An Econometric Analysis of Income Tax Evasion and Its Detection

CORRECTION

There is a mistake in the original paper in Equation (4) and the following Equation (5). These equations have in common an integral expression that refers to the likelihood of undetected evasion. A Jacobian term is missing from the integrand of this expression. Brian Erard and I discovered this a few years ago and I want to be make the correction publically known. I am indebted to Brian for helping to identify the problem and work out the solution.

In the paper, Y_i represents detected income for case i, y_i is true income, and s_i is the detection rate, a fraction between zero and one. The model is formulated in terms of y_i and s_i, but it is detected income Y_i that is observed, where

$$Y_i = s_i * y_i.$$

The integral expression in Equation (4) of the paper (and also in Equation (5)) is evaluated over all possible combinations of s_i (from zero to one) and y_i (greater than Y_i) that yield the observed value of Y_i. Specifically, s_i has been set equal to Y_i/y_i in the integrand. Formally, this represents the following change of variables:

$$s_i = Y_i/y_i$$

$$y_i = y_i.$$

The Jacobian matrix associated with this transformation is

$$\begin{pmatrix} 1/y_i & -Y_i/y_i{}^2 \\ 0 & 1 \end{pmatrix}.$$

The Jacobian is defined as the determinant of this matrix, which is just $1/y_i$. The integral expression in equations (4) and (5) of the paper should be multiplied by the (absolute value of) the Jacobian term, yielding

$$\int_0^1 \frac{1}{\sigma_1} \phi\left(\frac{y - X_{1i}\beta_1}{\sigma_1}\right) \frac{1}{\sigma_2 y} \phi\left(\frac{(Y_i/y) - X_{2i}\beta_2}{\sigma_2}\right) dy.$$

One additional comment. In the paper the integral is formulated over true income y_i. Brian has suggested that in many cases it may be preferable to formulate the integral in terms of s_i. The reason is that s_i runs over the bounded interval $(0, 1)$, whereas as formulated in the paper the integral runs over the unbounded interval (Y_i, ∞). Again the Jacobian must be worked out for this alternative formulation and it is just $1/s_i$. Under this formulation, the integral expression in equations (4) and (5) of the paper is replaced by

$$\int_0^1 \frac{1}{\sigma_1 s} \phi\left(\frac{(Y_i/s) - X_{1i}\beta_1}{\sigma_1}\right) \frac{1}{\sigma_2} \phi\left(\frac{s - X_{2i}\beta_2}{\sigma_2}\right) ds.$$

The paper follows.

An econometric analysis of income tax evasion and its detection

Jonathan S. Feinstein*

This article presents an econometric analysis of income tax evasion and its detection based on individual-level data drawn from the Internal Revenue Service 1982 and 1985 Taxpayer Compliance Measurement Programs. I specify a model consisting of two equations: the first measures the extent of evasion; the second, the fraction of evasion detected. The empirical analysis explores the effects of income, the marginal tax rate, and various socioeconomic characteristics on filer evasion behavior, and it assesses the variability in detection rates among IRS examiners. Finally, I use the empirical estimates to construct new estimates of the income tax gap; the new estimates are very close to the previous IRS estimates.

1. Introduction

■ In the United States, income tax evasion is one of the most widespread economic crimes. The economic costs of evasion are numerous; two of the most important costs are the lost government revenues, which must be recovered through less efficient tax programs, and the inequity between evaders and honest filers. Effective control of income tax evasion requires answering a number of empirical questions: How much total evasion exists? Which socioeconomic groups evade most? What is the relationship between income, the marginal tax rate, and evasion? How might Internal Revenue Service (IRS) examination practices be improved?

Answering these questions requires addressing a fundamental statistical problem of nondetection, which arises because not all evasion is detected and recorded in data. The problem of nondetection makes some evaders appear honest, and other evaders appear to cheat less than is actually the case. In turn, these misclassifications can bias estimates of the overall extent of evasion and, when some filer categories are assigned systematically better auditors than others, can bias estimates of the relative evasion propensities of different socioeconomic groups.

This article presents an econometric analysis of income tax evasion based on individual-level data drawn from the Internal Revenue Service (IRS) 1982 and 1985 Taxpayer Compliance Measurement Programs (TCMPs). The analysis controls for the problem of non-

detection by explicitly incorporating a detection process into the model and allowing for the variation (heterogeneity) in detection rates among IRS examiners. In addition to studying the IRS detection process, the analysis focuses on two topics closely related to the policy issues mentioned above. First, the empirical results provide estimates of the effects of income, the marginal tax rate, and various socioeconomic characteristics on filer evasion behavior. Second, the model results are used to develop a new method of estimating total evasion in the United States, based on multipliers (which derive from the prior statistical analysis) that extrapolate from the TCMP audit findings to the U.S. filer population as a whole.

I call the model developed in this article fractional detection; it belongs to the family of detection controlled models first discussed in Feinstein (1989, 1990) and applied to tax compliance by Alexander and Feinstein (1987) and Erard (1990). The fractional detection model consists of two equations. The first specifies taxpayer compliance behavior, which is assumed to follow a tobit form familiar from the work of Clotfelter (1983). The second is a model of the detection process. All of the earlier work on detection controlled models assumed the detection process to be "all or none"—that is, if evasion occurred it was assumed that either all of it or none of it was detected. In contrast, the fractional detection model makes the more realistic assumption that the IRS examiner can detect some fraction of the evasion—thus he may detect all, none, or any fraction between zero and one of the evasion present. I specify these two equations and derive the associated likelihood in the next section; I also develop a number of extensions of the basic model.

Empirical results are presented and discussed at length in Section 4, but they may be summarized as follows. First, the likelihood and magnitude of evasion increases with taxpayer income and the marginal tax rate in both 1982 and 1985 when data from these years are analyzed separately; however, it is difficult to separately identify the effect of the marginal tax rate from the overall income effect in these regressions. When a pooled model is run over both years, the two effects can be separated (because two filers with identical incomes filing in the different years face different marginal tax rates); in the pooled model, income per se exerts a very small and insignificant effect on evasion, while the marginal tax rate exerts a substantial negative effect.

Second, a number of socioeconomic characteristics significantly affect the evasion decision. Schedule C (own business) and F (farm) filers are much more likely to evade than the average taxpayer. Individuals who claim the over-65 exemption (who may or may not be retired) are less likely to evade and married individuals are more likely, while the effect of capital gains on evasion is slight in most models. Dummy variables are created for two occupational groups: high-visibility government administrators, judges, lawyers, and social workers/social scientists; and medical professionals. Interestingly, the high-visibility group's evasion behavior changes substantially between 1982 and 1985. Although in 1982 this group is significantly more likely than average to evade, in 1985 it is significantly less likely. One possible explanation of this finding is that tax law changes enacted during the mid-1980s reduced "gray areas" of evasion more open to members of this group than to other filers; alternatively, changes in the economic environment may have altered their psychological propensity to evade. In most models and in both years, medical professionals are generally less likely to evade than the average nonmedical taxpayer of comparable income and social status.

A third finding is that detection is imperfect, and variations in detection rates are at least as important a source of the variation in detected evasion across cases as variation in filer characteristics. To explore the variation in detection, individual effects are specified in the detection equation for those examiners in each year auditing 15 or more cases in the sample (43 examiners in each year). Collectively, these effects are highly significant, and they indicate considerable heterogeneity among examiners. Overall, the average examiner's detection rate is approximately 50% in both 1982 and 1985. These estimates of absolute

detection rates must be accepted with caution, largely because of statistical issues of identification, which are discussed in Section 6; if the estimates are correct, they indicate that for every dollar of evasion, approximately 50 cents is detected.

The final stage in the analysis uses the empirical results to construct multipliers, which provide a means of estimating total evasion in the U.S. filer population, called the tax gap. (For this purpose total evasion is defined as total net taxable income underreported.) The most reliable multiplier estimates are computed by forming a weighted (by caseload) average of the examiners' detection rates, and are 1.94 (with a 90% confidence interval of $\pm.42$) for the 1982 estimates and 1.96 ($\pm.49$) for the 1985 estimates. Thus, according to the 1985 estimates, if the TCMP audits uncovered the equivalent of $\$X$ of detected evasion in the population, the estimate of total evasion would be $\$1.96X$.

Based on these multipliers, the new tax gap estimates are $83.7 billion (with a 90% confidence interval of $\pm$$9.5 billion) in 1987, and $63.4 billion ($\pm$7.8 billion) in 1982. These new estimates are nearly identical to the previous IRS estimates.

The remainder of the article is organized as follows. The next section specifies models, and Section 3 describes the data and provides summary statistics. Section 4 presents the empirical results, Section 5 presents the multiplier calculations and tax gap estimates, and Section 6 discusses limitations and extensions of this work. The Appendix contains the derivation of a particularly complex extension of the basic model.

2. Models

■ The econometric model consists of two equations, one referring to the taxpayer's compliance decision, and the other to the IRS detection process. Before specifying this model, it may be helpful to remark briefly on the behavioral assumptions that underlie it.

The analysis of compliance rests on the assumption that individual taxpayers make the decision whether or not to evade income taxes in a rational and consistent fashion. Thus certain characteristics will tend to make evasion more attractive; for example, wealthier taxpayers may be less risk averse and more willing to evade, while older taxpayers may be more risk averse and less likely to evade. To quantify such effects, an evasion equation is specified in which the dependent variable is evasion, and the independent variables include whatever taxpayer attributes (such as income) are available in the data. Many filer characteristics that affect the evasion decision are not recorded in the data. As a result, we expect that among a group of observationally similar taxpayers, some will evade and some will not; this variation in behavior is captured by including a stochastic disturbance in the evasion equation.

The analysis of detection allows for the possibility that the ability of IRS tax examiners to detect evasion is imperfect and varies systematically from examiner to examiner. In particular, each examiner who is assigned a sufficiently large number of cases possesses an "individual effect," which is his or her average level of detection. On any particular return, the true extent of detection then fluctuates stochastically around this average rate. The model thus assumes that there is both a systematic and an unsystematic component to detection. In addition, the detection process is allowed to depend on other variables, such as the nature of the return or the examiner's GS (Government Service) grade.

□ **The extent of evasion.** Denote by Y_{1i} the extent to which total net income is underreported on return i in the sample. To specify a model of the extent of evasion, we adopt a conventional latent variables formulation of the tobit model. Define Y_{1i}^* to be a latent variable measuring "propensity to evade." Then

$$Y_{1i}^* = X_{1i}\beta_1 + \epsilon_{1i} \tag{1}$$

and

$$Y_{1i} = Y_{1i}^* \quad \text{if} \quad Y_{1i}^* > 0 \quad \text{(evasion in amount } Y_{1i})$$

$$Y_{1i} = 0 \quad \text{if} \quad Y_{1i}^* \leq 0 \quad \text{(no evasion)},$$

where X_{1i} includes any characteristics of filer i expected to affect his evasion decision; ϵ_{1i} is a stochastic disturbance assumed to be distributed $N(0, \sigma_1^2)$; and β_1 and σ_1^2 are parameters to be estimated. Under the distributional assumption on ϵ_{1i}, the probability of no evasion is $1 - \Phi(X_{1i}\beta_1/\sigma_1)$, and the probability of positive evasion in the amount y is $\frac{1}{\sigma_1} \phi[(y - X_{1i}\beta_1)/\sigma_1]$, where Φ and ϕ are the standard normal cumulative distribution and density function, respectively.

□ **A model of fractional detection.** If all evasion on a return were detected, model (1) would be a valid specification of tax compliance and could be estimated directly, as in Clotfelter (1983). However, this model assumes that data can be collected on which individuals evade their taxes, and how much. In fact, such data are not available. Instead, the data (discussed below) record which individuals are detected evading, and how much is detected. The distinction between total evasion and detected evasion is crucial whenever IRS examiners are less than perfect in their ability to detect evasion, for then Y_{1i} is itself unobserved. In particular, when detection is imperfect, some individuals who appear compliant really possess a nonzero Y_{1i} and have simply escaped detection; others may have evaded much more than is apparent (their Y_{1i} is larger than that recorded in the data).

To allow for the possibility of imperfect detection, Alexander and Feinstein (1987) introduced a detection controlled model of the tax audit (see also Feinstein (1990) and Erard (1990)).

In all of these earlier analyses, the detection process possessed the unrealistic property of being "all or none." Specifically, conditional on a given filer underreporting net income by an amount y, the detection process was assumed either to detect the full extent of the evasion or to fail to detect any evasion whatsoever. Put differently, the detection process had a binary choice character.

In contrast to this "all or none" specification, accumulated experience suggests that frequently the audit process detects some, but not all, of the evasion on a given return. To incorporate this observation into the econometric analysis, we introduce a model of fractional detection.

Let Y_{2i} be the extent of detection on return i (conditional on evasion having occurred). Conditional on $Y_{1i} > 0$, define

$$Y_{2i}^* = X_{2i}\beta_2 + \epsilon_{2i} \tag{2}$$

and

$$Y_{2i} = 1 \quad \text{if} \quad Y_{2i}^* \geq 1 \quad \text{(complete detection)}$$

$$Y_{2i} = 0 \quad \text{if} \quad Y_{2i}^* \leq 0 \quad \text{(no detection)}$$

$$Y_{2i} = Y_{2i}^* \quad \text{if} \quad 0 < Y_{2i}^* < 1 \quad \text{(detection of the fraction } Y_{2i}^* \text{ of evasion)},$$

where X_{2i} is a collection of explanatory variables expected to affect detection; ϵ_{2i} is a stochastic disturbance assumed to be distributed $N(0, \sigma_2^2)$; and σ_2^2 and β_2 are parameters to be estimated.

According to equation (2), detection may be of three types: (i) complete detection of all evasion on the ith return; (ii) failure to detect any of the evasion on the ith return; and (iii) detection of a fraction of the evasion on the ith return. Thus, for example, if Y_{2i}^* is

realized as 1.5, all evasion is detected, while if Y_{2i}^* is realized as 0.3, 30% of the evasion is detected.

Now suppose ϵ_{2i} to be distributed independently of ϵ_{1i}. Then the probabilities associated with these three categories are (i) complete detection:

$$\text{Prob } (\epsilon_{2i} \geq 1 - X_{2i}\beta_2) = \Phi\left(\frac{X_{2i}\beta_2 - 1}{\sigma_2}\right) ;$$

(ii) no detection:

$$\text{Prob } (\epsilon_{2i} \leq -X_{2i}\beta_2) = 1 - \Phi\left(\frac{X_{2i}\beta_2}{\sigma_2}\right) ; \quad \text{and}$$

(iii) detection of the fraction s, $0 < s < 1$:

$$\text{Prob } (\epsilon_{2i} = s - X_{2i}\beta_2) = \frac{1}{\sigma_2} \phi\left(\frac{s - X_{2i}\beta_2}{\sigma_2}\right) .$$

Notice that, unlike in usual probit analysis, category (i) and (ii) probabilities sum to less than one. Note also that ϵ_{2i}'s distribution has only one free parameter, which severely restricts flexibility. Future work might usefully explore a more flexible distribution: for example, one that specified separate parameters for the probabilities of no detection and complete detection. If $X_{2i}\beta_2$ lies between 0 and 1 and σ_2 is small, detection of a fraction of total evasion is most likely; conversely, if σ_2 is large, all three categories possess substantial probability of occurrence. If $X_{2i}\beta_2$ is larger than 1, complete detection is always more than 50% likely, and is more likely the larger is $X_{2i}\beta_2$ and the smaller is σ_2. If $X_{2i}\beta_2$ is less than 0, no detection is most likely.

Data on the detection process has not usually been collected by the IRS; hence in the empirical analysis that follows, X_{2i} includes only an overall constant, an indicator of the examiner's GS grade, a measure of the complexity of the return, and, for those examiners who perform a sufficient number of audits in the sample, an individual effect: if examiner j is assigned to the ith case, this effect is denoted μ_j.

Before continuing with the derivation, it is useful to point out a number of restrictions in model (1). First, ϵ_{1i} and ϵ_{2i} are presumed independent. This restriction is inessential and is relaxed below (a model allowing correlation between the errors is specified below after the basic likelihood is derived). Second, no possibility of false detection is allowed, in which an examiner falsely claims to detect evasion when none is present. This assumption accords with IRS "folklore," but it would be interesting to investigate the importance of this assumption to the findings obtained. Finally, identification of the parameter vectors β_1 and β_2 is an important statistical issue and is not always possible. This issue is discussed more fully in Section 6 (see also Feinstein (1990)).

□ **Likelihood.** Define the set A to consist of all those returns in the data for which some evasion is detected, and define A^c to be the remaining returns for which no evasion is detected. To derive the likelihood associated with observation i requires combining equations (1) and (2) and distinguishing between returns in these two sets.

Suppose first that i falls into A^c. Then either no evasion has occurred, or evasion has occurred but has not been detected. That is,

$$\text{likelihood} = P(\text{no evasion}) + P(\text{evasion})P(\text{not detected})$$

$$= 1 - \Phi\left(\frac{X_{1i}\beta_1}{\sigma_1}\right) + \int_0^\infty \frac{1}{\sigma_1} \phi\left(\frac{y - X_{1i}\beta_1}{\sigma_1}\right)\left[1 - \Phi\left(\frac{X_{2i}\beta_2}{\sigma_2}\right)\right]dy. \quad (3)$$

The term $1 - \Phi\left(\dfrac{X_{2i}\beta_2}{\sigma_2}\right)$ may be placed outside the integral (this relies on the assumed independence of ϵ_{1i} and ϵ_{2i}), so that

$$\text{likelihood} = 1 - \Phi\left(\frac{X_{1i}\beta_1}{\sigma_1}\right) + \left[1 - \Phi\left(\frac{X_{2i}\beta_2}{\sigma_2}\right)\right]\Phi\left(\frac{X_{1i}\beta_1}{\sigma_1}\right)$$

$$= 1 - \Phi\left(\frac{X_{1i}\beta_1}{\sigma_1}\right)\Phi\left(\frac{X_{2i}\beta_2}{\sigma_2}\right),$$

a form similar to the earlier "all or none" detection controlled models.

Next suppose that i falls in A. There are now two possibilities: either all evasion was detected, or some fraction was detected. Hence

$$\text{likelihood} = \frac{1}{\sigma_1}\phi\left(\frac{Y_i - X_{1i}\beta_1}{\sigma_1}\right)\Phi\left(\frac{X_{2i}\beta_2 - 1}{\sigma_2}\right)$$

$$+ \int_{Y_i}^{\infty} \frac{1}{\sigma_1}\phi\left(\frac{y - X_{1i}\beta_1}{\sigma_1}\right)\frac{1}{\sigma_2}\phi\left(\frac{Y_i/y - X_{2i}\beta_2}{\sigma_2}\right)dy, \quad (4)$$

where Y_i refers to the amount of evasion detected, and the integral represents the chances of evasion in amount y greater than Y_i, with the fraction Y_i/y being detected, so that detected evasion remains $(y)(Y_i/y) = Y_i$.

The log likelihood of the sample is then

$$L = \sum_{i \in A^c} \log\left[1 - \Phi\left(\frac{X_{1i}\beta_1}{\sigma_1}\right)\Phi\left(\frac{X_{2i}\beta_2}{\sigma_2}\right)\right] + \sum_{i \in A} \log\left[\frac{1}{\sigma_1}\phi\left(\frac{Y_i - X_{1i}\beta_1}{\sigma_1}\right)\Phi\left(\frac{X_{2i}\beta_2 - 1}{\sigma_2}\right)\right.$$

$$\left. + \int_{Y_i}^{\infty} \frac{1}{\sigma_1}\phi\left(\frac{y - X_{1i}\beta_1}{\sigma_1}\right)\frac{1}{\sigma_2}\phi\left(\frac{Y_i/y - X_{2i}\beta_2}{\sigma_2}\right)dy\right] \quad (5)$$

and is maximized over the parameters $(\beta_1, \beta_2, \sigma_1, \sigma_2)$.

Computational experience has indicated that equation (4) can be most accurately computed if, recognizing that

$$\phi\left(\frac{y - X_{1i}\beta_1}{\sigma_1}\right) = \frac{1}{\sqrt{2\pi}}e^{-(1/2\sigma_1^2)(y - X_{1i}\beta_1)^2} = \frac{1}{\sqrt{2\pi}}e^{-[(Y_i - X_{1i}\beta_1)^2/2\sigma_1^2 + (y - X_{1i}\beta_1)^2/2\sigma_1^2 - (Y_i - X_{1i}\beta_1)^2/2\sigma_1^2]},$$

it is rewritten

$$\frac{1}{\sigma_1\sqrt{2\pi}}e^{-[(Y_i - X_{1i}\beta_1)^2/2\sigma_1^2]}$$

$$\times \left[\Phi(X_{2i}\beta_2/\sigma_2) + \int_{Y_i}^{\infty} e^{-(1/2\sigma_1^2)[(y - X_{1i}\beta_1)^2 - (Y_i - X_{1i}\beta_1)^2]}\frac{1}{\sigma_2}\phi\left(\frac{Y_i/y - X_{2i}\beta_2}{\sigma_2}\right)dy\right],$$

for which the log is

$$-\log \sigma_1 - 1/2\log(2\pi) - \frac{(Y_i - X_{1i}\beta_1)^2}{2\sigma_1^2}$$

$$+ \log\left[\Phi(X_{2i}\beta_2/\sigma_2) + \int_{Y_i}^{\infty} e^{-(1/2\sigma_1^2)[(y - X_{1i}\beta_1)^2 - (Y_i - X_{1i}\beta_1)^2]}\frac{1}{\sigma_2}\phi\left(\frac{Y_i/y - X_{2i}\beta_2}{\sigma_2}\right)dy\right].$$

◻ **Extensions.** Let us briefly discuss two extensions of the basic model; both are estimated below. First, the model can be extended to allow an arbitrary correlation ρ between ϵ_{1i} and ϵ_{2i}. The likelihood for those observations i for which no evasion is detected is then

$$1 - \Phi\left[\frac{X_1\beta_1}{\sigma_1}\right] + \int_0^{+\infty} \frac{1}{\sigma_1} \phi((y - X_1\beta_1)/\sigma_1)\left[1 - \Phi\left(\frac{q}{\sigma_2\sqrt{1 - \rho^2}}\right)\right]dy,$$

while the likelihood for those observations i for which evasion (in the quantity Y) is detected is

$$\frac{1}{\sigma_1}\phi\left(\frac{Y - X_1\beta_1}{\sigma_1}\right)\Phi\left[\frac{q - 1}{\sigma_2\sqrt{1 - \rho^2}}\right] + \int_Y^{+\infty} \frac{1}{\sigma_1}\phi\left(\frac{y - X_1\beta_1}{\sigma_1}\right)\frac{1}{\sigma_2\sqrt{1 - \rho^2}}\phi\left[\frac{Y/y - q}{\sigma_2\sqrt{1 - \rho^2}}\right]dy,$$

where

$$q = X_2\beta_2 + \frac{\sigma_2}{\sigma_1}\rho(y - X_1\beta_1).$$

Second, we note that if true income Z_i is a component of X_{1i}, fully correct estimation is complicated because Z_i is only observed when detection is complete. As a result, the explanatory variables in X_{1i} (and X_{2i} in some specifications) are themselves not fully observed. This issue was first raised by Alexander and Feinstein (1987), who suggested that it be resolved by "projecting" true income based on the estimate of detection. In this article I follow through on that suggestion.

True income Z_i equals income reported plus Y_{1i}. If Y_{1i} were observed, Z_i could be computed readily. Since Y_{1i} is not observed, we must guess Z_i, and we must do so in a manner that is consistent with the remainder of the likelihood computation. Referring back to equation (4), for returns in the set A, the X_{1i} inside the integral simply estimates Z_i (one of its components) as y + *income reported* (where y is the variable of integration). Thus for each possible value of detection, the corresponding value of true income is computed and is substituted into the explanatory variables. For returns in the set A^c, the integral in equation (3) no longer collapses, so that for returns in this class, the likelihood becomes

$$1 - \Phi(X_{1i}\beta_1/\sigma_1) + [1 - \Phi(X_{2i}\beta_2/\sigma_2)]\int_0^\infty \frac{1}{\sigma_1}\phi\left(\frac{y - X_{1i}\beta_1}{\sigma_1}\right)dy,$$

where again the estimate of Z_i in X_{1i} inside the integral is replaced by y + *income reported*.

◻ **Disaggregation.** All of the models discussed above focus on the total amount of evasion. In fact, evasion occurs at the disaggregated level, line item by line item (see Klepper and Nagin (1989) for an interesting discussion on this and related points). I do not propose here to estimate evasion and detection rates by line item; however, the models can be readily extended to handle the first step in this direction, which disaggregates the data into its two principal components: underreported Adjusted Gross Income (AGI), and overstated deductions (based on the decomposition, total net income equals AGI minus deductions). The Appendix specifies the likelihood function for a model of this type that consists of four equations: an evasion equation for underreported AGI; an evasion equation for overstated deductions; a detection equation for detecting underreported AGI; and a detection equation for detecting overstated deductions. The model assumes that the stochastic disturbances in the two evasion equations are jointly normally distributed with correlation ρ_1 and that the stochastic disturbances in the two detection equations are jointly normally distributed with correlation ρ_2. Errors across the evasion and detection equations are assumed uncorrelated.

3. Data

■ The empirical analysis is based on data from the 1982 and 1985 TCMP datasets, which are described more fully in the IRS publications *Income Tax Compliance Research, Estimates for 1973–1981* (IRS, 1983) and *Income Tax Compliance Research, Gross Tax Gap Estimates and Projections for 1973 to 1992* (with Supporting Appendices) (IRS, 1988). The study data are from four districts chosen to be geographically and socioeconomically representative of the entire nation, and for each year all observations from each of these four districts were included (however, in 1985 a small number (100) of randomly chosen observations were dropped in order not to exceed computer internal memory limitations). All of the variables used in the study come from the TCMP computer database maintained by the IRS, except for the name of the examiner performing each examination, which was coded by hand from the original TCMP checksheets.

Table 1 provides summary statistics of the data for each year. The 1982 data contains 2267 observations, and the 1985 data 3050 observations. The IRS has constructed a weighting system that weights each observation according to the relative frequency of its taxpayer's characteristics in the U.S. filer population. All of the summary statistics, model estimates, and multiplier calculations presented below are based on the weighted sample, unless otherwise indicated. Some idea of the effect of weighting is gained by exploring the relative frequencies of two socioeconomic groups, schedule C (own business) filers and schedule F (farm) filers, both of which are viewed as higher-than-average evaders. For 1982, 32% of the unweighted sample of returns filed a schedule C (724), but only 11% of the weighted sample filed C. Similarly, 25% (570) of the unweighted sample filed schedule F, but only 3.2% of the weighted sample; similar figures apply for 1985. These findings are consistent with the IRS's stated intention of oversampling certain evasion-prone groups.

For 1982, 66% of the weighted sample was detected in understatement of total taxable income. Similarly, 64% of the weighted 1985 sample was detected in understatement. The average detected understatement of income per capita in 1982 was $1001 (1982 dollars); of this evasion, 73% occurred in the form of understatement of adjusted gross income (AGI), and the remainder in the form of overstatement of deductions. In 1985, average total per capita detected understatement (income understatement plus deductions overstatement) was $1141 (1985 dollars), of which 87% derived from income understatement; the fact that this fraction is larger in 1985 may well reflect clarifications and reductions in deduction opportunities available under the new tax law.

The detection controlled methodology focuses attention on the examiners performing TCMP audits. Further, consistency of the maximum likelihood procedure requires specifying

TABLE 1   Data Summary Statistics

	1982 TCMP	1985 TCMP
Number of Cases	2267	3050
Percentage of Cases Filing Schedule C	11%	11%
Percentage of Cases Filing Schedule F	3.2%	1.9%
Percentage of Cases Detected Understating Total Taxable Income	66%	64%
Per Capita Value of Detected Understatements of Total Taxable Income	$1001	$1141
Percentage of Per Capita Detected Understatement Derived from Income Understatement	73%	87%
Number of Examiners with 15 or More Cases	43	43
Their Weighted Caseload	44%	40%

All numbers except the first row refer to the weighted sample.

individual examiner effects only for those examiners with a sufficiently large caseload in the sample. For this study, individual effects were specified for examiners with 15 or more cases (unweighted). As Table 1 indicates, there are 43 examiners in each year who fulfill this criterion (however, across years they are not identical individuals), and they are responsible for 44% of the caseload (weighted) in 1982 and 40% in 1985.

In order to summarize the raw data evidence on detection, raw detection rates were computed for each of the 43 named examiners in each year. In this calculation an examiner's raw detection rate is defined to be the number of cases on which he detects evasion divided by his total caseload (these numbers are unweighted). In each year, approximately half of the examiners have raw detection rates above 60%, and half below. Further, the dispersion in rates is substantial, with examiner rates spread from 25% to above 90%. Finally, the rates are similar across the two years, with a slightly greater dispersion in 1982, and a slightly higher mean in 1985.

While such raw detection rates are of mild interest, they suffer from several drawbacks. Since not all filers evade, the raw detection rate may understate an examiner's true rate. Opposed to this effect, however, is the fact that the raw rate assumes that when an examiner detects any evasion on a return, he has detected all the evasion—this is clearly an unwarranted assumption (see the discussion of the fractional detection model above). While these two effects act in opposite directions, there is no guarantee that they cancel each other out. A further drawback to the raw rates is that they do not control for the socioeconomic characteristics of different filers—hence a relatively good examiner assigned a caseload with a low prevalence of evasion may appear to be relatively poor. In fact, a main point of the detection controlled models is to control for filer characteristics simultaneously with examiner effects. Overall, the raw rates suffer from significant drawbacks; the computed estimates presented shortly are likely to be far more reliable.

The TCMP data are used to construct the following variables used in the analysis:

CONSTANT

AGI = Adjusted gross income (per exam) divided by 100,000 (if AGI is less than zero, it is coded as zero).

AGI^2

$AGINEG$ = A dummy variable set to 1 if the filer has AGI (per exam) that is less than $-10,000$.

$MTAX$ = Marginal tax rate for that year for that filer category, evaluated at AGI (per exam).

$MARRIAGE$ = A dummy variable set to 1 if the filer is married.

$RETIRED$ = A dummy variable set to 1 if the filer checks the over-65 exemption status box.

$SCEDC$ = A dummy variable set to 1 if a schedule C form is filed.

$FARM$ = A dummy variable set to 1 if a schedule F form is filed.

IDM = A dummy variable set to 1 if the filer belongs to the occupational group Doctors (plus veterinarians in 1985).

IDH = A dummy variable set to 1 if the filer belongs to occupational groups A1 (Officials, Administrators, and Middle Managers in Public Administration (Government)), D (Social Scientists, Social Workers, Religious Workers, and Lawyers), or G (in 1982)/H (in 1985) (Entertainers, Artists, Writers, and Athletes).

$CAPGAINS$ = Dollar value of the amount of capital gains on the return (per exam).

$NUMFORMS$ = The number of extra forms (including schedule C, schedule F, and capital gains) filed.

$GS\ GRADE$ = A dummy variable set to 1 if the examiner's GS grade exceeds 11.

As this list makes clear, a number of important variables were not included in the study, either because they were not available on the TCMP database, were not able to be provided by the IRS, or were not considered germane to the analysis. In particular, more detailed socioeconomic variables, such as savings, home ownership, and the like, are not reported on the tax return and not available. The audit probability facing a given taxpayer is recorded in the data but is not available to researchers because the IRS considers it to be a sensitive piece of information. This is unfortunate, since most theoretical models of compliance (such as Graetz, Reinganum, and Wilde (1986)) stress the role of audit probabilities in deterring evasion. In addition, the unavailability of this data makes it more difficult to compare the micro-level results of this article to the aggregate analysis of Beron, Tauchen, and Witte (1988). In the absence of audit probabilities, the compliance equation estimates must be viewed as reduced-form coefficients that summarize both the taxpayer's proclivity for evasion and his fear of detection. I hope that audit probabilities will be made available to researchers, at least on a limited basis, in the future. Finally, I have not included data on whether or not a given filer hired a paid preparer to assist in filing his or her return. Recently, Erard (1990) has performed an extensive analysis of this topic, using methods similar to those presented here; I refer the reader to his work for a full discussion.

4. Estimates

■ Empirical results are presented as follows. Table 2 provides base case estimates (derived from the likelihood in equation (5)) separately for 1982 and 1985 and for the pooled sample. In combining the two years for the pooled estimation, 1985 dollar figures were deflated to be expressed in 1982 dollars (the consumer price index was used for the deflation). In each of these models the dependent variable is total net taxable income underreported. Table 3 provides two extensions of the basic model: the first extension introduces an arbitrary correlation between the evasion and detection disturbances; the second extension corrects for the problem of "true income" in the explanatory variables, using the income projection method discussed earlier in Section 2. Finally, Table 4 presents estimates of the disaggregated model, in which separate equations are specified for (i) underreported AGI, (ii) the detection of underreported AGI, (iii) overstated deductions, and (iv) the detection of overstated deductions; again, this model was discussed in Section 2 (and its likelihood is derived in the Appendix). (The decomposition used here follows from the identity, taxable income = AGI − deductions).

Referring to Table 2, consider first the estimates for years 1982 and 1985 separately. Results are quite consistent across the two years with regard to the sign of the coefficients, their significance, and their relative magnitudes; they do differ substantially in their absolute magnitudes, a point I shall return to below. The coefficient on AGI is positive in both models in both years, and highly significant. The coefficient on AGI^2 is much smaller in magnitude, though still significant at the 5% level; since the normalized AGI (divided by 100,000) is generally less than one, AGI^2 is quite small—combined with its small coefficient, it exerts little effect. In contrast, the marginal tax coefficient is positive, large, and significant in both models. Collectively, these results suggest a relatively sharp positive relationship between income and the frequency of evasion. However, these two models by themselves provide little evidence on the relative impact of the marginal tax rate and income; the pooled model to be discussed shortly does provide interesting evidence on this point. The final income-related variable, $AGINEG$, is positive and significant—most likely it is picking up farmers or others in special income situations with many opportunities to evade.

Coefficients for the socioeconomic groups accord with prior expectations and findings in Clotfelter (1983) and Alexander and Feinstein (1987). Thus schedule C and F filers are significantly more likely to evade, perhaps because their complicated tax forms provide

TABLE 2   Total Evasion and Its Detection: Fractional Detection Models
(Dependent Variable: Net Taxable Income Underreported)

Independent Variable	Model 1 1982	Model 2 1985	Model 3 Pooled
Evasion Equation:			
CONSTANT	−6175 (151)	−16160 (161)	−.106 (160)
AGI	6717 (342)*	4957 (659)*	.173 (382)
AGI²	−277 (83.1)*	−206.5 (94.4)*	.181 (170)
AGINEG	9669 (464)*	28010 (2136)*	11620 (3774)*
MTAX	12550 (245)*	25550 (618)*	−4872 (413)*
SCEDC	4400 (318)*	7554 (836)*	4408 (500)*
FARM	3247 (428)*	7243 (1985)*	3094 (877)*
RETIRED	−886.3 (315)*	488.4 (508)	−3219 (279)*
IDM	−4076 (2992)	−655 (4238)	3596 (3758)
IDH	312.9 (1233)	−8997 (1851)*	−4715 (925)*
CAPGAINS	−.00364 (.0193)	.0579 (0204)*	.0831 (0114)*
MARRIAGE	419.3 (210)*	3360 (240)*	2.63 (237)
Detection Equation:			
CONSTANT	.371 (.0170)	.281 (.00557)	.276 (.00610)
GS GRADE	−.151 (.0185)*	.134 (.0171)*	−.0366 (.00872)*
NUMFORMS	.153 (.0139)*	.0840 (.00986)*	.199 (.00905)*
σ	7450	1800	13000
σDET	.24	.24	.27
Log Likelihood	−3364	−5598	−9677

Standard errors are in parentheses.
* Statistically significant at 5% level.

more opportunities to conceal income and claim "gray area" deductions. Individuals who claim the 65 and over exemption (who may or may not be retired) are significantly less likely to evade in 1982 but insignificantly more likely to evade in 1985. Married individuals are significantly more likely to evade; this finding was also reported in Alexander and Feinstein (1987), who suggested that the married return may be more complex and allow more opportunities for evasion (it is also indicated there that married individuals commit more mistakes in filling out their return).

Coefficients for the high-visibility group (*IDH*) and medical professionals (*IDM*) are quite different across the two years and across models. In particular, the high-visibility group is more likely to evade in 1982 (though the effect is insignificant) and significantly less likely to evade than the average comparable (in terms of income and other socioeconomic characteristics) filer in 1985. Conversely, the medical professional group is less likely to evade in both 1982 and 1985, but the coefficient is small and insignificant. It would be interesting to investigate specific effects of the tax law changes enacted during this time that apply to these groups, in order to see whether their incentives to evade have changed in a manner consistent with the empirical results. Finally, the capital gains effect is significant only in 1985, at which time it is positive as expected.

Finally, consider estimates of the detection process. The *GS GRADE* coefficient is negative and significant in 1982 and positive and significant in 1985, suggesting that the relative quality of IRS senior examiners improved over the sample period. The coefficient on *NUMFORMS* is positive and significant in both years; this runs contrary to intuition, which would suggest that more complex returns (higher *NUMFORMS*) would pose more difficult detection. There are at least two possible explanations for this effect. On the one hand, examiners may "try harder" when faced with a more challenging return. On the other hand, *NUMFORMS* may belong in both the evasion and detection equations, and its positive

significance in the detection equation may be spurious, indicating greater evasion on more complex returns. Unfortunately, testing this second hypothesis is difficult due to identification problems (discussed in Section 6); hence the issue is not pursued further here.

Estimated at the mean value of NUMFORMS, the average detection rates are 50% in 1982 and 34% in 1985 for plain (GS below 12) examiners, and 37% and 52% for senior examiners. Thus the overall average detection rate is quite similar across the two years. The standard deviation in the detection process is estimated at .24 in both 1982 and 1985, indicating a relatively tight detection range; one implication is that examiners are unlikely to detect either all or none of the evasion on a return—they may be expected to detect some part.

A comparison of the fractional detection model fit to the fit of a tobit model that assumes complete detection provides a means of testing the hypothesis of complete and homogeneous detection across returns. Such a comparison (not reported here, but available on request from the author) strongly rejects the null hypothesis in both years.

Figure 1 displays the distribution of computed mean detection rates for the 43 examiners in each year with estimated individual effects. The upward-pointing histogram refers to 1982, the downward-pointing histogram to 1985. These histograms indicate substantial heterogeneity among the examiners in detection, even after controlling for the characteristics of the filers they have been assigned. In turn, this suggests that improved training or employment screening might substantially increase audit yields.

Finally, return briefly to the evasion equation estimates in models 1 and 2. An examination of the absolute levels of the evasion equation coefficients indicates a substantial difference between the two years, and it suggests problems in the tobit specification. The

FIGURE 1

DUAL HISTOGRAMS OF EXAMINER DETECTION RATES: TOTAL EVASION
COMPUTED BASED ON FRACTIONAL DETECTION ESTIMATES, 1982 AND 1985
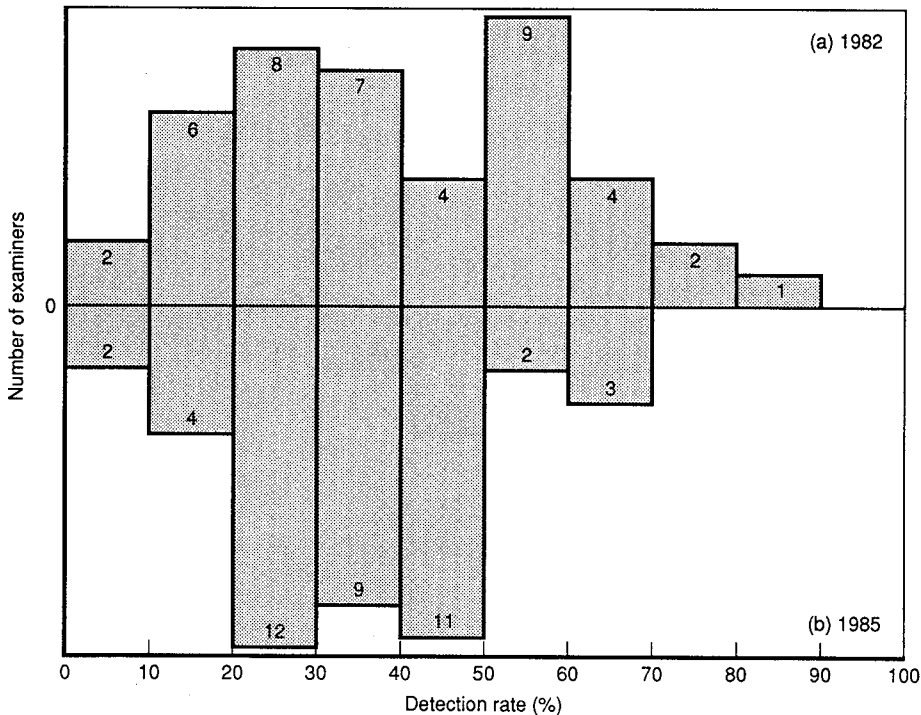
standard deviation σ is quite high in both years, and extremely high in 1985. The evasion equation coefficients essentially scale with σ, with the result that they are large, especially in 1985. The most likely explanation of the large estimated variances is that the evader population is highly skewed, with a few individuals evading very large amounts. Model fit might be improved in either of two ways: (i) by specifying an alternative asymmetric distribution for the disturbance in the evasion equation, for example log-normal, or (ii) positing a mixture of two distributions in the population. Hopefully the issue will be addressed in future work.

Consider next the pooled model reported as model 3 in Table 2. Overall, the pooled model fits the data significantly more poorly than the separate 1982 and 1985 models. While most coefficient estimates for this model are similar to those arising in the separate models, one important difference arises in the income and marginal tax coefficients in the evasion equation. Since model 3 pools data from two years with quite different marginal tax rate schedules, it allows some disentangling of two conceptually distinct but highly correlated (within a year) effects: the effect of income on evasion and the effect of the marginal tax rate on evasion. Specifically, two individuals possessing identical incomes but filing in different years face different marginal tax schedules. In the pooled model the effect of the marginal tax rate is negative and highly significant, while the effect of income is essentially zero. These results are sharply different from those in the separate years, and highly intuitive: a higher marginal tax rate discourages evasion, while income has only a very small effect on evasion. Figure 2 provides a histogram of computed examiner detection rates for the 86 examiners with individual effects in the pooled data (it was not possible to link examiners across years). These rates are similar to but somewhat lower than those for the separate years reported in Figure 1.

Table 3 provides estimates for two extensions to the basic fractional detection model. These extensions were estimated only for the 1982 dataset. Model 4 in Table 3 provides

FIGURE 2

HISTOGRAM OF EXAMINER DETECTION RATES: TOTAL EVASION
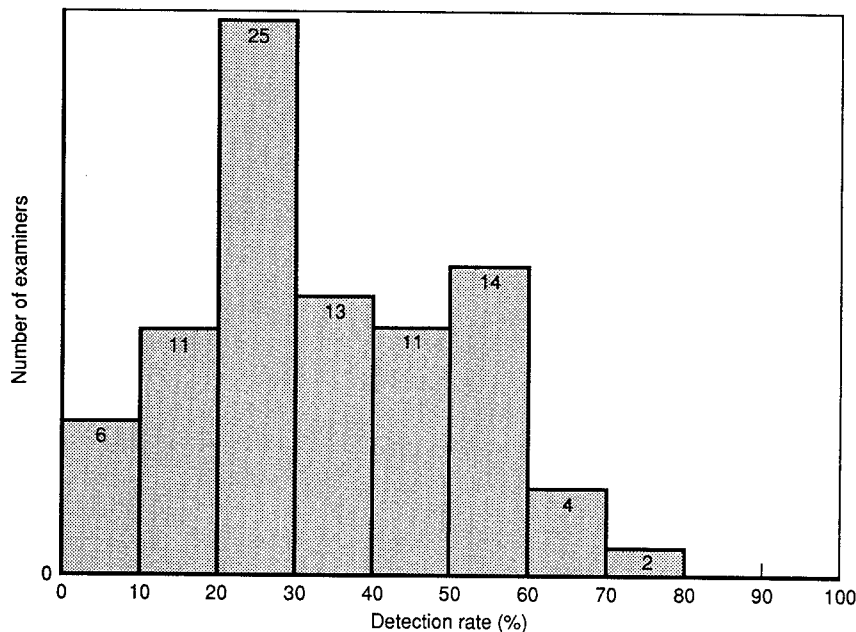COMPUTED BASED ON FRACTIONAL DETECTION ESTIMATES, POOLED SAMPLE

TABLE 3   Total Evasion and Its Detection: Extensions; 1982 Estimates
(Dependent Variable: Net Taxable Income Underreported)

| Independent Variable	Model 4 Correlation	Model 5 Income Projection
Evasion Equation:		
CONSTANT	−3261 (94.3)	−3056 (76.4)
AGI	1159 (421)*	3469 (318)*
AGI^2	−52.2 (78.1)	−308.3 (121)
AGINEG	3530 (4336)	3438 (3317)
MTAX	7270 (432)*	6136 (383)*
SCEDC	842 (356)*	2380 (224)*
FARM	2210 (1025)*	1382 (401)*
RETIRED	−1377 (297)*	4.705 (137)
IDM	−2645 (2103)	−2228 (1464)
IDH	221 (1141)	−33.78 (676)
CAPGAINS	.03521 (.0136)*	.00201 (.0238)
MARRIAGE	−189 (164)	495.7 (110)*
Detection Equation:		
CONSTANT	.131 (.0118)	.411 (.0133)
GS GRADE	.0544 (.0123)*	−.1487 (.0146)*
NUMFORMS	—	.1384 (.0138)*
σ	7300	3300
σDET	.27	.24
ρ	.8	
Log Likelihood	−3582	−3282

* Statistically significant at 5% level.

estimates for a fractional detection model that allows arbitrary correlation ρ between the stochastic disturbances in the evasion and detection equations. The parameter ρ maximizing the likelihood function was found via grid search, and it is estimated to be .8. Hence a greater propensity to evade is accompanied by a higher detection rate. This finding is intuitive: if, for example, the examiner observes certain characteristics of the filer that are not included in the evasion equation but correlate with evasion propensity, he is likely to increase the intensity of his examination. Feinstein (1990) discusses one possible model that allows for such "expectations simultaneity" between filer and examiner. Figure 3 provides computed examiner detection rates for model 4. One difference between these rates and the earlier rates in Figure 1 is that the dispersion among examiners has been substantially reduced: apparently, at least some of the dispersion derived from omitting the possible impact of compliance behavior on detection. Even in this case, however, the dispersion is substantial, ranging from 40% to 70%. The remaining coefficient estimates in model 4 are similar to those in model 1 and are not discussed further.

Model 5 in Table 3 presents estimates of a model incorporating the income projection principle for estimating true income that was discussed in Section 2; Figure 3 provides the accompanying histogram of examiner detection rates. Allowing for such projection has only a modest impact on the coefficient estimates; these estimates are not discussed further.

Finally, Table 4 provides estimates of the disaggregated model, which distinguishes between underreported AGI and overstated deductions; Figure 4 provides the accompanying histograms of computed detection rates. Again, these estimates are for the year 1982. A number of features of these results deserve comment. Consider first the two evasion equations. The overall magnitude of the evasion estimates is larger for underreported income than for overstated deductions; this is to be expected, since the majority of evasion falls in the first category. As in the earlier 1982 estimates in model 1, income is positive and significant for

FIGURE 3

DUAL HISTOGRAMS OF EXAMINER DETECTION RATES: TOTAL EVASION
COMPUTED BASED ON FRACTIONAL DETECTION ESTIMATES, EXTENSIONS
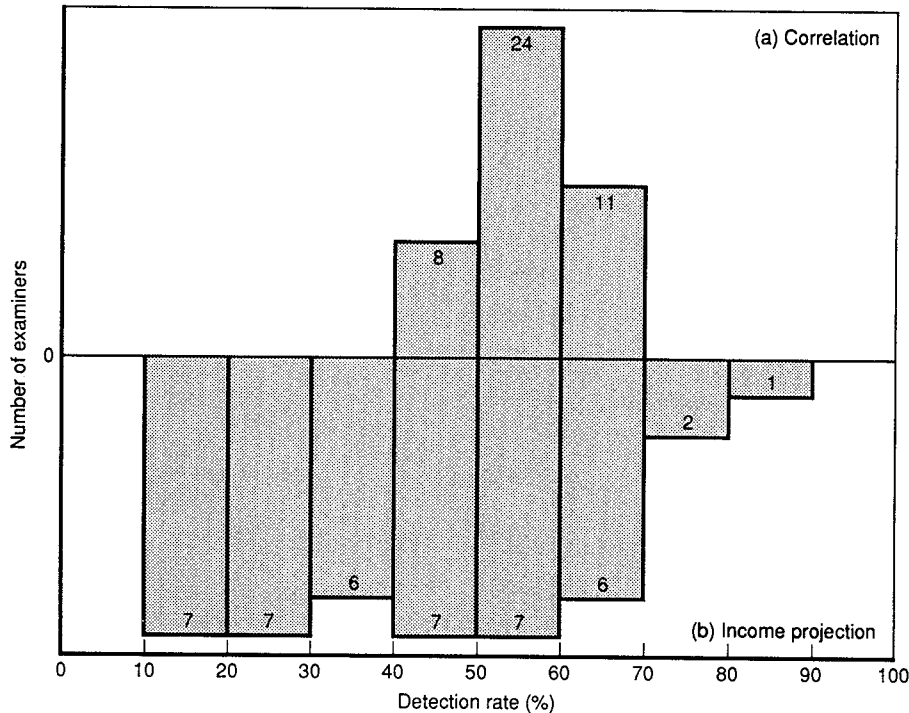
both types of evasion; however, the marginal tax rate is positive and significant in the equation for underreported AGI but negative (though insignificant) in the equation for overstated deductions. With two exceptions, the signs and relative magnitudes of the demographic variables are similar across the two evasion equations. The two exceptions are, first, the variable *RETIRED*, which is positive but insignificant in the underreported AGI equation but negative and significant in the overstated deductions equation; and second, the relative magnitude of the *FARM* coefficient is much larger in the underreported AGI equation, suggesting that farmer evasion opportunities are concentrated in this area (for example, income-smoothing techniques). The correlation between the disturbances in the two evasion equations is very large at .8.

The estimates of detection are somewhat troubling. It is a commonly accepted wisdom that overstated deductions are easier to detect than underreported income, because the former leave a "paper trail" behind. Unfortunately, the detection estimates in Table 4 and the detection rates in Figure 4 do not accord with this intuition. Overall, detection is actually somewhat lower in the overstated deductions equation, both on average and among the 43 examiners with individual effects. This result serves to highlight the care that must be taken in interpreting the absolute levels of computed detection rates (as opposed to the comparison of relative rates), discussed in Section 6 below. The correlation between the detection equations is estimated at .2.

5. The calculation of multipliers and tax gap estimates

■ I now address a central aim of the study, quantifying the aggregate extent of tax evasion (the "tax gap") in the population. I am interested in extrapolating data on the extent of detected evasion in the TCMP sample to estimates of the extent of total evasion, both

TABLE 4   **Underreported Income and Overstated Deductions: Joint Estimation for 1982**
(Dependent Variables: AGI Underreported and Deductions Overstated)

Independent Variable	Estimates	Independent Variable	Estimates
Evasion Equation for AGI Underreported:		Evasion Equation for Deductions Overstated:	
CONSTANT	−6008 (117)	CONSTANT	−1656 (81.4)
AGI	5329 (275)*	AGI	1655 (222.9)*
AGI^2	−1233 (259)	AGI^2	−348.9 (130)*
AGINEG	20380 (4821)*	AGINEG	−3405 (2605)
MTAX	9031 (236)*	MTAX	4505 (203)*
SCEDC	3781 (399)*	SCEDC	677.7 (212)*
FARM	2475 (702)	FARM	131.6 (612)
RETIRED	460.8 (288)	RETIRED	−377.8 (191)*
IDM	−22470 (952)*	IDM	−3904 (757)*
IDH	−11270 (864)*	IDH	−1154 (667)
CAPGAINS	−.01281 (.00377)*	CAPGAINS	−.0344 (.00735)*
MARRIAGE	975.5 (177)*	MARRIAGE	512.7 (98.1)*
Detection Equation for AGI Underreported:		Detected Equation for Deductions Overstated:	
CONSTANT	.479 (.00939)	CONSTANT	.353 (.0176)
GS GRADE	−.258 (.00857)*	GS GRADE	−.0917 (.0158)*
NUMFORMS	.138 (.0118)*	NUMFORMS	.0106 (.0194)
σ_1	6500	σ_2	2700
$\sigma_1 DET$	.25	$\sigma_2 DET$	.23
		ρ_1	.8
		ρ_2	.2
		Log Likelihood	−6146

* Statistically significant at 5% level.

detected and undetected, in the United States filer population at large. In focusing on the filer population, I am excluding several other important sources of evasion, including corporate evasion and nonfilers; future work may well be able to extend the results reported here to those categories.

Previous IRS estimates of filer evasion also use multipliers to scale up TCMP-detected evasion to estimates for the total filer population. The bulk of these multipliers derive from a 1976 study described in the IRS publication *1976 TCMP—IRP Study* (see IRS (1983, 1988)). This study consisted of two parts. First, the standard TCMP examination was performed, detecting a certain amount of filer evasion. In the second phase of the study, certain returns were reexamined by examiners who had access to information records about filers not normally available to TCMP examiners. This second round detected a considerable amount of evasion that had gone undetected during the usual TCMP audits, particularly in some income and deduction categories. The IRS has used the results of this study, together with a number of other special studies of certain line-item categories (such as tips), to compute multipliers with which to scale up different line-item categories of TCMP evasion; these multipliers range from 1.0 to 3.28 (see IRS (1988), especially the supporting appendices). It is important to note that the IRS methodology is fundamentally different from the approach of the current study.

My estimates of multipliers will be based on the statistical models presented above. Throughout the discussion, the term "multiplier" refers to a number a such that, if $\$X$ of evasion are detected in the weighted TCMP population, the estimate of total evasion in the U.S. filer population, both detected and undetected, is

$$\$aX.$$

FIGURE 4

DUAL HISTOGRAMS OF EXAMINER DETECTION RATES: UNDERREPORTED
INCOME AND OVERSTATED DEDUCTIONS
COMPUTED BASED ON FRACTIONAL DETECTION ESTIMATES
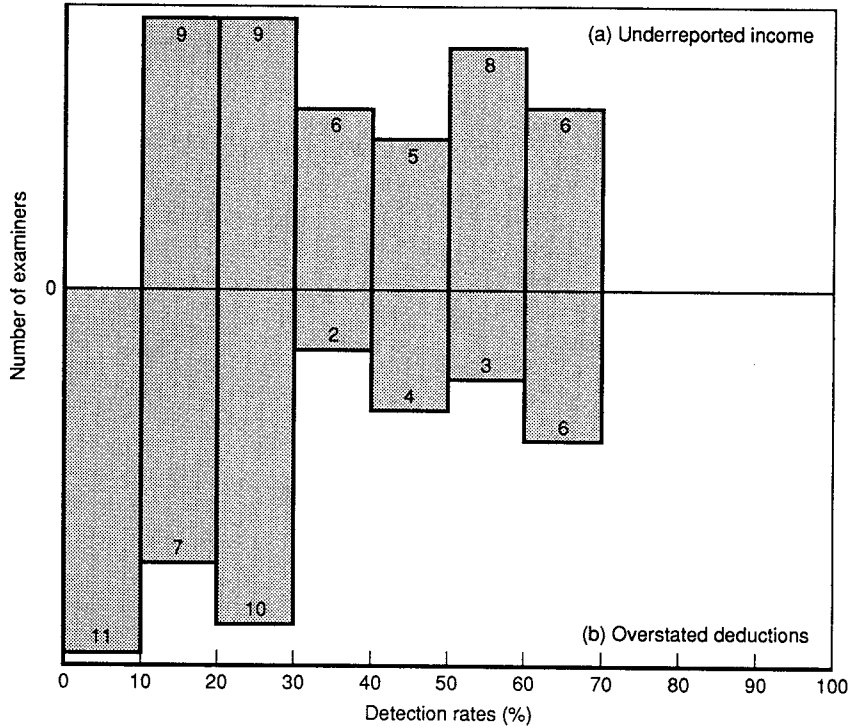
Since $\$X$ are detected, the estimate of undetected evasion is

$$\$(\alpha - 1)X,$$

which is the statistical estimate on which the discussion focuses. As discussed in the preceding paragraph, the IRS estimate of α is 3.28.

The proposed multipliers use the model estimates of examiner detection rates to create an estimate of the aggregate average detection rate, denoted T below.

I compute multipliers based on the basic models of total net taxable income evaded, presented in Table 2. Let w_i denote the weight of case i, and γ_i denote the detection parameter associated with the examiner assigned to case i. If the examiner is one of the 43 with individual effects, say the jth, $\gamma_i = \beta_{2j} + \beta_{GS}GS_j + \beta_{NUMFORMS}\overline{NUMFORMS_j}$; otherwise, $\gamma_i = CONSTANT + \beta_{GS}GS_j + \beta_{NUMFORMS}\overline{NUMFORMS}$. In these formulae β_{2j} is the examiner's detection effect in the relevant model, $CONSTANT$ the overall constant in the relevant detection equation estimates, β_{GS} the relevant GS grade coefficient, GS_j the examiner's GS grade dummy, $\beta_{NUMFORM}$ the relevant coefficient on the number of forms filed, and $\overline{NUMFORMS_j}$ (and $\overline{NUMFORMS}$) the average number of forms filed on an examiner's return. Let σ_i be the estimated standard error of γ_i; the true parameter γ_{i0} is then assumed to be normally distributed with mean γ_i and variance σ_i^2. Let f_i denote the expected detection rate on the ith case. Then, for each value γ_{i0}, the expected detection rate is

$$g(\gamma_{i0}) = \Phi\left(\frac{\gamma_{i0} - 1}{\sigma_2}\right) + \int_0^1 \frac{1}{\sqrt{2\pi}}\,\phi_2(s)e^{-(\gamma_{i0}-s)^2/2\sigma_2^2}s\,ds,$$

where $\phi_2(s)$ is the detection density at s; and

$$f_i = \int_{\gamma_{i0}} g(\gamma_{i0})N(\gamma_{i0}; \gamma_i, \sigma_i)d\gamma_{i0}.$$

We then define the statistic T as

$$T = \frac{\sum_i w_i f_i}{\sum_i w_i}.$$

T is an estimate of the average rate at which evasion (net taxable income underreported) is detected. The rate of nondetection is then $1 - T$. If $\$X$ of evasion is detected, a total $\$X/T$ of evasion is estimated to be present, hence a total $\dfrac{(1 - T)}{T} \$X$ remains undetected. Define $M = \dfrac{1 - T}{T}$; M refers to the ratio between detected and undetected evasion. Finally, then, the multiplier for total evasion is

$$\alpha = (1 + M).$$

Derivation of the standard error associated with α is straightforward. First, note that

$$\text{var}\,(\alpha) = \text{var}\,(M).$$

One can then derive

$$\text{var}\,(M) = \frac{\sum_i w_i^2 \, \text{var}\,(f_i)}{[\sum_i w_i]^2} \, \text{var}\,(T)$$

and

$$\text{var}\,(T) = \frac{\sum_i w_i^2 \, \text{var}\,(f_i)}{[\sum_i w_i]^2}.$$

□ **Multiplier and tax gap estimates.** Table 5 presents multiplier estimates. The most reliable multipliers were computed for total evasion, based on the model estimates for 1982 and 1985 presented in Table 2. These estimates are 1.94 for 1982 and 1.96 for 1985, with 90% confidence intervals stretching approximately from 1.60 to 2.4.

A second set of multipliers presented in Table 5 is based on estimates for underreported AGI similar to (but not identical with) those in Table 4. The rationale for these multipliers

TABLE 5　　　**Multiplier Estimates Based on Fitted Models**

	1982 TCMP Fractional Detection	1985 TCMP Fractional Detection
Multiplier Based on Total Detected Evasion	1.94	1.96
	(.42)	(.49)
Multiplier Based on Detected Underreported Income	2.38	3.03
	(.44)	(.50)
[Implied Multiplier for Total Evasion when all	2.01	2.77
Deduction Overstatements Detected]	(.32)	(.44)

The numbers in parentheses are estimated standard errors.

is that overstated deductions leave a "paper trail" leading to very high detection rates, which may be approximated as perfect. (Note however that the detection controlled estimates (DCE) for overstated deductions indicate surprisingly low detection rates.) Hence an alternative calculation applies a multiplier to the fraction of detected evasion due to underreported AGI and a multiplier of unity to the fraction due to overstated deductions. These multipliers comprise the bottom row of Table 5.

Overall, three of the four multiplier estimates, and the two most reliable estimates, are centered close to 2. According to these estimates, approximately one-half of all evasion is detected on TCMP returns.

Finally, an alternative means of computing these multipliers, which was not used, weights each examiner's contribution to the average detection rate not just by his weighted caseload, but also by the characteristics of the filers he has been assigned. Specifically, in each case i an estimate of the expected magnitude of evasion is computed, and multiplies the f_i defined above.

Table 6 presents the tax gap estimates that derive from the overall fractional detection multipliers described above, together with the previous IRS estimates. The striking feature of these estimates is their similarity: according to these estimates, the fractional detection methodology and the IRS methodology provide almost identical estimates of filer evasion. Thus, for example, for tax year 1987 the DCE estimates project the tax gap to be $83.7 billion (with a 90% confidence interval of ±$9.5 billion), and the IRS multiplier projects the tax gap to be $84.9 billion. Since the two methods are based on very different principles, and use somewhat different data sources, this result must be seen as substantially improving our confidence in the general level of these estimates.

In interpreting the new multiplier estimates, note that only the individual filers evasion category has been reestimated; it would be interesting to extend the present methodology to the other categories in future work.

In summary, we have used the fractional detection model estimates to derive multipliers with which to extrapolate from the TCMP audit findings to estimates of evasion for the total U.S. filer population. The new multiplier estimates are quite similar to the earlier IRS estimates.

6. Limitations and extensions

■ The models and estimates presented above depend on a number of assumptions for their validity; some of those assumptions are intrinsic to the method used, while others can

TABLE 6   Tax Gap Estimates

Source	Tax Year 1982 (Billions of 1982 Dollars)		Tax Year 1987 (Billions of 1987 Dollars)	
	New DCE Estimate	IRS[a] Estimate	New DCE Estimate	IRS[a] Estimate
Total Gap	63.4 (±7.8)	62.6	83.7 (±9.5)	84.9
Corporations	10.7	10.7	10.7	10.7
Individuals	52.7	51.9	73.0	74.2
Nonfilers	5.7	5.7	7.2	7.2
Filers	47 (±7.8)	46.2	84.6 (±9.5)	67

[a] Source: *Income Tax Compliance Research: Gross Tax Gap Estimates and Projections*, Department of the Treasury, Internal Revenue Service, Research Division, March 1988.

be relaxed in future work, particularly if larger and more powerful computer facilities are available. While many of these assumptions, limitations, and possible extensions have been alluded to above, additional insights can be gained by listing and comparing them here.

In a broad sense, the detection controlled method is statistical; hence parameter values and multipliers derived from this method represent estimates, not certain facts. As estimates these should be compared to alternative nonstatistical calculations where possible, and should always be interpreted cautiously.

In addition, the detection controlled models rest on certain specific statistical assumptions that in some cases are stronger than those required in other statistical procedures. The most important of these are the assumptions needed to identify the determinants of evasion separately from the determinants of detection, particularly in an absolute sense. To illustrate the identification problem, consider the following example, which refers to the detection controlled probit model discussed in Feinstein (1989, 1990). Suppose that the probability of filer i evading is $p_0 e^{x_{1i}\beta_1}$, where p_0 is the average frequency of noncompliance in the population and the term $e^{x_{1i}\beta_1}$ fluctuates above and below 1 depending on whether this filer's characteristics lead him to be more or less likely to evade than average. Similarly, conditional on evasion let the probability of detection be $q_0 e^{x_{2i}\beta_2}$, where q_0 is the average level of detection by IRS examiners and x_{2i} are the characteristics of the examiner assigned to filer i. All that is observed in the data is whether or not evasion is both committed and detected, the probability of which (under independence) is $e^{x_{1i}\beta_1} p_0 q_0 e^{x_{2i}\beta_2}$—hence statistical inference must be based on this product, as indeed the corresponding likelihood of equation (5) in Section 2 is. Notice, however, that in this example the factors p_0 and q_0 cannot be separately isolated; only their product $p_0 q_0$ can be determined. Put another way, a given level of average detected violation may be due to a high frequency of evasion and a low frequency of detection (high p_0 and low q_0), or to the opposite.

This example is extreme; it and related examples are discussed much more fully in Feinstein (1990, 1989), where it is shown that this double exponential case is the only parametric form in which identification of absolute levels formally fails. Nonetheless, it serves as a warning that absolute levels are more difficult to identify than the relative magnitudes of coefficients. To elaborate on this point, notice that even in the double exponential example, elements of x_1 that are excluded from x_2, and elements of x_2 excluded from x_1, possess estimable coefficients. Further, explanatory variables included in both x_1 and x_2 possess estimable reduced-form coefficients. (In all of the models presented in the previous section, x_1 and x_2 are disjoint other than their constants, the "p_0" and "q_0".)

The discussion of identification has two further implications. First, since the multipliers depend critically on estimates of the absolute (average) levels of detection, they are most sensitive to the identification problem. Second, if one or a group of particularly outstanding examiners could be found and assigned a set of reference cases, they would help determine p_0, and thus could be used to calibrate absolute detection rates among the remaining examiners (this point arose first in discussion with IRS researchers). Alternatively, the DCE estimates may be seen as tying down absolute detection rates by finding a set of "best" examiners in the data and assigning them the highest detection rates; all other examiner rates are then determined by comparing their performance to these top examiners. If such "best" examiners could be identified separately, confidence in the absolute level of the DCE estimates would be improved. As it is, the definition of "best" estimated here varies from model to model, ranging from approximately 75% to over 90%.

A second limitation of the current work is that the models are not fully structural. In particular, filer evasion equations have not been derived from utility maximization principles. Further, the probability of audit has not been explicitly introduced and allowed to depend on filer characteristics and reporting decisions; in a recent article, Beron, Witte, and Tauchen (1988) do address this issue, and they point out certain problems of endogeneity that result

and can bias estimates. In the fullest sense, the audit probability and evasion decision should be derived from a fully specified gaming relationship between the IRS and the taxpayer population. Theoretical models of this type have been proposed by Reinganum and Wilde (1986), Graetz, Reinganum, and Wilde (1986), and Scotchmer (1987); however, a satisfactory econometric model has not yet been developed and estimated. (See, however, Erard and Feinstein (1990) for a first step in this direction.)

The current work has taken one step toward estimating less aggregated models of evasion, by separating the AGI underreporting and deductions overstatement decisions. However, more work must be done in this area. (For a recent contribution see Klepper and Nagin (1989).)

Finally, as discussed in Section 4, the tobit specification does not fit the data well, and alternative specifications should be considered.

Appendix

■ Here I present the disaggregated model discussed in Section 2. Let Y_u refer to detected underreported income and Y_d refer to overstated deductions. There are four cases to consider: (i) $Y_u = 0$ and $Y_d = 0$; (ii) $Y_u = 0$ and $Y_d > 0$; (iii) $Y_u > 0$ and $Y_d = 0$; and (iv) $Y_u > 0$ and $Y_d > 0$. Let $\phi_{1u}(x) = \dfrac{1}{\sigma_{1u}} \phi(x/\sigma_{1u})$, and similarly define ϕ_{1d}, ϕ_{2u}, and ϕ_{2d}. In similar fashion define $\Phi_{1u}(x) = \Phi(x/\sigma_{1u})$ and Φ_{1d}, Φ_{2u}, and Φ_{2d}. Finally, define the conditional density and cumulative density function

$$\phi_{1u}(x|y) = \frac{1}{\sigma_{1u}\sqrt{1 - \rho_1^2}} \phi\left[\frac{x + \sigma_{1d}\rho_1/\sigma_{1u}y}{\sigma_{1u}\sqrt{1 - \rho_1^2}}\right]$$

and similarly define $\phi_{1d}(x|y)$, $\phi_{2u}(x|y)$, $\phi_{2d}(x|y)$, $\Phi_{1u}(x|y)$, $\Phi_{1d}(x|y)$, $\Phi_{2u}(x|y)$, and $\Phi_{2d}(x|y)$.

Case (i): the likelihood consists of four parts: (a) $Y_{1u} = 0$ and $Y_{1d} = 0$; (b) $Y_{1u} > 0$ (and $Y_{2u} = 0$) and $Y_{1d} = 0$; (c) $Y_{1u} = 0$ and $Y_{1d} > 0$ (and $Y_{2d} = 0$); and (d) $Y_{1u} > 0$ (and $Y_{2u} = 0$) and $Y_{1d} > 0$ (and $Y_{2d} = 0$). The likelihood of part (a) is

$$\int_{\epsilon_{1d}=-\infty}^{-X_{1d}\beta_{1d}} \phi_{1d}(\epsilon_{1d})[1 - \Phi_{1u}(X_{1u}\beta_{1u}|\epsilon_{1d})d\epsilon_{1d}.$$

The likelihood of part (b) is

$$\int_{\epsilon_{1d}=-X_{1d}\beta_{1d}}^{\infty} \phi_{1d}(\epsilon_{1d})[1 - \Phi_{1u}(X_{1u}\beta_{1u}|\epsilon_{1d})]d\epsilon_{1d}[1 - \Phi_{2d}(X_{2d}\beta_{2d})].$$

The likelihood of part (c) is identical to part (b), with the roles of d and u reversed. The likelihood of part (d) is

$$\int_{\epsilon_{1d}=-X_{1d}\beta_{1d}}^{\infty} \int_{\epsilon_{1u}=-X_{1u}\beta_{1u}}^{\infty} \phi_{1d}(\epsilon_{1d})\phi_{1u}(\epsilon_{1u}|\epsilon_{1d})d\epsilon_{1u}d\epsilon_{1d} \int_{\epsilon_{2d}=-\infty}^{-X_{2d}\beta_{2d}} \phi_{2d}(\epsilon_{2d})[1 - \Phi_{2u}(X_{2u}\beta_{2u}|\epsilon_{2d})]d\epsilon_2.$$

Case (ii): again, the likelihood consists of four parts: (a) Y_{2d} complete and $Y_{1u} = 0$; (b) Y_{2d} complete and $Y_{1u} > 0$; (c) Y_{2d} incomplete and $Y_{1u} = 0$; and (d) Y_{2d} incomplete and $Y_{1u} > 0$. The likelihood of part (a) is

$$\int_{\epsilon_{1u}=-\infty}^{-X_{1u}\beta_{1u}} \phi_{1u}(\epsilon_{1u})\phi_{1d}(\epsilon_{1d}|\epsilon_{1u})d\epsilon_{1u} \cdot \Phi_{2d}(X_{2d}\beta_{2d} - 1).$$

The likelihood of part (b) is

$$\int_{\epsilon_{1u}=-X_{1u}\beta_{1u}}^{\infty} \phi_{1u}(\epsilon_{1u})\phi_{1d}(\epsilon_{1d}|\epsilon_{1u})d\epsilon_{1u} \int_{\epsilon_{2u}=-\infty}^{-X_{2u}\beta_{2u}} \phi_{2u}(\epsilon_{2u})\Phi_{2d}(\epsilon_{2d}|\epsilon_{2u})d\epsilon_{2u}.$$

The likelihood of part (c) is

$$\int_{\epsilon_{1u}=-\infty}^{-X_{1u}\beta_{1u}} \int_{Y_{1d}=Y_d}^{\infty} \phi_{1u}(\epsilon_{1u})\phi_{1d}(\epsilon_{1d}|\epsilon_{1u})\phi_{2d}(Y_d/Y_{1d} - X_{2d}\beta_{2d})d\epsilon_{1u}d\epsilon_{1d},$$

where we implicitly have defined Y_{1d} and $\epsilon_{1d} = Y_{1d} - X_{1d}\beta_{1d}$. The likelihood of part (d) is

$$\int_{\epsilon_{1u}=-X_{1u}\beta_{1u}}^{\infty} \int_{Y_{1d}=Y_d}^{\infty} \phi_{1u}(\epsilon_{1u})\phi_{1d}(\epsilon_{1d}\mid \epsilon_{1u})\phi_{2d}(Y_d/Y_{1d} - X_{2d}\beta_{2d})[1 - \Phi_{2u}(X_{2u}\beta_{2u}\mid \epsilon_{2d})d\epsilon_{1u}d\epsilon_{1d}.$$

Case (iii) is identical to case (ii), with the roles of d and u reversed.

Case (iv): again, the likelihood consists of four parts: (a) Y_{2u} complete and Y_{2d} complete; (b) Y_{2d} complete and Y_{2u} incomplete; (c) Y_{2u} complete and Y_{2d} incomplete; and (d) both incomplete. The likelihood of part (a) is

$$\phi_{1u}(Y_u - X_{1u}\beta_{1u})\phi_{1d}(Y_d - X_{1d}\beta_{1d}\mid \epsilon_{1u}) \int_{\epsilon_{2u}=1-X_{2u}\beta_{2u}}^{\infty} \Phi_{2u}(\epsilon_{2u})\Phi_{2d}(X_{2d}\beta_{2d} - 1\mid \epsilon_{2u})d\epsilon_{2u}.$$

The likelihood of part (b) is

$$\phi_{1d}(Y_d - X_{1d}\beta_{1d}) \int_{Y_{1u}=Y_u}^{\infty} \phi_{1u}(\epsilon_{1u}\mid Y_d - X_{1d}\beta_{1d})\phi_{2u}(Y_u/Y_{1u} - X_{2u}\beta_{2u})\Phi_{2d}(X_{2d}\beta_{2d} - 1\mid \epsilon_{2u})d\epsilon_{1u}.$$

Part (c) is similar to (b), with the roles of d and u reversed. The likelihood of part (d) is

$$\int_{Y_{1u}=Y_u}^{\infty} \int_{Y_{1d}=Y_d}^{\infty} \phi_{1u}(\epsilon_{1u})\phi_{1d}(\epsilon_{1d}\mid \epsilon_{1u})\phi_{2u}(Y_u/Y_{1u} - X_{2u}\beta_{2u})\phi_{2d}(Y_d/Y_{1d} - X_{2d}\beta_{2d}\mid \epsilon_{2u})d\epsilon_{1d}d\epsilon_{1u}.$$

References

ALEXANDER, C. AND FEINSTEIN, J.S. "A Microeconometric Analysis of Individual Income Tax Evasion." Mimeo, MIT, 1987.

BERON, K., TAUCHEN, H., AND WITTE, A. "A Structural Equation Model for Tax Compliance and Auditing." Working Paper #124, Department of Economics, Wellesley College, 1988.

CLOTFELTER, C. "Tax Evasion and Tax Rates: An Analysis of Individual Returns." *Review of Economics and Statistics,* Vol. 65 (1983), pp. 363–373.

ERARD, B. *Tax Practitioners and Tax Compliance: A Microeconometric Analysis of the Decision to Engage a Tax Preparer and its Consequences.* Ph.D. dissertation, University of Michigan, 1990.

———— and FEINSTEIN, J.S. "An Econometric Model of the Tax Compliance Game." Mimeo, Stanford University, Graduate School of Business, 1990.

FEINSTEIN, J.S. "The Safety Regulation of U.S. Nuclear Power Plants: Violations, Inspections, and Abnormal Occurrences." *Journal of Political Economy,* Vol. 97 (1989), pp. 115–154.

————. "Detection Controlled Estimation." *Journal of Law and Economics,* Vol. 33 (1990), pp. 233–276.

GRAETZ, M.J., REINGANUM, J.F., AND WILDE, L.L. "The Tax Compliance Game: Toward an Interactive Theory of Law Enforcement." *Journal of Law, Economics and Organization,* Vol. 2 (1986), pp. 1–32.

Internal Revenue Service. *Income Tax Compliance Research, Estimates for 1973–1981.* 1983.

————. *Income Tax Compliance Research, Gross Tax Gap Estimates and Projections for 1973 to 1982.* 1988.

KLEPPER, S. AND NAGIN, D. "The Anatomy of Tax Evasion." *Journal of Law, Economics, and Organization,* Vol. 5 (1989), pp. 1–24.

REINGANUM, J.F. AND WILDE, L.L. "Equilibrium Verification and Reporting Policies in a Model of Tax Compliance." *International Economic Review,* Vol. 27 (1986), pp. 739–760.

SCOTCHMER, S. "Audit Classes and Tax Enforcement Policy." *American Economic Association Papers and Proceedings,* Vol. 77 (1987), pp. 229–233.