Alternative Methods for Modeling Income Tax Reporting Compliance

Alex Turk, Chih-Chin Ho, and Anne Steuer, Internal Revenue Service

Models of tax misstatement have typically relied on Tobit specifications to account for the large proportion of taxpayers who accurately report their tax liability (Clotfelter, 1983). Under these models, the observed tax misstatements are assumed to be censored from below at zero. Thus, the underlying implication is that tax overstatements are invalid realizations. However, a substantial number of taxpayers are observed via an IRS *Taxpayer Compliance Measurement Program* (TCMP) audit to be overstating their tax liability.[1] When the focus is modeling tax increases only, the censoring assumption may be viable. But, even though the Tobit estimation handles the disproportionate number of zero observations, the assumptions that go along with the model may not be appropriate for modeling audit-induced tax change.[2]

The Tobit specification is potentially problematic for two reasons. First, tax misstatement is assumed to be censored at zero rather than the zero occurring as a result of a distinct choice of compliance versus noncompliance. In the Tobit specification, the probability of observing a nonlimit value is essentially a monotonic transformation of level of misstatement. Thus, a given factor must have the same qualitative effect on the probability of observing a nonzero case as it does on the level of misstatement. Cragg (1971) proposed a more general specification that relaxes this restriction. Second, the Tobit specification requires the assumption that tax overstatements are pure measurement errors. There may be any number of reasons why a taxpayer might pay more tax than legally required. One possibility is that the overstatement was the unintentional result of confusion over the reporting requirements. Another is a deliberate overstatement because of additional costs that would be incurred to be completely accurate.[3] It is also possible the auditor estimated true tax liability with error. That is, there may be times when the auditor recommends a tax liability that is less than the actual (and less than reported) liability. While it is not hard to fathom measurement error existing in audit data, it is unlikely that measurement error exists only for tax overstatements. Using a Tobit specification would require transforming part of the distribution of observed tax misstatement on the basis of measurement error without dealing with the measurement error in the rest of the distribution.

The procedure described by Heckman (1979) can be used if one assumes that zero values for misreporting correspond to nonobservability of tax change. This procedure allows the parameters of the selection process (compliance versus noncompliance) to be estimated without regard to the parameters that determine the level. Thus, while a given factor may increase the expected level of misreporting, it may decrease the probability of observing tax misstatement. This is not possible in the standard Tobit specification. The Heckman procedure estimates separate parameters for the selection equation and misstatement equation. However, the observed zeros are essentially assumed to represent nonobservability of tax misstatement. If the cluster at zero is "real," then this model, as it is typically applied, is not appropriate. In this case, the Heckman procedure is applicable to the noncompliant taxpayers.

An alternative is to assume that the zero values are valid realizations of misstatement. That is, the assumption can be made that there is "real" clustering at zero. While this assumption causes little change in the specification of the model, it identifies which conditional mean is appropriate. In this case, taxpayers are being observed choosing between the compliant and noncompliant state. A value of zero tax change indicates that the taxpayer has chosen to remain in the compliant state. Furthermore, the probability of observing a zero value is not necessarily a monotonic transformation of the level of misstatement.

In this paper, a two-step model, which is a hybrid of the Tobit and Heckman estimations, is developed. The parameter estimates and predictions from this model are then compared with the Tobit estimates and with ordinary least squares. Conclusions and directions for further research conclude the paper.

■ Alternative Model Specification

Assume that a taxpayer must make two choices. First, the taxpayer makes the choice to comply (or not) with reporting requirements and thus accurately report tax liability. The choice between compliance and noncompliance is conditioned on an unobserved or latent variable y^*. While y^* is not observable, compliance versus noncompliance is. Let $y = 0$ represent compliance and let $y = 1$ represent noncompliance. Given that the taxpayer has chosen not to fully comply, the choice is then made as to the level of tax misstatement (TC). The model is formulated as follows:

$$y^* = \mathbf{X}\alpha + u. \qquad (1)$$

where $y = 1$ if $y^* > 0$, and 0 otherwise.

The probability that y^* is greater than 0 is

$$P(y^* > 0) = P(y = 1) = \Phi(\mathbf{X}\alpha). \qquad (2)$$

Tax misstatement, which is only relevant if $y=1$, is

$$TC = \mathbf{X}\beta + \varepsilon. \qquad (3)$$

where

$$(u,\varepsilon) \sim \text{bivariate normal } [0, 0, 1, \sigma^2, \rho].$$

X is a vector of k explanatory variables. α and β are kx1 vectors of the associated parameters.[4] ε and u are random error terms. Φ is a normal cumulative distribution function.

This model differs from the standard Tobit specification in that the latent variable y^* is not necessarily observed when TC is nonzero. In this formulation, as in the Heckman procedure, a given factor, x_i, can have differing impacts on the probability of observing a nonzero tax change and the level of misreporting. The assumption that makes this different from the procedure described by Heckman is that the zero values are assumed to be a valid value of misstatement chosen by the taxpayer. Thus, while in this context, the Heckman procedure provides unbiased estimates of β, the model is only useful for prediction on noncompliant taxpayers. Taxpayers who accurately report their tax liability are ignored in the conditional mean equation.[5] To provide accurate representation on expected compliance *ex ante*, one must incorporate the compliant taxpayers (those with TC=0) into the expectation. The expected value of TC conditional on **X** and $y^* > 0$ is

$$E(TC|X,y^* > 0) = \mathbf{X}\beta + \rho\sigma\lambda(\mathbf{X}\alpha) \qquad (4)$$

where $\lambda(\mathbf{X}\alpha) = \phi(\mathbf{X}\alpha)/\Phi(\mathbf{X}\alpha)$.

The expectation conditional on **X** in this context is

$$E(TC|\mathbf{X}) = \Phi(\mathbf{X}\alpha)(\mathbf{X}\beta + \rho\sigma\lambda(\mathbf{X}\alpha)). \qquad (5)$$

This is in essence an average of the expected values in each state (compliant and noncompliant) weighted by the probability of being in each state. Because tax change is by definition zero in the compliant state, that term is implicit.

Equation (4) is essentially what is estimated in the Heckman procedure. The product of ρ and σ is estimated in the second-stage Ordinary Least Square (OLS) regression. However, the Heckman procedure is typically used to control for "selection bias" that occurs because of ρ, the correlation between the error terms u and ε. The nonobserved cases are typically not of any further interest. However, in the model developed here, the choice of compliance is of interest and must be considered when examining comparative statics and predictions.

Equations (4) and (5) are actually a more general case of the standard Tobit specification. The Tobit model can be obtained by restricting[6]

$$\rho = 1 \text{ and } \alpha = \beta/\sigma.$$

The essence of this restriction is that there is only one source of random variation that taxpayers condition on. X and this error determine the optimal level of misstatement and whether or not the misstatement will be positive. Under these restrictions, the latent variable y^* is observed when TC >0. When these restrictions are relaxed, the latent variable is not observed, only its sign. Thus, the conditional expectation of y^* is not something that can be estimated in the context of the alternative model we are suggesting, as is typically done with the Tobit specification. The alternative model suggests that there is a "switching" between states of compliance and noncompliance that depends upon y^*.

Marginal Effects

The marginal effects of X on tax change, given the censoring, are

$$\partial TC/\partial x_i = \beta\Phi(\mathbf{X}\alpha). \quad \textbf{(6)}$$

Imposing the restrictions of the Tobit model, the marginal effects are

$$\partial TC/\partial x_i = \beta\Phi(\mathbf{X}\beta/\sigma). \quad \textbf{(7)}$$

■ Empirical Estimation

The estimation of the α and β vectors is accomplished using a methodology essentially identical to that described by Heckman (1979) As in Heckman, the procedure involves using a probit model in the first step and an ordinary least squares regression in the second step.

The procedure is as follows. First, using Maximum Likelihood, estimate α from (2) using a probit model. The probit results are then used to calculate an estimate of λ for each observation. Second, estimate β and $\rho\sigma$ from **(3)** via ordinary least squares regression of TC on **X** and the calculated λ.

Exclusion restrictions must be imposed to identify the system. Given the framework of the model, β and $\rho\sigma$ are estimated using those taxpayers with $TC \neq 0$ (the noncompliant taxpayers). The conditional tax change is obtained by replacing the α, β, and $\rho\sigma$ in **(5)** with the estimated parameters.

■ Empirical Comparison of Methodologies

We compare the parameter estimates and the estimated marginal effects in the context of a relatively simple model applied to a random sample of individual filers. We also report prediction error generated by the Tobit specification and the alternative model relative to OLS. The properties of the underlying parameters are certainly of interest. However, we do not deal with them here. We reserve this for future research.

Parameter Estimates

We estimate a model with ten continuous (x_1 through x_{10}) explanatory variables and nine indicator variables (x_{11} through x_{19}) corresponding to broad return type categories.[7] All the variables are derived from tax return characteristics for the TCMP data. The data are a stratified random sample of approximately 54,000 detailed audits of Form 1040 business and nonbusiness filers for 1988. The data contain taxpayer-reported and examiner-determined values for each line item on the return. The estimated parameters and the marginal effects[8] for the Tobit model and the alternative model[9] are reported in Table 1 along with the OLS estimates.

The parameter estimates are sensitive to the specification of the model. With the exception of x_5 and the constant, all the Tobit estimates of the β's are larger than the Alternative model estimates. Furthermore, the predicted proportion of "censored" observations is generally larger for the Tobit model than is actually observed (see Table 2). If the Tobit restriction that $\alpha=\beta/\sigma$ is invalid, this would tend to bias the Tobit estimates of β. Also, the β's for the continuous variables in the Alternative model are very similar to the estimates derived from OLS applied to the full sample. In addition, the estimated coefficient of the selection correction factor, λ, does not seem to be statistically different from zero.[10] This suggests that u and ε may be uncorrelated and that the selection and the misreporting equations could be estimated without estimating λ.

■ Prediction Error

Given the diversity of the parameter estimates, the question can be asked about the ability of the models to provide predictions in censored data. To this end, we split the data into five mutually exclusive categories

Table 1 - Model Estimates

Variable	Tobit Model β	Tobit Model Marginal	Alternative Model β	Alternative Model Marginal	OLS Model
x_1	7.9	3.49	4.66	3.54	4.45
x_2	28.35	18.31	27.84	24.03	22.18
x_3	6.8	3.07	3.13	2.44	3.26
x_4	5.49	3.24	4.55	4.10	4.45
x_5	26.47	20.20	28.76	26.17	24.48
x_6	17.85	2.75	1.41	0.34	2.16
x_7	6.74	1.71	1.17	0.55	1.22
x_8	4.86	1.54	0.86	0.53	1.14
x_9	26.15	9.73	12.25	9.20	12.20
x_{10}	40.89	22.98	38.20	30.14	31.77
x_{11}	4518.5	1147.70	189.31	89.17	479.33
x_{12}	5768.32	1828.56	148.57	92.41	612.01
x_{13}	-2846.1	-1058.75	-3855.80	-2895.71	-3355.01
x_{14}	-2732.1	-1535.44	-7894.73	-6228.94	-6061.14
x_{15}	5191.8	2341.50	-203.01	-158.55	219.40
x_{16}	7423.4	4379.81	827.87	746.74	1340.20
x_{17}	5717.4	4362.38	-1403.61	-1277.29	-419.95
x_{18}	4679.3	2068.25	-502.62	-381.49	-66.31
x_{19}	4503.0	2908.94	-1884.47	-1626.30	-839.25
Constant	-7030.9		458.68		-269.38
λ	4237.15		-380.01		

based on return characteristics. The categories can be defined by the dummy variables in the estimated model. Thus, the average of the OLS predictions will be equal to the average of the actual observed tax misstatement. The relative prediction (sum of the predicted tax changes/ sum of the actual tax changes) and the root mean square of the prediction relative to the OLS root mean square of prediction are reported in Table 2.

The results in Table 2 suggest that a Tobit specification for tax misstatement is inadequate. The model seems to overpredict the censored mean of tax misstatement. The Tobit model is especially problematic when the proportion of compliant taxpayers is very large. In Group 1, where 61 percent of the taxpayers are compliant, the Tobit prediction of expected tax change is more than five times larger than the actual observed tax change. The Tobit is a much closer approximation when the "censoring" is less severe. The alternative specification may have a slight tendency to "underpredict" average tax misreporting, but nowhere near the magni-

Table 2 - Relative Prediction for Censored Mean

	Group 1	Group 2	Group 3	Group 4	Group 5
Tobit Relative Prediction	5.394	2.221	1.708	1.142	1.694
Alternative Relative Prediction	0.984	0.997	1.000	1.000	1.000
Tobit Relative RMSE of Prediction	1.382	1.007	1.106	1.000	1.010
Alternative RMSE of Prediction	1.001	1.000	0.999	1.001	1.002
Observed Proportion Zero	0.606	0.248	0.156	0.091	0.213
Alternative Ave. Prob. of zero TC	0.607	0.249	0.156	0.091	0.213
Observed Prop. Zero or Negative	0.674	0.430	0.242	0.189	0.333
Tobit Ave. Prob. of Zero or Negative TC	0.777	0.597	0.471	0.269	0.504

tude of the Tobit model's overprediction.

As noted before, the Tobit specification also tends to overpredict the proportion of "censored" observations. The Tobit specification overpredicts the proportion of censored observations in all five groups reported in Table 2. An intuitive explanation is that this is the result of forcing the estimating β vector to play the role of the underlying α and β.

Conclusions

Traditional models that deal with limited dependent variables may not be desirable when modeling tax compliance. The fundamental assumptions about the censoring process may be invalid. This appears to cause dramatic inconsistencies in parameter estimates and predictions. Tax reporting compliance models will, in most cases, be empirically estimated with audit data. We have shown an instance with a very broad sample where the estimates and predictions vary widely, depending upon the methodology chosen. Thus, it would seem prudent to re-examine the assumptions imposed by the model. The alternative model we suggest specifies two distinct, although correlated, decisions about reporting.

While this paper has posed a potential problem and a potential solution, there are still many unanswered questions. The statistical properties of these estimates must be examined. Furthermore, the robustness of the parameters to model misspecification and distributional assumptions should be explored. The model could be extended from a binary decision (compliance/noncompliance) to differentiate between the overstatements and understatements.

Footnotes

1 In some return categories, as many as 18 percent of taxpayers are observed via audit overstating tax liability.

2 See Appendix A by Peter Schmidt in Roth *et al.* (1989) for a discussion of the statistical issues in modeling tax compliance.

3 More detailed records, additional forms, opportunity, cost of time to gather information, etc.

4 Some α's and β's could be zero.

5 Because they are assumed to be measurement error in the Heckman procedure. See Greene (1993) pp. 706-714 for a discussion of the Heckman estimation.

6 See Greene (1993) pp. 700-701 for a description of a methodology for testing $\alpha = \beta/\sigma$ in the context of the Tobit model.

7 The definition of each element of **X** will not be disclosed here because the model was developed in conjunction with an operational project. We instead focus on the methodology and modeling framework. The anonymity of **X** does not severely limit our ability to compare the estimates across methodologies.

8 The density and distribution functions for each model are evaluated on the average of the explanatory variables.

9 The alternative β estimates are identical to the Heckman procedure estimates.

10 Weighted OLS standard error = 474.2. Murray (1995) found a similar result when modeling sales tax compliance.

References

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

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

Greene, William H. (1993), *Econometric Analysis,* Second Edition, Prentice-Hall.

Heckman, J., "Sample Selection Bias as a Specification Error," *Econometrica* 47, pp. 153-161.

Murray, Matthew N. (1995), " Sales Tax Compliance and Audit Selection," *National Tax Journal* XLVIII, No. 4, pp. 515-530.

Schmidt, Peter (1989), "Statistical Issues in Modeling Taxpayer Compliance," Appendix A in *Taxpayer Compliance. Volume 1, An Agenda for Research,* edited by Roth, Scholz, and Witte, University of Pennsylvania Press, pp. 337-338.