Georgia State University
ScholarWorks @ Georgia State University

ExCEN Working Papers | Experimental Economics Center

1-1-2019

Audit State Dependent Taxpayer Compliance: Theory and Evidence from Colombia

James Alm
Tulane University

James Cox
Georgia State University

Vjollca Sadiraj
Georgia State University

Follow this and additional works at: https://scholarworks.gsu.edu/excen_workingpapers

Recommended Citation

Alm, James; Cox, James; and Sadiraj, Vjollca, "Audit State Dependent Taxpayer Compliance: Theory and Evidence from Colombia" (2019). *ExCEN Working Papers*. 3.
https://scholarworks.gsu.edu/excen_workingpapers/3

This Article is brought to you for free and open access by the Experimental Economics Center at ScholarWorks @ Georgia State University. It has been accepted for inclusion in ExCEN Working Papers by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.

Audit State Dependent Taxpayer Compliance: Theory and Evidence from Colombia

James Alm, James C. Cox, and Vjollca Sadiraj*

Forthcoming in *Economic Inquiry*

ABSTRACT. We develop and analyze a dynamic model of individual taxpayer compliance choice that predicts "audit state dependent taxpayer compliance," by distinguishing between the implications of forward-looking versus myopic versus naïve behavior. We then test experimentally the audit state dependent model by reporting the results from the first tax compliance experiment run in Colombia. Consistent with previous studies as well as theoretical predictions, we find that subjects' compliance rates increase with greater enforcement, especially the audit rate. We also find more novel results, both theoretically and empirically: fine rates should be increased after an audit to discourage otherwise-increased underreporting, and "nudging" myopic individuals toward reporting a constant rather than a fluctuating proportion of income would benefit both the taxpayer and the tax authority. (*JEL* H26, C91)

---

* Partial financial support for this research was provided by the National Science Foundation (grant number SES-1658743). We are grateful to individuals at Gandour Consultores in Bogota, Colombia, especially Carlos Mauricio Ortiz Nino and Diana Rocha. We also thank the editor and three anonymous referees for many helpful comments and suggestions. Please address all correspondence to James Alm.

Alm: Department of Economics, Tulane University, 6823 St. Charles Avenue, 208 Tilton Hall, New Orleans, LA 70118 (phone +1 504 862 8344; fax +1 504 865 5869; email jalm@tulane.edu).

Cox: Department of Economics and Experimental Economics Center (E*x*CEN), Andrew Young School of Policy Studies, Georgia State University, P.O. Box 3992, Atlanta, GA 30302-3992 (phone +1 404 413 0200; fax +1 404 413 0195; email jccox@gsu.edu).

Sadiraj: Department of Economics and Experimental Economics Center (E*x*CEN), Andrew Young School of Policy Studies, Georgia State University, P.O. Box 3992, Atlanta, GA 30302-3992 (phone +1 404 413 0193; fax +1 404 413 0205; email vsadiraj@gsu.edu).

I. INTRODUCTION

How should government design policies to improve tax compliance? Most answers to this question are based upon, or at least consistent with, the standard economics-of-crime model first applied to tax compliance by Allingham and Sandmo (1972), which assumes myopic and static behavior on the part of an individual who balances the expected costs of detected cheating with the expected benefits of successful cheating. Most answers also require estimating the impact of various policy tools on individual compliance decisions, a task made challenging by the absence of detailed and reliable information on individual compliance choices. By its very nature, people have an incentive to hide information on their evasion behavior, and this concealment makes empirical work quite difficult. In the United States, researchers have found increasingly creative ways to estimate the factors that motivate compliance by using naturally occurring field data (including administrative data), controlled field experiments, and laboratory experiments. These same empirical approaches have sometimes been applied elsewhere as well, but the extent of empirical work in other countries is far more limited.[1]

In this paper, we first develop a dynamic theoretical model of the individual compliance decision that makes it possible to identify "audit state dependent" compliance, in which an individual's compliance decision in the current period depends on the individual's previous audit history from earlier periods.[2] Our theoretical model distinguishes between compliance conditional on no previous audits and compliance conditional on previous audits, which makes it possible to discriminate in our empirical analysis among three types of behavior: forward-looking versus myopic versus naïve behaviors. A forward-looking taxpayer takes into account how current compliance choices affect future exposure to enforcement policies. One who is myopic is concerned only with

[1] For a recent and comprehensive survey of the tax compliance literature, see Alm (2019). For earlier discussions of much of this work, see Cowell (1990), Andreoni, Erard, and Feinstein (1998), Slemrod and Yitzhaki (2002), Kirchler (2007), Alm (2012), and Hashimzade, Myles, and Tran-Nam (2013).

[2] We are grateful to an anonymous referee for the suggestion that we emphasize more explicitly an individual's audit history in our discussion of the dynamic theoretical model.

instantaneous utility, but takes account of accumulated unreported income from earlier compliance decisions. A naïve taxpayer looks only at instantaneous utility; also this taxpayer "isolates" the current reporting decision and ignores any accumulated unreported income from earlier compliance decisions, as suggested by the isolation hypothesis of behavioral economics. Previous work applying static models to data analysis has not been able to identify either audit state dependent compliance or the impact of changes in audit rates on these three different types of decision-makers.

We then report the results of the first experiment to examine taxpayer compliance decisions in Colombia. The experiment includes treatments that incorporate changes in the full range of fiscal parameters (e.g., audit rates, fine rates, and tax rates), as well as a treatment in which tax revenue is donated to a charity (similarly to contributions to public good in other tax evasion experiments). The subject pool includes full-time students, part-time students, workers, and unemployed individuals, which makes it possible to examine subject pool effects. With these experimental data, in combination with our theoretical model, we are able to estimate for the first time the determinants of audit state dependent taxpayer compliance behavior, and to do so for a country (Colombia) with no previous history of experimental studies of tax compliance. The similarity of our experimental design to the tax compliance enforcement in most countries, as well as to the designs in many previous experimental studies raises the possibility that application of our dynamic model to data, rather than the usual static model applied in much of the literature, could lead to different conclusions about taxpayer responses to policy innovations.

Colombia has a long and rich tradition of sophisticated tax policy analyses and tax reforms, going back to several studies in the 1960s (Taylor and Richman, 1965; Bird, 1970; Gillis, 1971), continuing into the 21st century (Bird, Poterba, and Slemrod, 2005), and extending to the tax reform recently enacted by the Government of Colombia in December 2016. Such reforms can be informed by estimates of taxpayer compliance behavioral responses to policy innovations, estimates that would make any tax reform efforts more meaningful. We provide these estimates here.

There are of course reasons for caution in attempting to use our results to explain behavior outside the laboratory. Even so, the laboratory offers an opportunity to investigate in a controlled environment individual responses to audit, fine, and tax rate changes, as well as to other policies for which field data do not readily exist, while controlling for economic and demographic characteristics of the responders. Further, unlike much empirical work that relies on imperfect proxies for evasion, the measures that we use from the laboratory are accurate and unambiguous measures of individual noncompliance, derived in a setting that controls explicitly for extraneous influences on individual behavior.

Not surprisingly, we find that reported income increases with an increase in enforcement, especially the audit rate. However, we also find the reported income decreases following the event of an audit. Previous work applying static modeling has not been able to examine such distinctions, and concluded simply that taxpayer reporting increases with the audit probability. We explain why such audit state dependent compliance has new implications for tax policy design. We also find some evidence that taxpayer reporting is greater when payments are donated to a charity. Finally, we estimate the effects of all of these fiscal variables on two other variables of interest, tax revenues and earned income. In all of our results, we see no evidence of differences in behavior between subject types, regardless of whether they are students, workers, or unemployed.

Our paper makes several contributions. Our theoretical model of audit state dependent compliance is the first to differentiate between myopic, forward-looking, and naïve behavior. Unlike experimental papers with static theory but sequential subject responses, the dynamic theoretical model that we apply in our data analysis is consistent with the sequential choices in our experimental design. Finally, this paper is the first to use experimental methods to examine compliance behavior in Colombia, and to use these data to examine subject pool effects there.

II. THEORETICAL BACKGROUND

We present here a discussion of the dynamic theoretical model that is further developed in Appendix 1.

A. *Modeling Tax Compliance Behavior*

Consider an individual who earns income w and must choose the proportion R to report to the tax authorities. At time of the decision, let the cumulative amount of unreported income from previous time periods be x. The individual pays taxes at tax rate τ on each peso of reported income. Unreported income is not taxed, but the individual may be audited with a fixed audit rate p at which point the individual must pay taxes on unreported income $[x+(1-R)w]$ plus pay a fine at rate f on each peso of previously-unpaid taxes.[3]

Without any loss of generality, write accumulated unreported income as a multiple of current income, $x = zw$ to simplify notation. The instantaneous expected utility of the taxpayer from reporting R is [4]

$$\begin{aligned} &EU(R,z) = pU(\pi_1(R,z)) + (1-p)U(\pi_0(R,z)) \\ &\textit{where} \\ &\quad \pi_s(R,z) = w[1-\tau R - s\tau(1+f)(z+1-R)]; \quad s = \begin{Bmatrix} 0, & \textit{if not audited} \\ 1, & \textit{if audited} \end{Bmatrix} \end{aligned} \tag{1}$$

<u>A.1. A Forward-looking Decision-maker.</u> A "forward-looking" agent takes into account how current compliance choices affect future exposure to audits and fines, in addition to any accumulated unreported income. To allow for a dynamic process of individual choice, we denote the time-variant variables with a subscript *t* for the period. In each period *t*, the individual earns

---

[3] Note that we assume that the fine is assessed on previously unpaid *taxes*, as in Yitzhaki (1974), rather than on unreported *income*, as in Allingham and Sandmo (1972).

[4] Again, for simplicity we do not here introduce a public good that is financed by individual tax payments, although a public good is included in the model in Appendix 1.

income w and must choose the fraction R_t to report.[5] The individual plays the game that continues with probability $\delta \in (0,1)$.[6] The decision problem for a forward-looking individual is

$$\begin{aligned} &V(-\alpha z_0) = \max_{R_t \in [0,1]} E(\sum_{t=0}^{\infty} \delta^t [U(\pi_s(R_t, Z_t))] \\ &Z_t = X_{t-1}(1 - R_{t-1} + X_{t-2}(1 - R_{t-2})); X \sim Bernoulli(1-p) \end{aligned} \tag{2}$$

where X_t takes value 1 if not audited and 0 if audited at time t. We use state 00 to refer to the case where the individual is not audited in either of the two previous periods, state 1 to refer to the case when the individual is audited in the last period, and state 10 otherwise. At the time (t) of decision, the individual knows whether he was audited or not, so he knows the realization of Z_t. The Bellman equation in state 00 is

$$\begin{aligned} &V(-\alpha z_t) = \max_{R_t \in [0,1]} \left[EU(R_t, z_t) + \delta\left(pV(0) + (1-p)V(-\alpha z_{t+1})\right)\right], \quad \forall t \\ &z_{t+1} = 2 - R_{t-1} - R_t \end{aligned} \tag{3}$$

Optimal compliance rates are determined as solutions to the following optimality conditions (Euler equations, derived in Appendix 1)

$$\left[pfU'(\pi_{1t}) - (1-p)U'(\pi_{0t})\right] = -\delta(1-p)\left[pU'(\pi_{1t+1}) + (1-p)U'(\pi_{0t+1})\right], \quad \forall t \tag{4}$$

where $\pi_{1s} = \pi_1(R_s^d(z_s), z_s)), s = t, t+1$.

In the neighborhood of the steady state (z^o, R^o), the marginal effects of parameters of interest on optimal compliance rates $R_t^d = h(z_t)$ are derived using first order Taylor approximations at the steady state (z^o, R^o), or

$$h(z_t) \approx h(z^o) + h'(z^o)(z_t - z^o) \tag{5}$$

[5] For simplicity, income w is assumed to be constant over time in the theoretical derivation. Our experimental design allows income to vary, and this variation is included in the empirical estimation.

[6] The parameter δ captures various scenarios. If the stochastic game terminates with probability q and the individual's discount factor is β, then $\delta = \beta(1-q)$. If the game is played infinitely times but the individual discounts future payoffs, then δ is the discount factor. In our experiment, δ captures the subject's belief of the likelihood that the experiment continues.

We show in Appendix 1.B that in state 00: $h'(z^o)\in(0,1)$, there can be no more than one steady state, and the steady state compliance rate R^o is implicitly determined by

$$U'(\pi(R^o,2-2R^o))+\gamma U'(\pi(R^o))=0 \tag{6}$$

where $\gamma=\left(1-\frac{1}{p}\right)\frac{1-\delta(1-p)}{f+\delta(1-p)}<0$.

As further demonstrated in Appendix 1.B, the compliance rate after an audit is smaller than the steady state compliance rate in state 00 and the state 00 optimal compliance rate at time *t*

(1) increases with the fine rate;
(2) increases with the audit rate;
(3) increases with the tax rate;
(4) increases with income for CARA and IARA preferences;
(5) increases with probability of continuation of the game (or patience); and
(6) increases with the accumulated unreported income rate

A.2. A Myopic Decision-maker. A "myopic" model that incorporates only instantaneous utility (1) might be a good model for some taxpayers and for some experimental subjects. As discussed in Appendix 1.C for a myopic decision-maker, the signs of the marginal effects of the main determinants (e.g., accumulated unreported income rate; fine rate; audit rate; tax rate; and income) are the same as for a forward-looking decision-maker. However, in comparison to a forward-looking decision-maker, the optimal compliance rates R_t^m of a myopic decision-maker do not depend on the probability of continuation of the game and they also do not stabilize at an (interior) steady state. An implication of the former is oscillating patterns of compliance rates for a myopic decision-maker, as illustrated below in subsection II.B.1.

A.3. A Naïve Decision-maker. In the light of research in behavioral economics, we also look at a naïve decision-maker who maximizes instantaneous utility (1) but, unlike a myopic decision-maker, "isolates" individual decisions[7] and ignores any accumulated unreported income that is a consequence

[7] The "isolation hypothesis" was first stated by Kahneman and Tversky (1979), and subsequently it has become a recognized feature in behavioral economics.

of earlier compliance decisions. The sign of the marginal effects of the exogenous determinants (e.g., fine rate, tax rate, audit rate, and income) are the same as for the myopic model. However, in comparison to a myopic decision-maker, compliance rates of a naïve decision-maker at accumulated unreported income *z* do not depend on the accumulated unreported income rate.

B. *Illustrative Examples of Compliance in the Experiment*

In our experiment, an audited individual is responsible for current unreported income as well as for the unreported income during the two most recent periods. This is a simplified representation of government auditing practices in which taxpayers found to be underreporting in the current period may be subject to audits that examine previous years of reporting.[8]

B.1. Myopic Behavior. The dynamic model points to the cost of myopic decision-making. A myopic individual makes choices that at the interior satisfy optimality condition for problem (1)

$$\left[pfU'(\pi(R_t,z_t)) - (1-p)U'(\pi(R_t))\right] = 0 \qquad (7)$$

Because unreported income at time *t* will be included in accumulated unreported income at time *t*+1 if the individual is not audited, the optimal compliance rate (in state 00) is expected to oscillate with *t* for myopic decision-making. However, by concavity of $U(\cdot)$ the individual is better off reporting the same intermediate level of income in all rounds if he is not audited. We use CARA preferences to provide an example that illustrates this property.

Consider a myopic decision-maker with CARA preferences and absolute risk aversion, $\lambda > 0$, represented by $U(\pi) = (1 - e^{-\lambda\pi})/\lambda$. By equation (7) the optimal compliance rate is $\max\{0, \min\{R^*(z), 1\}\}$ where

[8] See Rickard, Russell, and Howroyd (1982) for theoretical analysis of this "conditional back audit" rule. Also, see Greenberg (1984) and Landsberger and Meilijson (1982) for theoretical analyses of a "conditional future audit" rule, in which detection of underreporting in the current period leads to audits in future periods. Alm, Cronshaw, and McKee (1993) use laboratory methods to examine these types of endogenous audit selection rules, and they find that endogenous audit selection rules are far more effective than random audit selection rules in deterring noncompliance.

$$R^*(z) = z + 1 - \frac{1}{\lambda w(1+f)\tau} \ln(\frac{1-p}{pf}) \tag{8}$$

Let a denote the last term (the expression after the minus sign), or $a = \frac{1}{\lambda w(1+f)\tau} \ln(\frac{1-p}{pf})$.

Suppose that $a \in (0,1)$.[9] At the beginning of the game, or having just been audited (state 1), the accumulated unreported income is 0 and therefore the optimal compliance rate is $1-a$. In the round after, if the individual is audited (the state is 1), then $z = 0$ so he will again report $1-a$. If he is not audited (the state becomes 10), then $z = a(=1-(1-a))$, and, by the optimality condition (9), his compliance rate is 1. Again, if he is audited (the state is 1), he goes back to compliance rate $1-a$. However, if he is not audited (i.e., state 00), his unreported income remains $a(=2-(1-a)-1)$, so it is optimal for him to report 1 again, after which he goes back to report $1-a$ (audited or not). Summarizing, the pattern of compliance rate is therefore $(1-a,1,1,1-a,1,1,1-a,1,1...)$. Given that $R \in [0,1]$, the patterns of compliance depend on the value of a (i.e., risk attitudes and fiscal parameters) and are as follows:

Case 1: $a < 0$. The compliance rate is 1, always.

Case 2: $a \in (0,1)$. The compliance rate is $1-a$ followed by 1 twice, or $(1-a,1,1,1-a,1,1,1-a,1,1,1-a,...)$.

Case 3: $1 < a < 2$. The compliance rate alternates between 0, 2-a, and 1, or $(0,2-a,1,0,2-a,1,0,2-a,1,0...)$.

Case 4: $a > 2$. The compliance rate is 0, always.

B.2. Numerical Illustration. For CARA preferences with w=240 and $\lambda = 0.026$, we find that it is optimal for the myopic individual to alternate between reporting a fraction of income 0.10 followed by full compliance for two subsequent rounds when in state 00. Such a pattern of compliance for

---

[9] This assumption is relaxed below.

myopic decision-making generates a sequence of expected utilities that alternate between 38.36 with a frequency of 1 and 37.90 with a frequency of 2. In this example, our individual's optimal compliance following two rounds of full compliance is the same as his optimal compliance following an audit. In contrast, the round expected utility from reporting, say, a constant fraction 0.76 of income is 38.13 (assuming that the state remains 00). It is straightforward to show that the stream of tax revenue from reporting a constant 0.76 fraction of income is also greater than the revenue stream from the reporting sequence {0.1, 1, 1}.

These results suggest that policies that remind taxpayers how current compliance choices affect future exposure to enforcement policies could nudge myopic individuals toward reporting more like forward-looking taxpayers, which would benefit the individual while also increasing tax revenue.

C. *Some Testable Implications*

The optimal fraction of income R_t that an individual chooses to report in period t depends upon $[p, f, \tau, w, z_t]$, of which the fiscal parameters are the audit rate p, the fine rate f, and the tax rate τ. This function can be written generically as $R_t=\phi(p, f, \tau, w, z_t)$.[10] Stated as hypotheses, the main theoretical implications for our experiment are as follows:

> *H_1: For any given state of the world, an individual's compliance rate is higher (lower) with a higher (lower) audit, fine, or tax rate.*
>
> *H_2: An individual's compliance rate is lower immediately following an audit than in state 00 (where the individual is not audited in either of the two previous periods).*
>
> *H_3: An individual's compliance rate is higher the larger is accumulated unreported income rate.*

The next section discusses our experimental design.

[10] As robustness checks, we use alternative specifications for $\phi(\cdot)$ in our empirical work.

III. EXPERIMENTAL DESIGN

A. *General Design Features*

Our experimental setting implements the fundamental elements of the voluntary reporting systems of Colombia's individual income tax, as well as of Colombia's social insurance program. Participants earn income, and they must choose how much income to report to the tax authority. The participant pays taxes on reported income but not on unreported income. However, the individual faces a fixed probability of audit. If the individual is audited, then any unreported income in the current period and in the previous two periods is discovered, and the individual must pay taxes on all unreported income and a fine on previously unpaid taxes.[11] Subjects are fully and accurately informed about the various features of the tax rate, audit rate, and fine rate used in their experimental session.

Participants were all adults, recruited from several major Colombian universities, from Colombian workers (both employed and self-employed), and from the unemployed.[12] No participant had prior experience in this experimental setting. Methods adhere to all guidelines concerning the ethical treatment of human subjects.[13]

Upon arrival at the laboratory, participants were reminded they had been recruited for two hours and then were assigned to a computer station with each station being situated in an isolation

[11] Note that the experimental and theoretical models are aligned on the feature that audits apply to the current and two previous rounds. This design feature was chosen to implement, in a simple, clear, and understandable way, tax return audit practice in which a single-year audit is often extended to include more than one year's return if underreporting is first detected. This is the common practice in Colombia, as well as world-wide. However, actual audit practices can be very complicated. For example, in the United States the Internal Revenue Service (IRS) typically has three years to audit returns after they are filed. If even modest underreporting is found in one return, then the IRS will often audit returns from the two preceding years; if underreporting has been found to exceed 25 percent, the audit can go back six or even more years; and if there is evidence of fraud, there is no limit to the number of years the IRS can go back.

[12] Subjects were recruited with flyers posted in universities and government agencies and also through social networks (e.g., Facebook). Interested individuals registered online in a Google format, which required them to report his or her personal information; over 300 individuals registered online. Participating subjects were then selected randomly from the registered individuals.

[13] All Colombian subjects provided written consent for their participation in the laboratory experiment, and the Institutional Review Board at Georgia State University approved the analysis of the experimental data.

carrel. The experimental laboratory consisted of networked computers, a server, and software designed for this series of experiments. Basic instructions were provided via a hardcopy and also via a series of screen images; a representative set of instructions is provided in Appendix 3. There was no interaction between the participants and the person running the experiment beyond the initial seating of the participants at terminals and the reading of the consent document. Decisions were made privately, and participants were not allowed to talk with one another during the session. Participants were informed that all responses were anonymous, that no individual identification would be collected, and that the only record of participation would be the receipt signed to receive payment at the end of the session. Participants were also told that payments would be made in private at the end of the session. Taken together, these experimental procedures effectively eliminate both subject-to-subject interaction and subject-to-experimenter interaction.

The detailed steps of the experiment can be briefly described. At the beginning of the session, a participant received an endowment of 500 experimental coins, equivalent to 5000 Colombian pesos. Then at the beginning of each round of the session, participants were presented with a simple task that required them to add numbers; their performance on this task determined their earned income. Subjects were then presented with a screen that provided the details of the treatment in effect. They were informed with certainty of the tax rate, the audit rate, and the fine rate. Each subject then chose the amount of income to report to the tax authority. The computer automatically calculated the resulting tax liability. Participants were able to experiment with different reports during the time allowed for filing, and they could observe the potential changes in their reported take home income.

The process of determining who among the filers is audited was generated by a computerized draw. After each subject reports income, the participant is presented on his or her computer screen with an animated (computerized) representation of a bucket from which a draw is made. In this bucket there is a fixed number of balls (either blue or white), with a white ball signifying no audit, a blue ball denoting an audit, and the number of blue balls relative to the number of white balls determining the audit rate. Each taxpayer is audited independently. The balls "bounce" in this bucket, and then a door opens and a ball exits the bucket through this door, with the color of the ball indicating whether the individual is audited. The audit applies to the current period declarations of taxable income and to the two previous periods. The computer automatically deducts taxes paid and penalties (if any are owed) from each audited participant's account. After the audit process has been completed, each subject is presented with a new screen that provides the earnings and audit outcome summary for the round.

Our experimental setting was very contextual in order to promote "parallelism" to the naturally occurring world (Smith, 1982; Plott, 1987). Our experimental interface and instructions used tax language throughout; the participants decided how much income to report in the same manner as on the typical tax form (e.g., entering income on a tax form). There was a time limit on the subjects in making the reporting decision. Of some importance, the main policy parameters (e.g., the tax, audit, and fine rates) were set in their baseline values to be in alignment with actual values in Colombia.[14] For example, the individual income tax rate in Colombia has ranged up to 39 percent in recent years, with an average of about 30 percent, and the baseline value of the tax rate in the experiment was set at 30 percent. Similarly, the fine rate on unreported taxes in Colombia is 60 percent (plus payment of unreported taxes), and the baseline fine rate in the experiment was set at 60 percent. The "audit coverage" (defined as audit actions/total audits) has varied from about 5 percent to 30 percent since around 2000, and the baseline audit rate in the experiment was set at 10 percent. The actual audit rate over this period was somewhat lower, between 1 and 3 percent (Vazquez-Caro and Slemrod, 2005), and in one of the experimental

[14] For details of the current Colombian income tax regime, see https://www.dian.gov.co/ . See Vazquez-Caro and Slemrod (2005) for a somewhat dated if still relevant discussion of Colombian tax administration procedures.

treatments the audit rate was set at 1 percent. The policy parameters varied across treatments, as discussed below.

Participants were not told the exact duration of the experimental session, which was predetermined to last for 20 real rounds. They had been told they were recruited for a session that could last at most two hours. Including instructions, 3 practice rounds, and the 20 real rounds, sessions actually took on average 70 minutes to complete. Each subject's payoff was realized at the end of each round, and each subject was paid the sum of payoffs in all 20 real rounds at the end of the experiment. Each session had either 15 or 16 subjects. In all, 122 subjects participated in these sessions, from a sample pool of over 300 potential participants. Salient payoffs averaged 35,390 Colombian pesos (about $17), which is more than 10 times the minimum hourly wage in effect at the time of the experiment.

B. *Experimental Treatments*

Recall that our main objective is to estimate the impact of independent changes in the main fiscal parameters (p, f, τ) on individual compliance choices. We do this by conducting treatments that differ from the baseline only in the value of a single fiscal parameter. In addition, we conduct one session in which the subjects' total tax payments are aggregated and donated to a Colombian nonprofit organization, *Fundación Cardioinfantil*, a medical institution that helps children with heart disease. This session is designed to examine the impact of public good provision on tax compliance decisions.[15] Table 1 summarizes the experimental treatments.

[15] There are various ways of introducing a public good in tax evasion experiments. An obvious way mimics standard voluntary contribution mechanism experiments, in which the taxes (or contributions) of all subjects are combined, increased by some multiple, and distributed to subjects in equal amounts; see Ledyard (1995) and Janssen and Ahn (2006) for comprehensive surveys of public good experiments. The same mechanism can be used in tax compliance experiments. For example, see Becker, Buchner, and Sleeking (1987), Alm, McClelland, and Schulze (1992), and Fochmann and Kroll (2016) for examples of this mechanism in tax compliance experiments; see Alm and Jacobson (2007) for a survey of this work. Another way that attempts to achieve more parallelism between subjects and the naturally occurring world is to make the public good a real-world charity, in which subject taxes are distributed to the charity. See Alm, Jackson, and McKee (1993) and Christian and Alm (2014) for examples of this approach. Using a real-world charity has also been used in other laboratory experiments, such as dictator game experiments (Eckel and Grossman, 1996). Note that there is a large literature that uses real-world charities in laboratory and field experiments on charitable giving. For a recent prominent example, see Filiz-Ozbay and Uler (forthcoming).

<Table 1 about here>

To establish a baseline, Treatment 1 (T1) sets the three fiscal parameters at baseline levels: p=10 percent, f=60 percent, and τ=30 percent. Note that a 60 percent fine rate means that the audited individual must pay unreported taxes plus an additional fine of 60 percent, so that the "effective" fine rate is 1.60. The other treatments differ by a single parameter while holding constant the others at their baseline values. In Treatment 2 (T2), the fine rate is increased to 120 percent (or an effective fine rate of 2.20); in Treatment 3 (T3), the fine rate is lowered to 30 percent (or an effective fine rate of 1.30). In Treatments 4 and 5, the tax rate is varied, increasing to 45 percent in T4 and decreasing to 10 percent in T5. In Treatments 6 and 7, the audit rate is varied, to 20 percent in T6 and to 1 percent in T7. Finally, Treatment 8 (T8) sets the fiscal parameters at their baseline (T1) levels, but donates all subjects' taxes to the nonprofit organization.

C. *Expected Value Calculations*

A risk-neutral individual will make choices so as to maximize the expected value of the compliance gamble. Accordingly, it is useful to calculate the expected value in the treatments. These calculations also provide support for the hypotheses from the earlier discussion.

For example, in the baseline session (T1), the expected value EV equals $0.7w$ when the individual reports fully and honestly (or R=100 percent). In contrast, the expected value from reporting zero income is $0.87w$.[16] In the baseline session, the optimal risk neutral strategy is therefore to report zero income. More generally, optimal individual decisions for any linear payoff function will tend to exhibit all-or-none behavior; only very large changes in parameter values alter

[16] To illustrate this calculation, recall that the effective fine rate is 1.60 (=1+0.60) because the individual must pay unreported taxes plus a fine of 60 percent. In state 1, expected income is $0.952w$ because the individual pays taxes plus a fine on w if audited; in state 10 expected income is $0.904w$ because the individual pays taxes plus a fine on $2w$ if audited; in state 00 expected income is $0.856w$ because the individual pays taxes plus a fine on $3w$ if audited. The expected value of strategy "always report 0 income" is $0.87w$ as our tax game is an ergodic Markov chain with steady-state probabilities 0.1 (state 1), 0.09 (state 10) and 0.81 (state 00).

this outcome. As we demonstrate below, our experimental results are largely consistent with all-or-none behavior. They are also largely consistent with lower compliance following an audit.

IV. Results

A. *Aggregate Data*

Table 2 shows the main summary statistics for compliance in each of the eight treatments. The overall mean *Compliance Rate* (calculated as the ratio of reported income to earned income) is 72.0 percent. The overall median *Compliance Rate* is 86.0 percent. Six subjects reported zero income in every round, and 30 subjects reported full income in every round. Full and zero compliance account for 57.8 percent and 17.5 percent of 2437 observations, respectively.

<Table 2 about here>

Following an audit, full compliance decreases to 49.5 percent while zero compliance increases to 22.2 percent. The demographic characteristics of the subjects are reported in Table 3, including dummy variables for whether the subject was *Born in Bogota*, *Lives in Bogota*, lives with a *Parent Household Head*, *Owns a House*, is a *Full-time Student*, and is a *Female*; asterisks denote significant differences in the variable across treatments, with Treatment 1 (T1) the baseline comparison.

<Table 3 about here>

B. *Tobit Estimates with Individual Data*

We look at individual data to test for the effects of variations in the experimental parameters. The individual data also allow us to examine the impact of demographic characteristics, including subject pool effects.

The common empirical approach used with data like ours involves estimation of a Tobit model because of bounds on the dependent variable, and we utilize this Tobit approach in several specifications. Table 4 presents Tobit regressions with random effects with an upper bound at 1 and a lower bound at 0 for the dependent variable, the observed *Compliance Rate* (defined as Reported Income/True Earned Income) for each individual in each round of the treatment; standard errors are calculated using bootstrap methods. Estimates reported in the left-most three columns use data from all 122 subjects. Estimates for 86 subjects reported in the right-most three columns do not include data from 36 subjects who either reported 0 or full income in every round, since responses to any changes in the environment are less informative for individuals who do not change their behavior.

<Table 4 about here>

Based upon the estimation results in Table 4, the impacts of the policy parameters are partly consistent with the theoretical implications summarized in hypothesis H_1. A higher fine rate has a weakly significant effect on increasing the compliance rate for the 86 subjects choosing in the interior. A higher audit rate has a strongly significant effect in increasing compliance in tests reported in all columns. Calculating elasticities at the baseline treatment values for the policy parameters (audit rate=0.1, fine rate=0.6, and tax rate=0.3) and at means of the other variables (*Accumulated Unreported Income Rate*, *Income*, *Round*), the fine rate has a significant effect on compliance (using the results for 86 subjects), with an estimated elasticity of 0.41. For the audit rate, the estimated elasticities are 0.66 (for all subjects) and 0.51 (for 86 subjects), which are consistent with most other empirical estimates derived either from field data or laboratory data (Alm, 2019). The tax rate estimate is not statistically different from 0 in any specification, and *Income* is also not statistically significant.[17] Of some note, the use of the tax payments (*Donation*) has a positive and statistically significant impact on compliance for the full sample of 122 subjects.

[17] It is not uncommon to find little or no impact of the amount of income on the compliance rate. See Alm and Malézieux (2019) for a meta-analysis of laboratory experiments on the impact of income (and other variables) on optimal compliance rates.

The dummy variable for *Audited Last Round* has a highly significant negative coefficient: the compliance rate is estimated to decrease by more than 32 percent following an audit event. The observed decrease in compliance after an audit is a result predicted by the dynamic model and stated in hypothesis H_2. Note that our empirical finding is consistent with the so-called "bomb-crater effect", in which an audited individual reduces his or her immediate post-audit compliance before recovering somewhat in succeeding rounds. This effect is often found in laboratory experiments (Mittone, 2006; Maciejovsky, Kirchler, and Schwarzenberger, 2007; Kastlunger et al., 2009), although previous experimental studies that observed the bomb-crater effect did not provide a theoretical explanation as we do here. Results reported in field studies are mixed: Mendoza, Wielhouwer, and Kirchler (2017) report bomb crater effects while Erard (1992) and Advani, Elming, and Shaw (2015) report the opposite. Weak bomb crater effects in field data are consistent with the theoretical model of naive decision-making in subsection II.A.3. They also could reflect taxpayers' belief that taxpayer-specific audit probability increases after a previous audit detects underreporting, whereas audit probability is typically held constant in an experiment. The possible implications of (weak or strong) bomb crater effects for tax policy are discussed in section V.

As reported in Table 4, demographic variables are largely insignificant, including dummy variables for whether the subject is *Born in Bogota*, is a *Full-Time Student*, owns a home (*Own House*), lives with a *Parent Household Head*, lives in Bogota (*Lives in Bogota*). Importantly, we find no evidence of subject pool effects; that is, there is no significant difference in behavior between students and other demographic groups, consistent with the results of Alm, Bloomquist, and McKee (2015). *Female* subjects seem more compliant than male subjects, a result that has generally been found in other laboratory experiments (Alm, Jackson, and McKee, 1992).

The positive estimated coefficient of *Accumulated Unreported Income Rate* is consistent with the theoretical prediction stated in hypothesis H_3.[18] The estimated elasticity for the basic model is 0.07 for all subjects and slightly larger 0.12 for selected subjects, indicating in both cases relatively inelastic compliance rates. We find no wealth effect from *Accumulated Payoff*, which is consistent with other experimental papers that accumulate payoff after each decision.[19] Negative *Round* effects are expected for the dynamic model with δ^t decreasing with t, but no effects are expected either for the myopic model or for the dynamic model with fixed δ (given that subjects did not know that the last round was round 20). While we see a significant negative estimate for the *Round* variable, the effect disappears when demographics are added.

C. *Robustness Tests (1): Tobit Estimates with Robust Standard Errors Clustered at the Individual Level*

A potential concern with the estimates in Table 4 is that the observations are not independently drawn; that is, although each subject went through 20 rounds of decisions in a given treatment, these rounds are not independent from one another. To examine this issue, we estimate the basic specifications reported in Table 4 but use robust standard errors clustered at the individual level. Our main results are unaffected.[20]

D. *Robustness Tests (2): Hurdle Estimates of Individual Data*

As suggested by Cragg (1971), it is possible to examine the individual as making two related but distinct compliance decisions: the first is the decision on whether to report all or some of the true earned income, and the second is the decision on how much of true earned income to

[18] We thank a reviewer for raising a potential issue of endogeneity of *Accumulated Unreported Income Rate* during the previous two rounds and suggesting exogenous variation of task difficulty as an instrument. We defined *Task Difficulty* as the minimum earned income by any subject in a given round across all treatments. The first stage (random effects) GLS estimate of *Task Difficulty* is -0.146 (significant at 1%). The second stage (random effects) Tobit regression produces estimated coefficients that are essentially the same as the ones reported in Table 4.

[19] Previous papers that used this "pay all sequentially" protocol and that performed careful analyses for possible wealth effects also found none that were significant (Cox and Epstein, 1989; Cox and Grether, 1996; Cox, Sadiraj, and Schmidt, 2015).

[20] These results are unreported but are available upon request. We are grateful to an anonymous referee for suggesting this estimation.

report (conditional upon reporting some income). We also estimate this "hurdle" model of Cragg (1971) as an alternative to the Tobit results in Table 4. These results are reported and discussed in detail in Appendix 2. The hurdle results are largely consistent with the Tobit results.

E. *Empirical Validity of Myopic Behavior*

As discussed in subsection II.B.1, compliance rates of a myopic individual would vary between lower and higher values in a periodic way, while a flat pattern is predicted for a forward-looking individual. As discussed in section II.A.2 and II.B.1, the compliance rates do not depend on the continuation of the game, so they do not directly depend on the time period. To get some insights on myopic behavior we look closely at individual data. Approach 1 focuses on consistency of "lambda" (risk attitudes, section II.B.1) while Approach 2 looks at alternating patterns.

E.1. Approach 1. We created a new variable, $\lambda_t = \dfrac{\ln((1-p)/pf)}{(z_t + 1 - R_t) w_t (1+f)\tau}$, using the optimal $R^*(z)$ specification (8) discussed earlier. This variable is undefined when full income is reported and accumulated unreported income is 0, or when $z_t = 0$ and $R_t = 1$. There are 73 (out of 122) subjects with a well-defined λ_t for at least 50 percent of the rounds. For a myopic decision-maker, λ_t does not vary with *t*, and we selected round 12 for our testing. For a myopic subject *i*, the null hypothesis is that $\lambda_t^i = \lambda_{12}^i$. If the *t-test* fails (at 10% significance level) to reject the null hypothesis, then the subject is classified as myopic. By this criterion, 37 percent of the subjects (or 27 out of 73) are classified as myopic decision-makers.

E.2. Approach 2. For each subject, we created a new variable defined as the difference in the compliance rate (dcr) of a subject's observed choices in two subsequent rounds when in state 00. For periodic compliance rates, the values of the difference variable are positive, 0, or negative. For each subject, we then conducted a *t*-test on the absolute value of the difference variable $(|dcr|)$. If the *t*-test does not reject the hypothesis that the mean of $|dcr|$ is larger than 0.1, then we classified the subject as "myopic". Using this criterion, we find that more than 35 percent (or 43 out of 122) of the subjects' compliance patterns in state 00 reveal myopic behavior.[21] The percentage figure increases to 50 percent (43 subjects out of 86) if we do not include 36 subjects whose compliance rate was the same for all 20 rounds.[22] A subject who reported full (or 0) income in all 20 rounds will not be classified as myopic by this criterion. We cannot classify these 36 subjects because a very risk averse subject (either myopic or forward-looking) would always fully comply, while a risk lover or risk neutral subject (either myopic or forward-looking) would always report zero income. The number of subjects classified as myopic by both approaches is 18 percent (13 out of 73). These results on myopic versus forward-looking behavior are generally similar to other papers that explore this behavior in areas other than tax compliance.[23]

F. *Additional Treatment Effects on Tax Revenues and Earned Income*

We also investigate the effect of treatments on tax revenues and on earned income by creating two variables, total tax revenue (*TaxRev*) and true earned income (*I*). For any given treatment and round, the value of true earned income is the total amount of income earned in that round by all subjects in the session. Similarly, *TaxRev* is the total amount paid in taxes in a given treatment and round by all subjects in the session. Note that *TaxRev* does not include the amount paid in penalties (tax plus fine times unreported income) in case the subject is audited. Because the Baseline and Higher Tax Rate treatments each had one more subject than the other treatments, we divide the values of these two variables by the number of subjects in the session. There are 20

[21] In the last column of Table 2 we report the percentages of myopic subjects across treatments.
[22] Compliance rates of a naive individual do not depend either on the accumulated unreported income or on the continuation of the game (see II.A.3). For each subject, we created a new variable defined as the difference between the upper and lower bounds of a 95 percent confidence interval of the compliance rate of the subject. Classifying a subject as naïve if this difference is less than 0.1, the estimated percentage of naïve subjects is 23 percent (or 20 out of 86).
[23] For some recent examples, see Hey and Lotito (2009), Hey and Panaccione (2011), Agranov, Caplin, and Tergiman (2015), all of whom use laboratory experiments to examine different types of behavior and all of whom find substantial evidence of myopic players. For example, Hey and Lotito (2009) find that the majority of their subjects are myopic, Hey and Panaccione (2011) find that the majority of subjects impose their first period preferences on their future choices, and Agranov, Caplin, and Tergiman (2015) estimate that 45 percent of the players in their "guessing games" are myopic.

observations of each variable for each of the 8 treatments, giving 160 total observations. Using OLS regression analysis, we estimate the impacts of the tax rate, the fine rate, the audit rate, and donations on each of these dependent variables, along with the demographic variable *Female* in an additional specification for each dependent variable. These results are shown in Table 5.

<Table 5 about here>

Results in Table 5 indicate that the tax rate has a negative effect on true earned income (paid in experimental coins, EC). The estimated tax rate effect is -97EC or -72EC, depending on the specification. (True earned income per subject in the Baseline treatment T1 was 235 EC.) The estimated elasticity is -0.13 or -0.09, revealing an inelastic income response to the tax rate. The effect of the tax rate on tax revenue is positive, with estimates for the marginal effect and the elasticity of 136EC and 0.86, respectively for the model with *Female* dummy variable. Both the fine rate and the audit rate have positive effects on true earned income as well as on tax revenues; the estimated elasticities are 0.25 for the fine rate and 0.35 for the audit rate.[24]

With respect to the Donation treatment (T8), tax revenues per capita are higher when subjects are informed that their taxes are being donated to a charity. This is consistent with a positive demand for giving to charities. Note that round estimates suggest that subjects get better in performing the income-earning task as the experiment proceeds. Also, the larger the fraction of females in a session the smaller is true income per capita but the larger is tax revenue. The negative income estimate comes from females earning on average less (218EC) than males (240EC). Even so, females pay more in taxes (49.86EC versus 47.76EC) because compliance rates for females versus males (0.754 for females versus 0.697 for males) are sufficiently large to compensate for lower income.

[24] We have also estimated an additional specification for tax revenue and true earned income that includes the proportion of subjects audited last round. This variable has no significant effect and other results are not affected, so we do not report this alternative specification. We find no evidence that occurrence of audits result in lower effort: the average earned income is 237EC in the rounds immediately following an audit and 230EC in other rounds.

V. Conclusions

The significance of changes in fiscal parameters in our experiment differs between effects on compliance and effects on basic objectives of tax policy, such as raising revenue with a small decrease in income from work effort. We find that increases in the audit rate have highly significant and positive effects on the compliance rate, while also increasing revenues and earned income. Increases in the fine rate increase revenues and earned income with somewhat variable effects on the compliance rate. As expected, increasing the tax rate has highly significant effects on increasing revenue but decreasing earned income, while having little impact on the compliance rate. Donating tax proceeds to charity has a highly significant positive effect on tax revenue and weakly significant positive (or no) effects on compliance.

Of perhaps more interest, our dynamic model of compliance predicts an idiosyncratic effect of tax return auditing that is supported by the experimental data. This audit state dependent model of taxpayer compliance predicts that income reporting decreases right after an audit if the audit probability and fine rate are constant over time. Indeed, our experimental results show a highly significant negative effect on compliance of having been audited in the last round. This result suggests that, in order to counter such a bounce in underreporting, tax authorities may want to consider informing taxpayers caught evading that on the next such occasion they will face conditionally higher fine rates.

We are also able to use the dynamic model to distinguish between myopic, forward-looking, and naïve behavior in tax compliance. Applying this distinction in analysis of the data, we find that about one-third of our subjects are myopic. Compliance rates for myopic individuals in the most likely scenario (e.g., no audit for adjacent reporting periods) fluctuate between a high and low fraction of income reporting. In contrast, forward-looking behavior requires a nearly constant fraction of income reporting. An implication for tax policy is that nudging taxpayers towards forward-looking behavior, by reminding them how their current compliance choices affect their future exposure to audits, could benefit both myopic taxpayers and the tax administration.

There are certainly reasons for caution in the use of and generalization from these estimates. They are based upon somewhat artificial laboratory experiments; they are derived from subjects whose experience with real-world taxation is uncertain; and they are generated from small samples. Still, it seems likely that the results can contribute to understanding of the compliance puzzle, especially in a country like Colombia where previous estimates of individual behavioral responses simply do not exist. The laboratory experiment provides accurate information on individual compliance choices for Colombian subjects. Further, there is much accumulating evidence that suggests that the "external validity" of tax compliance experiments is in fact significant (Alm, Bloomquist, and McKee, 2015).[25] Most importantly, there is now a large literature that argues convincingly that experimental methods can contribute significantly to policy debates, as long as some conditions are met: the payoffs to subjects are salient; better subject decisions yield higher subject payoffs; decision costs are dominated by payoffs; and the experimental setting captures the essential properties of the naturally occurring environment that is the subject of investigation (Smith, 1982; Plott, 1987). These conditions are met here.

[25] Alm, Bloomquist, and McKee (2015) present several types of evidence on the external validity of tax compliance experiments. First, they examine whether behavior by laboratory participants is replicated by behavior of individuals making a similar decision in the naturally occurring world, by utilizing a special data set assembled by the IRS as part of its National Research Program. Second, they also examine whether students behave differently than non-students in identical laboratory experiments, by using previously reported data from laboratory experiments that compare the decisions of a population of adults with those of undergraduate students, both of whom participate in the identical laboratory experiment. Alm, Bloomquist, and McKee (2015) find that there is an overall similarity between the behavior of individual taxpayers in the field and of student subjects making comparable decisions in the laboratory, so that data from the laboratory closely align with data from the field; in particular, they find that estimated responses of taxpayers in the field and students in the laboratory are quite similar. They also find that student and non-student subjects exhibit broadly similar behavior in the laboratory, even though there are some small differences in their responses to individual policy treatments. These results confirm that compliance behavior in the laboratory generalizes beyond the laboratory. See also Alm, Jackson, and McKee (1992) for earlier comparisons of estimated responses of subjects in laboratory experiments versus those of real-world taxpayers based on naturally occurring field data. See Choo, Fonseca, and Myles (2016) for an alternative view on student versus non-student behaviors.

How can the Government of Colombia use these results in its ongoing efforts to improve compliance? One obvious strategy is consistent with the "Enforcement Paradigm" of Alm and Torgler (2011): Increase audits and fines. However, our results also suggest additional strategies, such as ensuring that individuals can see the uses to which their taxes are put. These latter strategies are consistent with additional and emerging paradigms of tax administration, or the "Trust Paradigm" and the "Service Paradigm" (Alm and Torgler, 2011). Other novel and underutilized strategies, such as providing information designed to nudge individuals away from myopic decision-making, can also be beneficial when tax auditing protocols have dynamic features. In short, the Government of Colombia – just like governments elsewhere – should pursue a range of approaches in its efforts to promote compliance, approaches that are consistent with the "full house" of behaviors that our results demonstrate.

References

Advani, Arun, William Elming, and Jonathan Shaw. "How Long-lasting Are the Effects of Audits?" TARC Discussion Paper 011-15, Tax Administration Research Centre, University of Exeter, Exeter, UK, 2015.

Agranov, Marina, Andrew Caplin, and Chloe Tergiman. "Naive Play and the Process of Choice in Guessing Games." *Journal of the Economic Science Association*, 1 (2), 2015, 146-157.

Allingham, Michael G., and Agnar Sandmo. "Income Tax Evasion: A Theoretical Analysis." *Journal of Public Economics*, 1 (3-4), 1972, 323-338.

Alm, James. "Measuring, Explaining, and Controlling Tax Evasion: Lessons from Theory, Field Studies, and Experiments." *International Tax and Public Finance*, 19 (1), 2012, 54-77.

Alm, James. "What Motivates Tax Compliance?" *Journal of Economic Surveys*, 33 (2), 2019, 353-388.

Alm, James, Kim M. Bloomquist, and Michael McKee. "On the External Validity of Laboratory Tax Compliance Experiments." *Economic Inquiry*, 53 (2), 2015, 1170-1186.

Alm, James, Mark B. Cronshaw, and Michael McKee. "Tax Compliance with Endogenous Audit Selection Rules." *Kyklos*, 46 (1), 1993, 27-45.

Alm, James, Betty R. Jackson, and Michael McKee. "Estimating the Determinants of Taxpayer Compliance with Experimental Data." *National Tax Journal*, 45 (1), 1992, 107-114.

Alm, James, Betty R. Jackson, and Michael McKee. "Fiscal Exchange, Collective Decision Institutions, and Tax Compliance." *Journal of Economic Behavior & Organization*, 22 (4), 1993, 285-303.

Alm, James, and Sarah Jacobson. "Using Laboratory Experiments in Public Economics." *National Tax Journal*, 60 (1), 2007, 129-152.

Alm, James, and Antoine Malézieux. "40 Years of Tax Evasion Games: A Meta-analysis." Department of Economics Working Paper, Tulane University, New Orleans, LA, 2019.

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

Alm, James, and Benno Torgler. "Do Ethics Matter? Tax Compliance and Morality." *Journal of Business Ethics*, 101 (4), 2011, 635-651.

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

Becker, Winfried, Heinz-Jurgen Buchner, and Simon Sleeking. "The Impact of Public Transfer Expenditures on Tax Evasion: An Experimental Approach." *Journal of Public Economics*, 34 (2), 1987, 243-252.

Bird, Richard M. *Taxation and Development: Lessons from Colombian Experience*. Cambridge, MA: Harvard University Press, 1970.

Bird, Richard M., James M. Poterba, and Joel Slemrod, eds. *Fiscal Reform in Colombia – Problems and Prospects*. Cambridge, MA: The MIT Press, 2005.

Choo, C. Y. Lawrence, Miguel A. Fonseca, and Gareth D. Myles. "Do Students Behave Like Real Taxpayers in the Lab? Evidence from a Real Effort Tax Compliance Experiment." *Journal of Economic Behavior & Organization* 124, 2016, 102-114.

Christian, Roberta Calvet, and James Alm. "Sympathy, Empathy, and Tax Compliance." *Journal of Economic Psychology*, 40, 2014, 62-82.

Coursey, Don L., John L. Hovis, and William D. Schulze. "The Disparity Between Willingness to Accept and Willingness to Pay Measures of Value." *The Quarterly Journal of Economics*, 102 (3), 1987, 679-690.

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

Cox, James C., and Seth Epstein. "Preference Reversals Without the Independence Axiom". *The American Economic Review*, 79 (3), 1989, 408-426.

Cox, James C., and David M. Grether. "The Preference Reversal Phenomenon: Response Mode, Markets, and Incentives." *Economic Theory*, 7 (3), 1996, 381-405.

Cox, James C., Vjollca Sadiraj, and Ulrich Schmidt. "Paradoxes and Mechanisms for Choice under Risk." *Experimental Economics*, 18 (2), 2015, 215-250.

Cragg, John G. "Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods." *Econometrica*, 39 (5), 1971, 829-844.

Eckel, Catherine C., and Philip J. Grossman. "Altruism in Anonymous Dictator Games." *Games and Economic Behavior*, 16 (2), 1996, 181-191.

Engel, Christoph, and Peter G. Moffatt. "Dhreg, Xtdhreg, and Bootdhreg: Commands to Implement Double-hurdle Regression." *The Stata Journal* 14 (4), 2014, 778-797**.**

Erard, Brian. "The Influence of Tax Audits on Reporting Behavior," in *Why People Pay Taxes: Tax Compliance and Enforcement*, Joel Slemrod, ed. Ann Arbor, MI: The University of Michigan Press, 1992, 95-114.

Filiz-Ozbay, Emel, and Neslihan Uler. "Demand for Giving to Multiple Charities: An Experimental Study." *Journal of the European Economic Association*, forthcoming.

Fochmann, Martin, and Eike B. Kroll. 2016. "The Effects of Rewards on Tax Compliance Decisions." *Journal of Economic Psychology*, 52, 2016, 38-55.

Gillis, Malcolm, ed. *Fiscal Reform for Colombia: Final Report and Staff Papers of the Colombian Commission on Tax Reform*, Richard A. Musgrave, President. Cambridge, MA: International Tax Program, The Law School of Harvard University, 1971.

Greenberg, Joseph. "Avoiding Tax Avoidance: A (Repeated) Game-Theoretic Approach." *Journal of Economic Theory*, 32 (1), 1984, 1-13.

Hashimzade, Nigar, Gareth D. Myles, and Binh Tran-Nam. "Applications of Behavioural Economics to Tax Evasion." *Journal of Economic Surveys*, 27 (5), 2013, 941-977.

Hey, John D., and Gianna Lotito. "Naive, Resolute, or Sophisticated? A Study of Dynamic Decision Making." *Journal of Risk and Uncertainty*, 38 (1), 2009, 1-25.

Hey, John D., and Luca Panaccione. "Dynamic Decision Making: What do People Do?" *Journal of Risk and Uncertainty*, 42 (2), 2011, 85-123.

Holt, Charles A., and Susan K. Laury. "Risk Aversion and Incentive Effects." *The American Economic Review*, 92 (5), 2002, 1644-1655.

Janssen, Marco A., and T. K. Ahn. "Learning, Signaling, and Social Preferences in Public Good Games." *Ecology and Society*, 11 (2), 2006, Article 21.

Kahneman, Daniel and Amos Tversky. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, 47, 1979, 263-291.

Kastlunger, Barbara, Erich Kirchler, Luigi Mittone, and Julia Pitters. "Sequences of Audits, Tax Compliance, and Taxpaying Strategies." *Journal of Economic Psychology*, 30 (3), 2009, 405-418.

Kirchler, Erich. *The Economic Psychology of Tax Behavior*. Cambridge, UK: Cambridge University Press, 2007.

Landsberger, Michael, and Isaac Meilijson. "Incentive Generating State Dependent Penalty System: The Case of Income Tax Evasion." *Journal of Public Economics*, 19 (3), 1982, 333-352.

Ledyard, John O. "Public Goods: A Survey of Experimental Research," in *The Handbook of Experimental Economics*, John H. Kagel and Alvin E. Roth, eds. Princeton, NJ: Princeton University Press, 1995, 111-194.

List, John A. "Does Market Experience Eliminate Market Anomalies?" *The Quarterly Journal of Economics*, 118 (1), 2003, 41-71.

Maciejovsky, Boris, Erich Kirchler, and Herbert Schwarzenberger. "Misperceptions of Chance and Loss Repair: On the Dynamics of Tax Compliance." *Journal of Economic Psychology*, 28 (6), 2007, 678-691.

Mendoza, Juan P., Jacco L. Wielhouwer, and Erich Kirchler. "The Backfiring Effect of Auditing on Tax Compliance." *Journal of Economic Psychology*, 62, 2017, 284-294.

Mittone, Luigi. “Dynamic Behaviour in Tax Evasion: An Experimental Approach.” *The Journal of Socio-Economics*, 35 (5), 2006, 813-835.

Plott, Charles R. “Dimensions of Parallelism: Some Policy Applications of Experimental Methods,” in *Laboratory Experimentation in Economics: Six Points of View*, Alvin E. Roth, ed. New York, NY: Cambridge University Press, 1987, 193-219.

Rabin, Matthew. “Risk Aversion and Expected-utility Theory: A Calibration Theorem.” *Econometrica*, 68 (5), 2000, 1281-1292.

Rickard, J. A., A. M. Russell, and T. D. Howroyd. “A Tax Evasion Model with Allowance for Retroactive Penalties”. *Economic Record*, 58 (4), 1982, 379-385.

Slemrod, Joel, and Shlomo Yitzhaki. “Tax Avoidance, Evasion, and Administration,” in *Handbook of Public Economics*, Alan J. Auerbach and Martin Feldstein, eds. Amsterdam, London, and New York: Elsevier North-Holland, 2002, 1423-1470.

Smith, Vernon L. “Microeconomic Systems as an Experimental Science.” *The American Economic Review*, 72 (5), 1982, 923-955.

Taylor, Milton C., and Raymond L. Richman. *Fiscal Survey of Colombia: A Report Prepared Under the Direction of the Joint Tax Program of the Organization of American States and the Inter-American Development Bank*. Baltimore, MD: Johns Hopkins Press, 1965.

Vazquez-Caro, Jaime, and Joel Slemrod. “Issues in Colombian Tax Administration,” in *Fiscal Reform in Colombia – Problems and Prospects*, Richard M. Bird, James M. Poterba, and Joel Slemrod, eds. Cambridge, MA: The MIT Press, 2005, 139-152.

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

TABLE 1

Experimental Design

Session	Session Name	Tax Rate	Fine Rate	Audit Rate	Donation?
T1	Baseline	30%	60%	10%	No
T2	Higher Fine Rate	30%	120%	10%	No
T3	Lower Fine Rate	30%	30%	10%	No
T4	Higher Tax Rate	45%	60%	10%	No
T5	Lower Tax Rate	10%	60%	10%	No
T6	Higher Audit Rate	30%	60%	20%	No
T7	Lower Audit Rate	30%	60%	1%	No
T8	Donation of Taxes	30%	60%	10%	Yes

TABLE 2

a. Compliance Rates – Summary Statistics by Treatment for All Subjects

Session	Percent Zero Compliance Obs.	Percent Full Compliance Obs.	Median Compliance Rate	Mean Compliance Rate	Compliance Rate Standard Deviation	Nr. of Subjects	Percent Myopic Subjects (Criter. 1)[a]	Percent Myopic Subjects (Criter. 2)
T1	19	57	0.838	0.732	0.310	16	45	44 (58)
T2	4	63	0.925	0.797	0.251	15	50	33 (38)
T3	14	46	0.807	0.705	0.339	15	10	40 (55)
T4	14	66	0.832	0.775	0.257	16	50	50 (67)
T5	26	62	0.977	0.706	0.392	15	50	20 (43)
T6	12	63	0.859	0.810	0.200	15	40	40 (55)
T7	42	30	0.316	0.406	0.375	15	31	27 (33)
T8	9	75	0.950	0.830	0.242	15	20	27 (50)

Notes: [a] There are 73 (out of 122) subjects whose choices are well-defined to apply this criterion. Figures in parenthesis in the last column do not include 30 subjects who reported full income always and also do not include 6 subjects who always reported 0 income.

b. Compliance Rates – Summary Statistics by Treatment for Partial Compliers

Session	Median Compliance Rate	Mean Compliance Rate	Compliance Rate Standard Deviation	Number of subjects
T1	0.772	0.726	0.250	12
T2	0.876	0.766	0.256	13
T3	0.793	0.689	0.292	11
T4	0.663	0.700	0.256	12
T5	0.650	0.655	0.321	7
T6	0.832	0.741	0.191	11
T7	0.333	0.424	0.341	12
T8	0.734	0.681	0.250	8

Notes: 30 subjects who always reported full income and 6 subjects who always reported 0 income are not included.

TABLE 3

Demographic Characteristics of Subjects by Treatment

	Demographic Characteristic (Proportion of Subjects)					
Session	**Female**	**Student**	**Own House**	**Parent As Household Head**	**Lives in Bogota**	**Born in Bogota**
T1	0.563	0.563	0.813	0.938	0.875	0.813
T2	0.333	0.533	0.600	0.667*	0.867	0.600
T3	0.533	0.800	0.667	0.733	0.600*	0.600
T4	0.625	0.375	0.688	0.625**	0.813	0.750
T5	0.333	0.533	0.667	0.667*	0.733	0.733
T6	0.333	0.533	0.667	0.733	0.933	0.867
T7	0.333	0.667	0.667	0.600**	0.600*	0.800
T8	0.267*	0.667	0.867	0.867	0.733	0.800
Total	0.418	0.582	0.705	0.730	0.770	0.746

Notes: The demographic variable in each session is compared to the baseline (T1). p-values reported by the two-sample test of proportions are used for significance, where *** $p<0.01$, ** $p<0.05$, *$p<0.1$.

TABLE 4

Random-effects Tobit Regression Results for Compliance Rate

Explanatory Variable	Data from All 122 Subjects			Data from 86 Subjects		
	(1)	(2)	(3)	(1)	(2)	(3)
Donation [D]	0.950* (0.494)	0.915** (0.448)	1.115** (0.500)	0.126 (0.255)	0.120 (0.261)	0.125 (0.355)
Tax Rate	0.436 (1.664)	0.383 (1.748)	-0.435 (2.188)	0.685 (1.179)	0.592 (1.242)	0.375 (1.369)
Fine Rate	0.542 (0.446)	0.546 (0.449)	0.398 (0.542)	0.581 (0.363)	0.585* (0.299)	0.568* (0.345)
Audit Rate	8.311*** (2.938)	8.285*** (2.901)	7.785*** (2.841)	4.355** (1.748)	4.437** (1.736)	4.512** (1.764)
Accumulated Unreported Income Rate	0.229*** (0.072)	0.175** (0.083)	0.180** (0.080)	0.220*** (0.074)	0.163** (0.083)	0.168** (0.074)
Income	-0.073 (0.055)	-0.090 (0.060)	-0.089 (0.054)	-0.057 (0.054)	-0.074 (0.055)	-0.075 (0.054)
Round	-0.019*** (0.005)	-0.017*** (0.005)	0.001 (0.020)	-0.020*** (0.005)	-0.017*** (0.005)	-0.001 (0.020)
Audited Last Round [D]		-0.321*** (0.088)	-0.329*** (0.083)		-0.336*** (0.100)	-0.342*** (0.084)
Accumulated Payoff			-0.010 (0.012)			-0.009 (0.011)
Female [D]			0.516* (0.304)			-0.045 (0.171)
Full-time Student [D]			-0.119 (0.316)			-0.124 (0.189)
Own House			-0.058 (0.338)			-0.049 (0.194)
Parent Household Head [D]			-0.425 (0.483)			0.327 (0.304)
Lives in Bogota [D]			0.561 (0.518)			-0.174 (0.273)
Born in Bogota [D]			-0.415 (0.442)			-0.197 (0.277)
Constant	0.185 (0.696)	0.261 (0.721)	0.677 (0.853)	0.117 (0.507)	0.206 (0.521)	0.428 (0.667)
Observations	2,437			1,717		
Number of Subjects	122			86		
Number of clusters	122			86		
Censored observations (left, uncensored, right)	(427, 602, 1408)			(307, 602, 808)		

Notes: The dependent variable is *Compliance Rate* (=Reported Income/Full Income), with a lower bound of 0 and an upper bound of 1. In "Data from 86 Subjects", information from 30 subjects who reported full income in every round and information from 6 subjects who reported 0 income in every round are not included. The unit of measurement for *Income* (earned points in the current round) as well as *Accumulated Payoff* (sum of earnings at the beginning of the current round) is 100 EC; dummy variables are denoted with [D]; *Accumulated Unreported Income Rat*e is the ratio between the total unreported income during the previous two rounds and income in the current round. Bootstrap standard errors (clustered at the individual level) are in brackets; *** p<0.01, ** p<0.05, *p<0.1.

TABLE 5

OLS Estimates of Tax Revenues and True Earned Income

	Tax Revenue (per capita)		True Earned Income (per capita)	
Explanatory Variable	(1)	(2)	(1)	(2)
Donation [D]	9.658***	14.518***	1.155	-4.648
	(0.925)	(0.126)	(2.790)	(4.575)
Tax Rate	156.946***	135.983***	-96.724***	-71.694**
	(5.128)	(0.603)	(9.179)	(21.653)
Fine Rate	14.315***	20.070***	27.246**	20.375
	(0.919)	(0.159)	(11.351)	(11.554)
Audit Rate	165.718***	167.430***	143.345**	141.300***
	(25.533)	(0.635)	(56.305)	(26.785)
Round	-0.058	-0.054	1.211***	1.205***
	(0.039)	(0.042)	(0.152)	(0.158)
Female [D]		26.404***		-31.528
		(0.629)		(28.261)
Constant	-24.173***	-33.454***	214.769***	225.851***
	(3.047)	(0.399)	(10.197)	(11.940)
Observations	160	160	160	160
R-squared	0.894	0.906	0.485	0.499

Notes: Dummy variables are denoted with [D]. Robust standard errors (clustered at the session level) are in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

ONLINE APPENDICES

APPENDIX 1: THEORETICAL ANALYSIS

A. Donation Effects. If tax payments are used to finance a public good P that benefits others, then the utility function is $\varphi(\pi, P)$, which is a concave and increasing function of both own income (π) and the public good level (P). For a separable specification, $\varphi(\pi, P) = f(U(\pi), v(P))$, we have that $\partial V(\varphi(.)) / \partial R > \partial V(U(.)) / \partial R$, and therefore the optimal compliance rate for preferences represented by $\varphi(.)$ is larger than for preferences represented by $U(.)$.

B. A Forward-Looking Decision-maker. We show that the optimal (steady state) compliance rate in state 00 (where the individual is not audited in either of the two previous periods) for the dynamic decision problem in the text:

(1) increases with the fine rate f
(2) increases with the audit rate p
(3) increases with the tax rate τ
(4) increases with income w for CARA and IARA preferences
(5) increases with the probability of continuation of the game (or patience) δ; and
(6) increases with the accumulated unreported income rate z in the neighborhood of the steady state.

Also, the compliance rate after an audit is smaller than the steady state compliance rate in state 00 (absent audits).

Proof. To simplify notation, write the accumulated unreported income x as $x = zw$ and $\alpha = w\tau(1+f)$. We derive the Euler equation, the steady state compliance rates, and the first order Taylor approximation for the optimal control rule and use this information to predict signs of the estimates in the empirical analysis.

The decision problem for a forward-looking individual is

$$V(-\alpha z_0) = \max_{R_t \in [0,1]} E(\sum_{t=0}^{\infty} \delta^t [U(\pi_s(R_t, Z_t))] \quad \text{`}$$
$$Z_t = X_{t-1}(1 - R_{t-1} + X_{t-2}(1 - R_{t-2})); X_t \sim Bernoulli(1-p) \tag{A1.1}$$

where $\pi_s(R, z) = w[1 - \tau R - s\tau(1+f)(1-R+z)], s \in \{0,1\}$. Since $\pi_0(R,z) = \pi_0(R,0)$ for all accumulated unreported income zw, for simplicity we abuse the notation and write $\pi(R,z)$ when referring to $\pi_1(R,z)$ and $\pi(R)$ when referring to $\pi_0(R,z)$. The dynamic programming (Bellman) equation of problem (A1.1) in state 00 is

$$V(-\alpha z_t) = \max_{R_t \in [0,1]} \left[EU(\pi_s(R_t, z_t)) + \delta\left(pV(0) + (1-p)V(-\alpha z_{t+1}) \right) \right], \quad \forall t$$
$$z_{t+1} = 2 - R_{t-1} - R_t \tag{A1.2}$$

Differentiate the expression within the square brackets with respect to R_t and divide by $w\tau$ to get

$$G(\theta) = [pfU'(\pi(R_t, z_t)) - (1-p)U'(\pi(R_t))] + \delta(1-p)(1+f)V'(-\alpha z_{t+1}) \tag{A1.3}$$

where $\theta = (z_t, p, f, \tau, w, \delta)$. The (interior) optimal $R_t^d = h(z_t)$ is implicitly determined by $G(\theta) = 0$. Differentiate (A1.2) with respect to R_{t-1} and divide by α to get

$$V'(-\alpha z_t) = pU'(\pi(R_t^d, z_t)) + \delta(1-p)V'(-\alpha z_{t+1}) \tag{A1.4}$$

Write (A1.3) at time t and t+1 and (A1.4) at time $t+1$ absent an audit to get

$$\begin{aligned}
&[pfU'(\pi_{1t}) - (1-p)U'(\pi_{0t})] = -\delta(1-p)(1+f)V'(-\alpha z_{t+1}) \\
&[pfU'(\pi_{1t+1}) - (1-p)U'(\pi_{0t+1})] = -\delta(1-p)(1+f)V'(-\alpha z_{t+2}) \\
&V'(-\alpha z_{t+1}) = pU'(\pi_{1t+1}) + (1-p)\delta V'(-\alpha z_{t+2})
\end{aligned} \tag{A1.5}$$

where $\pi_{1(t+1)} = \pi_1(R_{t+1}^d(z_{t+1}), z_{t+1})$, $\pi_{1t}(z_t) = \pi_1(R_t^d(z_t), z_t)$. Use the first and the second equations to replace $V'(.)$ in the third equation of (A1.5) to get the system of difference equations of optimal compliance rates

$$[pfU'(\pi_{1t}) - (1-p)U'(\pi_{0t})] = -\delta(1-p)[pU'(\pi_{1t+1}) + (1-p)U'(\pi_{0t+1})], \quad \forall t \tag{A1.6}$$

In a steady state (z^o, R^o), the accumulated income and compliance rate do not change.

The marginal effects of parameters of interest on optimal compliance rates $R_t^d = h(z_t)$ are derived using first order Taylor approximation at a (stable) steady state, (z^o, R^o)

$$h(z_t) \approx h(z^o) + h'(z^o)(z_t - z^o) \tag{A1.7}$$

B1. Comparative Statics of Steady State Compliance in State 00

By the equation of motion (see A1.2), $z^o = 2(1 - R^o)$, which together with (A1.6) imply that

$$U'(\pi(R^o, 2 - 2R^o)) + \gamma U'(\pi(R^o)) = 0 \tag{A1.8}$$

where $\gamma = \left(1 - \frac{1}{p}\right)\frac{1 - \delta(1-p)}{f + \delta(1-p)} < 0$. Let *H* denote the expression on the left hand side, or

$$H = U'(\pi(R, 2 - 2R)) + \gamma U'(\pi(R)) \tag{A1.9}$$

Note that there can be no more than one steady state because *H*(.) is monotonic in R, or

$$H_R = w\tau\left[(2+3f)U''(\pi_{1o}) - \gamma U''(\pi_o)\right] < 0$$

The steady (00) state compliance rate R^o:

(1) increases with the fine rate f as $\partial R^o / \partial f = -H_f / H_R$, where

$$\frac{\partial}{\partial f}H(.) = \left[-\tau w 3(1-R)U''(\pi_{1o}) + (\frac{1}{p} - 1)\frac{(1 - \delta + \delta p)}{(f + \delta - \delta p)^2}U'(\pi_o)\right] > 0$$

(2) increases with the audit rate p as $\partial R^o / \partial p = -H_p / H_R$ where

$$\frac{\partial}{\partial p}H(.) = \frac{f[1 - \delta(1 - p^2)] + (1-p)^2(1-\delta)\delta}{p^2(f + (1-p)\delta)^2}U'(\pi_{0o}) > 0$$

(3) increases with the tax rate τ as $\partial R^o / \partial \tau = -H_\tau / H_R$ where

$$\begin{aligned}\frac{1}{w}\frac{\partial}{\partial \tau}H(.) &= \left[-\big(3(1+f) - (2+3f)R\big)U''(\pi_{1o}) - RU''(\pi_{1o})\right] \\ &= \left[-\big(1 + (2+3f)(1-R)\big)U''(\pi_{1o}) - RU''(\pi_{1o})\right] > 0\end{aligned}$$

(4) increases with income for IARA/CARA preferences as $\partial R^o / \partial w = -H_w / H_R$ and

$$w\frac{\partial}{\partial w}H(.)=\left[\pi_{1o}U''(\pi_{1o})+\gamma\pi_{0o}U''(\pi_{0o})\right]=U'(\pi_{1o})\left[-\pi_{1o}ARA(\pi_{1o})+\pi_{0o}ARA(\pi_{0o})\right]$$
$$>U'(\pi_{1o})ARA(\pi_{1o})\left[\pi_{0o}-\pi_{1o}\right]>0$$

where the second equality follows from (A1.9).

(5) increases with the probability of continuation δ as $\partial R^o/\partial\delta=-H_\delta/H_R$ where

$$\frac{\partial}{\partial\delta}H(.)=-(1-\frac{1}{p})\frac{(f+1)(1-p)}{(f+\delta(1-p))^2}U'(\pi_{0o})>0$$

(6) increases with the accumulated unreported income rate z in the neighborhood of the steady state. Letting b denote $h'(z^o)$, the first order approximation of the optimal compliance function (A1.7) is $h(z)\approx R^o+b(z-z^o)$. The compliance rate increases in z if $b>0$. To find b, replace $R_t^d=h(z_t)$ in the Euler equation (A1.6), differentiate with respect to z_t, and evaluate it at the steady state to show that b solves $a_2b^2+a_1b+a_0=0$ (*), where

$$\begin{aligned}
&a_0=\left[-pf(1+f)U_1^{''}\right]>0\\
&a_1=U_1^{''}\left[pf^2+\delta p(1-p)(2f+1)\right]+(1-p)(1-\delta(1-p))U_0^{''}<0\\
&a_2=-\delta(1-p)\left[pfU_1^{''}-(1-p)U_0^{''}\right]\\
&\quad U_1^{''}=U''(\pi(R^o,2-R^o));\quad U_0^{''}=U''(\pi(R^o))
\end{aligned}$$

For $pfU_1^{''}-(1-p)U_0^{''}<0$, then the quadratic equation (*) can only have positive solutions. For $pfU_1^{''}-(1-p)U_0^{''}>0$, it has a positive and a negative solution. Stability conditions for (*) (where R_t^d converges to R^o) are that $b<1$, $|b|<1+b$, and $|-b\pm\sqrt{b^2-4b}|<2$. The negative solution of (*) does not satisfy all these conditions while the positive one does. As a result, it must be that $h'(z^o)=b\in(0,1)$, so that the optimal compliance rate R_t^d increases in accumulated unreported income rate z_t.

B2.The compliance rate after an audit is smaller than the steady state compliance R^o

in state 00 (where the individual is not audited in either of the two previous periods).

Suppose by contradiction that the individual reports wR^o after an audit and in state 10. Then the individual's accumulated unreported income is 0, and if not audited again, it is $1-R^o$. Therefore the expression in the Euler equation (A1.6) is

$$\begin{aligned} G &= \begin{bmatrix} pfU'(\pi(R^o,0) \\ -(1-p)U'(R^o) \end{bmatrix} + \delta(1-p)\begin{bmatrix} pU'(R^o,1-R^o) \\ +(1-p)U'(R^o) \end{bmatrix} \\ &< \left[f+\delta(1-p)\right]pU'(R^o,2-2R^0) - \left[1-\delta(1-p)\right](1-p)U'(R^o) = 0 \end{aligned}$$

where the inequality follows from concavity of *U(.)* and the last equality follows from (A1.8). It follows from $G<0$ that R^o is too large to be optimal after an audit.

C. A Myopic Decision-maker. A myopic decision-maker when making a decision on compliance rate at time *t* takes into account the accumulated unreported income but ignores the future effects of current compliance rate. Such a decision-maker is a special case of a forward-looking decision-maker with $\delta = 0$. We show that the optimal compliance rate for myopic decision-makers:

(1) increases with the accumulated unreported income, and is smallest in state 1 (after an audit);
(2) increases with the fine rate;
(3) increases with the audit rate;
(4) increases with the tax rate for CARA and DARA preferences;
(5) increases with income for CARA and IARA preferences; and
(6) does not depend on the probability of continuation of the game.

Proof. At any time *t*, the decision problem of our individual is

$$V(\alpha z_t) = \max_{R\in[0,1]}\left[(1-p)U(\pi_0(R,z_t)) + pU(\pi_1(R,z_t))\right] \tag{A1.10}$$

when the accumulated unreported income is wz_t. Differentiating the expression within the square brackets with respect to *R*, we get

$$\frac{1}{\tau w}F(R,\theta_{t0}) = pfU'(\pi_1(R,z_t)) - (1-p)U'(\pi_o(R,z_t))$$

The interior optimal compliance rate R^* solves[26]

$$F(R,\theta_\delta) = pfU'(\pi_1(R,z_t)) - (1-p)U'(\pi_o(R,z_t)) = 0 \qquad \text{(A1.11)}$$

Note that (as expected) $F(.) = G(.)$ stated in (A1.3).

It is straightforward to verify that optimal compliance rate R_t^m:

(1) increases with accumulated unreported income as $sign(R_z^*) = sign(F_z)$, and

$F_z(.) = -\alpha U''(\pi_{1t}) > 0$. In state 1, the accumulated unreported income, *z* takes the smallest possible value (0), and by positive monotonicity of $R^*(z)$, the chosen fraction $R^*(0)$ takes its smallest value.

(2) increases with *f* as $sign(R_f^*) = sign(F_f)$, and $F_f(.) = pU'(\pi_{0t}) - (z+1-R)w\tau pfU''(\pi_{1t}) > 0$.

(3) increases with the audit rate p as $sign(R_p^*) = sign(F_p)$, where $F_p(.) = fU'(\pi_{1t}) + pU'(\pi_{0t}) > 0$.

(4) increases with the tax rate τ for CARA and DARA preferences. Verify that

$sign(R_\tau^*) = sign(F_\tau)$, and $F(R,\theta_\delta) = pfU'(\pi_1(R,z_t)) - (1-p)U'(\pi_o(R,z_t)) = 0$

$$\frac{1}{w}\frac{\partial F(.)}{\partial \tau} = \left[\frac{1-p}{pf}RU''(\pi_o) - \big((1+f)z + f(1-R)+1\big)U''(\pi_1)\right]$$

It follows from (A1.11) and the Arrow-Pratt measure of absolute risk aversion that

$$sign(F_\tau) = sign\left[-A(\pi_{0t})R_t^* + \left((1+f)(z+1) - fR_t^*\right)A(\pi_{1t})\right]$$

For CARA preferences or DARA preferences, $A(\pi_{0t}) \leq A(\pi_{1t})$ the expression within the square brackets is positive.

(5) increases with income *w* for CARA and IARA preferences. Verify that

$$w\frac{\partial F(.)}{\partial w} = pf\pi_{1t}U''(\pi_{1t}) - (1-p)\pi_{0t}U''(\pi_{0t})$$

[26] The optimal compliance rate, R^* is at the corners, 0 and 1 if $F(0,.) < 0$ and $F(1,.) > 0$, respectively.

Use similar steps as in the previous to show that $sign(F_w) = sign\left[-\pi_1 A(\pi_{1t}) + \pi_0 A(\pi_{ot})\right]$. For CARA and IARA preferences, $A(\pi_o) \geq A(\pi_1)$, and the expression within the square brackets is positive.

(6) does not depend on the probability of continuation of the game. This follows from the optimality condition (A1.11) not depending on δ

D. A Naive Decision-maker. A naïve decision-maker isolates and does not take into account either the accumulated unreported income or the future effects of current decision. The decision problem for such a decision-maker is

$$\max_{R\in[0,1]}\left[(1-p)U(\pi_0(R,0)) + pU(\pi_1(R,0)\right]$$

The optimal compliance rate R increases in ($f, p, \tau_{CARA/DARA}, w_{CARA/IARA}$) but does not vary with previously unreported income, does not depend upon whether an audit took place, and does not depend upon with the time of the game.

APPENDIX 2: DOUBLE HURDLE PANEL ESTIMATES

Observations with full compliance account for 58 percent of our data, and full compliance remains the mode choice even in state 1 (following audit). Estimation of treatment effects requires that we address this aspect of our data. Out of 122 subjects, 30 subjects reported full income in every round. One way to think about this is to categorize subjects as "truthful" types (e.g., subjects are averse to misreporting, they think of paying taxes as a moral duty) or not; this is the first hurdle. In the main text we dealt with this problem by getting Tobit estimates with and without these subjects. Here we report double hurdle models that allow for estimating "truthful" types of subjects, while at the same time subjects who cleared the "truth" hurdle are allowed to report true income at any time if they wish to do so (e.g., when previously unreported income becomes large). The first hurdle is a probit equation, and the Tobit equation is estimated only with subjects who clear the first hurdle. The double hurdle model allows for differences in the variables that affect the different decisions of whether to report all or some of true earned income versus the decision on how much to report conditional upon clearing the first hurdle (e.g., deciding to misreport). The estimation approach we use here is a double-hurdle panel model (Cragg, 1971; Engel and Moffatt, 2014).

More precisely, let y_i denote the observed compliance rate of individual i (defined as Reported Income/True Earned Income) and y_i^* denote the continuous latent compliance rate that is observed only if the subject has elected to misreport income. The model specification is:

$$y_i = D_i y_{it}^*$$
where
$$D_i = \begin{cases} 1 & if \quad \beta Z_i + \delta_i < 1 \\ 0 & if \quad \beta Z_i + \delta_i \geq 1 \end{cases}; \quad \delta_i \sim N(0,1)$$

$$y_{it}^* = \begin{cases} \alpha X_{it} + u_i + \varepsilon_{it}, & if \quad \alpha X_{it} + u_i + \varepsilon_{it} < 1 \\ 1, & if \quad \alpha X_{it} + u_i + \varepsilon_{it} \geq 1 \end{cases}; \quad \varepsilon_{it} \sim N(0,\sigma^2),\ u_i \sim N(0,\sigma_u^2)$$

where Z_i and X_{it} are explanatory variables for the first and the second equations. The first equation (for D_i) does not depend on task t because a "truthful" type would never report less than full income under no circumstances. As a result, in the list of the regressors for the first stage probit equation we include only variables that are fixed across rounds, such as the tax rate, the audit rate, the fine rate, and gender. Subjects who report true income in round will fail the hurdle. The second equation (y_{it}) is a random effects Tobit regression, which allows for choices to depend on the task that differs across rounds because of the variation on unreported income from previous rounds. Therefore, we add time-varying variables (such as unreported income rate, true earned income, being audited last) in the list of regressors.

Appendix Table A.2 reports average marginal effects for the double hurdle panel model, estimated using all individual data by round and by treatment. The columns labeled *Hurdle* present the marginal effect on the decision to report *True Earned Income*; the columns labeled *Linear* report the results for the subject's *Compliance Rate* conditional upon the subject misreporting income. The impacts of the policy parameters are again largely consistent with the theory and with the Tobit estimation results of Table 4 in the main text. For the basic model specification (1), knowing that the tax payment goes to charity decreases the probability of clearing the hurdle by 23 percent while a one-unit change in the tax rate increases it by almost 87 percent. The signs are consistent with Tobit estimates in the first column of Table 4 (which can be seen as the first hurdle that applies to all subjects). Focusing on the linear estimates in Appendix Table A.2, or the estimates of individual behavior conditional upon clearing the "hurdle" of choosing some noncompliance, compliance increases with an increase in the audit rate, an increase in the fine rate, and with an increase in accumulated unreported income; compliance decreases with income. The use of the tax payments (or *Donation*) and the *Tax Rate* are not statistically significant determinants of the individual compliance rate, once the hurdle has been passed. *Female* decreases the probability of clearing the hurdle by 19 percent, but has no effect on compliance rate conditional on passing. These estimates are comparable to the Tobit estimates for selected subjects (the 30 subjects not included there would have failed the hurdle).

APPENDIX TABLE A.2

Panel Double Hurdle Model: Average Marginal Effects for Compliance Rate

Explanatory Variable	Specification 1		Specification 2		Specification 3	
	Hurdle	**Linear**	**Hurdle**	**Linear**	**Hurdle**	**Linear**
Donation [D]	-0.227*** (0.086)	-0.055 (0.135)	-0.223** (0.089)	-0.068 (0.133)	-0.259*** (0.092)	-0.016 (0.138)
Tax Rate	0.869* (0.498)	1.233 (0.834)	0.822* (0.488)	0.961 (0.610)	0.835* (0.445)	0.926 (0.732)
Fine Rate	0.216 (0.259)	0.527*** (0.185)	0.205 (0.249)	0.502*** (0.174)	0.110 (0.216)	0.442** (0.200)
Audit Rate	-0.678 (0.920)	3.500*** (0.885)	-0.714 (0.927)	3.651*** (0.841)	-0.903 (0.943)	3.333** (1.408)
Accumulated Unreport. Income Rate		0.097*** (0.022)		0.088*** (0.024)		-0.016 (0.138)
Income		-0.080** (0.032)		-0.055* (0.032)		0.926 (0.732)
Audited Last Round [D]				-0.209*** (0.045)		-0.264*** (0.042)
Accumulated Payoff				-0.004*** (0.001)		-0.004*** (0.001)
Female [D]					-0.190** (0.077)	0.016 (0.130)
Observations	2437		2437		2437	
Number of Subjects	122		122		122	

Notes: The dependent variable is *Compliance Rate* (=Reported Income/Full Income), with a lower bound of 0 and an upper bound of 1. The first hurdle is full compliance. The unit of measurement for *Income* (earned points in the current round) as well as *Accumulated Payoff* (sum of earnings at the beginning of the current round) is 100 EC; dummy variables are denoted with [D]; *Accumulated Unreported Income Rat*e is the ratio between the total unreported income during the previous two rounds and income in the current round. Standard errors are in brackets; *** p<0.01, ** p<0.05, *p<0.1.

APPENDIX 3: SUBJECT INSTRUCTIONS FOR TREATMENT 1

Welcome and thank you for participating in today's experiment.

Introduction
This is an experiment in the economics of decision-making. Your earnings will be determined by your answers and decisions and the chance of audit*, as described in the following instructions.

* In this context, "audit" means that the system can choose you in order to verify if you are behaving correctly and telling the truth.

IT IS IMPORTANT THAT YOU READ THESE INTRUCTIONS CAREFULLY.

This experiment is structured in a way that only you can know your earnings. All of the money that you earn will be paid to you privately and in cash immediately at the end of today's experiment.

> *If you have any questions, RAISE YOUR HAND and an experimenter will approach you and answer your questions in private. Please feel free to ask as many questions as you like.*

A. Time
This experiment will last no more than two hours.

B. Show up fee
You will be paid an amount of 25,000 pesos (USD 13) for showing up for today's experiment. This is in addition to what you will earn from the decisions you make during today's experiment. You will not receive any payment if you decide to leave the experiment before it finishes.

C. Earning money during the experiment
You earn money in Experimental Coins (EC) in each decision period. This amount will be displayed on your computer screen at the completion of each decision period. At the end of today's experiment, your total accumulated earnings in Experimental Coins will be converted into Colombian Pesos at the below mentioned conversion rate.

Conversion Rate: 1 Experimental Coin (EC) = 10 Pesos

How to earn money in the experiment?

D. Task and decision-making process
Beginning
Each subject will be given an endowment of 500 EC at the beginning of the experiment.
In this experiment you will have to pass through an indefinite number of decision-making periods. Each of the periods has the following sequence of events.

Event 1

During each period, each subject is given a work task of adding numbers together in order to earn income. You will be given 90 seconds to conduct the task. You will earn 30 Experimental Coins for each correct answer to an addition question. This income you earn will be displayed on your screen at the completion of the task.

Event 2
Your Earned Income is what you earn in Event 1. You will have to make the choice of how much of this Earned Income to report for tax purposes using the sliding scale on your computer screen. There is an income tax at rate 30% that you need to pay on the income you report. As you move the slide to determine how much income you will report, you can see the consequences of your choice in terms of your net income if you are audited or not.

You can choose to report all of your income, part of your income, or none of the income earned from the work task in this experiment. Since the income tax rate is 30%, the amount of tax you will owe is equal to:

30% X Reported Income

Event 3
Once you choose the amount of your Earned Income to report, a random audit will be performed. The probability that you will be randomly selected for audit is 10% in each and every period, which means that you would be selected for audit about 10 times in 100 periods. If you are chosen for the random audit in some period, <u>your earned income in that period and in the preceding two (2) periods</u> will be disclosed for audit. If the audited individual's reported income is <u>less than</u> the earned income in the current period or any of the preceding two (2) periods, then the individual pays, <u>in addition to the tax of</u> 30% of the earned income, a fine of 60% of the unpaid tax in the current and two (2) preceding periods.

You pay a tax fine only if you are audited and if your reported income is less than the earned income in the current period or any of the 2 preceding periods.

Section E below shows your total payoff in each decision period resulting from Events 1 to 3 explained above.

E. Earnings in each decision period

<u>Scenario I: If you are **no**t audited</u>
Total earnings equals Earned Income minus Tax Payment on Reported Income.

<u>Scenario II: If you **are** audited</u>
Total earnings equals Earned Income minus Tax Payment on Earned Income minus Fine paid on unreported income in the current period and the preceding 2 periods.

(Note: The fine is equal to zero if your Reported Income is equal to your Earned Income in every period.)

F. Practice Periods and Real Periods
Practice Periods: There will be 3 Practice Periods in the experiment. In the Practice Periods, the audit results are not randomly selected. You will be audited in the last Practice

Period but not in any of the others. No money can be made or lost in the practice periods; they are only for practice.

Real Periods: There will be a predetermined number of Real Periods in the experiment. Real money, paid in cash at the end of the experiment, can be made and lost in the Real Periods. The audit probability is 10% in every Real Period.