Given this structure, there are several options for analyzing the data. We have elected to utilize a panel estimation technique that allows us to address certain characteristics of the data at the expense of foregoing some other factors.¹⁵ We collect subject characteristic data, and we find that some of these systematically affect estimated compliance behavior. However, we acknowledge that we may be missing some unobservable effects that could be addressed with a subject random effects estimation approach. The distributional assumptions required of the random effects estimation do not seem to be justified here. Hence, we opt for the use of the cross-section time series estimation utilizing a generalized least squares estimator incorporating panel-specific heteroskedastic error terms.¹⁶

The experimental design suggests that the amount of income declared by a taxpayer in each round depends upon such factors as the individual's actual (or "true") earned *Income* in the round, the *Audit Rate*, the penalty imposed on the individual (if audited and noncompliant) in the previous round (*Lagged Penalty*), variables that represent the experimental treatments, and individual characteristics. The experimental treatment variables are discussed later. The individual characteristics include a dummy variable for whether the individual has experience with preparing his or her own tax return (*Prepare Own Taxes*) and a dummy variable for gender

¹⁴ For a detailed discussion of this model, see Wooldridge (2002).

¹⁵ Because of the experimental nature of the data, there is perfect correlation between subjects and experimental treatment variables, so that we cannot use a subject fixed effects estimation method. A subject random effects approach has a certain appeal, but such estimation imposes considerable structure on the error terms. In the end we opted for the less restrictive approach, and addressed subject effects through observed subject characteristics.

¹⁶ All estimations are undertaken using the *xtgls* estimation in STATA release 8.

(*Male*) to control for any systematic effects across subjects due to gender. The dependent variable is measured as *Compliance Rate*, or the ratio of income declared by the subject to the tax authority to true income. Summary statistics for the variables used are reported in Table 2.

The baseline model is summarized as:

$$\begin{aligned} \text{Compliance Rate}_{i,t} = & \beta_0 + \beta_1 \text{Income}_{i,t} + \beta_2 \text{Audit Rate}_{i,t} + \beta_3 \text{Lagged Penalty}_{i,t-1} \\ & + \beta_4 Y + \beta_5 Z_i + \varepsilon_{it}, \end{aligned} \quad (3)$$

where i and t are individual and round indices and $\varepsilon_{it} = u_i + w_{it}$. The traditional error term is denoted by w_{it} , and is assumed to meet all of the usual requirements. The individual-specific effect is denoted by u_i , and controls for individual level heterogeneity. The variable Y denotes the different experimental treatments, and Z_i represents individual-specific demographic variables. It should be noted that we have estimated a wide variety of alternative specifications. For example, we have investigated models in which the dependent variable is simply declared income, and we have included other explanatory variables (such as the age of the subject and controls to denote the round). Our results are quite robust to alternative specifications. All estimation results are available upon request.

RESULTS

The average compliance rates for each treatment are reported in Table 1. Holding the productivity (*Audit Productivity*) of the audits fixed, we see that pre-announcing who will be audited actually decreases overall compliance. Comparing T2 and T4 (when the audit success rate is 65 percent), we find that T2 (where audit selection is not pre-announced) has higher compliance rates than T4 (where audit selection is pre-announced); the t-statistic is 2.36, with a p-value of 0.01. Comparing T1 and T3 (when the audit success rate is 100 percent), we find again that compliance is lower when audit selection is pre-announced, and this difference is

statistically significant (the t-statistic is 1.76, with a p-value of 0.04). While we expect those who know that they will be audited to increase their compliance, those who are expecting not to be audited will lower their compliance to near zero; this latter effect requires that individuals trust the tax authority to not audit them.

Figures 1 and 2 show the individual subject compliance rates for those individuals who were told that they would be audited with certainty. The compliance rates for these individuals are in most all cases 100 percent, and there is a tendency for slightly lower compliance rates when the audit productivity rate is 65 percent versus 100 percent. Even so, an interesting anomaly in the aggregate results is that the compliance rate apparently increases as the audit productivity falls. Compliance in T2 is greater than in T1 (t-statistic=1.64; p-value=0.05), and the compliance rate is not different when we compare T3 and T4 (t-statistic=0.74).

We report the results of our GLS estimations for various specifications in Table 3. The simplest specification (Model 1) incorporates only subject factors (*Male* and *Prepare Own Taxes*), basic economic factors (*Income*), and audit factors (*Audit Rate*) and the amount the individual was penalized in the previous round (*Lagged Penalty*) conditional on being audited. The results demonstrate that the subjects in these experiments exhibit behavior similar to that reported in previous research investigating compliance behavior.¹⁷ For example, the compliance rate is increasing in the probability of audit, with an estimated *Compliance Rate – Audit Rate* elasticity of 0.34 in Model 1. Also, the compliance rate is decreasing in income, and is lower for males. Interestingly, compliance is (weakly) lower for individuals reporting that they prepare their own tax returns, a result likely due to our design parallelism and the fact that few people are actually audited. Our subjects also incorporate past audit outcomes into their compliance

¹⁷ See, for example, Alm, Jackson and McKee (1992, 1993).

decisions. The impact of *Lagged Penalty* on compliance is negative, suggestive of “catching up” or “gamblers’ fallacy” behavior.

Some more interesting results emerge as we investigate the effects of the treatments in the experimental design. In Model 2 we include a combined enforcement effectiveness variable (*Audit Productivity X Audit Rate*), and find that subjects respond to increased enforcement by increasing compliance. Separating these two enforcement parameters yields the results reported in Model 3. The subjects respond to the audit probability as expected (e.g., the coefficient on *Audit Rate* is positive and significant, with an estimated elasticity again of 0.34), but the coefficient on *Audit Productivity* is negative, although only very weakly so.

In Model 4 we include the other treatment variables. We find that pre-announcing audits lowers compliance (e.g., the coefficient on *Audit Before* is negative and significant). This is to be expected because we never audit more than forty percent of the individuals. Importantly, in our lab setting we are able to capture the behavior of those not receiving the Slemrod, Blumenthal, and Christian (2001) “letter”. Consistent with the results in the Minnesota experiment, we find that those who are told that they will be audited with certainty do in fact increase significantly their compliance, as reflected in the positive and significant coefficient on *Known Audit*. We do not investigate the high income effect reported by Slemrod, Blumenthal, and Christian (2001). Even so, however, we still find that pre-announcement lowers compliance, as reflected in the negative and significant coefficient on *Audit Before*.

The result for *Audit Productivity* in Model 3 is puzzling. To further address the effects of enforcement effort, we include in Model 5 a variable that captures the interaction between audit productivity and knowledge that one is to be audited. The coefficient on this variable, *Audit Productivity X Known Audit*, is positive and significant. Increasing enforcement (e.g., the

“closely examined” phrase in the Minnesota letter) has the intended effect on those receiving the letter. Even so, it appears that our taxpayer subjects respond to increased audit effectiveness only when they are expecting to be audited. More productive audit procedures alone are insufficient inducements to comply; instead increased compliance seems to require a heightened expectation of the audit itself.

Finally, from the above results, it seems unlikely that a tax authority could credibly use the type of letter sent by the Minnesota Department of Revenue as an enforcement mechanism. The coefficient on *Audit Before* is consistently negative and significant. Overall, compliance falls under this program (e.g., compare the compliance rates in T1 versus T3 and in T2 versus T4), and the effect of pre-announcement is too large to be offset by the behavior of those being told their returns will be reviewed; that is, the negative effect arising from the non-recipients will offset the positive effect on those who are informed they will be audited. Over time, it is to be expected that the information that such letters are being sent will become widely known, so that individuals will learn that they are less likely than usual to be audited if they do not receive such advance warning. If the tax authority does not audit those receiving the letter, then its credibility will fall. Absent additional resources, the tax authority will not have the personnel to audit non-recipients sufficiently to prevent the overall fall in compliance that we observe in our data.

CONCLUSIONS AND DISCUSSION

Our experimental design allows us to investigate the response to the pure information regarding whether an individual’s return is to be audited. We are able to measure the response to the information in ways not available to other researchers like Slemrod, Blumenthal, and Christian (2001). Further, our experiments generate information on actual compliance rates for

all the taxpayers in the setting, and we are also able to control directly for changes in audit rates and audit productivity.

We find that our taxpayer subjects respond as predicted to higher audit probability information that is provided prior to the filing decision. Those subjects who know that they are to be audited increase their compliance when so informed, and they are also responsive to increases in the audit productivity. Overall, we also find that taxpayers respond to higher audit probabilities and to the overall (compound lottery) structure of the audit. Further, we find that some of this increased compliance is offset by the responses of those who learn that they will not be audited in the current round. These results complement the findings of Slemrod, Blumenthal, and Christian (2001).

From a policy perspective, our results provide useful insights as to the potential tradeoffs between higher audit probabilities and more productive audits. Earlier we stated that, if taxpayers can correctly evaluate compound lotteries, then the compliance effect of changing the audit *probability* is the same as the effect of an equivalent change in audit *productivity*. If this holds, then tax authority can increase compliance via the less costly strategy. However, our results suggest that increasing audit productivity alone is not effective. It is only when greater audit productivity is combined with a higher audit probability that the overall effect on compliance is positive.

Our results also have implications for informational campaigns intended to increase compliance. Such campaigns may well generate some unintended consequences. It is certainly possible for the tax authority to target individuals to receive *ex ante* information on their audit status. However, the remaining taxpayers may well form conjectures that are harmful to the tax authority. Since the tax authority cannot credibly commit to auditing all or even a substantial

number of taxpayers, informing some taxpayers that their returns will be subject to additional scrutiny is in effect informing the other taxpayers (who are by far more numerous) that their returns will be ignored. If taxpayers have rational expectations, the net effect of such a program is bound to be a fall in overall compliance. This result may be surprising, but it is to be expected, and it should be taken into consideration by the tax authority when devising policies to increase compliance.

Table 1 – Experimental Design and Aggregate Results^a

	Audit Productivity?	
Audit Selection Timing?	1.00	0.65
Audit Selection Not Pre-Announced	T1 Subjects = 48 Compliance Rate = 0.515	T2 Subjects = 32 Compliance Rate = 0.546
Audit Selection Pre-Announced	T3 Subjects = 32 Compliance Rate = 0.480	T4 Subjects = 32 Compliance Rate = 0.496

Notes: Sessions in all treatments last 30 rounds, and in all treatments the tax rate is 0.35, the fine rate is 1.5, and subjects are organized into groups of eight persons. The income range is the same for all sessions; the maximum is 100 lab dollars and the minimum is 60 lab dollars, in increments of 10 lab dollars, with one person at 100, one person at 60, and two persons at all other income levels. Audit rates are 0.05, 0.10, 0.30, and 0.40 in all treatments.

Table 2 – Variable Definitions and Summary Statistics

Variable	Definition	Mean (Standard Deviation)
Compliance Rate	Declared income divided by actual income	0.509 (0.462)
Income	Income (scaled)	0.803 (0.121)
Audit Rate	Probability of audit – 0.05, 0.1, 0.3, 0.4	0.230 (0.161)
Lagged Penalty	Value of penalty imposed in previous period	0.487 (2.35)
Prepare Own Taxes	Dummy variable equal to 1 if the individual reports filing own income tax return and 0 otherwise	0.289 (0.453)
Male	Dummy variable equal to 1 if male and 0 otherwise	0.556 (0.497)
Known Audit	Dummy variable equal to 1 if the individual is told his or her return will be audited in advance and 0 otherwise	0.111 (0.314)
Audit Before	Dummy variable equal to 1 if audit selection is prior to tax filing and 0 otherwise	0.451 (0.497)
Audit Productivity	Dummy variable equal to 1 if the audit productivity is 1.0 and 0 if audit productivity is 0.65	0.563 (0.496)

Table 3 – Estimation Results

Independent Variables	Model 1	Model 2	Model 3	Model 4	Model 5
Constant	0.632*** (0.042)	0.642*** (0.042)	0.649*** (0.043)	0.737*** (0.039)	0.747*** (0.039)
Income	-0.106** (0.051)	-0.102** (0.051)	-0.111** (0.051)	-0.092** (0.046)	-0.094** (0.045)
Audit Rate	0.761*** (0.038)		0.757*** (0.013)	0.365*** (0.035)	0.365*** (0.035)
Lagged Penalty	-0.031*** (0.003)	-0.032*** (0.003)	-0.030*** (0.003)	-0.028*** (0.003)	-0.028*** (0.003)
Male	-0.347*** (0.012)	-0.347*** (0.012)	-0.347*** (0.012)	-0.340*** (0.012)	-0.339*** (0.012)
Prepare Own Taxes	-0.019 (0.013)	-0.027* (0.013)	-0.016 (0.010)	-0.023* (0.012)	-0.015 (0.012)
Audit Before				-0.273*** (0.012)	-0.274*** (0.012)
Known Audit				0.689*** (0.018)	0.650*** (0.025)
Audit Productivity			-0.021* (0.012)	-0.015 (0.011)	-0.027** (0.012)
Audit Productivity X Audit Rate		0.832*** (0.043)			
Audit Productivity X Known Audit					0.073** (0.032)
Wald Statistic	1394.6***	1352.9***	1395.3***	3259.7***	3271.6***
Log-likelihood	-2122.73	-2137.83	-2121.55	-1640.97	-1638.30

Note: The dependent variable in all estimations is the *Compliance Rate*. These estimations are panel models using feasible generalized least squares estimators (Stata 8). In all estimations, the number of observations is 4176, the number of subjects (panels) is 144, and the number of time periods is 29 (omitting period one for the lag operator). The numbers in parentheses are z-statistics. Significance levels are denoted as: * 0.10, ** 0.05, *** 0.01.

Figure 1 – Individual Compliance Rates when Audit Success Rate = 0.65

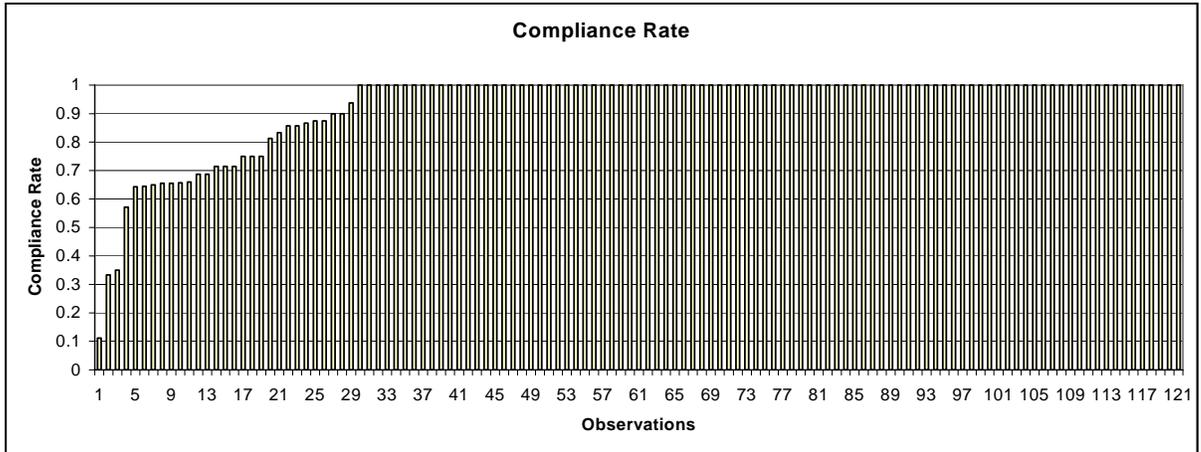
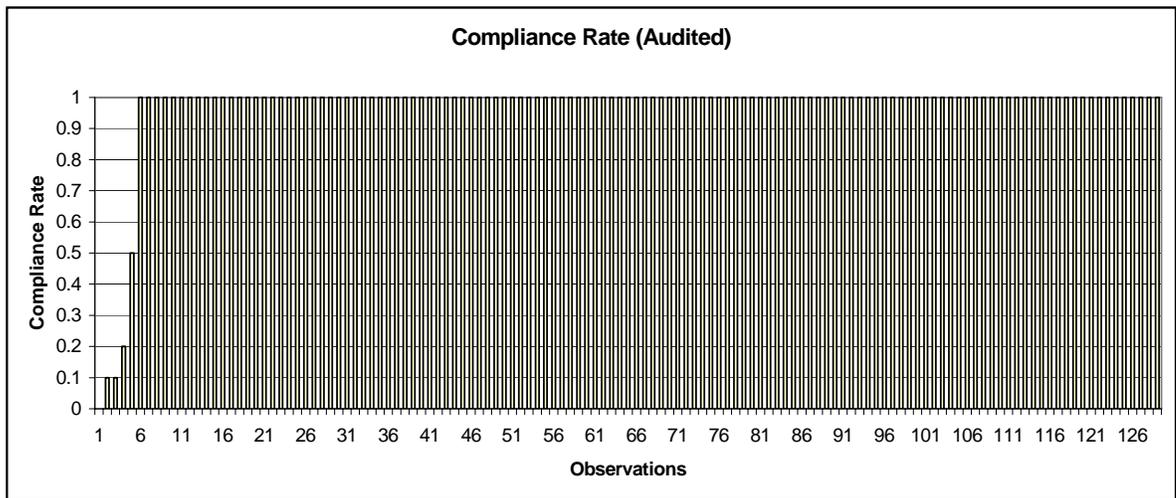


Figure 2 – Individual Compliance Rates when Audit Success Rate = 1.00



Acknowledgements

We are grateful to Kim Bloomquist and John Deskins for many helpful comments and discussions during this research. Michael Jones programmed the experiments, and David Bruner, John Deskins, and Zach Richards provided very able assistance in running the experiments. An earlier version of this paper was presented at the conference “Experimental Public Economics”, held in Atlanta, GA in May 2005 and sponsored by the International Studies Program, Andrew Young School of Policy Studies, Georgia State University. We thank conference participants for helpful comments. Funding for this research was provided by the Internal Revenue Service (TIRNO – 03 – R – 00027). All results and interpretations are our own.

References

Allingham, Michael and Agnar Sandmo. "Income Tax Evasion: A Theoretical Analysis".

Journal of Public Economics 1, No. 4 (1972): 323-228.

Alm, James. "Tax Compliance and Administration". In *Handbook on Taxation*, edited by W.

Bartley Hildreth and James A. Richardson, 741-768. New York, NY: Marcel Dekker, Inc., 1999.

Alm, James, Betty R. Jackson, and Michael McKee. "Estimating the Determinants of Taxpayer

Compliance with Experimental Data". *National Tax Journal* 65, No. 1 (1992): 107-114.

Alm, James, Betty R. Jackson, and Michael McKee. "Fiscal Exchange, Collective Decision

Institutions, and Tax Compliance". *Journal of Economic Behavior and Organization* 22, No. 4 (1993): 285-303.

Alm, James, Gary H. McClelland, and William D. Schulze. "Why Do People Pay Taxes?"

Journal of Public Economics 48, No. 1 (1992): 21-38.

Alm, James and Michael McKee. "Tax Compliance as a Coordination Game". *Journal of*

Economic Behavior and Organization 54, No. 3 (2004): 297-312.

Andreoni, James, Brian Erard, and Jonathan Feinstein. "Tax Compliance". *Journal of Economic*

Literature 36, No. 2 (1998): 818-860.

Becker, Gary S. "Crime and Punishment – An Economic Approach". *The Journal of Political*

Economy 76, No. 2 (1968): 169-217.

Becker, Werner, Hans Buchner, and Simon. Sleeking. "The Impact of Public Transfer

Expenditures on Tax Evasion: An Experimental Approach". *Journal of Public*

Economics 34, No. 2 (1987): 243-52.

Bernasconi, Michele. "Tax Evasion and Orders of Risk Aversion". *Journal of Public*

- Economics* 67, No. 2 (1998): 123-134.
- Cowell, Frank. *Cheating the Government: The Economics of Evasion*. Cambridge, MA: The MIT Press, 1990.
- Cummings, Ronald, Jorge Martinez-Vazquez, and Michael McKee. "A Portfolio Approach to Tax Evasion: Theory and Experimental Evidence". Georgia State University Working Paper (2005).
- Dubin, Jeffrey A. "Criminal Investigation Enforcement Activities and Taxpayer Noncompliance". California Institute of Technology Working Paper (2004).
- Dubin, Jeffrey A., Michael J. Graetz, and Louis L. Wilde. "The Effect of Audit Rates on the Federal Individual Income Tax, 1977-1986". *National Tax Journal* 43, No. 4 (1990): 395-409.
- Elffers, Henk. *Income Tax Evasion: Theory and Measurement*. Amsterdam, The Netherlands: Kluwer Academic Publishers, Amsterdam, 1991.
- Feinstein, Jonathan S. "An Econometric Analysis of Income Tax Evasion and its Detection". *The RAND Journal of Economics*, 22 (1) (1991): 14-35.
- Frey, Bruno. "Tertium Datur: Pricing, Regulating, and Intrinsic Motivation". *Kyklos* 45, No. 2 (1992): 161-184.
- Frey, Bruno. *Not Just For the Money – An Economic Theory of Personal Motivation*. Cheltenham, United Kingdom: Edward Elgar Publishing Limited, 1997.
- Friedland, Nehemiah, Shlomo Maital, and Aryeh Rutenberg. "A Simulation Study of Income Tax Evasion". *Journal of Public Economics*, 10, No. 2 (1978): 107-116.
- Kinsey, Karyl A. "Deterrence and Alienation Effects of IRS Enforcement: An Analysis of Survey Data". In *Why People Pay Taxes*, edited by Joel Slemrod, 259-285. Ann Arbor,

- MI: University of Michigan Press, 1992.
- Plott, Charles R. “Dimensions of Parallelism: Some Policy Applications of Experimental Methods”. In *Laboratory Experimentation in Economics: Six Points of View*, edited by Alvin E. Roth, 193-217. New York, NY: Cambridge University Press, 1987.
- Sheffrin, Steven M. and Robert K. Triest. “Can Brute Deterrence Backfire? Perceptions and Attitudes in Taxpayer Compliance”. In *Why People Pay Taxes*, edited by Joel Slemrod, 193-218. Ann Arbor, MI: University of Michigan Press, 1992.
- Slemrod, Joel, Marsha Blumenthal, and Charles Christian. “Taxpayer Response to An Increased Probability of an Audit: Evidence from a Controlled Experiment in Minnesota”. *Journal of Public Economics* 79, No. 4 (2001): 455-483.
- Smith, Vernon L. (1982). “Microeconomic Systems as an Experimental Science”. *The American Economic Review* 72, No. 4 (1982): 923-955.
- U.S. Department of the Treasury Inspector General for Tax Administration (TIGTA) (2002). Management Advisory Report: The Internal Revenue Service’s Response to the Falling Level of Income Tax Examinations and Its Potential Impact on Voluntary Compliance, Reference Number: 2002-30-092. Washington, D.C.: 2002.
- Viscusi, W. Kip and Richard Zeckhauser. “Recollection Bias and the Combat of Terrorism”. *Journal of Legal Studies* 34, No. 1 (2005): 27-56.
- Webley, Paul, Henry Robben, Henk Elffers, and Dick Hessing. *Tax Evasion: An Experimental Approach*. Cambridge, United Kingdom: Cambridge University Press, 1991.
- Wooldridge, Jeffrey M. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: The MIT Press, 2002.