A MicroEconometric Analysis of Income Tax Evasion.

Craig Alexander, M.I.T.

Jonathan S. Feinstein, M.I.T.

April, 1987

The first author gratefully acknowledges the financial support of the Alfred P. Sloan Foundation, the second that of the National Science Foundation. We thank Franklin Fisher, Bruce Meyer, James Poterba and Michael Weisbach for useful comments. We especially wish to thank members of the IRS Compliance Analysis Group, particularly Dennis Cox and Mike Udell. We are jointly responsible for all remaining errors.

Introduction

Individual income tax evasion is probably the most widespread economic crime in the United States: the IRS estimates that in 1981 a total of $75.3 billion in individual income taxes owed to the Federal Government went unpaid, an amount comparable to that year's Federal deficit.

In this paper we present a microeconometric analysis of income tax evasion. We propose three new econometric methods for analyzing evasion data. To begin with we extend the traditional tobit model of evasion, estimated for example by Clotfelter (1983), to a more general bivariate tobit model of the type proposed by Heckman (1979), in which filers' decisions whether or not to evade are separate from (though correlated with) their decisions of how much to evade. The tobit model of evasion is derived from Allingham and Sandmo's (1972) classic analysis, under the assumption that, whenever the extent of evasion is small, the penalty assessed if the filer is caught is also small. The bivariate tobit model is also derived from the Allingham and Sandmo analysis, but under the more general assumption that filers also suffer a fixed cost if they are caught evading, for example because they face a higher probability of audit in future tax years.

The second method we propose addresses the problem of taxpayer errors. In our data sample 13.5% of all filers overpay their taxes, behavior which is most sensibly interpreted as an error either in calculation or in interpretation of the tax code. We suspect that a comparable number also underpay due to error, which confounds the analysis of evasion, since we can never be sure whether an underpayment is the result of deliberate intent or error. We present a simple two-stage technique which corrects for taxpayer errors.

Finally, we propose methods which take into account the IRS' inability to detect all instances of evasion. We develop detection controlled estimation techniques which explicitly model the detection process, and which allow the probability of detection to vary amongst IRS examiners. Our analysis follows a more general treatment in Feinstein (1986a). A byproduct of these detection controlled techniques is an estimate of the incidence and extent of undetected evasion, which is of considerable policy interest.

We estimate our models using the 1982 IRS Taxpayer Compliance Measurement Program (TCMP) dataset. The TCMP dataset contains approximately 50,000 individual tax returns (of which we use 2,267), each of which has been thoroughly audited by an IRS examiner; it provides detailed information on the filer's socioeconomic and income characteristics, and identifies the IRS examiner assigned to audit his return.

Our results suggest three broad conclusions about tax evasion. First, the source from which income is derived (capital gains, farm, etc...) exerts a strong impact on the filer's evasion behavior. Second, together the level of income and the marginal tax rate have a positive and significant effect on the filer's decision whether or not to evade taxes, particularly in models which control for nondetection and taxpayer errors, but exert only a weak positive effect on the extent of evasion; in addition, we are unable to untangle the income and marginal rate effects from one another. Third, detection rates differ substantially amongst IRS examiners, and much evasion appears to go undetected -- we calculate that in 1982 filers' undetected understatement of taxable income was approximately $62 billion.

The econometric methods which we propose represent an improvement over earlier efforts in several regards. Much previous work has used aggregated data. A recent example is Witte and Woodbury (1985), who focus on the impact of audit rates on compliance, and use data aggregated to the three digit zip code level. Since the evasion decision is made at the individual level, aggregation is likely to introduce biases into their estimates.

Clotfelter (1983) does use individual data, drawn from the 1969 TCMP, and fits a tobit model of evasion. He finds a significant positive effect of the marginal tax rate on evasion, which he takes as evidence in favor of the view that reduced marginal rates would increase compliance. However, tobit models seem to fit the data poorly (which is why we have generalized them to a bivariate framework), and we have found that Clotfelter's estimates predict unrealistically low levels of evasion (between $20 and $100). In addition, Clotfelter does not address the problems of taxpayer errors and nondetection.

Previous IRS studies have focused primarily on estimating the magnitude of evasion. In a study (1983) of the 1976 individual TCMP a group of experienced examiners was assigned to reaudit a subsample of the returns, using income-related documents from the Information Returns Program (IRP) which were not available to the original TCMP examiners. These more detailed audits uncovered 3.28 times as much evasion as the original TCMP audits, which lead the IRS to estimate the true extent of evasion as 3.28 times that uncovered by the TCMP. Our detection controlled techniques are a statistical analogue of this approach, which allow for imperfect detection by even the most experienced examiners.

The remainder of the paper is organized as follows. The first section describes our models f evasion and its detection. Section II contains a short description of the data set. Statistical methods for dealing with taxpayer errors and imperfect IRS detection of evasion are discussed in section III. Section IV presents our empirical results and discusses their interpretation. Section V contains estimates of the magnitude of undetected evasion in the 1982 taxpaying population, and a policy analysis of the effect of several recent U.S. tax reforms on evasion levels. Finaly, a conclusion summarizes our findings and suggests directions for further research, and an appendix contains technical details.

Section I: Model Development

Filer Behavior

In an important paper Allingham and Sandmo (1972) pioneered the study of income tax evasion as a rational economic decision, drawing upon earlier work on the economics of crime by Becker (1968). In the Allingham and Sandmo framework each filer possesses a concave utility function U( ), and true income I. The filer also possesses characteristics x which enter U primarily through their effect on his degree of risk aversion. The filer reports his income to the IRS and is taxed based on his report; if he reports less than his true income, he faces the probability p of being caught evading, in which case he pays a penalty on the difference between his reported and true income. Let Y denote this difference; Y is then the extent of evasion, and is assumed to be nonnegative. The filer chooses Y to solve:

$$\max_{Y} \quad (1-p)U(I-t(I-Y)) + pU(I-t(I-Y)-\Theta(Y)) \qquad (1)$$

$$\text{subject to: } 0 \le Y \le I$$

where t(I-Y) is the tax payable on the filer's declared income, and Θ(Y) is the penalty function specifying how much the filer must pay if detected evading in the amount Y. The first order condition associated with (1) is:

$$(1-p)U'(I-t(I-Y))t'(I-Y) + pU'(I-t(I-Y)-\Theta(Y))(t'(I-Y)-\Theta'(Y)) \qquad (2)$$

$$= 0 \text{ if } 0<Y<I$$
$$< 0 \text{ if } Y=0$$
$$> 0 \text{ if } Y=I$$

In the United States tax system the usual penalty Θ( ) imposed for detected evasion equals 1.25[t(I)-t(I-Y)], which has the property that it is small whenever the extent of evasion is small. If Θ is assumed to possess this property, the first order condition in (2) is both a necessary and sufficient condition for a solution to (1). Linearizing (2) about the filer's true income I, we derive a tobit model of evasion:

$$Y^* = x\beta - \varepsilon \tag{3}$$

$$Y = Y^* \text{ if } Y^* \geq 0$$
$$Y = 0 \text{ if } Y^* < 0$$

where x includes the filer's characteristics, mentioned above, as well as his true income I and the marginal tax rate t' evaluated at his true income, and ε is a mean 0 disturbance drawn from the distribution F.

Clotfelter (1983) has estimated the tobit model in equation (3), but his results predict unrealistically low levels of evasion. We also have estimated the tobit model in (3), but have found it to fit the data poorly. In particular, it predicts extremely low probabilities of evasion, $F(x\beta)$, and conditional (on evasion having occurred) expected levels of evasion, $E(Y| \varepsilon > -x\beta)$, which are essentially constant across a wide range of characteristics x, including all relevant incomes. We believe the reason the tobit model fits poorly is because it models both the decision whether or not to evade and the decision of how much to evade (given that evasion occurs) as determined by the same model $x\beta$ and same error distribution F. Nearly half the TCMP population does not evade, which indicates that $F(-x\beta/\sigma)$ (the probability of no evasion) is large. However, amongst those who do evade the extent of evasion is substantial: the average underpayment amongst those who underpay is $5,000, which indicates that $x\beta$ is large. These two effects pull in opposite directions and confound estimation.

To surmount this difficulty we generalize the tobit model in equation (3) to a bivariate system which explicitly distinguishes between the filer's decison whether or not to evade at all, and his subsequent decision of how much to evade. The model we propose is related to the bivariate selection models pioneered by Heckman (1979) and discussed in Amemiya (1985). To motivate this system we assume that a filer who is caught evading not only suffers the penalty Θ, which as we discussed above is small whenever the extent of evasion is small, but also suffers a fixed cost which is independent of the level of evasion. We suggest two justifications for this fixed cost. First, a filer who is caught evading may face a higher probability of reaudit in the future. Second, when a filer is caught evading it signals the community that he is dishonest, which affects the way he is treated. Of course both of these effects are likely to be larger when the extent of evasion is larger. What is important for our purposes is that they are substantial even when the amount of evasion is small (they are discontinuous at the zero evasion level).

Under the fixed cost assumption the first order condition in (2) is no longer sufficient for a solution to (1), since the objective function is no longer continuous at Y equal to zero. Equation (2) must be supplemented by a second equation, which compares the filer's utility at the Y value which solves (2) (and which is given by equation (3)) to his utility at Y equal to zero. We reverse this ordering and formulate the system:

$$Y^*_1 = x_1\beta_1 + \varepsilon_1 \qquad (4)$$

$$Y_1 = 1 \text{ (evasion) if } Y^*_1 > 0$$
$$Y_1 = 0 \text{ (no evasion) if } Y^*_1 < 0$$

and conditional on $Y_1 = 1$,

$$Y_2 = x_2\beta_2 + \varepsilon_2 \qquad \text{(the extent of evasion)} \qquad (5)$$

$$Y_2 > 0$$

where x_1 and x_2 are filer characteristics, and β_1 and β_2 are parameter vectors to be estimated. The errors ε_1 and ε_2 are drawn from a bivariate normal distribution; they possess correlation ρ, ε_2 has variance σ_2^2, and ε_1's variance is normalized to one for identification. Equation (4) is a binary choice selection equation, while (5) is a model of the extent of evasion.

Since (5) is conditional on $Y_1 = 1$, it is conditional on $\epsilon_1 > -x_1\beta_1$, which affects (5) because ϵ_1 and ϵ_2 are correlated. Heckman shows that (5) can be rewritten as:

$$Y_2 = x_2\beta_2 + \rho\sigma_2\phi(x_1\beta_1)/\Phi(x_1\beta_1) + \epsilon'_2 \qquad (5)'$$
$$= (x_2\beta_2)' + \epsilon'_2$$

where $Var(\epsilon'_2)$ equals

$$\sigma_2^2[(1-\rho^2)+\rho^2(1-\phi(x_1\beta_1)/\Phi(x_1\beta_1)(x_1\beta_1+\phi(x_1\beta_1)/\Phi(x_1\beta_1)))]$$

which we denote σ'^2_2. The fact that Y_2 must be positive imposes further restrictions on (5)'. In particular, we must treat (5)' as a truncated tobit model, in which we explicitly recognize that ϵ'_2 is truncated below at -$(x_2\beta_2 + \rho\sigma_2\phi(x_1\beta_1)/\Phi(x_1\beta_1))$. Working with $(x_2\beta_2)'$ and σ'_2 as defined, we use conditions for a truncated tobit model presented in Amemiya (1985) to rewrite (5)' as:

$$Y_2 = (x_2\beta_2)' + \sigma'_2\phi((x_2\beta_2)'/\sigma'_2)/\Phi((x_2\beta_2)'/\sigma'_2) + \epsilon''_2 \qquad (5)''$$
$$= (x_2\beta_2)'' + \epsilon''_2$$

where ϵ''_2 is mean zero, and possesses a truncated normal distribution with variance

$$\sigma''^2_2 = \sigma'^2_2(1 - \phi((x_2\beta_2)'/\sigma'_2)/\Phi((x_2\beta_2)'/\sigma'_2)\ [(x_2\beta_2)'/\sigma'_2) + \phi((x_2\beta_2)'/\sigma'_2)/\Phi((x_2\beta_2)'/\sigma'_2)]) \qquad (6)$$

Equation (5)" is the variant of (5) which we use for estimation. For later reference we define $\phi^*_2 = (1/\sigma''_2)\phi((Y_2 - (x_2\beta_2)'')/\sigma''_2)$, the probability of observing evasion in the amount Y_2, conditional on evasion occurring ($Y_1 = 1$).

We concentrate on estimating two types of models. First, we estimate the binary choice probit model of equation (4), which is a model of the incidence of evasion amongst the population. As Heckman shows (see also Amemiya) this equation can be estimated consistently over the entire sample of both evaders and nonevaders, without estimating (5). Second, we estimate (5)", using the estimates of ε_1 obtained from estimating the binary choice model (4) as weights. We call this second estimation a bivariate tobit model of the extent of evasion. It is a nonlinear least squares form.

The Detection Process

The detection of tax evasion is difficult, and by all accounts much evasion goes undiscovered by the authorities. The IRS itself has admitted the difficulties it faces in identifying evaders, and has repeatedly asked Congress for more money with which to increase detection efforts. While previous empirical work has implicitly assumed that the IRS detects all evasion (at least within the TCMP sample), recent theoretical contributions by Reinganum and Wilde (1985) and Graetz, Reinganum and Wilde (1986) have focused attention on the problem of detection, and especially on the interdependence between the detection process and filers' evasion decisions. We will follow the approach of Reinganum and Wilde and of Feinstein (1986a, 1986b) in explicitly modeling the detection process.

Reinganum and Wilde view the evasion decision and the detection process as the outcome of a sequential move game of incomplete information. Abstracted to our puposes, this game proceeds as follows. The filer moves first: given a true income I the filer chooses how much to evade, Y, reports income I-Y to the tax authorities, and remits a tax t(I-Y). We have modeled this part of the game in the previous subsection.[1] Next the IRS tax

[1] Notice that the filer's true income I is hidden from the authorities' view, which is what makes the game one of "incomplete information".

authorities move: an examiner is assigned to the return, decides how much time and efffort to devote to auditing, and either detects or fails to detect evasion. This game is a simplification of the true situation on two counts. First, it ignores a prior stage in the game during which the authorities choose the tax function t, the penalty function Θ, the aggregate resources to devote to detection, and the audit rule. Second, it does not split the detection process into its two components, the decision whether or not to audit, and the subsequent decision of how much effort to devote to the audit. In the TCMP tax sample which we use every return is audited, so we focus exclusively on the second aspect of detection.

We define the latent variable W^* to be the intensity of the detection process, and specify a linear detection technology for W^*:

$$W^* = z\Pi - \mu - \varepsilon_3 \tag{7}$$ [2]

where z denotes variables which determine the intensity of detection, such as the examiner's GS (Government Service) grade, μ denotes the ability level of the examiner assigned to the case, and ε_3 is a mean zero error drawn from the cdf G.[3] The μ variables are examiner fixed effects designed to capture heterogeneity amongst examiners. If we were able to obtain good information about each examiner's socioeconomic characteristics, such as education and age, we might be able to dispense with the μ's. Since the μ's are fixed effects, for the techniques which we propose to be consistent requires a reasonable number of cases for each examiner for whom we specify a fixed effect (see Heckman (in Manski and McFadden (1983)) for a discussion) -- hence we specify examiner effects only for those examiners with at least 15 cases in the sample. Conditional on the filer having committed evasion, detection

[2] It would be preferable to derive (7) from a theoretical model in which the examiner optimally chooses the time and effort to devote to a particular case, taking into account all the relevant costs and benefits. While we have developed such a model, we do not present it here as we were unable to obtain the data necessary to its estimation.

[3] Within the sample examiner's will be indexed by the subscript j, so that μ_j will denote the jth examiner's ability level.

occurs if $W^* \geq 0$, an event which occurs with probability $G(z\Pi - \mu)$. We define:

$$
\begin{aligned}
L &= 1 \text{ iff } W^* \geq 0 \\
L &= 0 \text{ iff } W^* < 0
\end{aligned}
\tag{8}
$$

We assume that if detection occurs it is complete and the entire amount Y of evasion is detected, whereas when detection fails no evasion is found. This is an extreme model, and a more general formulation would allow for "fractional detection" in which a variable amount of the evasion is detected.[4] Equations (7) and (8) focus exclusively on the problem of nondetection, which might be called a "type I error" committed by the examiner, and ignore the complementary problem of a "type II error" which arises when an examiner falsely accuses an honest filer of evasion.[5]

Viewing evasion and its detection as parts of a larger encompassing system leads naturally to a consideration of the interdependence which arises between the filer's evasion decision and the IRS examiner's willingness to devote time and effort to the audit process. A filer will be more likely to evade if she believes the probability of detection to be small, and an examiner will devote more time to a particular return if she believes the probability of evasion to be high.

Typically the filer possesses little information about the examiner whom she will be assigned (though she may possess quite good information about the audit process, which we do not model); hence we do not explore the dependence of the evasion decision on the filer's expectation's about the

[4] Extending (7) and (8) to allow for fractional detection would be empirically useful because IRS examiners tend to be better at detecting some types of evasion, such as over-expensing, than others, such as non-reporting of an income source . Alternatively, evasion and detection could be estimated separately (and probably jointly) over a variety of income types.

[5] Though type II errors do arise occasionally during the initial audit phase, there is substantial evidence (IRS [...]) that they are generally corrected during follow-up investigations.

detection process.[6] When the examiner also possesses little information about the filer, expectations do not enter the detection equation either. We call this the "base case" and focus most of our attention on it. When the examiner does possess good information about the filer, this information should be included in the detection function (8). If for example the examiner observes the same variables x_1 as the econometrician, then we can specify the examiner's assessment of the probability of evasion as

$$\phi(x_1\beta_1) \tag{9}$$

from above. We then amend the detection equation (7) to

$$W^* = z\Pi + \mu + \delta\phi(x_1\beta_1) + \varepsilon_3 \tag{10}$$

which can be interpreted as a linear rational expectations formulation. The parameter δ is to be estimated and measures the impact of the examiner's information on the detection process -- we expect δ to be positive. If the examiner possesses better information than the econometrician (10) remains appropriate, but the errors ε_1 (and ε_2 in most cases) and ε_3 will become correlated due to the examiner's extra information which derives from sources other than x_1 (and x_2).

[6] In the most general formulation, the filer's likelihood of being detected should depend both on whether or not she evades, and how much. This leads to a simultaneous (recursive) system in which Y_1 enters the detection equation (8). This model is beyond the scope of the current paper, but remains an important topic for future research.

Section II : The Data.

Our data are drawn from the 1982 Taxpayer Compliance Measurement Program (TCMP III-8) data set maintained by the Internal Revenue Service. The TCMP consists of tax information on a stratified random sample of 50,653 individual income tax returns filed for tax year 1982. Each individual in the TCMP sample was subject to an intensive audit by an IRS examiner; the examinations were carried out over a period of approximately two years beginning in August 1983. Since every return is audited, the TCMP allows a direct comparison of the taxpayer's report of each line item to the IRS examiner's assessment of the item -- hence its usefulness in studying evasion behavior. For each line item, we denote the taxpayer's report of the item as "per return" (subscript $_R$), and the IRS examiner's assessment of the item as "per exam" (subscript $_E$). For example, Capital Gains reported is $CAPGAINS_R$, whereas Capital Gains assessed by the examiner is $CAPGAINS_E$.

We use a subsample of the 1982 TCMP consisting of 2267 cases from four districts.[7] We restrict our sample to all cases from these four districts, as opposed to randomly sampling a few cases from each of the 57 districts, so as to ensure that the sample includes a reasonably large number of cases for each IRS examiner for whom we specify a fixed effect and calculate an individual detection rate.[8] Since each examiner audits cases in only one district, random sampling from all districts would not fulfill this criterion. The four districts were selected according to the following procedure: first we chose at random four of the seven geographic regions into which the IRS divides the country; then we randomly chose one district from each of these four regions. The districts constitute a representative cross-

[7] The IRS has requested that we not reveal the names of these districts and that we make it clear that we did not have direct access to the data. Rather, we prepared computer programs which the IRS subsequently ran.

[8] Recall that consistency of the examiner fixed effects requires a large number of cases for each examiner for whom an effect is specified.

section of the national TCMP population.[9] For each case, we have the following variables: Adjusted Gross Income (AGI), Capital Gains Income, Total Deductions, Total Taxes, Occupation Category, Marital Status, Schedule C (i.e. income from a sole proprietorship) Income, Farm Income, and whether or not the filer is aged over 65. Each variable is present in both the "per return" and "per exam" forms. We summarize Schedule C and Farm Income as dummy variables which take the value one for a filer if he has nonzero income in that category. Dummy variables are also specified for Marital Status (one if married), aged over 65 (one if over 65), and for two occupation groups: (i) Government officials, religious and legal personnel and entertainment workers; and (ii) Medical personnel. The occupation group dummies are estimated in the binary choice model, since we believe that more visible individuals such as those in (i) and (ii) are likely to suffer a larger fixed cost if caught evading. Finally, we also specify a dummy variable for those filers possessing AGI_E greater than $40,000, and use this variable in our bivariate tobit models.

Table 1 reports summary income and tax statistics for our sample. "AGI_E" refers to Adjusted Gross Income as calculated by the IRS examiner, "$AGI_E - AGI_R$" refers to the difference between the examiner's assessment of the filer's Adjusted Gross Income and the Income originally reported by the filer, and "$TAX_E - TAX_R$" refers to the difference in tax due. We note that our sample (and the TCMP as a whole) contains heavy oversampling at higher levels of income. For example, 8.7% of the filers in our sample have AGI in excess of $100,000 while the figure from Statistics of Income 1982 is a mere 0.6%. This oversampling does not jeopardize the validity of our estimation techniques (since it is based on explanatory variables which we observe). However, to insure that the results of our estimation are fully representative of the U.S. taxpaying population, we weight each observation to reflect its

[9] In particular, the income distribution ("per exam") for our four districts is essentially identical to the income distribution for the TCMP sample as a whole.

frequency in the overall population.[10] The weights were made available to us by the IRS.

The detection-controlled estimation procedures require data on the IRS examiner assigned to each case. For all cases we know the GS grade of the examiner assigned to the case, and summarize this information in the dummy variable GS, which takes the value one if the grade is 11 or higher. In addition, for 20% of the cases the IRS was able to identify the name of the examiner by matching the TCMP file with the original audit checksheets, which the examiner is required to sign. A total of 353 examiners are identified in the data. Of these, 44 audited 15 or more cases, and we assign fixed effects to each of these; they are responsible for approximately half of the cases in our sample.[11]

Table 1: Summary Statistics.

Statistic	Cases Positive	Mean of Positive Cases	Cases Zero	Cases Negative	Mean of Negative Cases
AGI_E	2153	$52,536	0	114	$49,522
$AGI_E - AGI_R$	1411	$5,327	531	325	$1.779
$TAX_E - TAX_R$	1337	$1,707	624	306	$573

Note: Subscript "E" refers to item as "per exam" and subscript "R" refers to item as "per return". All "negative cases" represent overstatements of liability.

[10] Formally, we view the weights as heteroscedasticity corrections which arise because filers with unusual characteristics (very high or low income) are likely to exhibit more variable behavior.

[11] For the remaining cases detection depends upon the overall constant and the GS grade (and an expectations term in one of the model extensions).

Section III: Statistical Techniques

Estimation of the binary choice and bivariate tobit models presented in Section I is complicated by two issues: (1) taxpayer errors; and (2) the inability of the IRS to detect all instances of evasion, behavior which we have modeled in Section I by specifying an explicit detection process. In this section we present statistical techniques which address each of these topics.

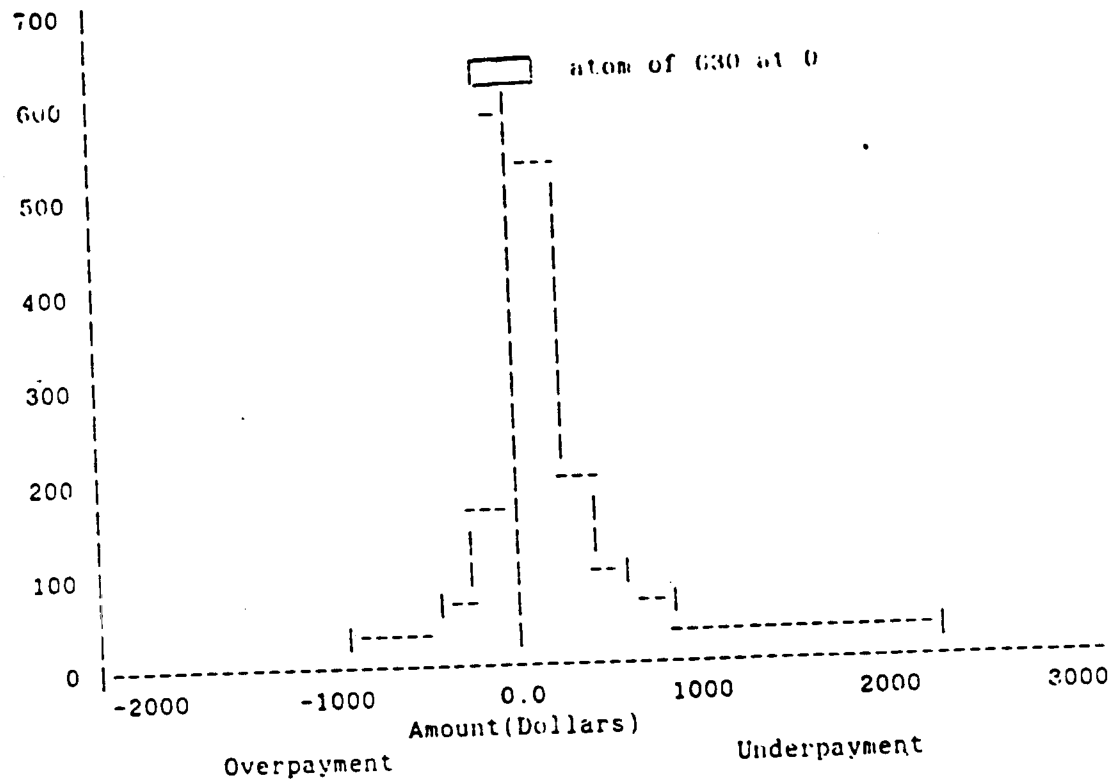
Error Corrected Estimation

Taxpayers not only underpay their taxes but also frequently overpay them (see Clotfelter (1983)), presumably because they commit errors either in interpreting the tax code or in calculating their tax liability. Taxpayer errors seriously confound the analysis of evasion because they can lead to underpayment of taxes as well as overpayment. Ignoring errors and assuming that all underpayments are the result of deliberate intent biases estimates, leading us to predict too high a probability of evasion for error prone filers.

Figure 1 depicts the distribution of over and underpayment of taxes in our sample: 13.5% of the filers overpay, as compared with 59% who underpay, and the overpayments are of considerably smaller magnitude than the underpayments (the average overpayment is \$573, and the average underpayment \$1707). If taxpayer errors are distributed symmetrically around zero (with a zero error resulting in correct payment), then we would expect about 13.5% of the taxpayers to understate taxes as a result of error. Subtracting this figure from the 59% who actually do underpay, we would conclude that 45.5% of

Figure 1

The Difference Between Reported and True Tax Liability
(As Calculated by the IRS)



all taxpayers evade. In the terminology which we adopt 45.5% is an "error-corrected" estimate of the evasion rate.

We present a two-stage procedure for taking account of errors which formalizes this principle of error correction. Consider the binary choice model of evasion, in which the filer either does or does not decide to evade, and concurrently either does or does not commit an error. From (4) the probability of evasion is $\Phi(x_1\beta_1)$. Define the probability of error to be $H(v\Psi)$, where v includes filer characteristics which affect the likelihood of error and H is a cumulative distribution function.[12] We assume that the probabilities Φ and H are independent, and that taxpayers who commit errors are equally likely to overpay as to underpay their taxes (the density h corresponding to H is symmetric about 0).

Each return now falls into one of five categories:

(1) No evasion, no error, which occurs with probability $(1-\Phi)(1-H)$

(2) Evasion, no error, which occurs with probability $\Phi(1-H)$

(3) Evasion, error, which occurs with probability ΦH

(4) No evasion, positive error, which occurs with probability $(1-\Phi)H/2$

(5) No evasion, negative error, which occurs with probability $(1-\Phi)H/2$.

Taxpayer returns fall into 3 classes: understatement of tax liability, correct statement, and overstatement. We make one additional simplifying assumption: that when both evasion and error occur the magnitude of the evasion is always larger than that of the error, so that on net the

[12] Lying behind this formulation is a "two-sided" tobit model of the error process, in which the probability of no error is 1-H, and the probability of a positive or negative error in the amount x is $(1/\sigma_h)h((v\Psi-x)/\sigma_h)$. It is easy to show that the binary choice version of this model, in which only whether or not a postive or negative error has been committed is observed, and not its magnitude, produces consistent estimates of Ψ up to a scaling factor (see Amemiya (1985) for an analogous proof that probit estimates are consistent for a tobit model, up to a scaling factor). See the conclusion of this subsection and the appendix.

filer underpays his taxes.[13] Under this assumption the probabilities associated with the three classes are:

$$
\begin{aligned}
P(\text{understatement}) &= \phi - (1-\phi)H/2 \\
P(\text{correct statement}) &= (1-\phi)(1-H) \\
P(\text{overstatement}) &= (1-\phi)H/2
\end{aligned}
\tag{11}
$$

Indexing returns by i and using C_1, C_2, and C_3 to denote these three classes, the log likelihood of the sample is:

$$L = \sum_{i\epsilon C_1} \log[\phi_i+(1-\phi_i)H_i/2] + \sum_{i\epsilon C_2} \log[(1-\phi_i)(1-H_i)] + \sum_{i\epsilon C_3} \log[(1-\phi_i)H_i/2] \tag{12}$$

Equation (12) could be estimated jointly over the parameters β_1 of ϕ and Ψ of H. Instead we propose a simpler two-stage procedure. Stage 1 uses only data from the last two classes of returns, which consist of filers who have either correctly stated or overstated their tax liability. The conditional likelihood of a return falling into each of these classes (given that it has fallen into one of the two, but not into C_1) is then:

$$
\begin{aligned}
P(\text{correct}|\text{correct or overstatement}) &= [(1-\phi)(1-H)]/[(1-\phi)(1-H/2)] \\
&= (1-H)/(1-H/2) \\
P(\text{over}|\text{correct or overstatement}) &= H/[2(1-H/2)]
\end{aligned}
\tag{13}
$$

and the conditional log likelihood is

$$L = \sum_{i\epsilon C_2} \log[(1-H_i)/(1-H_i/2)] + \sum_{i\epsilon C_3} \log[H_i/(2(1-H_i/2))] \tag{14}$$

Notice that these probabilities are independent of ϕ. Therefore we can use this first stage to estimate the parameter vector Ψ by applying maximum likelihood to (14), without having specified ϕ or β_1. In the second stage we return to (12) and estimate the parameter vector β_1 of ϕ over the entire sample, using the fitted values for H from the first stage, $H = H(v\Psi)$, as

[13] This assumption is, to some extent, supported by the data: overpayments of tax are on average much smaller than underpayments.

weights. These second round estimates of β_1 are then "error corrected" in that consistent estimates of the likelihood of error have been explicitly included in the model. The first stage error corrector is itself of interest since it provides estimates of the relationship between taxpayer characteristics and error proneness.

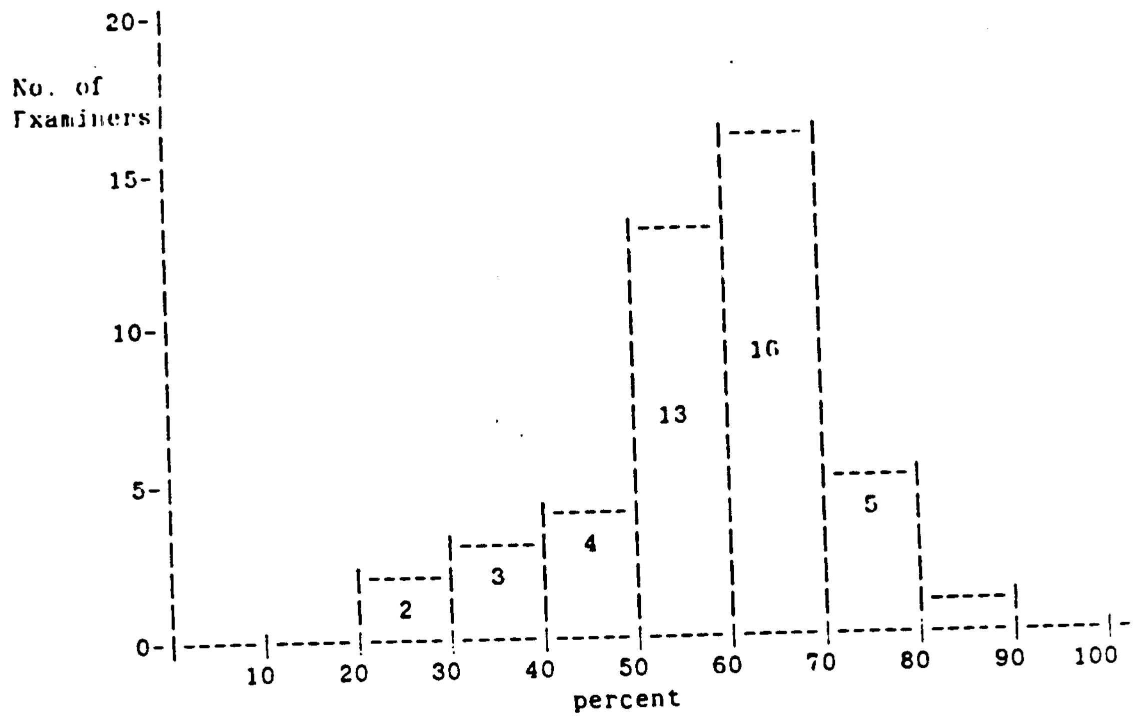
The error correction principle extends readily to the bivariate tobit model of evasion. First, the binary choice error correction model (based on equations (12) and (14)) is estimated, producing error corrected estimates of β_1 and corresponding error corrected weights ϕ_1. Next, the tobit analogue of the error corrector in equation (14) is estimated: estimation is restricted to returns with correct statement or overstatement of liability, with the probability of no error being set to $1-H(v\psi/\sigma_h)$, and the probability of a positive error x being set to $(1/2)(1/\sigma_h)h((v\psi-x)/\sigma_h)$, where h is the density of H. Finally, the bivariate tobit analogue of equation (12) is estimated, using both the error corrected binary choice weights ϕ_1 and the tobit error correcting weights ψ and σ_h. In the appendix we present a more formal development of this procedure.

Detection Controlled Estimation

We now relax the assumption of complete detection and take the more realistic view that IRS examiners fail to detect evasion in some case. Our discussion draws on a more extensive treatment in Feinstein (1986a).

Figure 2 illustrates the variation in detection rates amongst the 44 examiners with 15 or more cases in our sample. Amongst these examiners evasion was detected on 55% of all returns audited. As can be seen the variation in detection rates amongst the examiners is substantial, which suggests that modeling the detection process is important. Of course figure 2 is too simple, because it does not control for the differences in the types of

Figure 2 : Histogram of Detection Rates - Raw Data.



taxpayers assigned to different examiners. The Detection Controlled Estimation (DCE) method which we now present does control simultaneously for variations in taxpayer and examiner characteristics.

We first derive the binary choice version of the DCE estimator. As a simplification we assume that when detection does occur it is complete, so that the examiner uncovers the full extent of evasion.

Taxpayer returns fall into two disjoint sets: the set A, which consists of returns for which evasion has been fully detected; and the set A^c, which consists of returns for which evasion has not been detected. To derive the estimator, consider first returns in A. For return i to fall in A two events must have occurred in succession. Firstly, filer i must have chosen to commit evasion, an event which occurs with probability $\Phi(x_{1i}\beta_1)$. Secondly, the IRS examiner assigned to return i, whom we will call examiner j, must have successfully detected the evasion. Previous analysts have implicitly assumed that this detection probability is one, but the detection model of section I allows us to relax this assumption and to calculate the detection probability explicitly as $G(x_{2i}\Pi+\mu_j)$. Assuming that the errors ε_{1i} of the evasion equation and ε_{3i} of the detection process are independent of one another, the probability of the two events occurring in succession is:

$$\Phi(x_{1i}\beta_1)G(z_i\Pi+\mu_j) \qquad (15)$$

Next consider a return in the set A^c. Evasion has not been detected on this return, but that does not mean that evasion has not occurred. Returns in A^c fall into two classes: those for which evasion was not committed, and those for which evasion was committed but not detected. The inability to distinguish between these two classes of returns is at the heart of the nondetection problem. Fortunately, the technique of Maximum Likelihood can still be applied to returns in A^c. First the probability of the filer not committing evasion is calculated. Then the probability of the filer committing evasion which is not detected is calculated. Finally these two disjoint probabilities are summed to give the overall probability of the return falling in A^c. The probability of no evasion is just $1-\Phi(x_{1i}\beta_1)$ and the probability that evasion was committed and not detected is:

$$\Phi(x_{1i}\beta_1)[1-G(z_i\Pi-\mu_j)] \tag{16}$$

Collecting terms, the probability of the return falling into A^c is:

$$\begin{aligned} &1-\Phi(x_{1i}\beta_1) + \Phi(x_{1i}\beta_1)[1-G(z_j\Pi-\mu_j)] \\ = \quad &1 - \Phi(x_{1i}\beta_1)G(z_i\Pi-\mu_j) \end{aligned} \tag{17}$$

Combining (15) and (17) together yields the DCE log likelihood of the entire sample:

$$L = \sum_{i\epsilon A} \log[\Phi(x_{1i}\beta_1)G(z_i\Pi+\mu_j)] + \sum_{i\epsilon A^c} \log[1 - \Phi(x_{1i}\beta_1)G(z_i\Pi+\mu_j)] \tag{18}$$

The DCE method extends to the bivariate tobit model of evasion in a straighforward manner. We assume that ϵ_{3i} is independent of both ϵ_{1i} and ϵ_{2i}. For returns in the set A, the probability of detected evasion in the amount Y_{2i} is $\Phi(x_{1i}\beta_1)\phi^*_{2i}G(z_i\Pi-\mu_j)$, where ϕ^*_{2i} was defined in Section I and depends on Y_{2i}. For returns in the set A^c, the probability of no evasion is 1-$\Phi(x_{1i}\beta_1)$, while the probability of undetected evasion is:

$$\int_0^{+\infty} \Phi(x_{1i}\beta_1)\phi^*_{2i}[1-G(z_i\Pi+\mu_j)]\, dY_{2i} \tag{19}$$

where ϕ^*_{2i} again depends on Y_{2i}. The integral in (19) arises because the quantity Y_{2i} of evasion is unobservable (since it is not detected). As long as the probability of detection $G(z_i\Pi-\mu_j)$ is independent of the amount of evasion Y_{2i} (notice that the detection probability may still depend upon examiner j's expectation of Y_{2i}), an assumption which we will maintain, (19) simplifies to $\Phi(x_{1i}\beta_1)[1-G(z_i\Pi+\mu_j)]$, since the integral of ϕ^*_{2i} is one.

Collecting terms, the probability of return i falling in A^c is

$$1 - \Phi(x_{1i}\beta_1)G(z_i\Pi-\mu_j)$$

just as in the binary choice case. The log likelihood of the entire sample for the bivariate tobit DCE is then:

$$L = \sum_{i\epsilon A} \log[\Phi(x_{1i}\beta_1)\phi^*_{2i}G(z_i\Pi-\mu_j)] + \sum_{i\epsilon A^c} \log[1-\Phi(x_{1i}\beta_1)G(z_i\Pi-\mu_j)] \quad (20)$$

Equation (20) could be estimated directly. Instead, we exploit the fact that the binary choice estimates of β_1, Π, and the μ_j's based on equation (18) are also consistent estimates of these parameters in the bivariate tobit model (20). The proof of this proposition follows directly upon integrating each term of the log likelihood in (20) over positive quantities of evasion -- ϕ^*_{2i} integrates to one, and the remaining terms are not affected by the integral, hence (20) becomes exactly (18).[14] Therefore we use the binary choice estimates of β_1 and of the parameters of the detection process as weights in (20), and estimate only over the parameters β_2, σ_2, and ρ of ϕ^*_{2i}.

Detection controlled estimation has several substantive benefits. First, it corrects for biases which may arise when nondetection is ignored. For example, if higher income taxpayers are systematically assigned better quality IRS examiners, their detected evasion rate will be a larger fraction of their true evasion rate than is the case for lower income filers. Non-detection controlled procedures, which implicitly set the detection rate G to 1 and use the detected evasion rate as a proxy for the true rate, will overestimate the relationship between income and evasion.

[14]This result is similar to the earlier well know result quoted in Section I, that the binary choice estimates of β_1 are consistent for this same parameter in the full bivariate tobit model of evasion. See Amemiya for a proof of the related proposition that probit estimates are consistent for the parameters of a tobit model, up to a scaling factor.

Second, the DCE method allows us to investigate a number of hypotheses about the IRS audit process. Likelihood Ratio Tests which compare the fit of the DCE models in equations (19) and (20) to the original non-detection controlled versions of these models allow us to determine whether or not nondetection is a serious problem in the data. Likelihood Ratio Tests which compare DCE models which include examiner fixed effects μ_j to DCE models which omit these effects test for the presence of significant statistical variation amongst examiners in detection abilities.

Lastly, an important byproduct of the DCE method is an estimate of the incidence and extent of undetected evasion. Let N be the total number of returns in the sample, and N^A be the number in A. Bayes' Law allows us to calculate the probability that evasion has been committed but not detected on return i in A^c as:

$$w_i = \Phi(x_{1i}\hat{\beta}_1)[1-G(z_i\hat{\Pi}+\hat{\mu}_j)]/[1-\Phi(x_{1i}\hat{\beta}_1)G(z_i\hat{\Pi}+\hat{\mu}_j)] \qquad (21)$$

where hats denote the Maximum Likelihood Estimates. A consistent estimate of the incidence of undetected evasion in the population is then:

$$(1/N) \sum_{i \in A^c} w_i \qquad (22)$$

and the total incidence of evasion is the sum of (22) plus the rate of detected evasion, N^A/N. To estimate the dollar value of undetected evasion we must calculate

$$r_i = (x_{2i}\hat{\beta}_2)" \qquad (23)$$

the conditional expectation of the extent of evasion given that evasion has occured, as defined in (5)". A consistent estimate of the extent of undetected evasion is then

$$(1/N) \sum_{i \in A^c} w_i r_i \qquad (24)$$

The DCE models which we have presented can be extended in a number of ways. Following the discussion in Section I, the detection equation can be expanded to include a term reflecting the examiner's assessment of the probability of evasion. To estimate the linear rational expectations version of this model, the detection probability $G(z_i\Pi-\mu_j)$ is modified to $G(z_i\Pi+\mu_j+\delta\Phi(x_{1i}\beta_1))$, but otherwise equations (18) and (20) remain unchanged.

Equations (18) (and (20)) become slightly more complex when x_1 (and x_2) includes variables which evaders misrepresent. In particular, x_1 will typically include the taxpayer's true income, and may well include his marginal tax rate, which is calculated at his true income. For taxpayer's who commit evasion reported income falls short of true income. Of course, when the IRS detects the evasion true income is fully observable, and previous analysts, relying on the assumption that all evasion is detected, have been content to use the IRS' estimate. However, when evasion goes undetected true income remains hidden. As a rudimentary solution to this problem we have estimated a binary choice DCE model in which the ordinary bivariate tobit estimates are used to scale up income for those taxpayers who are not caught evading: we add to reported income a term which represents the expectation of evasion given that it occcurs. Equation (17) is then modified to:

$$1-\Phi(x_{1i}\beta_1) + \Phi(x_{1i}^{*}\beta_1)[1-G(z_i\Pi+\mu_j)] \tag{25}$$

where x_{1i}^{*} includes the projected estimate of true income given that evasion has occured, and the original x_{1i} remains appropriate for the first term, which refers to the probability of no evasion.

The DCE method can also be extended to the case in which the detection error is correlated with the evasion errors ε_{1i} and ε_{2i} (see Feinstein (1986a) for related methods). However, even in the case of joint normality this leads to a trivariate normal likelihood density, which is computationally burdensome to evaluate. Hence we have not explored estimation of these models.

Combining Error Correction and Detection Controlled Estimation

We can combine the techniques discussed above to produce estimates which are both error corrected and detection controlled. We describe the steps necessary to produce these estimates here in the text, and provide a more formal development in the appendix. We assume that, within the TCMP sample of examined returns, errors are always detected, since they are computational mistakes made directly on the taxpayer's return, with no attempt at concealment. Returns then fall into 3 categories: apparent understatement of tax liability, due either to error or detected evasion; apparently correct statement, meaning that no errors were committed and no evasion detected; and apparent overstatement of liability, due to a committed error, with no evasion detected. In the binary choice case the error corrector weights are estimated via equation (14), and are used as weights in an expanded detection controlled likelihood which allows for all three categories of returns. The bivariate tobit DCE uses as weights the tobit error corrector estimates and the binary choice error corrected and detection controlled estimates, and maximizes a likelihood only over the parameters of ϕ^*_2.

Section IV - Empirical Results

We present our results in the following order. First we present estimates of the binary choice (probit) models of evasion, considering four principal models in Table 2 and a few extensions in Table 4. We discuss the estimates in Table 2 in some detail, and also pay a good deal of attention to the detection controlled estimates of both Tables 2 and 4, providing histograms of the examiners' implied detection rates and reporting test statistics which demonstrate that nondetection is a significant problem in the data. Finally, Table 5 provides estimates of the evasion probabilities predicted by our models for typical filers. Second, we present estimates of the bivariate tobit models. Table 6 reports estimates of the 4 models analogous to those in Table 2, and Table 7 provides predictions of the likely extent of evasion for typical filers who choose to evade.

Table 2 presents results for the four principle binary choice models: (1) Standard Probit, (2) Error Corrected Probit (EC), (3) Detection Controlled Probit (DCE), and (4) Simultaneously Error Corrected and Detection Controlled Probit (ECDCE). Coefficient signs are similar across all four models, though exact magnitudes vary. Table 3 presents estimates of the error corrector functions used to form weights in EC and ECDCE; both the probit and tobit error correctors are displayed.

Collectively, the four models in Table 2 suggest several broad conclusions about the determinants of tax evasion. First, source of income is apparently a more important factor than the absolute level of income. Filing of a Schedule C form, which indicates that some income is derived from self-employment, has a strong positive impact on the probability of evasion. The Schedule C effect increases slightly when we correct for taxpayer errors, and increases dramatically (70%) when we control for nondetection, probably

Table 2: Binary Choice Models

	Model I (Basic)	Model II (EC)	Model III (DCE)	Model IV (ECDCE)
Constant1	-1.18 (0.0345)	-1.45 (0.0469)	-1.19 (0.0446)	-1.21 (0.0729)
AGI	-2.18E-06* (2.52E-07)	-4.37E-06* (2.63E-07)	-3.96E-05* (6.55E-06)	-1.74E-06 (4.69E-06)
Marginal Tax	3.65* (0.150)	4.51* (0.195)	5.86* (0.455)	5.87* (0.506)
Capital Gains	7.54E-07* (3.50E-06)	1.38E-06 (1.25E-06)	2.15E-05 (4.31E-05)	3.48E-05 (6.26E-05)
Schedule C	0.973* (0.0719)	1.02* (0.0872)	1.65* (0.354)	15.7* (.134)
Married	0.315* (0.0294)	0.00726 (0.0386)	0.519* (0.0639)	-0.199* (0.0723)
Farmer	0.315* (0.111)	0.316* (0.155)	0.427 (0.237)	0.693* (0.339)
Age 65+	-0.232* (0.0458)	-0.261* (0.0566)	-0.353* (0.0732)	-0.481* (0.0878)
Occ. Group I	3.20E-03 (0.653)	.0480 (0.172)	-.127 (1.20)	0.189 (1.74)
Occ. Group II	0.596* (.195)	0.674* (.143)	25.0* (1.00)	30.0* (1.00)
Detection Eq.:				
Constant2			0.549* (0.052)	-0.0186 (0.0437)
GS Grade			0.233 (0.148)	0.168 (0.166)
Log Likelihood	-1365	-1840	-1275	-1756

Standard errors are in parentheses.

* indicates significance at 5% on a two-sided test.

because Schedule C filers are better able to conceal income and claim invalid business expenses. Capital gains income is also associated with a heightened probability of evasion, and its effect increases in the EC and DCE models.

The coefficient on AGI (Adjusted Gross Income) is uniformly negative. The magnitude of the AGI coefficient is such that its numerical impact is small over most of the range of AGI. However, the marginal tax rate coefficient is always positive, and increases in all the more sophisticated models. There are two possible interpretations of the marginal tax rate effect. One interpretation is that high marginal rates do encourage taxpayers to evade. A second interpretation is that the marginal tax rate, which is a convex function of income, is proxying for a nonlinear relationship between income and the probability of evasion. To investigate this second possibility we have fit models which include higher order polynomial terms in AGI (AGI^2 and AGI^3). In the probit models these terms are insignificant and are not presented. In the bivariate tobit models presented below, however, these terms are significant, and the coefficient on AGI itself is positive and significant, whereas the coefficient on the marginal tax rate is negative and significant. Given this pattern of sign switches between the two models, we are not confident that the income and marginal rate effects can be untangled from one another. Overall, the combined effect is for increasing income to significantly increase the probability of evasion in the binary choice models, but to exert little effect on the extent of evasion, conditional on evasion occuring, in the bivariate tobit models.

A second conclusion we draw from Table 2 is that certain socioeconomic groups are considerably more likely to evade than others. Farm filers are more likely to evade, and are also more likely to commit errors (see Table 3). This could be due in part to the difficulty of determining a farm's income stream for a given tax year, which would suggest that it is the nature of farmers' income sources which matters. The farm coefficient increases in the

DCE and ECDCE models. However, error correction lowers its significance -- compare columns 1 and 2, which show a drop of about 50% in the value of the farm coefficient's t-ratio, leaving it only marginally significant.

Older (age 65+) filers seem to be uniformly less likely to evade than younger filers. This may be because older filers are inherently more honest (a view held by many IRS personnel), or simply because older individuals, especially those who are retired, possess less complex income patterns which make evasion more difficult to conceal. This second view is reinforced by the estimates of Table 3, which indicate that older filers are also less likely to commit errors.

Married taxpayers seem to be significantly more likely to evade than single taxpayers in the basic probit model. EC, however, lowers the effect of married status, rendering it statistically insignificant. While DCE raises the marriage effect, the combination of DCE and EC is to make married filers less likely to evade. This evidence suggests that married filers make more mistakes than single filers, perhaps because of the added complexity of a married person's return. This hypothesis is further supported by Table 3 -- married filers are significantly more likely to commit errors than single filers.

The results of Table 2 also indicate that occupation groups 1 and 2 exhibit different evasion behavior than the general population. However, the signs and magnitudes of the occupation group dummies vary substantially across the four models, which prevents us from reaching any strong conclusions about the nature of these groups' evasion tendencies.

A third conclusion which emerges from Table 2 is that controlling for nondetection, and allowing for variation in detection rates across IRS examiners, is important. A Likelihood Ratio Test based on Models I and III of Table 2 allows a test of the null hypothesis that detection of evasion is complete amongst all IRS examiners. The Test Statistic is 180.0, which

Table 3: Error Correction Equations

	Model I (Probit)	Model II (Tobit)
Constant	-1.147	-2483
	(0.0480)	(109)
AGI	-4.253E-07	0.0107*
	(9.08E-07)	(0.00144)
Schedule C	0.571*	1184*
	(0.136)	(252)
Married	1.159*	1551*
	(0.0532)	(131)
Farmer	0.337*	1437*
	(0.149)	(259)
Age 65+	-0.228*	331*
	(0.0776)	(154)
Sigma	-	1805*
		(57)
Log Likelihood	-472.2	-3724

Standard errors in parentheses.

* indicates significance at the 5% level for a two-sided test.

strongly rejects the null (the 95% critical value of a $\chi^2(46)$ is 62.8) in favor of the alternative conclusion that detection is imperfect.

Table 4 presents a series of extensions of the models in Table 2. Model I is a detection controlled probit without examiner effects. A comparison of Model I of Table 4 with Table 2's Model III allows us to test for the joint significance of the examiner fixed effect dummies. The log likelihood for Model III of Table 2 is -1275, while that of Model I of Table 4 is -1382. The Likelihood Ratio Test Statistic for the joint significance of the 44 examiner effects is thus 214, while the 5% critical value for the $\chi^2_{(44)}$ distribution is 60.481. We therefore strongly reject the hypothesis that all examiners have the same efficiency in detecting evasion in favor of the hypothesis of differing detection abilities.

Figures 3 and 4 present the estimated examiner detection rates. The upper half of each figure is the same as Figure 1, for the purposes of comparison. The lower half of Figure 3 gives the estimated detection probabilities for the 44 examiners assigned dummies, computed as $\Phi(x_{2i}\Pi+\mu_j)$. Notice that the estimated histogram has shifted to the right, as compared with the raw histogram, with 12 examiners being placed in the 90-100% detection probability class. This rightward shift is to be expected, because controlling for nondetection can only raise an examiner's detection rate, not decrease it. If we divide the N cases audited by an examiner into three groups -- N_1 cases in which evasion is detected, N_2 cases in which evasion occurs but goes undetected, and N_3 cases in which there is no evasion -- then the raw detection rate is N_1/N and the prediction from the DC model estimates $N_1/(N_1+N_2)$. Apart from sampling error, the second ratio must be at least as large as the first. Figure 4 gives the estimated detection probabilities from Model IV of Table 2. Again, there is a shift to the right in the distribution, this time more pronounced than in Figure 3.

Table 4: Additional Probit Models

	Model I (DCE-no) (fixed effects)	Model II (DCE with) (projection)	Model III (EC-DCE with) (projection)	Model IV (DCE with) (expectation)
Constant1	-1.49 (0.05)	4.32 (172.7)	-1.53 (0.11)	-0.956 (0.052)
AGI	1.01E-06 (1.32E-06)	-6.47E-06 (7.19E-04)	4.32E-05* (4.25E-07)	-9.28E-07* (3.55E-07)
Marginal Tax	12.22* (0.46)	2.49* (1.00)	6.93* (0.65)	9.66* (0.58)
Capital Gains	1.76E-05 (1.14E-04)	1.86E-06 (1.70E-03)	-6.11E-05* (7.77E-06)	1.28E-06 (2.73E-04)
Schedule C	1.38* (0.42)	2.27 (30.7)	1.43* (0.31)	1.47* (0.43)
Married	0.648* (0.146)	0.456 -	-1.02* (0.12)	0.412* (0.192)
Farmer	0.531 (0.795)	6.99 -	0.896* (0.435)	1.27 (1.80)
Age 65+	-0.753* (0.147)	-0.0124 (13.37)	-0.188 (0.147)	-0.171 (0.118)
Occ. Group I	2.99* (1.12)	0.714 (3397)	-21.0 -	-1.08* (0.520)
Occ. Group II	1.92* (0.466)	25.0* -	29.0* -	-.725* (0.156)
Detection Eq.:				
Constant2	0.0651* (0.0199)	-2.11* (0.009)	-0.396* (0.036)	6.84* (1.04)
GS Grade	0.609* (0.096)	0.262* (0.043)	0.688* (0.167)	0.352 (0.545)
Evader Probability				-5.89* (1.04)
Log Likelihood	-1382	-2823	-3632	-256

Standard errors are in parentheses.
* indicates significance at 5% on a two-sided test.

Figure 3 : Histograms of Detection Rates
Before and After Correction.
(Model II)

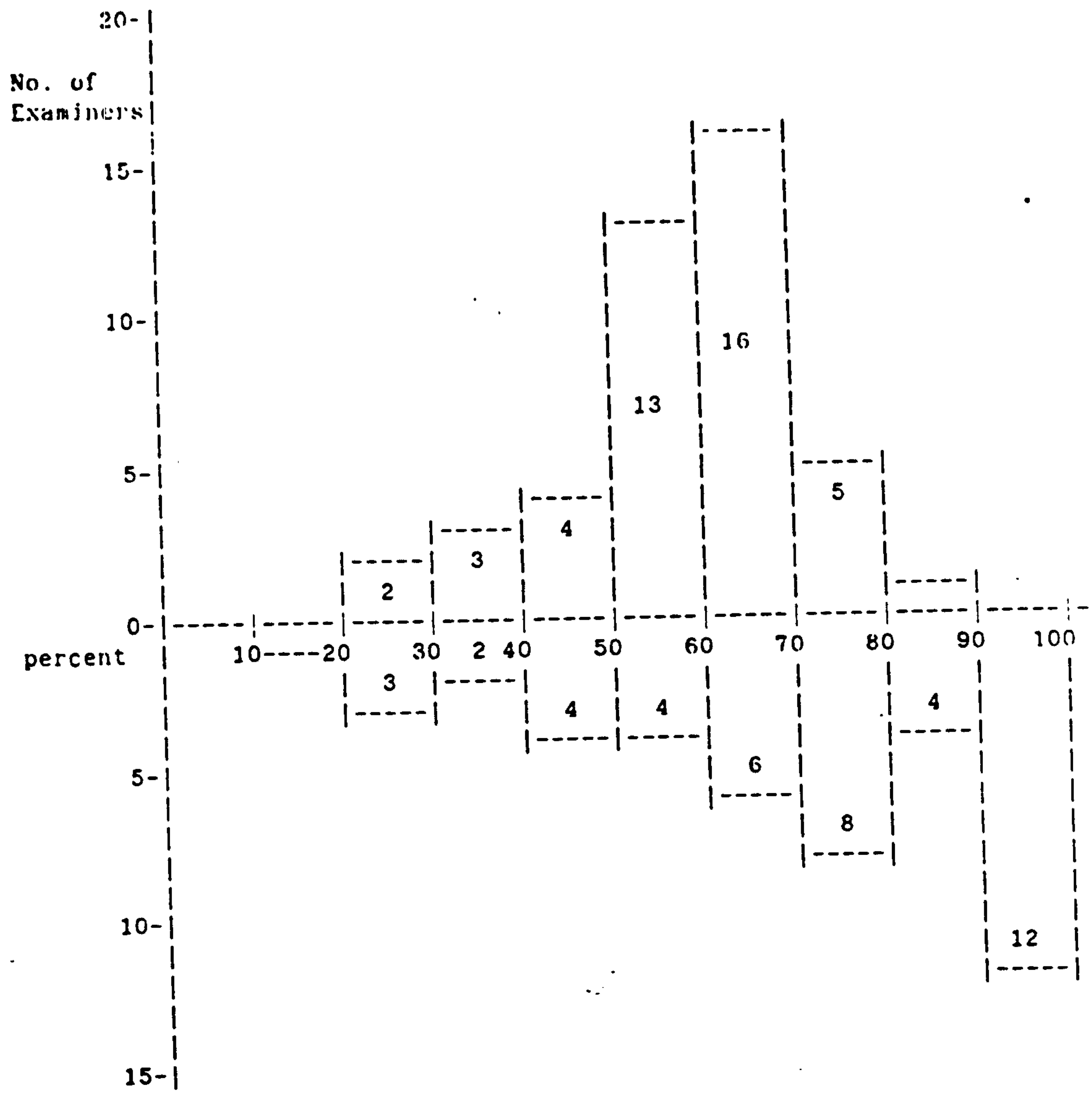
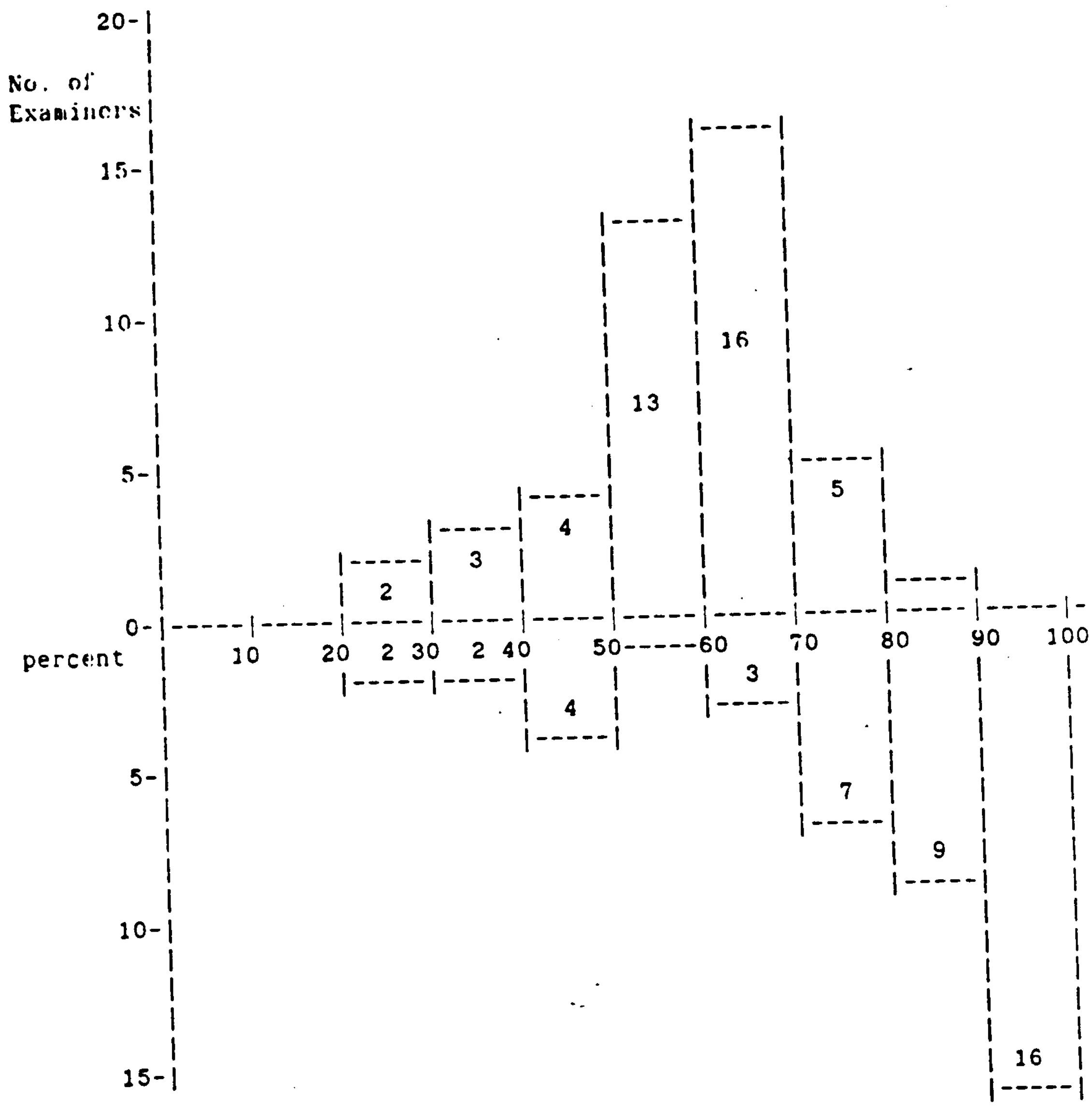


Figure 4 : Histograms of Detection Rates
Before and After Correction.
(Models IV)



The detection controlled estimates also allow us to evaluate the relationship between an examiner's GS grade and his detection rate. The coefficient on GS grade is positive in all the detection controlled models, which indicates that examiners with higher GS grades do perform better than their lower-ranking colleagues; but the coefficient is statistically insignificant most of the time.

The remaining models in Table 4 fall into two categories. The first category consists of models II and III, which correct for the fact that the AGI (and marginal tax rate) recorded for undetected evaders will be too low (see Section III) by projecting the extent of understatement using the standard bivariate tobit model estimates of Table 6. As can be seen, these models are not well-identified, the likelihood surface being locally flat in all directions around several of the coefficients. The second category, Model IV, includes the predicted evasion probability of the filer as an argument of the examiner's detection equation. Unfortunately, the estimated coefficient is of the wrong sign.

Table 5 contains estimated evasion probabilities for a variety of "typical" filers, using the coefficients from Table 2. Comparing the first two columns, we find that, for married filers, error correction significantly lowers the probability of evasion, while it makes much less difference for single filers. The Schedule C filer has a very high probability of evasion in all models, as discussed above. And older filers have a uniformly low probability of evasion, also as discussed above.

We now discuss estimates of the bivariate tobit models of evasion. Table 6 presents estimates of four principle models, which are analogous to those presented in Table 2: (1) Standard Bivariate Tobit, (2) Error Corrected Bivariate Tobit (EC), (3) Detection Controlled Bivariate Tobit (DCE), and (4) Simultaneously Error Corrected and Detection Controlled Bivariate Tobit (ECDCE). The variable AGI has been divided by 100,000 as a

Table 5 : Implied Probabilities of Evasion for Representative Taxpayers.

	Model I (Simple Probit)	Model II (EC Probit)	Model III (DCE Probit)	Model IV (ECDCE Probit)
Married Filers:				
$25000 AGI	0.497	0.336	0.422	0.506
$50000 AGI	0.736	0.627	0.471	0.861
$75000 AGI, Sched "C" Filer	0.959	0.929	0.838	1.000
$30000 AGI, Sched "F" Filer	0.671	0.520	0.605	0.825
$10000 AGI, Age 65+	0.296	0.152	0.315	0.166
Single Filers:				
$25000 AGI	0.517	0.508	0.449	0.788
$50000 AGI	0.704	0.721	0.405	0.949
$75000 AGI, Sched "C" Filer	0.928	0.934	0.701	1.000
$30000 AGI, Sched "F" Filer	0.702	0.705	0.653	0.962
$10000 AGI, Age 65+	0.230	0.185	0.204	0.277

Table 6: Bivariate Tobit Models

	Model I (Basic)	Model II (EC)	Model III (DCE)	Model IV (ECDCE)
Constant1	1052.0 (80.1)	1466.0 (19.2)	973.2 (146.7)	2166.0 (58.6)
AGI	3942.0* (219.4)	5862.0* (41.4)	4008.0* (653.9)	10640.0* (199.9)
AGI^2	-1147.0* (78.7)	-1899.0* (2.7)	-1308.0* (282.2)	-3387.0* (61.5)
AGI^3	65.7* (4.8)	103.5* (0.4)	77.6* (16.8)	188.1* (19.8)
AGI>$40,000 (Dummy Var.)	137.8* (53.8)	208.7* (27.4)	222.1 (141.1)	-49.6 (32.9)
Marginal Tax	-5167.0* (256.1)	-7654.0* (96.0)	-5681.0* (712.7)	-12970.0* (359.7)
Capital Gains	-6.97E-03 (3.76E-03)	-4.52E-03* (5.04E-04)	-6.23E-03 (1.20E-02)	-6.98E-03 (7.39E-03)
Schedule C	527.8* (51.7)	432.6* (19.2)	986.6* (103.5)	749.3* (59.9)
Married	-48.2 (32.6)	-765.2* (16.7)	73.1 (92.1)	-1285.0* (18.2)
Farmer	430.0* (69.7)	53.5 (37.7)	516.8* (179.9)	228.7* (93.9)
Age 65+	32.8 (42.6)	-101.3* (23.5)	0.93 (32.3)	-174.9* (58.4)
Sigma	441.6 (79.6)	235.4 (4.9)	948.9 (86.9)	442.1 (14.0)
Rho	0.6* (1.34E-03)	0.6* (7.17E-03)	0.7* (.0187)	0.8* (.0143)
SSR	1.76E+09	3.10E+09	1.80E+09	3.01E+09
No. Obs.	1337	1337	1337	1337

Standard errors are in parentheses.

* indicates significance at 5% on a two-sided test.

normalization and the occupation dummies were found to be insignificant and are omitted.

Just as in Table 2, coefficient signs are similar across all four models in Table 6, though magnitudes vary somewhat. Coefficient signs and relative magnitudes are also comparable to those in the corresponding probit models, except for the income and marginal tax rate effects. The coefficient on AGI is now positive and highly significant, while that on the marginal tax rate is negative and strongly significant. Higher order AGI terms are also significant, suggesting a complex nonlinear relationship between income and evasion in our sample. As we discussed earlier, these results in tandem with the probit results indicate that income and marginal rate effects cannot be separately identified. As Table 7 (to be presented next) makes clear the combined effect is for increasing income to increase the extent of evasion, just as it also increases the probability of evasion (Table 5). We note that the correlation ρ between the probit and tobit models is positive and significantly different from zero in all models.

Table 7 provides estimates of the likely extent of evasion for typical filers who choose to evade taxes, based on the numbers in Table 6. Notice that income exerts a mild effect and occasionally negative effect on the extent of evasion: a married filer with $50,000 income evades $719, while his counterpart with $25,000 evades $855. Source of income exerts a considerably stronger impact on the extent of evasion. Thus a married Schedule C filer with net income of $75,000 evades $1467, which is roughly twice the corresponding amount for the non-Schedule C filer. The magnitude of evasion also varies considerably amongst different socioeconomic groups; it is especially high for farmers, and is higher for single filers than for married filers.

Table 7 : Conditional Expectation of the Extent of Evasion for Representative Taxpayers.

	Model I (Simple Tobit)	Model II (EC Tobit)	Model III (DC Tobit)	Model IV (ECDC Tobit)
Married Filers:				
$25000 AGI	$855	$321	$1271	$457
$50000 AGI	$719	$208	$1165	$227
$75000 AGI, Sched "C" Filer	$1467	$962	$1964	$1346
$30000 AGI, Sched "F" Filer	$1162	$284	$1494	$474
$10000 AGI, Age 65+	$911	$233	$1354	$331
Single Filers:				
$25000 AGI	$484	$292	$838	$390
$50000 AGI	$533	$400	$937	$399
$75000 AGI, Sched "C" Filer	$1478	$1649	$1963	$2500
$30000 AGI, Sched "F" Filer	$677	$220	$966	$386
$10000 AGI, Age 65+	$849	$682	$1290	$986

Section V : Projections.

We use the estimates in Table 2 to investigate the likely incidence of tax evasion under several recent and anticipated U.S. tax systems. We do not present estimates of the extent of evasion under these tax regimes because we are not confident that separate effects of the marginal tax rate and income are identified in the bivariate tobit models of Table 6.[15] Projections based on the standard probit model are presented in Table 8, while those based on the more sophisticated ECDCE model are presented in Table 9. In each case the income levels in the leftmost column are in 1982 dollars. Looking at the third line of Table 8, we ask "What is the probability that a married joint filer with an income equivalent to $75000 at 1982 prices and some income ($5000) derived from a sole proprietorship (Schedule C income) would evade taxes in 1980, 1985 and 1988?" We see that the probability falls from 0.97 to 0.87. The three years in the tables were chosen to represent three distinct stages in the recent history of the U.S. tax system. First, in 1980, marginal rates ranged from 14% to 70%. The tax reform of 1981 cut rates progressively, and the 1985 column represents the conclusion of this reform. In that year, marginal rates ran from 11% to 50%. In 1986, the Tax Reform Act was passed. Its provisions will come fully into force in 1988, which is why 1988 was chose as the third column. In 1988 there will be only three marginal tax rates: an initial rate of 15%, a higher rate of 28%, and a penalty rate of 33% that will apply to income in a certain range -- $43,150 to $89,560 of taxable income for a single filer -- and which is designed to bring the average tax rate up towards the marginal rate of 28%.

In both tables the incidence of evasion falls from 1980 to 1988. Table 9 presents estimates which have been corrected for taxpayer errors and

[15] The estimated extent of evasion increases as the tax system moves towards lower marginal rates --- contrary to what we expect.

Table 8 : Predicted Probabilities of Evasion under Different Tax Systems
Basic Probit Models

	1980	1985	1988
Married Filers:			
$25000 AGI	0.49	0.49	0.35
$50000 AGI	0.73	0.66	0.51
$75000 AGI,"C" Filer	0.97	0.93	0.87
$30000 AGI,"F" Filer	0.66	0.66	0.47
$10000 AGI, Age 65+	0.30	0.27	0.28
Single Filers:			
$25000 AGI	0.51	0.44	.41
$50000 AGI	0.77	0.59	0.46
$75000 AGI,"C" Filer	0.97	0.91	0.79
$30000 AGI,"F" Filer	0.69	0.62	0.53
$10000 AGI, Age 65+	0.23	0.20	0.19

Table 9 :Predicted Probabilities of Evasion under Different Tax Systems
ECDC Probit Models

	1980	1985	1988
Married Filers:			
$25000 AGI	0.49	0.50	0.28
$50000 AGI	0.85	0.77	0.55
$75000 AGI,"C" Filer	1.00	1.00	1.00
$30000 AGI,"F" Filer	0.81	0.81	0.54
$10000 AGI, Age 65+	0.17	0.14	0.15
Single Filers:			
$25000 AGI	0.77	0.69	0.65
$50000 AGI	0.97	0.88	0.73
$75000 AGI,"C" Filer	1.00	1.00	1.00
$30000 AGI,"F" Filer	0.96	0.92	0.86
$10000 AGI, Age 65+	0.28	0.22	0.20

nondetection. A married filer with a $50000-equivalent income experiences a decline from 0.94 to 0.70 in his evasion probability and his unmarried counterpart experiences a decline from 0.97 to 0.77. At very low incomes, the effect of the reforms is much less marked, reflecting the comparatively small changes in marginal rates faced.

The detection controlled estimates of evasion also allow us to calculate the incidence and extent of undetected evasion in the 1982 U.S. filing population. These projections are calculated using the formulae presented in the subsection "Detection Controlled Estimation" of Section III. The projections are based on the weighted sample and have been scaled-up by the ratio of the size of our sample to the overall TCMP sample; hence they fully reflect the U.S. taxpaying population. The DCE probit model predicts 15.5 million undetected evaders (16% of all filers), while the ECDCE probit model predicts 14.7 million (15%).

The bivariate tobit ECDCE model predicts the dollar value of unreported taxable income by filers to be $62 billion, a number which is large but comparable to previous IRS estimates.[16] The assumptions under which this estimate were derived are: (1) that all taxpayer errors are detected -- these errors are screened out of this estimate; and (2) that when evasion is detected, its full extent is uncovered, which suggests that the $62 billion is something of a lower bound on actual unreported income. As a point of comparison, the aggregate value of detected unreported income is $81 billion, so that the ratio of undetected to detected unreported income is .77 to 1. Our multiplier of .77 is one piece of the IRS multiplier of 3.28 estimated in the IRP study (see the introduction).

[16] The DCE model yields an estimate of undetected evasion of only $9.8 billion.

Section VI : Conclusions.

In this paper we have presented a comprehensive microeconometric analysis of income tax evasion. We have paid particular attention to two important econometric issues: taxpayer errors, and the fact that not all evasion is detected. We have also extended the traditional tobit analysis of evasion to a more general bivariate model in which the filer's decision whether or not to evade is evaluated separately from his decision of how much to evade.

We draw three main conclusions from our work. First, the sources of income -- as opposed to the level of income -- are major determinants of evasion, presumably because different types of income offer very different possibilities for evasion. Second, the combined effect of income and the marginal tax rate is a positive and significant effect on the probability of evasion, but only a weak effect on the conditional extent of evasion. Third, there is substantial variation amongst IRS examiners in detection rates, and there is reason to believe that much evasion goes undetected - perhaps as many as 15 million filers evaded taxes in 1982 but were not caught.

We hope, in the future, to extend our work, particularly to allow for fractional detection - the case in which IRS examiners only detect a fraction of the true evasion when they find any at all. We also look forward to research that estimates separate evasion and detection equations for different types of income (e.g. capital gains), and to models which are more explicitly derived from expected utility theory.

References

[illegible]lingham, M.G. and A. Sandmo. "Income Tax Evasion: A Theoretical Analysis". Journal of Public Economics, 1972.

Amemiya, T. "Regression Analysis When the Dependent Variable is Truncated Normal". Econometrica, 1973.

Amemiya, T. Advanced Econometrics. Harvard U.P. 1985.

Clotfelter, C.T. "Tax Evasion and Tax Rates: An Analysis of Individual Returns", Review of Economics and Statistics, 1983.

Feinstein, J.S. "Detection Controlled Estimation", mimeo, M.I.T., 1986.

Graetz, M.J., J.F. Reinganum, and L.L. Wilde. "The Tax Compliance Game: Toward an Interactive Theory of Law Enforcement". Caltech Social Science Working Paper 589, 1985.

Groves, H.M. "Empirical Studies of Income-Tax Compliance", National Tax Journal, 1958.

Heckman, James J. "Sample Bias as a Specification Error". Econometrica, Vol. 47, 1979.

Internal Revenue Service. Estimates of Income Unreported on Individual Income Tax Returns. 1979.

Internal Revenue Service. Income Tax Compliance Research: Estimates for 1973-81. 1983.

Reinganum, J.F. and L.L. Wilde. "Income Tax Compliance in a Principal-Agent Framework", Journal of Public Economics, 1985.

Sandmo, A. "Income Tax Evasion, Labour Supply, and the Equity-Efficiency Tradeoff", Journal of Public Economics, 1981.

Witte, A. and D. Woodbury. "The Effect of Tax Laws and Tax Administration on Tax Compliance: The Case of the U.S. Individual Income Tax", National Tax Journal, 1985.

Appendix

In this appendix we derive a number of statistical models discussed in Section III.

Error Corrected Bivariate Tobit

Let the probability of a taxpayer committing an error in the amount x (either positive or negative) be $(1/2)(1/\sigma_h)h((x-v\psi)/\sigma_h)$, and the probability of no error be 1-H. We continue to assume that when evasion occurs, it is always of substantially larger magnitude than any error which may occur, so that cross-terms allowing for both evasion and error can be subsumed into terms allowing only for evasion (equivalently we could assume that the probability of both an error and evasion is small). As in the binary choice case returns fall into 3 classes: understatement of tax liability, correct statement, and overstatement. The probability of a misstatement x or correct statement is then:

$$P(\text{understatement } x) = \Phi_1\phi^*_2 + (1-\Phi_1)(1/2)(1/\sigma_h)h \quad \text{(A1)}$$

$$P(\text{correct statement}) = (1-\Phi_1)(1-H)$$

$$P(\text{overstatement } x) = (1-\Phi_1)(1/2)(1/\sigma_h)h$$

where Φ_1 is the binary choice probability of evasion, ϕ^*_2 is, as defined in Section I, the conditional probability of evasion in the amount x, and h is the probability of an error in the amount x. Restricting analysis to only the last two classes of returns, it is easy to see that

$$P(\text{correct}|\text{correct or overstatement}) = (1-H)/[1-H/2] \quad \text{(A2)}$$

$$P(\text{over}|\text{correct or overstatement}) = (1/2)(1/\sigma_h)h/[1-H/2]$$

since the integral of $(1/2)(1/\sigma_h)$ over all positive x is H/2. Therefore just as in the binary choice case the error probabilities can be estimated in a first stage without estimating evasion rates, although in this case the correct estimation procedure is a modified tobit rather than a modified probit (note that the probit error corrector estimates from the binary choice procedure are consistent for ψ up to a scale facor σ_h).

Now return to (A1). Integrating the probability of understatement over all understatements greater than 0, and the probability of overstatement over all overstatements greater than 0, we recover the binary choice version of (A1), which is equation (14) in the text (this follows from two facts: that the integral of ϕ^*_2 is 1; and that the integral of $(1/2)(1/\sigma_h)h$ is H/2). This demonstrates that the error corrected estimates of Φ_1 from the binary choice case remain consistent for the error corrected bivariate tobit, just as in the usual case.

Finally, then, the error corrected bivariate tobit can be estimated using the tobit error corrector weights h and the binary choice weights Φ_1, estimating (A1) only over the parameters of ϕ^*_2.

Combining Error Correction and Detection Controlled Estimation

We suppose that errors are always detected, as discussed in the text. We continue to summarize the probability that evasion is detected by $G(z_j\Pi+\mu_j)$. We also continue to assume that if an error is committed and evasion is committed and detected, the magnitude of the evasion outweighs that of the error. Returns then fall into 3 classes: detected understatement of liability, no detected misstatement, or detected overstatement. The binary choice version of this model associates the following probabilities with these three outcomes:

$$P(\text{understatement}) = G\Phi_1 + [(1-G)\Phi_1 + (1-\Phi_1)]H/2 \qquad (A3)$$

$$P(\text{correct statement}) = (1-\Phi_1)(1-H) + (1-G)\Phi_1(1-H)$$

$$P(\text{overstatement}) = (1-\Phi_1)H/2 + (1-G)\Phi_1 H/2$$

It is easy to show that the conditional likelihood of a return falling into each of the last two classes (conditional on it having fallen into one of these two) is independent of G and Φ_1; therefore the original probit error correcting weights continue to be valid for this model. The parameters of Φ_1 and G are then estimated jointly over all three classes.

Estimation of the bivariate tobit version of this model is analogous and we do not discuss it.