THE EFFECT OF PERFECT MONITORING OF MATCHED INCOME ON SALES TAX COMPLIANCE: AN EXPERIMENTAL INVESTIGATION

Cathleen Johnson, David Masclet, and Claude Montmarquette

Tax noncompliance is a quantitatively important phenomenon that significantly affects government revenues, and thus raises challenging questions about the determinants of tax reporting and the appropriate design of a tax system. This paper provides empirical insights regarding the nature of tax noncompliance, using an experimental approach to evaluate the effects of systematic sales tax monitoring and identify the determinants of sales tax compliance. The results suggest that if perfect monitoring of a single revenue source is introduced without other complementary policies, an increase in tax revenues is not the likely outcome as evasion increases for other revenue sources. That is, the data suggest that once taxpayers have chosen their level of tax compliance, they will try to recover their losses following any policy changes, even if it implies assuming more risks.

Keywords: tax, perfect monitoring, experimental economics, reference dependent effect

JEL Codes: C9, H9, H71

I. INTRODUCTION

Tax noncompliance is a quantitatively important phenomenon that significantly affects revenue sources for governments. This phenomenon raises challenging questions about the determinants of tax reporting, the appropriate design of a tax system, and the magnitude of resources that should be devoted to auditing.

Numerous theoretical and empirical studies have examined the impact of audits on income tax compliance (Jackson and Milliron, 1986; Roth and Scholtz, 1989; Slemrod, 1992; Andreoni, Erard, and Feinstein 1998). Using random surveys (Fisher, Goddeers,
and Young 1989) and available tax databases (Clotfelter, 1983; Dubin, Graetz, and Wilde 1990; Erard and Ho, 2001), researchers have identified the characteristics of noncompliant taxpayers and the factors that are likely to motivate tax compliance. Clotfelter (1983) provided an empirical analysis of taxpayer compliance with information from the Taxpayer Compliance Measurement Program (TCMP) of the Internal Revenue Service (IRS) in the United States. He concluded that noncompliance is strongly positively related to the marginal tax rate. Dubin, Graetz, and Wilde (1990) investigated the impact of audit rates and tax rates on tax compliance with state-level time series data from 1977–1985. The authors observed that the continual decline in the audit rate over that period led to a significant decrease in IRS collections. Many experimental economic studies have examined the effects of audit rates and penalties on fraud and tax evasion, and the nature of tax compliance when income is either earned or endowed (Friedland, Shlamo, 1978; Webley et al., 1991; Alm, Jackson, and McKee, 1992; Boylan and Sprinkle, 2001; Gërxhani and Schram, 2006; Cadsby, Maynes, and Trivedi 2006; Alm and McKee, 2006; and Alm, Deskins, and McKee 2009). Most studies on tax compliance have examined the personal income tax. Despite the importance of the sales tax in state and local government budgets, surprisingly little academic research has focused on the subject of sales tax compliance, with some exceptions (Mikesell, 1985; Murray, 1995; Alm, Blackwell, and McKee, 2004). Both the magnitude and determinants of sales tax noncompliance remain elusive. Murray (1995) has shown that taxpayers with greater opportunities to reduce their tax liabilities exploit these opportunities to their advantage. He also concluded that there is no obvious or easy-to-implement policy to combat sales tax noncompliance.

The purpose of this paper is to provide empirical insights on sales tax compliance, using an experimental approach to evaluate the effects of systematic sales tax monitoring and the determinants of sales tax compliance. The experimental approach makes it possible to measure exactly the rate of tax compliance. We investigate the extent to which taxpayers would alter their compliance behavior in response to a change in the audit environment. In particular, we study whether perfect monitoring of electronic retail sales transactions might improve tax compliance and raise the level of tax revenues. Perfect monitoring of sales is analogous to increasing the income tax audit rate on matched income to 100 percent. It is technically possible to match individual declarations of income to relevant third-party information. In the United States, this type of matching is reserved for audits and taxpayer compliance studies (Internal Revenue Service, 1996). The Canadian federal government matches random individual tax returns with third-party information on earnings as part of an ongoing monitoring system of tax compliance.\(^1\)\(^2\) The primary objective of instituting direct and automatic capture of

---

2 At present, financial institutions credit electronic purchases (including the levied taxes) to the merchant within 24 hours following the purchase. The merchant acts as an agent for the government and returns the levied taxes according to the frequency of his remittance agreement (either annually, quarterly or monthly). It is now possible for sales tax to be directly and immediately captured from every electronic transaction. Just as the value of the purchase can be credited to the merchant’s account, the corresponding taxes can be credited simultaneously to the taxing authority’s account.
the tax portion of sales paid electronically is to increase revenues by reducing tax evasion. Third-party reporting on matched income severely limits an individual’s ability to evade taxes. Higher detection probabilities reduce the marginal benefit of evasion, and therefore make evasion less attractive. For example, in the United States, small businesses and farms, which have less matched income than large businesses, have significantly higher rates of evasion (Internal Revenue Service, 1996). In Canada, the household repair work, renovation and construction sectors are particularly affected by underreporting of sales receipts (Fortin et al., 1996). The Canada Revenue Agency justifies higher audit rates on sole proprietorships by the fact that wage and salary earners present relatively few compliance problems. Salaried employees’ taxes are collected through payroll deductions and their contributions are readily verified by reference to information filed by their employers.3

However, several studies have shown that an increase in audit probability does not necessarily lead to improved tax compliance. Whether increased audits and penalties are the best way to deal with noncompliance depends on the reasons that taxpayers fail to comply. If taxpayers are “playing the audit lottery,” increasing penalties and audit rates should improve compliance. However, if their objective is to maintain a certain level of income, increased audits might not necessarily induce higher compliance. For example, Slemrod, Blumental, and Christian (2001) show that an increase in auditing does not necessarily mean an increase in voluntary compliance.4 Even if perfect monitoring increases tax revenues paid by some taxpayers since there is less room for cheating, this effect may be offset by reduced revenues from other taxpayers. For example, knowing that all non-cash transactions will be automatically reported, merchants may tend to underreport more on unmonitored income (e.g., by reporting fewer cash transactions) in order to maintain a net expected income comparable to what they had before the introduction of perfect monitoring. Alternatively, merchants can offer discounts for cash payment for goods and services rather than accept traceable on-cash payments. As a result, increasing the audit rate may contribute to the underground economy.5 In addition to examining the extent to which increasing the audit rate affects compliance, our study also illustrates the negative effect on revenues of announcing such a change in the tax system. Indeed, we anticipate that taxpayers may try to offset the eventual consequences of the future introduction of a more closely monitored tax system by taking greater advantage of their current, liberal environment.

4 To test the impact on compliance behavior of increased awareness of the likelihood of an audit, the Minnesota Department of Revenue carried out a controlled field study, described in Slemrod, Blumental, and Christian (2001). A stratified sample was selected based on three income levels and split into a treatment and a control group. The treatment group was informed by mail that their tax returns would be “closely examined.” The comparison between the treatment and control groups showed that the threat of examination increased reporting compliance among low- and middle-income taxpayers, but had the opposite effect among high-income taxpayers.
5 Several studies have shown that taxes are undeniably an important factor in the underground economy. There also appears to be a rare degree of unanimity on the empirical proposition that the underground economy has grown substantially as a percentage of GDP since early 1991. One key piece of evidence for this is the large increase in cash in circulation relative to reported incomes (Lippert and Walker, 1997).
Several treatments were conducted in our experiment in order to isolate these different effects. In the baseline treatment, subjects received income from two sources. To the participants, these were represented as Source A (i.e., resulting from electronic transactions) and Source B (resulting from cash transactions). Subjects received income in each period and were asked to voluntarily report their income. Participants paid tax on the reported income. They were subject to an audit with some probability. The second treatment, referred to as the announcement treatment, was identical to the first, except that after a specified period of play, an announcement was posted that a change in policy would take effect in the next treatment. Subjects were told that a change in policy would institute perfect monitoring of Source A income. The amount of Source B income remained private. Participants were informed of the change in policy before the policy was instituted to ascertain whether behavior changed significantly due to the expectation of imminent monitoring. The last two treatments examine the effects of the announced policy of perfect monitoring, conditional on whether there is an option to alter the source of income at a cost. The third treatment allows perfect monitoring of Source A income without the option to alter the sources of income, while the fourth treatment allows perfect monitoring of Source A income with the opportunity to alter the sources of income, that is, it allows income earners to pay a premium to move income from Source A to Source B.

Our work is related to several previous laboratory experiments that examined compliance decisions. Alm, Jackson, and McKee (1992) used laboratory experiments to estimate individual responses to audit probability, penalties and taxes. They showed that reporting rates increase with audit and penalty rates, but decline with taxes. Alm and McKee (2006) also investigated how individual tax reporting decisions may be affected by audit probabilities and also by audit announcements. The authors found that audit announcements increase the compliance rates of those who are told that they will be audited, while the compliance rate declines for those who are told that they will not be audited. In this study, we evaluate the potentially negative impact induced by the announcement of a policy change when such policy is not immediately implemented. Most experimental studies have utilized a single category of income, with the exceptions of Gërxhani and Schram (2006) and Alm, Deskins, and McKee (2009). In Gërxhani’s and Schram’s paper, the participants choose between income that is automatically audited and unregistered income that is subject to different probabilities of being audited. They found that subjects choose unregistered income more frequently. In our paper, the two sources of income were assigned exogenously, except for one treatment. In that treatment, a participant could transfer the monitored source of income to the unmonitored category (but at a cost), which was not the case in Gërxhani and Schram. The paper by Alm, Deskins, and McKee is more similar to ours in its focus on the design of tax collection institutions. They use an experimental protocol to determine the impact of

6 Gërxhani and Schram (2006) set the problem of tax evasion in the context of contributing to a public good, which is not considered in our experiment. They do not address directly the differential in tax compliance when the monitoring of income increases.

7 We thank a referee for this reference.
audit and tax rates on subjects whose incomes are not reported by their employers. Their results suggest that evasion is, in part, contingent on the source of income, with the amount of evasion being higher for individuals with non-matched income.

Our study focuses on sales taxes, but is equally applicable to labor or capital income taxes with different sources of income. However, sales taxes offer interesting insights on evasion of tax on sales that are difficult to trace, such as cash transactions. Merchants generally accept two forms of payment on sales transactions: conventional electronic payments and cash. While electronic payments will generally be subject to sales taxes, cash payments may go unreported and escape taxation. We investigate the extent to which taxpayers reduce the impact of monitoring by transferring sales from electronic to cash transactions, even if such recharacterization is costly, e.g., because merchants are expected to accept lower cash amounts for unrecorded sales (Gordon, 1990).

Our study seeks to examine the negative effects of announcing a change in auditing policy. In particular, assuming that a policy cannot be implemented immediately, we investigate the extent to which such an announcement may provoke individuals to report less income in the current period in order to counteract the future effects of this policy.

We find that taxpayers’ noncompliance is related to opportunities for cheating. Subjects report less of Source B income when monitored perfectly and when presented with the opportunity to get rid of their observed income. Subjects more than compensate for the increase in taxes on perfectly monitored Source A income by paying less tax on Source B income.

In Section II, we discuss the research objectives and the institutional setting for our analysis. In Section III, we describe the experimental design and protocol. Section IV presents theoretical predictions and behavioral conjectures about the expected treatment effects. In Section V, the experimental results are presented and discussed. The last section concludes and presents our policy recommendations.

II. RESEARCH OBJECTIVES AND INSTITUTIONAL SETTING

Using the Canadian tax system as reference point, the objective of the research is to examine which factors influence tax compliance by merchants acting as a collection agent for the government, with respect to the federal goods and services tax (GST) and provincial sales tax (PST). The basic goal is to establish whether perfect monitoring of matched sales receipts will increase tax revenues. There are two ways a merchant can frustrate the goal of this type of policy: the merchant can conceal or underreport unobserved sales tax revenue (i.e., cash), and the merchant can shift sales from traceable and monitored forms of payment to cash.

---

8 To some extent, such tax evasion is close to those reported in self-employment activity. Self-employment activity is well known to exhibit lower rates of compliance than taxpayers whose primary source of income are wages or salaries. A number of studies investigated differences in the reporting rates between salaried workers and self-employed (Clotfelter, 1983; Feinstein, 1991; Joulfaian and Rider, 1998; Bruce, 2000).
It is likely that many merchants hide part of their sales and possibly even do not pass levied taxes on to the government. For such merchants, perfect monitoring of electronic sales changes the environment in two important ways.

First, the merchant cannot avoid taxes levied on sales using an electronic payment method. There are costs and benefits to the merchant from generating automatic sales tax payment at the point of sale. Automatic taxation saves the vendor the cost of accounting for and remitting the required taxes at a later date. On the downside, it disallows the use of those funds until tax is remitted and eliminates any opportunity to evade sales taxes on electronic sales.

Compulsory payment of taxes for all electronic transactions requires financial institutions to separate out the tax collected from electronic purchases and send it directly to the relevant agency rather than allowing the merchant to keep the revenue until the time the tax payment is due. This policy aims to persuade consumers to pay for their purchases using electronic payment methods as opposed to cash payments. When an electronic payment is made by credit card or by using a bank debit card, the GST and the PST levied on the purchases are drawn automatically and almost instantly. At present, the financial institutions who issue the employed credit or debit card return the entire amount of the transaction (payment and the levied taxes) to the merchant within 24 hours after the purchase. The merchant acts as an agent for the government and returns the levied taxes according to the frequency of required remittance (either annual, quarterly or monthly). The agent can, therefore, benefit from the use of these sums until the actual tax payment due date. It is worth noting that, following the establishment of immediate GST and PST collection on all electronic transactions, the current tax collection delay will continue to apply for cash payment purchases.

Second, without perfect monitoring of sales-matched income, the merchant could have hidden most sales if he wished to do so. But, once the taxes are automatically transferred to the government agency, the merchant’s sales are automatically revealed.

Faced with these new rules, how will the fraudulent merchant react? The only way left to defraud the government is to evade tax on cash payments. Will the merchant conceal the part of his sales paid in cash and, in addition, find a way to transfer income from electronic transactions to cash? More specifically, in order to maintain income at the pre-regime change level, will the merchant choose to evade tax even more on the unmonitored cash component of sales, given that the room for cheating on monitored electronic sales is limited?

According to contracts with financial institutions, a merchant cannot favor cash payment as opposed to electronic payment, but customers are most likely unaware of this fact. Furthermore, it is not clear whether customers would resist discounts offered by recalcitrant merchants on purchases paid with cash. Since one of the objectives of perfect monitoring of electronic sales is to convince buyers to use electronic payments as much as possible instead of cash payments, buyers’ behavior is critical. In the present study, however, we ignore this aspect of buyer behavior. We also do not examine the impact of monitoring on the timing of sales and taxes paid, as all taxes in our experiment are paid on a period-by-period basis, with the exception of payment of back taxes resulting from an audit of a fraudulent merchant.
The participants in the experiments are paid according to the decisions they make. They face a stylized incentive structure that is comparable to perfect monitoring of electronic sales and imperfect monitoring of cash sales. As a result, we can analyze and understand the potential differences that exist between the theoretical predictions of evasive behavior and the experimental outcomes.

III. EXPERIMENTAL DESIGN AND PROTOCOL

Experimental economics allows us to reproduce, in a controlled environment, a system of revenue declaration under different monitoring schemes. The experimenter has the advantage of observing actual income from different sources — corresponding, for example, from cash sales or electronic sales in the case of the sales tax — as well as reported income. The primary question posed by this experiment is, “Does the institution of perfect monitoring on matched income affect tax compliance?” To focus on this question, other interesting but complicating factors have not been incorporated into the experimental design, such as the redistribution of taxes in the form of a public good, using earned income rather than endowed income, and the application of imperfect audit rules. Because the payment of taxes can be seen as a repeated action by the taxpayer, participants are fed a stream of income in the experiment, randomly realized one period at a time, and are asked to make a reporting decision each period.

The experiment consists of 48 periods of income declaration. In each period, subjects receive income from two sources, choose to report voluntarily the amount of their total income they desire, pay tax on the income reported, are subject to an audit with some probability, and are able to examine their own income, income declaration, tax payment, audit history and penalty history. Participants are told that, for each period, their assigned income is randomly (from a uniform distribution) drawn between 10–110 experimental units (eu) of currency. The study uses two categories of income, simply designated to the participants as Source A and Source B.9

To examine the influence of different proportions of matched income, each participant is assigned randomly one of three types at the beginning of the experiment. The endowment of Source A income is fixed at either 80 percent, 50 percent, or 20 percent of total income for types I, II and III, respectively. All three types are present in each treatment. Participants are told that the government does not know their true earned income or the fraction of their income from Source A.10 A tax rate $t$ of 40 percent is assessed on reported income. Audits are always successful at exposing unreported income. The penalty $f$ for underreporting is an additional 50 percent on unpaid taxes.

---

9 Participants are drawn from a database of 4000 individuals. They are mainly university students, but about 10 percent are from the general population. A filtering system selects only those who have not been involved in similar experiments.

10 Participants are differentiated by their type of income from Source A. The differences in income level come from the draws in each round. To avoid too much variation and to better control the income conditions in the experiment, we have imposed the same level of income for each participant in each round. This last point was not known to the participants.
and an automatic audit of income reported for the previous two periods. The chances of being audited in any one period are 10 percent and increase to 20 percent if a participant reports income less than the median reported income. Using these parameters and a maximum audit rate of 20 percent, we still cannot induce an expected income maximizing (risk-neutral) agent to report positive income. However, this may not be the case for individuals who are risk averse. Even though audit rates of 10 percent and 20 percent might seem high, they are not unlikely in some economic sectors. Using lower audit rates would increase the incentive to cheat. Moreover, the levied taxes are not returned to the participants in any way, and knowledge of such money burning may add to the motivation to underreport as well.

To examine the effects on reporting behavior of instituting perfect monitoring on a portion of income, the participants learn after several periods that Source A will be perfectly monitored by the government. Participants are informed of the change in policy before it is instituted to ascertain whether behavior changes significantly due to the expectation of imminent monitoring.

The experiment includes several treatments. The baseline treatment consists of 21 periods. In each period, subjects receive income, report the amount of their total income they desire, pay tax on the reported income, and are subject to an audit with some probability. Immediately following the baseline treatment, subjects are subject to the announcement treatment for six periods. In this treatment, participants are exposed to the announcement that the change to the new Source A income monitoring system will begin with period 28. The announcement of the new tax collection policy is simply implemented by posting an announcement at the beginning of period 22. Participants are told it will take six periods to implement the new policy. During the announcement phase, we attempt to observe whether facing the imminent implementation of perfect monitoring of matched income induces some participants to change their behavior. This question is relevant since we know that perfect monitoring of matched income will not come into effect without a realistic delay.

Lastly, participants play 21 periods under the new policy (perfect monitoring of Source A income). Once perfect monitoring of Source A income is implemented, participants are told that any income received from Source A that is not reported will result in an automatic audit of total income. In the absence of an audit, the government does not know each participant’s true total income or player type. The two perfect monitoring treatments (with or without the opportunity to alter the sources of income, that is, with

11 This feature of the experiment borrows from current practices of revenue agencies in Canada.
12 For example, the Transactional Records Access Clearinghouse (2005) documents three years of IRS data and finds that corporate audit rates vary depending on the nature of business activities.
13 Two pilot sessions of 12 participants each were conducted to ascertain whether declaration rates were stable for participants during 48 periods of receiving and reporting income. After 10 periods of play, there was no detectable change in the pattern of reporting rates conditional on income received. For this reason, the initial or basic declaration and tax collection regime was set for 21 periods, well beyond any anticipated adjustment by the participants.
or without recharacterization of income) differ in the ability to convert income from Source A to Source B income. In the case of perfect monitoring treatment without recharacterization, participants are not allowed to transfer units of income from Source A to Source B. Perfect monitoring with recharacterization allows participants to transfer income from perfectly monitored Source A to unmonitored Source B at the rate of six units for five units. The goal of this treatment is to observe a costly shift towards cash payments in response to perfect monitoring. The other components of the protocol are the same as for the previous monitoring treatment.

Prior to taking part in any experimental sessions, all participants are required to go through a series of computerized instructions that simulates each potential event during the experiment: a reporting period without audit, a reporting period with audit and no repercussions, and a reporting period with a successful audit and resulting penalty. The participants are informed that the experimenter does not know their total income or player type. In fact, the only information the experimenter has is the declared income for the group, which is used to assign the audits and the value of the total income after an audit.

The actual payoff to each subject for participation in the experiment is exactly his or her income from one period of play. The total income appears to participants as the sum of income from Source A and Source B, randomly drawn from a uniform distribution, and ranged from 10–110 experimental units. At the conclusion of the experiment, one period is randomly drawn by the experimenter for payment and is compensated at the rate of CAN$0.50/eu. Participants earned an average of $25 including a $10 show-up fee for less than 90 minutes of participation.\(^{14}\)

Sixteen sessions with 12 participants each were conducted, resulting in a total of 192 participants.\(^{15}\) Each session had four of each type of player (20 percent, 50 percent and 80 percent Source A income). Eight sessions were conducted with the perfect monitoring treatment without recharacterization of income and eight sessions were conducted with recharacterization of income.

\(^{14}\) The external validation of laboratory experiments is always a challenge in the field of experimental economics. Our risk-taking experiments occur in a very small-stakes environment. How will the taxpayers behave in the analogous “real world” situation that involves large amounts of tax? One might expect that higher stakes may induce higher risk aversion. Some experimental studies have shown, however, that if risk aversion increases as the scale of payoff increases, subjects still exhibit risk aversion even for low payments (Holt and Laury, 2002). Furthermore, in our experiment, the stakes vary substantially between 10–110 experimental units of currency. Social and moral considerations may also matter in the “real world.” Such considerations would probably induce people to refrain from tax evasion (McCaffery and Slemrod, 2006). Note, however, that the behavior of our participants may be subject to social and moral considerations. Furthermore, such considerations may have played a role since our instructions specifically referred to taxation. The instructions to participants are available upon request.

\(^{15}\) Two participants over-reported their income in more than 60 percent of the periods. They were dropped from the initial database, which means that the data analysis was conducted with 190 players. Keeping those observations does not significantly affect the results.
IV. THEORETICAL CONSIDERATIONS

We provide, in this section, a brief theoretical discussion of the standard expected utility analysis of the individual compliance decision and then examine the implications of behavioral economics for the analysis.

A. Theoretical Predictions of the Expected Utility Model

Our experiment borrows from the seminal work of Allingham and Sandmo (1972) and Yitzhaki (1974), which is based on the expected utility model, or A-S-H model. Consider first the baseline treatment, letting $y$ denote the subject’s gross income drawn between 10–110, the expected value of which is 60. This is common knowledge to all participants. The expected probability of being audited is:

$$p = 0.5 \times A_{\text{Low}} + 0.5 \times A_{\text{High}},$$

where $A_{\text{Low}}$ is the audit rate for those reporting income below the median and $A_{\text{High}}$ is the audit rate for those reporting income above the median. Recall that $t$ is the tax rate and $f$ is the penalty rate. For simplicity, consider the decision of a risk-neutral participant who decides to report an amount $R$ of gross income $y$. Expected net income (ENI) is:

$$ENI = (1 - p)(y - Rt) + p[y - Rt - (y - R)t - (y - R)ft]$$

$$= y - p\left[t(y - R) + ft(y - R)\right] + tR.$$

If an individual chooses to report full income ($R = y$), then net income ($NI$) is simply:

$$NI = y(1 - t).$$

When the choice is to report no income ($R = 0$), expected net income is:

$$ENI = (1 - p)y + p(1 - t - ft)y = (1 - pt)(1 + f)y$$

The expected rate of return from reporting no income, relative to reporting all income, is:

$$\frac{ENI\left(with\ R = 0\right) - NI}{NI}.$$

With $t = 40$ percent, $f = 50$ percent, $p = 0.15$ (with $A_{\text{Low}} = 10$ percent and $A_{\text{High}} = 20$ percent), and $y = 60$, an individual who reports full income receives an expected payoff of 36 units, but receives 54.6 units if he or she reports nothing in period 1, a return on evasion of 51.7 percent. For periods 3–48, the rate of return for cheating falls to 21.67 percent since, if the taxpayer is audited, automatic audits for the previous two periods
occur. Based on this simple model, consider now the theoretical predictions of the expected utility model of the effects of the announcement treatment. Since the parameters remain unchanged in the announcement treatment, the theoretical predictions are exactly the same as in the baseline treatment, that is, individuals should not report their incomes.

Turning to the perfect monitoring treatment, we can easily see that the overall compliance rate should be higher in this treatment, compared to the baseline, since individuals should still evade taxes from income of source B (as the expected rate of return of such income is unchanged) but should always report their income from source A, as the probability of an audit is equal to 1 for underreporting income of source A.

Finally, consider the theoretical predictions of the effects of the perfect monitoring treatment with recharacterization of income. In this case, the overall compliance rate should be close to the baseline treatment. The reason is that it is always optimal for a participant to buy the maximum possible number of units of income of source B with units of A. Although purchasing additional units of B income costs 1.20 units of income A, such recharacterization allows the player to avoid automatic auditing. The expected rate of return of the purchased units of income B is, therefore, 1.67 percent (i.e., 21.67 percent minus 20 percent). Thus, it remains optimal to buy extra units of B and not report them. The expected rate of return for tax evasion of income from source B already owned by the participants is unchanged at 21.67 percent.

The above predictions are derived from simple versions of the Allingham-Sandmo and Yitzhaki models under the assumption of risk neutral individuals. For all treatments, these predictions suggest that individuals will totally avoid tax by not reporting unmonitored incomes. To escape those corner solutions, a more realistic assumption is to presume that the taxpayers are risk averse. Depending on the degree of risk aversion, various levels of unmonitored income will be reported as long as the expected penalty rate is less than the regular tax rate. Social and moral considerations will also induce people to refrain from tax evasion. This point is discussed by McCaffery and Slemrod (2006) in the context of behavioral public finance.

In addition, several authors have recently used non-expected utility models to explain why people pay or decide not to pay taxes (for example, see the analysis of prospect...
B. Predictions of Alternative Behavioral Economics Models

A possible objection to the expectation of a positive relationship between tax compliance and perfect monitoring without recharacterization of income is that individuals may counteract the consequences of the implementation of perfect monitoring by taking on more risk and declaring less of their imperfectly monitored income. Such an effect is akin to the idea of a reference point in the prospect theory. According to this theory, individuals will try to recover their losses following any policy change, even if it means taking on more risk. This is a familiar finding among gamblers and traders who set specific financial targets. Our conjecture is stated more precisely in hypothesis H1: The increase in tax revenue due to the implementation of perfect monitoring of income of one class (Source A) may be offset by reporting less income from another class (Source B).

Our second conjecture is that, for reasons similar to those presented above, some participants may report less income in the announcement treatment than in the baseline treatment in order to offset the future tax burden of the perfect monitoring policy. This is summarized as hypothesis H2: Participants may be willing to counteract the future consequences of the perfect monitoring policy by reporting less income in the current announcement treatment than in the baseline.

Audit experience may also influence the perception of the likelihood of future audits, which is theoretically incorrect since the instructions of the experiment made it clear that audits were randomly assigned in each period, and not conditioned on past behavior. A possible reason for this misperception is that a participant may focus on the bad outcome and neglect the fact that such an event is very unlikely to occur. This effect is generally termed the “probability neglect” effect. The idea is that, “When intense emotions are engaged, people tend to focus on the adverse outcome, not on its likelihood, which may lead to significant distortions in both private and public arenas” (Sunstein, 2002, 61). Moreover, people may feel a disproportionate fear of risk when the associated risks are hard to control (Slovic, 2000; Sunstein, 2002). Finally, the occurrence of an audit may also induce the opposite effect, by provoking individuals to underestimate the likelihood of future audits. This effect, generally termed the “gambler’s fallacy” or “Monte Carlo

18 It is outside the scope of this paper to formalize those models.
19 The reference point is of the individual’s benchmark against which each alternative is compared. In our case, this can be the initial outcome obtained in the baseline treatment. According to the prospect theory, individuals will “code” each alternative as a gain or a loss in utility relative to this reference point rather than using the absolute value of the outcome. The resulting losses or gains are then weighed by their perceived probabilities of occurrence, forming a non-linear value function. Because people typically approach gains and losses differently, generally acting risk averse on gains and risk seeking on losses (Kahneman and Tversky, 1979), the result is a non-linear utility function.
20 A similar situation exists with the concept of a target income and the behavioral model of labor supply developed by Altman (2001). See also Camerer et al. (1997) on income targets in the case of the supply of taxis.
fallacy,” relies on the impression that a certain stochastic event is less likely to happen following the occurrence of an event or a series of events. Again, this is obviously a fallacy since the occurrence of past events does not impact the probability that similar random events will occur in the future. According to Kahneman and Tversky, the gambler’s fallacy may be induced by a psychological heuristic. This is stated as hypothesis H3: Audit experience may influence perceptions of risk and therefore influence current decisions.

Finally, in addition to risk perception, other behavioral factors may also influence tax compliance. For example, some individuals may desire to conform to a norm of truth-telling. Such individuals may avoid the emotional cost associated with deviation from this norm by reporting some or all of their income (Coricelli et al., 2007). This is summarized as hypotheses H4: Irrespective of pecuniary incentives, non-pecuniary factors such as norms of truth-telling may motivate subjects to report some or all of their income in both the baseline and the announcement treatments.

V. EXPERIMENTAL RESULTS

A. Descriptive Statistics

Figures 1a and 1b show the time path of the average reporting rates by period in the different treatments for sessions 1-8 (without recharacterization of income) and sessions 9-16 (with recharacterization of income), respectively. To facilitate comparisons across treatments, average reporting rates are depicted graphically on the same time scale. Average reporting rates are significantly different from zero, in contradiction to the prediction of the simple version of the A-S-H models. Both figures indicate that, in most treatments, reporting rates are rather stable over time. Comparing treatments, both figures indicate important differences. Figure 1a shows that the announcement phase induces a significant reduction of overall reporting rate. Introducing the monitoring policy without recharacterization of income does not lead to an improvement of reporting rates. In contrast, Figure 1b shows that introducing perfect monitoring with recharacterization of income induces a negative effect on the reporting rate level. These results seem to confirm the idea that introducing a perfect monitoring policy does not necessarily lead to an improvement of tax compliance because agents may be willing to counteract the effects of such a policy by reporting less income from other sources, which translates into a lower overall reporting rate. In general, these results support the predictions of the behavioral model.

Table 1 presents descriptive statistics on the behaviors exhibited by the participants in the experiments (definitions of all variables are presented in the Appendix). It shows that the reporting rate significantly decreases with perfect monitoring and recharacterization of income. The reporting rate is 76.59 percent in the baseline

---

21 The intuition behind the gambler’s fallacy is that people may expect that deviations from average should balance out. For example, if a fair coin is tossed repeatedly and tails comes up a larger number of times than is expected, a gambler may incorrectly believe that this means that heads is more likely in future tosses.
Figure 1a
Overall Reporting Rate by Treatment Over Time (Sessions 1–8)

Figure 1b
Overall Reporting Rate by Treatment Over Time (Sessions 9–16)
### Table 1
Descriptive Statistics by Treatment

<table>
<thead>
<tr>
<th>Variables</th>
<th>Baseline Treatment (Both Sessions 1–8 and 9–16)</th>
<th>Announcement Preceding Monitoring Without Recharacterization of Income (Sessions 1–8)</th>
<th>Announcement Preceding Monitoring with Recharacterization of Income (Sessions 9–16)</th>
<th>Perfect Monitoring Treatment Without Recharacterization of Income (Sessions 1–8)</th>
<th>Perfect Monitoring Treatment with Recharacterization of Income (Sessions 9–16)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All income reported in €</td>
<td>Average 28.74</td>
<td>Average 27.99</td>
<td>Average 30.40</td>
<td>Average 27.16</td>
<td>Average 28.48</td>
</tr>
<tr>
<td>Total income</td>
<td>60.96</td>
<td>52.62</td>
<td>62.45</td>
<td>57.93</td>
<td>59.38</td>
</tr>
<tr>
<td>Reported income</td>
<td>45.54</td>
<td>34.64</td>
<td>45.63</td>
<td>43.52</td>
<td>40.15</td>
</tr>
<tr>
<td>Overall reporting rate (%)</td>
<td>76.59</td>
<td>67.67</td>
<td>76.56</td>
<td>76.94</td>
<td>70.30</td>
</tr>
<tr>
<td>Reporting rate on B (%)</td>
<td>63.66</td>
<td>52.05</td>
<td>59.30</td>
<td>54.22</td>
<td>54.91</td>
</tr>
<tr>
<td>Taxes paid on the reported income</td>
<td>17.95</td>
<td>13.69</td>
<td>17.99</td>
<td>17.18</td>
<td>15.88</td>
</tr>
<tr>
<td>Amount B purchased</td>
<td>6.22</td>
<td>4.91</td>
<td>4.91</td>
<td>4.91</td>
<td>4.91</td>
</tr>
<tr>
<td>Reporting no income (%)</td>
<td>5.51</td>
<td>9.82</td>
<td>4.91</td>
<td>0.05</td>
<td>1.55</td>
</tr>
<tr>
<td>Reporting total income (%)</td>
<td>47.54</td>
<td>39.12</td>
<td>42.81</td>
<td>39.90</td>
<td>38.45</td>
</tr>
<tr>
<td>Reporting no source B income (%)</td>
<td>26.44</td>
<td>31.75</td>
<td>25.44</td>
<td>25.06</td>
<td>23.21</td>
</tr>
<tr>
<td>Auditing (%)</td>
<td>14.31</td>
<td>13.51</td>
<td>13.51</td>
<td>14.04</td>
<td>16.14</td>
</tr>
</tbody>
</table>

(1) This variable is defined as (reported income—income from Source A) / (income from Source B)
treatment. This percentage falls to 70.30 percent in the monitoring treatment with recharacterization of income and is significantly lower than the baseline result ($z = 4.452; p = 0.0001$, Wilcoxon signed-rank test). In contrast, we find no significant difference between the baseline treatment and the monitoring treatment without recharacterization of income ($z = -0.490; p > 0.1$). We also observe that the reporting rate is significantly lower in the announcement stage preceding the monitoring treatment without recharacterization of income (67.67 percent), compared to the baseline treatment ($z = 3.456; p = 0.0005$, Wilcoxon signed-rank test). We speculate that the difference in behavior between the two treatments is the result of efforts of the participants to counteract the expected consequences of the perfect monitoring policy prior to its enactment. In contrast, no significant difference is found between the baseline treatment and the announcement treatment with the opportunity for recharacterization of income ($z = 0.941; p > 0.1$).

A possible explanation is that participants anticipate that they can avoid future taxes under the implementation of perfect monitoring when recharacterization is permitted.

In addition to the differences in the overall reporting rate across the treatments discussed above, we also observe that the zero-compliant participants, although small in number, are more numerous in the monitoring treatment with recharacterization of income (1.55 percent) than in the monitoring treatment without recharacterization of income (0.05 percent). Figures 2a and 2b offer information about the effects of announcement and monitoring with respect to the level of income for sessions 1–8 (perfect monitoring without recharacterization) and sessions 9–16 (perfect monitoring with recharacterization), respectively.

As illustrated in Figures 2a and 2b, reporting in the baseline treatment is greater than 70 percent for every income range. Given the substantial incentives to underreport, positive and relatively high reporting rates suggest that the participants are either generally quite risk averse or that alternative models of tax compliance describe their behavior. We also observe that, in all treatments, the reporting rate decreases with the level of income. Figure 2a shows that the announcement induces a significant reduction of the overall reporting rate. In contrast, perfect monitoring without the opportunity to recharacterize income seems to have two opposing effects depending on the level of income. An increase in monitoring seems to have a negative effect on the reporting rate for lower incomes, whereas it has a positive effect on this ratio for higher incomes. Figure 2b shows that perfect monitoring with recharacterization of income induces a reduction in the reporting rate level. Consistent with our previous results, the negative effect of the announcement is lower in this case, compared to Figure 2a, which indicates that subjects are aware that the introduction of perfect monitoring is less dramatic when they have the opportunity to recharacterize income. Finally, the negative effect of perfect monitoring (significant at the five percent level), compared to the baseline treatment, indicates that increasing monitoring does not necessarily improve reporting rates when individuals have the opportunity to recharacterize income.

Figures 3a and 3b provide information concerning reporting rates for the three types of players across all three experimental phases, with and without recharacterization.

22 Note that a Mann-Whitney-Wilcoxon test indicates no significant difference between the two baseline treatments in sessions with and without recharacterization of income ($z = 1.131; p > 1$).
Figure 2a
Ratio of Reported Income per Level of Income (Sessions 1–8)

Figure 2b
Ratio of Reported Income per Level of Income (Sessions 9–16)
Figure 3a
Overall Reporting Rate by Treatment and Type of Player (Sessions 1–8)

Figure 3b
Overall Reporting Rate by Treatment and Type of Player (Sessions 9–16)
of income, respectively. Figure 3a shows that without recharacterization of income, highly monitored participants who had 80 percent of their income from Source A tend to increase their reporting rate on average in the monitoring treatment and in the announcement treatment, compared to the baseline treatment. In sharp contrast, for the other two types of participants (e.g., with 50 percent and 20 percent of income from Source A), the reporting rates decline through the announcement stage and move back upward throughout the perfect monitoring treatment. When future recharacterization is possible (Figure 3b), highly monitored participants tend to stabilize their reporting rates, while the other two types of participants significantly reduce them in the announcement and the monitoring treatments.

Figures 4, 5 and 6 compare the two monitoring treatments with and without recharacterization of income for each participant’s income type. For those with 80 percent and 50 percent of income from Source A (Figures 4 and 5), the reporting rates are generally lower for higher income levels, especially when participants had the opportunity to transfer income from Source A to Source B. For the least monitored participants, those with only 20 percent of total income derived from Source A, the difference between the two treatments is insignificant. This presumably occurs because participants do not want to face a 20 percent auditing rate. Indeed, since buying extra units of B is costly, it would be useless to do so if one had to report it. The participant knows that, having

---

Figure 4

Ratio of Reported Income for Those with 80% of Income Automatically Audited (Source A)

---

23 Alm and McKee (2006) found that announcement increases compliance of those told they will be audited, but reduces compliance of those knowing they will not be audited; the net effect is that overall compliance falls.
Figure 5
Ratio of Reported Income for Those with 50% of Income Automatically Audited
(Source A)

Figure 6
Ratio of Reported Income for Those with 20% of Income Automatically Audited
(Source A)
80 percent of his or her total income already not systematically audited, part of this income will have to be reported to try to avoid the higher auditing rate.

**B. Econometric Results**

In this section, the data are analyzed with parametric regressions. Table 2 consists of two panels. The left panel displays the result of five regressions in which the dependent variable is the percentage of total income reported to the tax authority by individual \( i \) in period \( t \). It shows the estimates of the determinants of reporting rates by treatment with random effect Tobit models.\(^{24}\) The use of Tobit models is justified by the number of left and right censored observations in the sample. The right panel (column 6) presents the results of the determinants of tax revenue from a random effect feasible Generalized Least Squares regression.\(^{25}\) We control for demographics (not reported in the estimates, but available upon request).\(^{26}\)

The estimates support earlier observations that income reporting declines with total income level. We also observe that the participants reduce their reported income in the period following an audited period. This reaction to previous audit is in accordance with our discussion of optimal strategy in the expected utility model, and is also consistent with the reference-dependence effect in prospect theory which implies that participants may be willing to compensate for the losses suffered in the audit by taking more risks (Kahneman and Tversky, 1991). This occurrence could also be explained by the gambler’s fallacy concept, which states that the participants believe that after being audited once, the probability of being audited again soon after is smaller than the described probability. This is of course incorrect, as auditing probabilities are independent across

---

\(^{24}\) In a panel Tobit model, the error component splits into a time-invariant individual random effect, \( \alpha \), and a time-varying idiosyncratic random error \( \epsilon \).

\(^{25}\) The variable \( RT_{it} \) measures the taxes paid (tax revenues) by participant \( i \) in period \( t \). Tax revenues are explained by a vector of exogenous variables \( z \) and the corresponding vector with parameters \( \delta \), and a random variable that can be divided into a random individual effect \( \alpha \), and a pure random variable \( \epsilon \):

\[
RF_{it} = z_{it} \delta + \alpha_i + \epsilon_{it}, \quad i = 1, \ldots, n, \quad t = 1, \ldots, T.
\]

Assuming uncorrelated error terms and that the \( z \) are also uncorrelated with the error terms, an appropriate estimation technique is the feasible Generalized Least Squares regression discussed in Greene (2008).

\(^{26}\) Demographic variables include various participant characteristics: age (mean: 25.4; standard deviation: 5.95); male (53.1 percent); previous participation is a binary variable for whether the participant has already taken part in an experiment other than the current one (1 = participated in a previous experiment (38.5 percent)), 0 = otherwise); gamble indicates whether the participant chose to earn a guaranteed $5.00 show-up fee or a gamble with a 50 percent probability of earning $11.00 and a 50 percent probability of earning $0, with participants who chose the gamble considered to be relatively less risk averse than those who chose the $5.00 show-up fee (73.4 percent ); instruction feedback describes the participant’s assessment of the clarity of the instructions on a scale of 0 to 10, 10 being “very clear” (mean: 8.29; standard deviation:1.46); and binary variables indicating whether the participant is a worker (8.85 percent), unemployed (6.25 percent), a student (80.2 percent), a graduate student (22.4 percent) or a student with prior mathematical training (66.1 percent). The introduction of demographic variables does not affect the estimated coefficients of the experimental variables. Most of these variables are not statistically significant.
### Table 2
Determinants of Reporting Rates and Tax Revenue

<table>
<thead>
<tr>
<th>Treatment:</th>
<th>Perfect Monitoring</th>
<th>All Treatments</th>
<th>Tax Revenue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent variable Models</td>
<td>Baseline</td>
<td>Announcement</td>
<td>Without Recharacterization</td>
</tr>
<tr>
<td></td>
<td>RE Tobit¹</td>
<td>RE Tobit¹</td>
<td>RE Tobit¹</td>
</tr>
<tr>
<td>Total income</td>
<td>–0.0057***</td>
<td>–0.0055***</td>
<td>–0.0033***</td>
</tr>
<tr>
<td></td>
<td>(0.0003)</td>
<td>(0.0004)</td>
<td>(0.0002)</td>
</tr>
<tr>
<td>Previous period audit</td>
<td>–0.0640***</td>
<td>–0.0961**</td>
<td>–0.0766***</td>
</tr>
<tr>
<td></td>
<td>(0.0243)</td>
<td>(0.0374)</td>
<td>(0.0163)</td>
</tr>
<tr>
<td>Type II</td>
<td>0.0222</td>
<td>–0.0918</td>
<td>–0.1367**</td>
</tr>
<tr>
<td></td>
<td>(0.0776)</td>
<td>(0.0816)</td>
<td>(0.0699)</td>
</tr>
<tr>
<td>Type III</td>
<td>0.0747</td>
<td>–0.1045</td>
<td>–0.2703***</td>
</tr>
<tr>
<td></td>
<td>(0.080)</td>
<td>(0.0844)</td>
<td>(0.070)</td>
</tr>
<tr>
<td>Dummy, Sessions 7–16</td>
<td>0.112</td>
<td>0.217***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0687)</td>
<td>(0.0724)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Baseline</th>
<th>Ref.</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Announcement (both with and without recharacterization)</td>
<td>–0.153***</td>
<td>–1.539***</td>
</tr>
<tr>
<td></td>
<td>(0.0198)</td>
<td>(0.318)</td>
</tr>
<tr>
<td>Announcement and recharacterization</td>
<td>0.105***</td>
<td>0.651</td>
</tr>
<tr>
<td></td>
<td>(0.0284)</td>
<td>(0.447)</td>
</tr>
<tr>
<td>Type I (monitoring both with and without recharacterization)</td>
<td>0.170***</td>
<td>3.498***</td>
</tr>
<tr>
<td></td>
<td>(0.0223)</td>
<td>(0.349)</td>
</tr>
<tr>
<td>Type I and monitoring with recharacterization</td>
<td>–0.242***</td>
<td>–0.405***</td>
</tr>
<tr>
<td></td>
<td>(0.0314)</td>
<td>(0.492)</td>
</tr>
<tr>
<td>Type II (monitoring both with and without recharacterization)</td>
<td>–0.0183</td>
<td>–0.0145</td>
</tr>
<tr>
<td></td>
<td>(0.0221)</td>
<td>(0.349)</td>
</tr>
<tr>
<td>Type II and monitoring with recharacterization</td>
<td>–0.172***</td>
<td>–3.119***</td>
</tr>
<tr>
<td></td>
<td>(0.0311)</td>
<td>(0.491)</td>
</tr>
<tr>
<td>Type III (monitoring both with and without recharacterization)</td>
<td>–0.0854***</td>
<td>–1.608***</td>
</tr>
<tr>
<td></td>
<td>(0.0221)</td>
<td>(0.355)</td>
</tr>
<tr>
<td>Type III and monitoring with recharacterization</td>
<td>–0.111***</td>
<td>–0.979*</td>
</tr>
<tr>
<td></td>
<td>(0.0320)</td>
<td>(0.500)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>1.170***</td>
<td>1.182***</td>
<td>1.391***</td>
<td>0.928***</td>
<td>1.101***</td>
</tr>
<tr>
<td></td>
<td>(0.188)</td>
<td>(0.198)</td>
<td>(0.235)</td>
<td>(0.220)</td>
<td>(0.151)</td>
</tr>
<tr>
<td>Observations</td>
<td>3,800</td>
<td>1,140</td>
<td>1,995</td>
<td>1,995</td>
<td>8,930</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>–2,410</td>
<td>–689</td>
<td>–553</td>
<td>–811</td>
<td>–4,740</td>
</tr>
<tr>
<td>Left censored observations</td>
<td>214</td>
<td>84</td>
<td>1</td>
<td>31</td>
<td>330</td>
</tr>
<tr>
<td>Right censored observations</td>
<td>1,886</td>
<td>481</td>
<td>814</td>
<td>778</td>
<td>3,959</td>
</tr>
</tbody>
</table>

Notes: Standard errors are in parentheses. Asterisks denote significance at the 1% (***) , 5% (**), and 10% (*) levels. The first period of observation was eliminated due to the construction of the lagged audit variable. The correlation coefficient ρ is statistically significantly different from zero according to the Lagrange multiplier test with a 5 percent critical threshold.

1) RE Tobit = Random Effect Tobit.
2) RE FGLS: Random Effect Feasible Generalized Least Squares.
Columns (1) and (2) indicate that there is no significant difference among player types in the baseline and announcement treatments. In contrast, as shown by column (3), the perfect monitoring treatment without recharacterization of income reveals that players of types II and III report significantly less income than players of type I, which is consistent with our previous findings. The lower constant term suggests that, on average, people report less income in the perfect monitoring treatment with recharacterization.

With respect to the reference variable — that is the baseline treatment preceding the announcement and policy implementation treatments — column (5) reveals that income reporting decreases for all types of players under perfect monitoring when transfers from Source A to Source B income are permitted. If the participants are not able to shift income from Source A to Source B, income reporting increases for those participants with 80 percent of their initial income coming from Source A under perfect monitoring, when compared to the baseline treatment. Income reporting is stable for participants with 50 percent of their income coming from Source A, and decreases for those with 20 percent of their income coming from Source A.

In the announcement treatment preceding perfect monitoring without recharacterization of income, income reporting decreases with respect to the baseline treatment. There is no diminution in income reporting, as already noted earlier, in the announcement treatment preceding the introduction of perfect monitoring with the opportunity to recharacterize income.

Turning to the determinants of tax revenues, column (6) reports a positive and significant coefficient associated with the “total income” variable. Earlier results indicated that reporting rates decrease with income level. Nonetheless, column (6) shows that tax revenues increase with income, since the increase in income compensates for the reduction in reported income. Consistent with our previous results, column (6) shows that tax revenues decrease under perfect monitoring with recharacterization, compared to the baseline. Tax revenues decrease by 0.907 eu, 3.134 eu, and 2.587 eu with respect to the baseline treatment for participants with 80 percent, 50 percent and 20 percent of their initial income coming from Source A, respectively. If the participants are not allowed to shift income from Source A to Source B, tax revenues increase by 3.498 eu for those participants with 80 percent of their initial income coming from Source A under perfect monitoring, when compared to the baseline treatment. Tax revenues are stable for participants with 50 percent of their income coming from Source A, and they decrease by 1.607 eu for those with 20 percent of their income coming from Source A.

Behavioral economics is an explanation of the results and not the basis for a policy recommendation. McCaffery (2008) has discussed the marriage of behavioral economics and fundamental tax reform in the context of policies aimed at increasing the savings rate in the United States.

For example, to obtain tax revenue for the type II participant, add the coefficient –0.0145 of type II (both treatments) to the coefficient –3.119 of “Type II and recharacterization.”
VI. Conclusion

This study examines the impact on tax compliance of perfect monitoring of a single source of income in the presence of a second unmonitored source, using a laboratory experiment to observe actual and reported incomes in a repeated-decision framework. The background application is sales taxes. Two treatments were used to observe the differences in responses to monitoring — with and without the opportunity to alter the sources of income from the perfectly monitored source to the alternative unmonitored source.

Our experiment contributes to the existing literature on tax compliance by showing that increasing the probability of audit does not necessarily lead to a decline in tax evasion, because taxpayers may offset the revenue gain from an increased probability of audit of one source of income by reporting less income from other sources that are less monitored, and by recharacterizing closely monitored income as an alternative less monitored source of income when such recharacterizing is possible. This result is consistent with Gërxhani and Schram (2006), who find a significant shift from perfectly monitored income towards less monitored income when tax evasion is made possible. It is also in line with Alm, Deskins, and McKee (2009), who suggest that the amount of tax evasion is greater for individuals with non-matched income.

We have three key findings. First, on average, the less monitored income types (50 percent and 20 percent of income coming from the perfectly monitored Source A), and all income types when the option to alter the sources of income was allowed, had higher rates of noncompliance. The only group that had the same rate of compliance was the group of participants who derived 80 percent of their income from Source A and had no opportunity to recharacterize their incomes. Second, when participants had the opportunity to recharacterize perfectly monitored income, they more than compensated for the unavoidable taxes on Source A income by reporting less unmonitored income (from Source B). Third, a significant decrease in tax revenues (approximately a 15 percent) occurred if income could be recharacterized (at a cost). We are of course not claiming that a 15 percent tax reduction in tax revenues would be observed with such a reform, but our experiment suggests that if perfect monitoring is instituted without some other complementary policies, an increase in tax revenues is not the likely outcome.

ACKNOWLEDGEMENTS

The authors thank Nathalie Viennot-Briot for her technical assistance in preparing this paper. The comments and suggestions of the referees and the editor were very helpful in revising the paper. The usual disclaimer applies.

REFERENCES


### Appendix Table A1

**Definitions of Variables**

<table>
<thead>
<tr>
<th>Name</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic (baseline)</td>
<td>1 if Baseline treatment is played and 0 otherwise</td>
</tr>
<tr>
<td>Announcement</td>
<td>1 if Announcement treatment is played and 0 otherwise</td>
</tr>
<tr>
<td>Announcement with recharacterization</td>
<td>1 if Announcement treatment preceding monitoring with the option to alter is played and 0 otherwise</td>
</tr>
<tr>
<td>Announcement without recharacterization</td>
<td>1 if Announcement treatment preceding monitoring without the option to alter is played and 0 otherwise</td>
</tr>
<tr>
<td>Monitoring with recharacterization</td>
<td>1 if Monitoring with the option to alter treatment is played and 0 otherwise</td>
</tr>
<tr>
<td>Monitoring without recharacterization</td>
<td>1 if Monitoring without the option to alter treatment is played and 0 otherwise</td>
</tr>
<tr>
<td>Perfect monitoring</td>
<td>1 if Monitoring is possible (i.e. in monitoring treatments with and without option to alter) and 0 otherwise</td>
</tr>
<tr>
<td>Type I</td>
<td>Player’s Source A income is 80% of his or her total income</td>
</tr>
<tr>
<td>Type II</td>
<td>Player’s Source A income is 50% of his or her total income</td>
</tr>
<tr>
<td>Type III</td>
<td>Player’s Source A income is 20% of his or her total income</td>
</tr>
<tr>
<td>Type i and monitoring</td>
<td>Interaction variables</td>
</tr>
<tr>
<td>Total income</td>
<td>Player’s income randomly drawn between 10 and 110 experimental units</td>
</tr>
<tr>
<td>Reported income</td>
<td>Reported income by each player</td>
</tr>
<tr>
<td>Overall reporting rate</td>
<td>Reported income / total income</td>
</tr>
<tr>
<td>Audit</td>
<td>1 if the participant is audited and 0 otherwise</td>
</tr>
<tr>
<td>Age</td>
<td>Player’s age</td>
</tr>
<tr>
<td>Men</td>
<td>1 if participant is a male and 0 otherwise.</td>
</tr>
<tr>
<td>Previous participation</td>
<td>1 if the player has already taken part in an experiment other than the current one and 0 otherwise</td>
</tr>
<tr>
<td>Gamble</td>
<td>1 if the participant chose a gamble (a 50% probability of earning $11 and a 50% probability of earning $0) and 0 if he or she chose to earn a guaranteed $5 show-up fee.</td>
</tr>
<tr>
<td>Instruction feedback</td>
<td>Describes the participant’s assessment of the clarity of the instructions (from 0–10)</td>
</tr>
<tr>
<td>Worker</td>
<td>1 if the participant is a worker and 0 otherwise</td>
</tr>
<tr>
<td>Unemployed</td>
<td>1 if the participant is unemployed and 0 otherwise</td>
</tr>
<tr>
<td>Student</td>
<td>1 if the participant is a student and 0 otherwise</td>
</tr>
<tr>
<td>Graduate student</td>
<td>1 if the participant is a graduate student and 0 otherwise</td>
</tr>
<tr>
<td>Mathematical training</td>
<td>1 if the student has prior mathematical training and 0 otherwise</td>
</tr>
</tbody>
</table>