Online appendix for "Endogenous audits, uncertainty, and taxpayer assistance services: Theory and experiments" by Christian A. Vossler and Scott M. Gilpatric.

This document includes theoretical proofs (Appendix A), representative materials describing the experimental setting (Appendix B), and supplemental econometric analysis (Appendix C). The representative materials are for Treatment 8 , which reflects the most complicated setting (partial audits, both audit and liability information services).

## Appendix A: Proofs

The following Lemma will be useful for proofs below.

Lemma 1: The cost-minimizing choice of reported income and deductions is such that, if an audit occurs, the probability of being found to have underreported income equals the probability of being found to have over-reported deductions, $\int_{R}^{b} f(x) d x=\int_{j}^{D} h(\delta) d \delta$.

Proof of Lemma 1: Satisfaction of the FOCs, equations [2] \& [3], imply that

$$
\begin{aligned}
& t+\gamma \frac{\partial G(y-R+D)}{\partial R} \beta\left[\int_{R}^{b}(x-R) f(x) d x+\int_{j}^{D}(D-\delta) h(\delta) d \delta\right] \\
&-\beta[c+\gamma G(y-R+D)] \int_{R}^{b} f(x) d x \\
&=t-\gamma \frac{\partial G(y-R+D)}{\partial D} \beta\left[\int_{R}^{b}(x-R) f(x) d x+\int_{j}^{D}(D-\delta) h(\delta) d \delta\right] \\
&-\beta[c+\gamma G(y-R+D)] \int_{j}^{D} h(\delta) d \delta
\end{aligned}
$$

Noting that $\frac{\partial G(y-R+D)}{\partial R}=-\frac{\partial G(y-R+D)}{\partial D}$, the above equation reduces to $\int_{R}^{b} f(x) d x=\int_{j}^{D} h(\delta) d \delta$.

## Proof of PROPOSITION 1:

We employ Cramer's Rule to sign comparative statics. We will denote the objective function (the taxpayer's expected cost function) $K$ and use subscripts to denote first and second partial derivatives in the usual manner. The first derivatives, corresponding to the first-order conditions given by equations [2] and [3], are

$$
\begin{aligned}
& K_{R}=t+\gamma \frac{\partial G(y-R+D)}{\partial R} \beta\left[\int_{R}^{b}(x-R) f(x) d x+\int_{j}^{D}(D-\delta) h(\delta) d \delta\right] \\
& -\beta[c+\gamma G(y-R+D)] \int_{R}^{b} f(x) d x \\
& K_{D}=-t+\gamma \frac{\partial G(y-R+D)}{\partial D} \beta\left[\int_{R}^{b}(x-R) f(x) d x+\int_{j}^{D}(D-\delta) h(\delta) d \delta\right] \\
& \quad+\beta[c+\gamma G(y-R+D)] \int_{j}^{D} h(\delta) d \delta
\end{aligned}
$$

To sign the various second derivatives, note that $\frac{\partial G(y-R+D)}{\partial R}<0, \frac{\partial G(y-R+D)}{\partial D}>0, \frac{\partial^{2} G(y-R+D)}{\partial R^{2}} \geq$ $0, \frac{\partial^{2} G(y-R+D)}{\partial D^{2}} \geq 0, \frac{\partial^{2} G(y-R+D)}{\partial R \partial D} \leq 0$. Then, $K_{R R}=\gamma \frac{\partial^{2} G(y-R+D)}{\partial R^{2}} \beta\left[\int_{R}^{b}(x-R) f(x) d x+\int_{j}^{D}(D-\delta) h(\delta) d \delta\right]$ $+\beta[c+\gamma G(y-R+D)] f(R)-\gamma \beta \frac{\partial G(y-R+D)}{\partial R} \int_{R}^{b} f(x) d x>0$ $K_{D D}=\gamma \frac{\partial^{2} G(y-R+D)}{\partial D^{2}} \beta\left[\int_{R}^{b}(x-R) f(x) d x+\int_{j}^{D}(D-\delta) h(\delta) d \delta\right]$ $+\beta[c+\gamma G(y-R+D)] h(D)+\gamma \beta \frac{\partial G(y-R+D)}{\partial D} \int_{j}^{D} h(\delta) d \delta>0$
$K_{R D}=\gamma \frac{\partial^{2} G(y-R+D)}{\partial R \partial D} \beta\left[\int_{R}^{b}(x-R) f(x) d x+\int_{j}^{D}(D-\delta) h(\delta) d \delta\right]+\gamma \beta \frac{\partial G(y-R+D)}{\partial R} \int_{j}^{D} h(\delta) d \delta-$ $\gamma \beta \frac{\partial G(y-R+D)}{\partial D} \int_{R}^{b} f(x) d x<0$
$K_{R D}=K_{D R}<0$ (by Young's theorem)
$K_{R c}=-\beta \int_{R}^{b} f(x) d x<0$
$K_{D c}=\beta \int_{j}^{D} h(\delta) d \delta>0$
$K_{R \gamma}=\frac{\partial G(y-R+D)}{\partial R} \beta\left[\int_{R}^{b}(x-R) f(x) d x+\int_{j}^{D}(D-\delta) h(\delta) d \delta\right]<0$
$K_{D \gamma}=\frac{\partial G(y-R+D)}{\partial D} \beta\left[\int_{R}^{b}(x-R) f(x) d x+\int_{j}^{D}(D-\delta) h(\delta) d \delta\right]>0$
The second-order conditions for cost minimization require, in addition to $K_{R R}>0$ and $K_{D D}>0$ (as noted above), that the determinant of the Hessian matrix is also positive:
$\left|\begin{array}{ll}K_{R R} & K_{R D} \\ K_{D R} & K_{D D}\end{array}\right|=\left(K_{R R}\right)\left(K_{D D}\right)-\left(K_{R D}\right)^{2}$. Note that satisfaction of the SOC requires $\left(K_{R R}\right)\left(K_{D D}\right)>$ $\left(K_{R D}\right)^{2}$. Further, note that $\lim _{\gamma \rightarrow 0} K_{R D}=0$, whereas $\lim _{\gamma \rightarrow 0} K_{R R}=\beta c f(R)>0$ and $\lim _{\gamma \rightarrow 0} K_{D D}=$ $\beta c h(D)>0$. Therefore the SOC will be satisfied in general so long as the endogenous audit function is not too steep, i.e. $\gamma$ is not too large.

Assuming the SOC is satisfied, then by Cramer's Rule
$\operatorname{sign} \frac{\partial R}{\partial c}=\operatorname{sign}\left|\begin{array}{ll}-K_{R c} & K_{R D} \\ -K_{D c} & K_{D D}\end{array}\right|=\left(-K_{R c}\right)\left(K_{D D}\right)-\left(K_{R D}\right)\left(-K_{D c}\right)$, and similarly
$\operatorname{sign} \frac{\partial D}{\partial c}=\operatorname{sign}\left|\begin{array}{ll}K_{R R} & -K_{R c} \\ K_{D R} & -K_{D c}\end{array}\right|=\left(K_{R R}\right)\left(-K_{D c}\right)-\left(-K_{R c}\right)\left(-K_{D R}\right)$.
First note that by Lemma 1 we have $K_{R c}=-K_{D c}$. Therefore $\frac{\partial R}{\partial C}>0$ iff $\left|K_{D D}\right|>\left|K_{R D}\right|$ and similarly $\frac{\partial D}{\partial c}<0$ iff $\left|K_{R R}\right|>\left|K_{R D}\right|$. Satisfaction of the SOC requires $\left(K_{R R}\right)\left(K_{D D}\right)>\left(K_{R D}\right)^{2}$, which implies either $\left|K_{D D}\right|>\left|K_{R D}\right|,\left|K_{R R}\right|>\left|K_{R D}\right|$, or both. Equivalently, either $\frac{\partial R}{\partial C}>0, \frac{\partial D}{\partial c}<$ 0 , or both for any interior minimum. Finally, Lemma 1 implies that $\operatorname{sign} \frac{\partial R}{\partial C}=-\operatorname{sign} \frac{\partial D}{\partial c}$ in order to hold $\int_{R}^{b} f(x) d x=\int_{j}^{D} h(\delta) d \delta$ for all values of $c$. Therefore it must be the case that both $\frac{\partial R}{\partial C}>0, \frac{\partial D}{\partial c}<0$ when the SOC is satisfied.

## Proof of PROPOSITION 2:

This closely parallels the proof of Proposition 1. Again applying Cramer's Rule we have:
$\operatorname{sign} \frac{\partial R}{\partial \gamma}=\operatorname{sign}\left|\begin{array}{ll}-K_{R \gamma} & K_{R D} \\ -K_{D \gamma} & K_{D D}\end{array}\right|=\left(-K_{R \gamma}\right)\left(K_{D D}\right)-\left(K_{R D}\right)\left(-K_{D \gamma}\right)$, and similarly
$\operatorname{sign} \frac{\partial D}{\partial \gamma}=\operatorname{sign}\left|\begin{array}{ll}K_{R R} & -K_{R \gamma} \\ K_{D R} & -K_{D \gamma}\end{array}\right|=\left(K_{R R}\right)\left(-K_{D \gamma}\right)-\left(-K_{R \gamma}\right)\left(K_{D R}\right)$.
Because $\frac{\partial G(y-R+D)}{\partial R}=-\frac{\partial G(y-R+D)}{\partial D}$ we have $K_{R \gamma}=-K_{D \gamma}$. Therefore $\frac{\partial R}{\partial \gamma}>0$ iff $\left|K_{D D}\right|>\left|K_{R D}\right|$ and similarly $\frac{\partial D}{\partial \gamma}<0$ iff $\left|K_{R R}\right|>\left|K_{R D}\right|$. Satisfaction of the SOC requires $\left(K_{R R}\right)\left(K_{D D}\right)>\left(K_{R D}\right)^{2}$, which implies either $\left|K_{D D}\right|>\left|K_{R D}\right|,\left|K_{R R}\right|>\left|K_{R D}\right|$, or both. Equivalently, either $\frac{\partial R}{\partial \gamma}>0, \frac{\partial D}{\partial \gamma}<$ 0 , or both. Finally, Lemma 1 implies that $\operatorname{sign} \frac{\partial R}{\partial \gamma}=-\operatorname{sign} \frac{\partial D}{\partial \gamma}$ in order to hold $\int_{R}^{b} f(x) d x=$ $\int_{j}^{D} h(\delta) d \delta$ for all values of $\gamma$. Therefore it must be the case that both $\frac{\partial R}{\partial \gamma}>0, \frac{\partial D}{\partial \gamma}<0$ when the SOC is satisfied.

## Proof of PROPOSITION 3:

This follows directly from Lemma 1. If income and deductions are identically and symmetrically distributed around their respective means, so $f(x-\bar{x})=h(\delta-\bar{\delta})$ for all $x, \delta$, then $\int_{R^{*}}^{b} f(x) d x=\int_{j}^{D^{*}} h(\delta) d \delta$ implies that underreporting of income is equal to over-reporting of deductions: $\bar{x}-R^{*}=D^{*}-\bar{\delta}$.

## Proof of PROPOSITION 4:

From Proposition 3, when income and deductions are identically and symmetrically distributed around their respective means, so $f(x-\bar{x})=h(\delta-\bar{\delta})$ for all $x, \delta$, then reported taxable
income can be modeled as a single line because the two reporting decisions respond oppositely but with identical magnitudes to any change in the reporting environment. For a single line item, income, drawn from the density $f(x)$ on $[a, b]$ the optimal report, $R^{*}$, is defined by
[A1] $\quad K_{R}=t+\gamma \frac{\partial G\left(y-R^{*}\right)}{\partial R} \beta\left[\int_{R^{*}}^{b}\left(x-R^{*}\right) f(x) d x\right]-\beta\left[c+\gamma G\left(y-R^{*}\right)\right] \int_{R^{*}}^{b} f(x) d x=0$.
Now suppose the density from which income is drawn shifts to $\tilde{f}(x)$ on $[a+\mu, b+\mu]$ such that $\tilde{f}(x)=f(x-\mu)$ for all $x$. The FOC would no longer be satisfied at $R^{*}$. Specifically, the derivative of the objective function evaluated at $R^{*}$ is:
[A2] $\quad K_{R}=t+\gamma \frac{\partial G\left(y-R^{*}\right)}{\partial R} \beta\left[\int_{R^{*}}^{b+\mu}\left(x-R^{*}\right) \tilde{f}(x) d x\right]-\beta\left[c+\gamma G\left(y-R^{*}\right)\right] \int_{R^{*}}^{b+\mu} \tilde{f}(x) d x<0$.
To see that the sign of this expression is negative, note that $\left[\int_{R^{*}}^{b+\mu}\left(x-R^{*}\right) \tilde{f}(x) d x\right]>$ $\left[\int_{R^{*}}^{b}\left(x-R^{*}\right) f(x) d x\right]$ while these expression are both multiplied by $\gamma \frac{\partial G\left(y-R^{*}\right)}{\partial R} \beta<0$ in [A1] and [A2] respectively. Similarly $\int_{R^{*}}^{b+\mu} \tilde{f}(x) d x>\int_{R^{*}}^{b} f(x) d x$, and both terms are multiplied by $-\beta\left[c+\gamma G\left(y-R^{*}\right)\right]$ in [A1] and [A2] respectively. Therefore since the expression in [A1] equals 0 at $R^{*}$ by definition, and the expression in [A2] is strictly less than 0 . This shows that the optimal reported income when income is drawn from $\tilde{f}(x)$ is greater than $R^{*}$.

Now consider whether the optimal reported income is $R^{*}+\mu$. The derivative of the objective function evaluated at $R^{*}+\mu$ is:
[A3] $\quad K_{R}=t+\gamma \frac{\partial G\left(y-\left(R^{*}+\mu\right)\right)}{\partial R} \beta\left[\int_{R^{*}+\mu}^{b+\mu}\left(x-\left(R^{*}+\mu\right)\right) \tilde{f}(x) d x\right]-\beta[c+\gamma G(y-$

$$
\left.\left.\left(R^{*}+\mu\right)\right)\right] \int_{R^{*}+\mu}^{b+\mu} \tilde{f}(x) d x>0
$$

To see that the sign of this expression is positive, note that $\left[\int_{R^{*}}^{b}\left(x-R^{*}\right) f(x) d x\right]=$ $\left[\int_{R^{*}+\mu}^{b+\mu}\left(x-\left(R^{*}+\mu\right)\right) \tilde{f}(x) d x\right]$ and $\int_{R^{*}}^{b} f(x) d x=\int_{R^{*}+\mu}^{b+\mu} \tilde{f}(x) d x$, but $G\left(y-R^{*}\right)>$
$G\left(y-\left(R^{*}+\mu\right)\right)$ and $\frac{\partial G\left(y-R^{*}\right)}{\partial R} \geq \frac{\partial G\left(y-\left(R^{*}+\mu\right)\right)}{\partial R}$. These differences imply that since equation [A1] equals 0 at $R^{*}$ by definition, and equation [A3] is greater than zero at $R^{*}$. This shows that the optimal reported income when income is drawn from $\tilde{f}(x)$ is less than $R^{*}+\mu$. Having shown that the optimal report given income is drawn from $\tilde{f}(x)$ is $\tilde{R}$ such that $R^{*}<\tilde{R}<\tilde{R}+\mu$, note that this implies the taxpayer underreports by more when income is drawn from $\tilde{f}(x)$ relative to the mean of $\tilde{f}(x)$, which is of course greater than the mean of $f(x)$ by $\mu$.

Having established the result for the single line-item case, we can apply it to the two line item case when income and deductions are identically and symmetrically distributed around their respective means (with the mean of the distribution of income increasing by $\mu$ ). As before, a shift in the location of the density of income by $\mu$ increases reported income, but by less than $\mu$, so underreporting of income increases relative to the new mean. By Proposition 3, since underreporting of income equals over-reporting of deductions relative to their respective means, the change in the density from which income is drawn results in an increase in reported deductions.

## Proof of PROPOSITION 5:

As discussed in the article, here we simplify the analysis by assuming only one line item, income, and assuming the distribution from which income is drawn, represented by the density $f(x)$ on $[a, b]$ is symmetric around its mean $\bar{x}=\frac{a+b}{2}$. The optimization problem can then be stated (analogously to [1]):

$$
\min _{R, D} t(R)+[c+\gamma G(y-R)] \beta\left[\int_{R}^{b}(x-R) f(x) d x\right]
$$

The first order condition in this case is
$K_{R}=t+\gamma \frac{\partial G(y-R)}{\partial R} \beta\left[\int_{R}^{b}(x-R) f(x) d x\right]-\beta[c+\gamma G(y-R)] \int_{R}^{b} f(x) d x=0$.
As discussed in the article, if the audit mechanism is not endogenous (simply a random audit probability $c$, with $\gamma=0$ ), then the condition reduces to $t-\beta c \int_{R}^{b} f(x) d x=0$, which can be stated as $\int_{R^{*}}^{b} f(x) d x=\frac{t}{c \beta}$. Note that underreporting, $R<\bar{x}$ is optimal when $\frac{t}{c \beta}>\frac{1}{2}$. Consider holding the parameters $t, c, \gamma$ constant such that $\frac{t}{c \beta}>\frac{1}{2}$ but reducing income uncertainty such that income is now drawn from density $\hat{f}(x)$ on $[a+\mu, b-\mu]$ where $\hat{f}(x)$ and $f(x)$ belong to the same location-scale family, have a common mean, but differ in variance. When the income uncertainty is reduced in this manner then optimal reported income $R$ increases to maintain the equality. That is, given income is drawn from $\hat{f}(x)$, then $R^{*}$ is no longer optimal:
$\int_{R^{*}}^{b-\mu} \hat{f}(x) d x>\frac{t}{c \beta}$. The report $\hat{R}$ that satisfies $\int_{\hat{R}}^{b-\mu} \hat{f}(x) d x=\int_{R^{*}}^{b} f(x) d x=\frac{t}{c \beta}$ is such that $\hat{R}>$ $R^{*}$.

Now consider optimal income reporting with the endogenous audit mechanism, $\gamma>0$. Suppose that the parameters $t, c, \gamma$ are such that the report that satisfies the FOC for optimality given income is drawn from the initial density $f(x)$ on $[a, b]$ is $\bar{x}$, i.e. $R^{*}=\bar{x}$. This is the case when $t, c, \gamma$ are such that

$$
\begin{equation*}
t+\gamma \frac{\partial G(y-\bar{x})}{\partial R} \beta\left[\int_{\bar{x}}^{b}(x-\bar{x}) f(x) d x\right]-\beta[c+\gamma G(y-\bar{x})] \int_{\bar{x}}^{b} f(x) d x=0 \tag{A4}
\end{equation*}
$$

Note that because $\gamma>0$ it is not the case that $\frac{t}{c \beta}=\frac{1}{2}$ for $R^{*}=\bar{x}$. Instead, $R^{*}=\bar{x}$ must correspond to $\frac{t}{c \beta}>\frac{1}{2}$ because at $\frac{t}{c \beta}=\frac{1}{2}$ if $R=\bar{x}$ then $t-\beta c \int_{\bar{x}}^{b} f(x) d x=0$. Therefore if $\frac{t}{c \beta}=\frac{1}{2}$ then $t+\gamma \frac{\partial G(y-\bar{x})}{\partial R} \beta\left[\int_{\bar{x}}^{b}(x-\bar{x}) f(x) d x\right]-\beta[c+\gamma G(y-\bar{x})] \int_{\bar{x}}^{b} f(x) d x<0$. Thus it must be the case that $R^{*}=\bar{x}$ when $\frac{t}{c \beta}>\frac{1}{2}$.

Now holding $t, c, \gamma$ constant with $\frac{t}{c \beta}>\frac{1}{2}$ such that $R^{*}=\bar{x}$, consider the change in the optimal report when income uncertainty is reduced, now being drawn from density $\hat{f}(x)$ as defined earlier. It must be the case that the FOC is not satisfied at $R=\bar{x}$. Specifically:
[A5] $t+\gamma \frac{\partial G(y-\bar{x})}{\partial R} \beta\left[\int_{\bar{x}}^{b-\mu}(x-\bar{x}) \hat{f}(x) d x\right]-\beta[c+\gamma G(y-\bar{x})] \int_{\bar{x}}^{b-\mu} \hat{f}(x) d x>0$.
To see why this is the case, note that because $f(x)$ and $\hat{f}(x)$ have a common mean and median, $\bar{x}, \int_{\bar{x}}^{b} f(x) d x=\int_{\bar{x}}^{b-\mu} \hat{f}(x) d x=\frac{1}{2}$. Therefore the only difference between the expressions in [A4] and [A5] comes from the fact that $\gamma \frac{\partial G(y-\bar{x})}{\partial R} \beta\left[\int_{\bar{x}}^{b}(x-\bar{x}) f(x) d x\right]<$ $\gamma \frac{\partial G(y-\bar{x})}{\partial R} \beta\left[\int_{\bar{x}}^{b-\mu}(x-\bar{x}) \hat{f}(x) d x\right]$. Both terms are negative, but the former is larger in magnitude because the greater variance of $f(x)$ relative to $\hat{f}(x)$ yields higher possible penalties when reporting $R=\bar{x}$.

Since the expression in [A4] equals 0 at $\bar{x}$ by construction, in must be the case that the expression in [A5] is strictly greater than 0 . That is, at $R=\bar{x}$ the objective function is increasing in reported income when income is drawn from $\hat{f}(x)$, so it is optimal to report $\hat{R}<\bar{x}$. In this case, the effect of reducing uncertainty is to reduce reported income. Finally, suppose $t, c, \gamma$ were negligibly different such that if income were drawn from $f(x)$ then $R^{*}=\bar{x}-\varepsilon$, where $\varepsilon>$ 0 is arbitrarily small. By continuity of all terms in the objective function the preceding logic would hold with $\hat{R}<R^{*}$. This establishes that under the endogenous audit mechanism, differently from random audits, for some conditions of underreporting a reduction in income uncertainty will reduce reported income.

## Auditing a single line item (partial audit)

As stated in the article, in this case the taxpayer's objective is
[A6] $\min _{R, D} t(R-D)+s(\gamma(y-R+D)+c) \beta\left[\theta \frac{(b-R)^{2}}{2(b-a)}+(1-\theta) \frac{(D-j)^{2}}{2(k-j)}\right]$
The first order conditions identifying cost-minimizing reporting and deductions are respectively
[A7] $t-\beta s \gamma\left[\theta \frac{(b-R)^{2}}{2(b-a)}+(1-\theta) \frac{(D-j)^{2}}{2(k-j)}\right]-s c \beta \theta\left(\frac{b-R}{b-a}\right)-\frac{s \gamma \beta \alpha(y-R+D)(b-R)}{(b-a)}=0$, and
[A8] $t-\beta s \gamma\left[\theta \frac{(b-R)^{2}}{2(b-a)}+(1-\theta) \frac{(D-j)^{2}}{2(k-j)}\right]-s c \beta(1-\theta)\left(\frac{D-j}{k-j}\right)-\frac{s \gamma \beta(1-\alpha)(y-R+D)(D-j)}{(k-j)}=0$.
Consider the case when $s=2$ and $\theta=\frac{1}{2}$, corresponding to the notion that the auditor can audit twice as many taxpayers with the same resources if only one of the two lines is audited. Note that these terms then cancel and equations [A6], [A7], and [A8] in that case reduce to equations [4], [5], and [6]. That is, if the audit probability doubles for any given choices, $R$ and $D$, but only one line-item is audited rather than both and the taxpayer believes either is equally likely to be audited, then the optimal reported income and deductions are identical. But suppose instead the taxpayer may acquire some informative signal regarding which line is more likely to be audited. How does behavior compare when $s=2$ and $\theta \gtrless \frac{1}{2}$ to that when $s=2$ and $\theta=\frac{1}{2}$ ? To answer this question it is helpful to begin by assuming the variance of the income and deduction distributions is identical, $(b-a)=(k-j)$, and recall that in the case when $s=2$ and $\theta=\frac{1}{2}$ optimal behavior is symmetric in that the underreporting of income (relative to the expected value) is equal to over-reporting of deductions: $E[x]-R^{*}=D^{*}-E[\delta]$. This symmetry implies that the expected penalty if either line is audited is identical: $\frac{(b-R)^{2}}{2(b-a)}=\frac{(D-j)^{2}}{2(k-j)}$.

Let $\hat{R}, \widehat{D}$ be the optimum when $s=2$ and $\theta=\frac{1}{2}$. Then for $\theta>\frac{1}{2}$ it must be the case that the left-hand-side expression in equation [A7] is negative at $\widehat{R}, \widehat{D}$, and similarly the left-hand-side
expression in equation [A8] is also negative at $\widehat{R}, \widehat{D}$. That implies that the optimum with $\theta>\frac{1}{2}$ is to report $R>\hat{R}$ and $D>\widehat{D}$. That is, when it is more likely income will be audited, it is optimal to underreport income by less and over-report deductions by more than under otherwise equal conditions when either line is equally likely to be audited. Similarly, if $<\frac{1}{2}$, then the optimum is to report $R<\hat{R}$ and $D<\widehat{D}$, that is, to underreport income by more and over-report deductions by less. If the variance of income exceeds the variance of deductions, then the same results hold for $\theta>\frac{1}{2}$, because at $\widehat{R}, \widehat{D}$ the expected penalty is higher if income is audited than if deductions are audited, so if income becomes more likely to be audited the incentive to report more income and more deductions are both strengthened. However, if the variance of deductions exceeds the variance of income the results do not hold unambiguously.

## Appendix B: Experiment Screenshots and Instructions

Figure B1. Subject screen for risk elicitation task

| scenario | Lottery | certain amount | Your choice |
| :---: | :---: | :---: | :---: |
|  | Choice A | Choice B |  |
| 1 | $10 \%$ chance of $\$ 4$ and $90 \%$ chance of $\$ 0$ | \$2 for sure | Choice A $\subset \subset$ Choice B |
| 2 | $20 \%$ chance of \$4 and $80 \%$ chance of \$0 | \$2 for sure | Choice A $\ll$ choice B |
| 3 | $30 \%$ chance of $\$ 4$ and $70 \%$ chance of \$0 | \$2 for sure | Choice A $\subset \subset$ Choice B |
| 4 | $40 \%$ chance of $\$ 4$ and $60 \%$ chance of $\$ 0$ | \$2 for sure | Choice A $\subset \subset$ Choice B |
| 5 | $50 \%$ chance of $\$ 4$ and $50 \%$ chance of $\$ 0$ | \$2 for sure | Choice $A \subset C$ choice B |
| 6 | 60\% chance of 54 and | \$2 for sure | Choice $\mathrm{A} \subset$ C Choice B |
| 7 | $70 \%$ chance of $\$ 4$ and $30 \%$ chance of $\$ 0$ | \$2 for sure | Choice A $\subset \subset$ Choice B |
| 8 | $80 \%$ chance of $\$ 4$ and $20 \%$ chance of $\$ 0$ | \$2 for sure | Choice A $\subset \subset$ Choice B |
| 9 | $90 \%$ chance of $\$ 4$ and $10 \%$ chance of \$0 | \$2 for sure | Choice A $\subset \subset$ Choice B |
| 10 | $100 \%$ chance of $\$ 4$ and $0 \%$ chance of \$0 | \$2 for sure | Choice A $\ll$ Choice B |

On the left are 10 scenarios which allow you to choose between receiving On the leff are 10 scenario
$\$ 2.00$ or playing a lottery.

Please choose either $A$ or $B$ for each scenario.
At the end of the experiment the computer will randomly select ONE of these At the end of th
10 scenarios.

If you selected the lottery, choice A, for the randomly selected scenario, the computer will determine the outcome based on the chances associated with computer will determin

Otherwise you will receive $\$ 2.00$.

## Figure B2. Subject screen for income group determination task



Figure B3. Subject tax reporting screen, Treatment 8 (waiting for services)


Figure B4. Subject tax reporting screen, Treatment 8 (services displayed)


Figure B5. Audit determination screen (animated)


Figure B6. Subject screen for end of round summary, Treatment 8


## Experiment Instructions (Treatment 8)

You are about to participate in an experiment in economic decision making. Please follow the instructions carefully, as the amount of money you earn in the experiment will depend on your decisions. At the end of today's session, you will be paid your earnings privately and in cash. Please do not communicate with other participants during the experiment unless instructed. Importantly, please refrain from verbally reacting to events that occur during the experiment.

Today's experiment will involve several decision "rounds". You will not know the number of rounds until the end of the experiment. The rounds are arranged into multiple series. After all decision rounds are finished, we will ask you to complete a questionnaire.

Aside from decisions in "training" rounds, each decision impacts your earnings, which means that it is very important to consider each decision carefully prior to making it. Each decision round is separate from the other rounds, in the sense that the decisions you make in one round will not affect the outcome or earnings of any other round. All money amounts are denominated in lab dollars, and will be exchanged at a rate of 900 lab dollars to US $\$ 1$ at the end of the experiment.

As we read the instructions, we will work through one training round to help our understanding of the procedures. Here is the set up. In each round, you fill out and file a tax form. Then, there is a process for selecting whether your tax form is audited. Last, a summary of your earnings for the decision round, including the outcome of the audit process, is provided.

## Tax Reporting

On the tax form, located on the right side of the screen, you will report income and deduction amounts.

## Reporting your income

You will not know your income for sure. Instead, on the left side of your screen, you will be shown a range of possible income amounts. Any number in this range has an equal chance of being your actual income. On the tax form, you can report any amount within the income range.

The more you report in income, the higher your tax payment will be.

## Reporting your deduction

You will not know the amount you are allowed to claim in deductions. Instead, on the left side of your screen, you will be shown a range of possible deduction amounts. Any number in this range has an equal chance of being your actual deduction. On the tax form, you can report any amount within the deduction range.

The more you report in deductions, the lower your tax payment will be.

## Your tax payment

Your tax payment is determined by multiplying your taxable income by a tax rate of $50 \%$ :
Tax Payment $=\mathbf{5 0 \%} \times($ your reported income - your reported deduction)

On the tax form, after you choose income and deduction amounts to report, click on the "Do the Math" button. When you do this, you will see relevant tax calculations appear below the tax form, including the Tax Payment. At this time, for practice, please enter income and deduction amounts on the tax form and click the "Do the Math" button.

## Audit Procedures

There is a chance that the tax agency will audit your tax form.
The chance you are audited depends on the tax payment you report on the tax form. The audit chance decreases as you increase your tax payment.

On the left part of the tax reporting screen you will see a table that shows your audit chance based on different reported tax payment amounts. You will also be shown the rate of change: the increase in tax payment associated with a $1 \%$ decrease in the audit chance. You will notice that there is a $10 \%$ range of possible audit chances for each tax payment amount. Each audit chance within this range will be equally likely.

If you are selected for audit, EITHER your reported income OR your reported deduction will be checked for unpaid taxes (NOT both). You will not know in advance which amount will be checked. There is an equal chance that either amount will be checked, if you are audited.

If you are not audited, however, no unpaid taxes will be found.

Any taxes you overpaid will not be refunded to you. In this sense, the audit process can never increase your earnings.

## Unpaid taxes

If audited, you will have unpaid taxes if you reported too little in income or too much in deductions. Unpaid taxes are calculated as the difference between your actual and reported amounts multiplied by the tax rate. If you underreported your taxes, only the unpaid taxes on the item selected by the audit (income or deductions) will be found. Any unpaid taxes found must be paid back.

## Penalty

If you have unpaid taxes, a penalty of $300 \%$ will be assessed. What this means is that, if you are audited, for every lab dollar in unpaid taxes you will have to pay back the 1 dollar you owed and in addition pay 3 lab dollars in penalties.

## Tax Information Service

On the middle of your screen, towards the top, you may be provided tax information from a thirdparty; i.e. this information is not provided by, nor is it known to, the tax agency. In particular, you may be provided better (i.e. more precise) information about your actual income and/or deduction. Please know that the third-party information is accurate. For example, if the service provides you with a range of possible income (deduction) amounts, your actual amount is contained within the interval. Any amount within the interval has an equal chance of being your actual income (deduction).

If provided with better information about your income or deductions, you do not have to report an amount within the specified range(s).

## Tax Information Service Guarantee

On the middle of your screen, towards the bottom, you may see that a third party has offered you a guarantee.

When available, if you follow the requirement for a particular reporting item (income or deduction), any audit penalties that result for that reporting item will be paid by the Service.

You will still be responsible for any unpaid taxes.

## Audit Information Service

On the middle of your screen you will see audit information provided by a third-party service.

The service has provided you with better information about which item (income or deductions) will be checked for unpaid taxes, in the event you are selected for audit.

Also, the service has provided you with better information about your chance of audit for different reported tax payment amounts.

Please know that the third-party information is accurate.

## Filing the tax form

When you are ready to record the particular income and deduction amounts you wish to report, you must first click the "Do the Math" button. Once you see the choices recorded, click the "FILE TAXES" button.

There is a timer on the tax reporting screen. If you do not file the tax form before time runs out, this will be treated the same as if you submitted a form that reported 0 in income and 0 in deductions. In addition, your tax form will automatically be audited. In other words, it is NOT in your best interest to let the tax reporting screen time out!

After you file the tax form, you will see an audit screen. While you are on this screen the tax agency is determining whether to audit your tax form, using the audit chance associated with your particular tax payment. At this time, for practice, please click the "FILE TAXES" button.

## Round Summary

After the tax reporting decision, three things can happen: (1) you are not audited; (2) you are audited but did not underreport your taxes for the item selected for audit; or (3) you are audited and you did underreport your taxes for the item selected for audit. Your earnings are, of course, the same for the first two scenarios. The computer will calculate earnings for you, but it is important that you understand how your earnings are determined. The relevant earnings calculations are given below.

Your earnings (you are not audited OR you are audited but did not underreport taxes)
In both cases, there is no adjustment to your earnings based on the audit process. Your earnings for the round are equal to your actual income minus your tax payment.

|  | Income | Your actual income (not your reported income) |
| :--- | :--- | :--- |
| - | Tax Payment | (Reported Income - Reported Deduction) $\times 50 \%$ |

Your earnings (You are audited and you underreported your taxes for the selected item) In this case, all unpaid taxes are found for the item selected for audit, and a penalty is assessed. Income Your actual income (not your reported income)

- Tax Payment
(Reported Income - Reported Deduction) $\times 50 \%$
- Unpaid Taxes

Difference between what you owed and what you paid

- Penalties* (Unpaid Taxes) $\times 300 \%$
$=\quad$ Earnings
*When available, if you follow the requirement for a particular item (income or deduction), any audit penalties for that item will be paid by the Service.

At this time, please click the "Finished" button on the Round Summary screen.

## Examples

Before we continue, let us work through some examples to make sure we all understand some basic concepts. You will need to refer to the "Your Audit Chance" information on the computer screen for the first two examples. Please ignore information provided by information services at this time.

Example 1. Suppose you report 2000 in income and 1000 in deductions.
$\begin{array}{lllll}\text { What is your tax payment? } & 500 & 1000 & 1500 & 2000\end{array}$
What is your audit chance? $15 \%$ to $25 \% \quad 35 \%$ to $45 \% \quad 43 \%$ to $53 \% \quad 55 \%$ to $65 \%$

Example 2. Suppose you report 2000 in income and 800 in deductions.

| What is your tax payment? | 400 | 600 | 1200 | 2000 |
| :--- | :--- | :--- | :--- | :--- |
| What is your audit chance? | $35 \%$ to $45 \%$ | $45 \%$ to $55 \%$ | $51 \%$ to $61 \%$ | $55 \%$ to $65 \%$ |

Example 3. Suppose your reported income is audited. You reported 2200 in income and your actual income is 2300.

| What unpaid taxes would be found? | 0 | 50 | 100 | 200 |
| :--- | :--- | :--- | :--- | :--- |
| What penalties would you pay? | 0 | 50 | 150 | 300 |

Example 4. Suppose that your reported deduction is audited. You reported 250 in deductions and your actual allowed deduction is 400 .
What unpaid taxes would be found? $0 \quad 75 \quad 150 \quad 300$
$\begin{array}{llllll}\text { What penalties would you pay? } & 0 & 150 & 300 & 450\end{array}$

## Second training round

We will now continue on to a second training round. As with the first, your decisions in the second training round will not affect your earnings. After the training round you will have a final opportunity to ask questions. At this time, please fill out and file the tax form for the second training round.

## Beginning the experiment

Going forward, before we begin each series of paid rounds, you will first be asked to complete an earnings task. Your score for the task, relative to others in the room, will determine whether you are in the high, medium or low income group for the series of rounds. Roughly one-third of the players will be placed in each group. From time to time a new series will begin and you will be asked to complete a new earnings task.

At the beginning of a new series some of the tax settings will change. When a new series begins please pay close attention to any information that has changed prior to making any decision.

Before we proceed to the paid decision rounds, are there any questions?

## Appendix C: Additional Econometric Models

Table C1. Tax Reporting Models: Full Audit Treatments (Without Additional Controls)

|  | Model 1: <br> Reported Taxable <br> Income | Model 2: <br> Reported Income | Model 3: <br> Reported <br> Deduction |
| :---: | :---: | :---: | :---: |
| Liability Service Effects |  |  |  |
| Income and Deduction Info | $-75.01^{* *}$ (33.06) | $-50.52^{* *}$ (20.26) | 24.49 (20.53) |
| Income Information Only | $-71.27^{* *}$ (31.40) | $-38.58^{* *}$ (19.51) | 32.69 (20.27) |
| Deduction Information Only | $-70.86^{* *}$ (31.13) | $-64.00^{* *}(19.80)$ | 6.87 (20.10) |
| Expected Income Change | $0.54 * * * 0.03)$ | 0.69 ** (0.02) | $0.15{ }^{* *}$ (0.02) |
| Expected Deduction Change | $-0.51^{* *}(0.03)$ | 0.13 ** (0.02) | 0.63 ** (0.02) |
| Liability Service Guarantee |  |  |  |
| Guarantee Available | $100.01^{* *}$ (43.00) | $66.51{ }^{* *}$ (24.79) | -33.49 (25.98) |
| Audit Service Effects |  |  |  |
| Audit Service | -8.47 (39.01) | -10.20 (27.28) | -1.73 (25.65) |
| Audit Intercept Increase | $10.65^{* *}$ (4.57) | 3.48 (3.36) | $-7.17^{* *}$ (3.01) |
| Audit Intercept Decrease | $-21.74 * *$ (4.49) | $-9.39^{* *}(3.32)$ | $12.35{ }^{* *}$ (2.78) |
| Service Interaction Effects |  |  |  |
| Audit Service $\times$ Liability Service | 86.36** (41.64) | 61.86 ** (27.63) | -24.50 (27.26) |
| Audit Service $\times$ Guarantee Available | $-116.89^{*}(60.76)$ | $-81.59^{* *}(36.46)$ | 35.30 (35.26) |
| Other Experiment Treatments |  |  |  |
| High Audit Slope | 60.71 ** (13.69) | $33.61{ }^{* *}$ (9.54) | $-27.09^{* *}(8.51)$ |
| High Income Group | 919.92** (29.72) | 970.71*** (19.68) | 50.79 ** (17.56) |
| Middle Income Group | 444.50 ** (26.59) | $469.63 * *$ (17.06) | 25.14 (15.36) |
| Constant | 969.84** (33.11) | $1501.53^{* *}$ (21.83) | 531.68** (20.88) |
| Number of Observations | 8427 | 8427 | 8427 |
| F | $130.86{ }^{* *}$ | $270.79^{* *}$ | 107.46** |
| $R^{2}$ | 0.503 | 0.691 | 0.202 |

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

Table C2. Tax Reporting Models: Partial Audit Treatments (Without Additional Controls)

|  | Model 4: <br> Reported Taxable <br> Income | Model 5: <br> Reported Income | Model 6: <br> Reported <br> Deduction |
| :--- | :---: | :---: | :---: |
| Liability Service Effects |  |  |  |
| Income and Deduction Info | $-17.48(17.32)$ | $-28.05^{* *}(12.14)$ | $-10.56(11.48)$ |
| Income Information Only | $-22.23(17.54)$ | $-35.32^{* *}(12.42)$ | $-13.09(12.54)$ |
| Deduction Information Only | $-5.98(18.43)$ | $-22.77^{*}(12.85)$ | $-16.79(12.11)$ |
| Expected Income Change | $0.63^{* *}(0.02)$ | $0.74^{* *}(0.02)$ | $0.12^{* *}(0.01)$ |
| Expected Deduction Change | $-0.51^{* *}(0.02)$ | $0.15^{* *}(0.02)$ | $0.66^{* *}(0.02)$ |
| Liability Service Guarantee |  |  |  |
| Guarantee Available | $-4.03(25.63)$ | $8.74(15.99)$ | $12.76(15.50)$ |
| Audit Service Effects |  |  |  |
| Audit Targets Income | $-56.00^{* *}(27.73)$ | $33.65^{* *}(17.62)$ | $89.65^{* *}(20.34)$ |
| Audit Targets Deduction | $-53.70^{*}(27.98)$ | $-96.34^{* *}(21.33)$ | $-42.64^{* *}(16.42)$ |
| Audit Intercept Increase | $-3.77(3.41)$ | $-4.84^{*}(2.75)$ | $-1.07(2.61)$ |
| Audit Intercept Decrease | $-5.70(3.79)$ | $-0.66(3.01)$ | $5.04^{*}(3.04)$ |
| Other Experiment Treatments |  |  |  |
| High Audit Slope | $49.74^{* *}(10.94)$ | $21.92^{* *}(8.15)$ | $-27.82^{* *}(7.49)$ |
| High Income Group | $890.73^{* *}(26.12)$ | $930.37^{* *}(16.27)$ | $39.64^{* *}(16.23)$ |
| Middle Income Group | $397.92^{* *}(22.39)$ | $443.53^{* *}(14.89)$ | $45.61^{* *}(14.46)$ |
| Constant | $1066.30^{* *}(24.90)$ | $1576.81^{* *}(17.33)$ | $510.51^{* *}(16.73)$ |
| Number of Observations | 8559 | 8559 | 80.859 |
| F | $154.80^{* *}$ | $365.75^{* *}$ | $134.36^{* *}$ |
| $R^{2}$ | 0.557 | 0.700 | 0.243 |

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

Table C3. Tax Reporting Models: Full Audit Treatments (Two-Limit Tobit)

|  | Model 1: Reported Taxable Income | Model 2: <br> Reported Income | Model 3: <br> Reported <br> Deduction |
| :---: | :---: | :---: | :---: |
| Liability Service Effects |  |  |  |
| Income and Deduction Info | $-76.79^{* *}$ (34.30) | $-51.75{ }^{* *}$ (20.79) | 24.33 (21.26) |
| Income Information Only | $-74.16^{* *}$ (32.46) | $-43.14 * *(19.85)$ | 36.19* (20.93) |
| Deduction Information Only | $-74.12^{* *}$ (32.36) | $-69.73{ }^{* *}$ (20.08) | 6.50 (20.85) |
| Expected Income Change | $0.54^{* *}$ (0.03) | $0.61{ }^{* *}$ (0.02) | $0.14 * * * 0.02)$ |
| Expected Deduction Change | -0.50 ** (0.03) | $0.12 * *$ (0.02) | $0.57{ }^{* *}$ (0.02) |
| Liability Service Guarantee |  |  |  |
| Guarantee Available | $106.04{ }^{* *}$ (43.94) | $68.39^{* *}$ (25.47) | -36.55 (26.84) |
| Audit Service Effects |  |  |  |
| Audit Service | -11.63 (40.09) | -12.31 (28.21) | -1.83 (25.92) |
| Audit Intercept Increase | $10.45{ }^{* *}$ (4.76) | 2.99 (3.47) | -6.43 ** (2.99) |
| Audit Intercept Decrease | $-22.08^{* *}$ (4.62) | $-9.61{ }^{* *}$ (3.45) | $12.79^{* *}$ (2.84) |
| Service Interaction Effects |  |  |  |
| Audit Service $\times$ Liability Service | 93.52** (43.27) | $70.37 * * * 28.51)$ | -29.34 (28.20) |
| Audit Service $\times$ Guarantee Available | $-122.84^{* *}$ (62.44) | $-85.55^{* *}$ (38.04) | 38.32 (36.26) |
| Other Experiment Treatments |  |  |  |
| High Audit Slope | $64.33^{* *}$ (14.13) | $36.76{ }^{* *}$ (9.72) | $-29.10^{* *}$ (8.66) |
| High Income Group | 919.82** (30.79) | 969.89** (20.60) | $47.04 * *$ (17.91) |
| Middle Income Group | $443.49^{* *}$ (27.50) | $466.21{ }^{* *}$ (17.92) | 22.93 (15.69) |
| Constant | 961.91** (35.02) | 1501.61** (28.03) | 549.94** (24.61) |
| Number of Observations | 8427 | 8427 | 8427 |
| F | $57.88^{* *}$ | 86.50** | 102.31** |
| Log-likelihood | -59994.34 | -49722.22 | -52271.80 |

Notes: Table entries are marginal effects, treating limit observations as corner solutions. * and ${ }^{* *}$ denote estimates that are statistically different from zero at the $10 \%$ and $5 \%$ significance levels, respectively. Standard errors (in parentheses) are clustered at the participant-level.

Table C4. Tax Reporting Models: Partial Audit Treatments (Two-Limit Tobit)

|  | Model 4: <br> Reported Taxable <br> Income | Model 5: <br> Reported Income | Model 6: <br> Reported <br> Deduction |
| :--- | :---: | :---: | :---: |
| Liability Service Effects |  |  |  |
| Income and Deduction Info | $-17.14(17.55)$ | $-29.37^{* *}(12.50)$ | $-11.93(11.55)$ |
| Income Information Only | $-22.00(17.78)$ | $-35.34^{* *}(12.91)$ | $-11.31(12.45)$ |
| Deduction Information Only | $-5.90(18.66)$ | $-25.83^{* *}(13.22)$ | $-16.81(12.28)$ |
| Expected Income Change | $0.62^{* *}(0.02)$ | $0.66^{* *}(0.02)$ | $0.11^{* *}(0.01)$ |
| Expected Deduction Change | $-0.51^{* *}(0.02)$ | $0.14^{* *}(0.02)$ | $0.62^{* *}(0.02)$ |
| Liability Service Guarantee | $-3.40(25.68)$ | $10.13(16.66)$ | $11.16(15.85)$ |
| Guarantee Available |  |  |  |
| Audit Service Effects | $-54.95^{* *}(27.77)$ | $30.20^{*}(17.86)$ | $100.02^{* *}(21.37)$ |
| Audit Targets Income | $-53.01^{*}(28.09)$ | $-107.73^{* *}(22.50)$ | $-38.94^{* *}(16.22)$ |
| Audit Targets Deduction | $-3.71(3.42)$ | $-4.98^{*}(2.84)$ | $-1.42(2.64)$ |
| Audit Intercept Increase | $-5.70(3.80)$ | $-0.89(3.13)$ | $4.77(3.09)$ |
| Audit Intercept Decrease |  |  |  |
| Other Experiment Treatments | $49.81^{* *}(11.08)$ | $22.57^{* *}(8.39)$ | $-26.69^{* * *}(87.45)$ |
| High Audit Slope | $891.26^{* *}(26.26)$ | $928.08^{* *}(16.58)$ | $39.24^{* *}(16.59)$ |
| High Income Group | $398.08^{* *}(22.44)$ | $442.56^{* *}(15.20)$ | $43.51^{* *}(14.58)$ |
| Middle Income Group | $1065.12^{* *}(25.38)$ | $1595.46^{* *}(21.76)$ | $516.88^{* *}(19.14)$ |
| Constant | 8559 | $85.50^{* *}$ | $150.72^{* *}$ |
| Number of Observations | -61454.01 | -51052.27 | -53775.68 |
| F |  |  | 8559 |
| Log-likelihood |  |  | $127.23^{* *}$ |

[^0]
[^0]:    Notes: Table entries are marginal effects, treating limit observations as corner solutions. ${ }^{*}$ and ${ }^{* *}$ denote estimates that are statistically different from zero at the $10 \%$ and $5 \%$ significance levels, respectively. Standard errors (in parentheses) are clustered at the participant-level.

