Abstract

This article provides estimates of the probable saving in the resource cost of complying with the tax law that would result from simplifying the individual income tax law. These estimates are based on an econometric analysis of the tax-filing behavior in 1982 of a sample of Minnesota taxpayers. A simple model of tax-compliance behavior based on utility maximization is first presented in order to suggest the important determinants of compliance behavior. The empirical model treats the discrete choices of whether to itemize deductions and whether to hire professional tax advice, and the choice of how much time and money to spend, conditional on the discrete choices made. Simulations based on the econometric results suggest that significant resource saving could be expected from eliminating the system of itemized deductions, although no significant saving from changing to a single-rate tax structure can be confidently predicted. Results suggest that the Tax Reform Act of 1986 will, in the long run, decrease the use of professional tax assistance, but its net effect on the use of taxpayer's own time is unclear.

THE RETURN TO TAX SIMPLIFICATION: AN ECONOMETRIC ANALYSIS

JOEL SLEMROD
University of Michigan

1. INTRODUCTION

An analysis of any tax simplification proposal ought to consider its effect on the cost of operating the tax system, including both the administrative and compliance costs, along with the distributional and efficiency implications of the proposal. Brannon (1979) and Slemrod (1983) have argued that, in principle at least, the resource costs of operation can be quantified and

AUTHOR'S NOTE: I am grateful for financial support from the Sloan Foundation grant to the Applied Microeconomics Workshop at the University of Minnesota, and the Office of Tax Analysis of the U.S. Department of Treasury. I would also like to thank Tom Mroz for many helpful suggestions. This research was begun while I was a National Fellow at the Hoover Institution of Stanford University.

PUBLIC FINANCE QUARTERLY, Vol. 17 No. 1, January 1989 3-28

considered with the more standard equity and efficiency effects. The principal obstacle to implementing this framework for analysis has been the dearth of quantitative information about the resource cost implications of simplification plans. Although recent work by Slemrod and Sorum (1984) has provided useful estimates of the total resource cost of the current U.S. income tax system, what is more important for policy purposes is the expected *change* in the resource cost due to a proposed tax reform.

The purpose of this article is to provide information on the determinants of the compliance cost of income taxation that can be used to predict the probable resource cost savings from tax simplification. These estimates are based on an econometric analysis of the tax-filing behavior in 1982 of a sample of Minnesota taxpayers. The article is arranged as follows. In Section 2, a simple theoretical model of tax-compliance behavior based on utility maximization is presented. The model is used to suggest the important determinants of compliance behavior to be considered in an empirical study and what the likely direction of influence is. Section 3 develops an empirical model that estimates the determinants of the choice of whether to itemize deductions and whether to hire professional tax advice, and the choice of how much time and money to spend on tax matters, conditional on whether itemization was chosen and/or professional assistance was purchased. Section 4 describes the data for the analysis. In Section 5, the results of the econometric analyses are presented and discussed, and Section 6 briefly describes their implications for the resource saving from some radical tax simplification proposals and the Tax Reform Act of 1986. Section 7 concludes.

2. A SIMPLE THEORETICAL MODEL OF COMPLIANCE BEHAVIOR

The goal of this section is to construct a simple model of the taxpaying household's choices concerning its tax-filing behavior. The household must choose how much of its own time to spend

on tax matters and how much, if any, professional tax advice to purchase. It is assumed that both of these activities uncover legitimate ways to reduce taxable income, and thereby reduce tax liability. The model thus abstracts from the use of illegal reductions in or deductions from taxable income.[1] The decision problem for a representative household can be stated as

$$\underset{I, C, L, H, B}{\text{Maximize}} \quad U(C, L + \gamma H) \tag{1a}$$

subject to

$$C = A + wL - T(A + wL - E - R - (1 - I)S - I(D + P_B B)) - P_B B \tag{1b}$$

where the notation is defined as follows:

A: nonlabor income
B: hours of professional tax assistance purchased
C: consumption of composite good
D: deductions from taxable income (not including payments for professional tax assistance)
E: exemptions
H: hours of own time spent on preparing tax return
I: dummy variable equal to one for households that itemize deductions and equal to zero otherwise
L: labor supply
P_B: price per hour of professional tax assistance
R: reductions in taxable income
S: standard deduction
T: tax function
w: wage rate
γ: labor-equivalent of one hour spent on tax compliance

According to Equation 1a, utility is a function of consumption of a composite good, labor supply, and hours spent on tax matters. Equation 1b states that consumption equals gross income, which consists of exogenously given nonlabor income and labor income, minus taxes paid and payments for profes-

sional tax assistance. Taxable income is gross income minus exemptions, which require no effort to uncover, reductions in taxable income, and deductions from taxable income, which equal $(1 - I)S + I(D + P_BB)$.

Consider a general compliance technology that embodies the relationship between the level of tax liability that can be justified and the inputs to the process. We represent this technology as

$$R = (H, B, Z) \qquad [2a]$$

$$D = (H, B, Z), \qquad [2b]$$

so that both the amount of reductions in taxable income and the amount of deductions from taxable income depend on the amount of own time spent on uncovering either, the amount of professional assistance purchased, and a vector Z of other factors including the type of tax return the household files, personal characteristics such as age, level of education completed, and attitude toward tax matters, and the tax law itself. Time and money spent on uncovering and documenting deductions are treated distinctly from resources spent on finding reductions in taxable income because the former activities are worthwhile only if the household chooses to itemize deductions, while the latter may be worthwhile to any household that has a nonzero marginal tax rate.

The household is faced with a joint decision problem. It must make a discrete choice between itemizing deductions or not. Conditional on that choice, it must decide how much labor to supply and how much to expend on compliance, including its own time and payments for professional tax preparers.

The household will choose whichever itemization status yields the higher level of utility. Without considering compliance costs this decision involves a comparison of S and $D + P_BB$, where $D + P_BB$ is to be interpreted as the amount of available deductions. If $D + P_BB$ exceeds the standard deduction, S, then the household would itemize deductions. However, in the presence of compliance costs the choice is more complicated because it is no longer costless to uncover and document the deductions.

The first-order conditions of the Kuhn-Tucker optimization problem for the conditional choice of H, B, assuming that the choice of L is interior, are

$$t(R_H + ID_H) - \gamma w(1-t) \leqslant 0 = 0 \text{ if } H > 0 \qquad [3a]$$

$$t(R_B + ID_B) - (1 - It)P_B \leqslant 0 = 0 \text{ if } B > 0 \qquad [3b]$$

where t is the marginal tax rate, R_H and R_B are the marginal reduction in taxable income generated by an hour of own time and professional time, respectively, and D_H and D_B are the marginal amount of deductions generated by own and professional time, respectively. Equation 3 states that (when $H > 0$) the marginal payoff to own time, in terms of lower tax liability, must equal the opportunity cost of that time. Equation 3b states that (when $B > 0$) the marginal payoff to hiring professional assistance must equal its cost. Note that the value of I enters the first-order conditions in two separate ways. First, the marginal benefit of uncovering deductions can be positive only if I is one. In addition, the after-tax cost of professional tax advice depends on the itemization decision because these expenses are deductible from taxable income.

Differentiation of (3a) – (3b), detailed in an appendix to this article available from the author, allows us to investigate the effect on compliance behavior of changes in the marginal tax rate, the net wage rate, and the itemization status when both H and B are positive. An increase in the marginal tax rate directly increases the payoff from reducing taxable income by a dollar. For itemizers, it also decreases the cost of professional tax assistance. The theory implies that both H and B will necessarily increase only if H and B are complementary inputs in both compliance technologies. Otherwise, it is conceivable that an increase in t will induce a change in technique, so that either H or B, but not both, will increase.

An increase in the after-tax wage rate, holding t constant, increases the opportunity cost of own time spent on tax matters. This leads to an unambiguous decline in H, but the effect on B is not determinate. For nonitemizers, B will decline if it is comple-

mentary with H in the taxable income reduction technology, and increase otherwise. For itemizers, the response of B depends on its complementarity with H in both the "deduction" and "reduction" technologies.

A change to itemizing status has an effect similar to that of an increase in the marginal tax rate. First of all, the payoff to uncovering deductions (but not reductions) from taxable income increases in proportion to the marginal tax rate. Second, the cost of professional tax assistance declines, again in proportion to the marginal tax rate. The impact of a change in I on H and B depends, as in the case of the marginal tax rate, on whether H and B are complementary inputs.

As we will see later, for many observations B equals zero and (3b) does not hold as an equality. In this case, differentiation of (3a) alone reveals that, as long as own time spent is subject to locally diminishing returns, an increase in the marginal tax rate will increase H, an increase in $w(1-t)$ will decrease H, and a small increase in I will cause H to increase.

3. ESTIMATION STRATEGY

We consider two distinct estimation approaches. The first approach is to estimate reduced form equations for H and B using ordinary least squares (OLS). We include all exogenous variables as regressors plus a dummy variable for itemization status, and, in an alternative version, also an interactive dummy variable. This model has the advantage of transparency and direct interpretability. At the same time, it ignores the statistical problems that arise because of the endogeneity of the itemization decision and the existence of a significant fraction of households that do not hire any professional tax assistance. This first problem is especially significant because one of the goals of this research is to investigate the effect of eliminating the option of itemizing deductions.

The second estimation approach is decidedly more ambitious. We proceed by constructing a general linear model that can accommodate all of the important aspects of the theoretical

model and data structure. The model has three parts. In the first part, we introduce an unobservable variable I_i^*, which is a measure (of arbitrary scale) of the difference between the maximum utility attainable in the event that taxpayer i itemizes deductions and the maximum utility attainable if deductions are not itemized. We posit that $I_i^*$is a linear function of the exogenous variables of the system, Z_i, and append an additive error term, ϵ_i, which encompasses both optimization error and unobservable variables that affect I_i^*, such as intelligence and unknown characteristics of the taxpayer's tax situation, so that

$$I_i^* = \gamma' Z_i + \epsilon_i \qquad [4]$$

$$I_i = 1 \text{ iff } I_i^* > 0$$

$$I_i = 0 \text{ iff } I_i^* \leqslant 0.$$

We make the standard assumption that ϵ_i is distributed as a standard normal variable. Equation 4 indicates that, although I_i^* is unobserved, the itemization decision is observed, so that we know whether I_i^* is positive or not.

The second part of the system models the decision of whether to purchase professional tax advice, conditional on the itemization status chosen. As with the itemization decision, it is posited that there is a latent variable (called either J^*_{Ii} or J^*_{Ni} for itemizers and nonitemizers, respectively), which represents the propensity to pay for professional advice.

As before,

$$J_{Ii}^* = \delta_I' Z_i + \nu_{Ii} \qquad \nu_{Ii} \sim N(0, 1) \qquad [5a]$$

$$J_{Ii} = 1 \text{ if } J_{Ii}^* > 0$$

$$J_{Ii} = 0 \text{ if } J_{Ii}^* \leqslant 0$$

$$J_{Ni}^* = \delta_N' Z_i + \nu_{Ni} \qquad \nu_{Ni} \sim N(0, 1) \qquad [5b]$$

$$J_{Ni} = 1 \text{ if } J_{Ni}^* > 0$$

$$J_{Ni} = 0 \text{ if } J_{Ni}^* \leqslant 0$$

The error terms ϵ_i, ν_{Ii}, and ν_{Ni} will, in general, be correlated.

The third part of the system models the choice of H and B, conditional on the outcome of the two discrete choices. A linear representation of the first-order conditions (3a) and (3b) is

Itemizing Regime ($I_{Ii} = 1$)

$$H_i = \alpha_I B_i + \alpha_I' X_{Hi} + u_{HIi} \qquad [6a]$$

$$B_i = \beta_I H_i + \beta_I' X_{Bi} + u_{BIi} \qquad [6b]$$

Not Itemizing Regime ($I_{Ii} = 0$)

$$H_i = \alpha_N B_i + \alpha_N' X_{Hi} + u_{HNi} \qquad [6c]$$

$$B_i = \beta_N H_i + \beta_N' X_{Bi} + u_{BNi} \qquad [6d]$$

Here X_H and X_B refer to the vectors of exogenous variables in the H and B equations, respectively. The error terms in Equations 6a-6d may be correlated with the error terms in Equations 4, 5a, and 5b. Note that when J_{Ii} is zero (no professional assistance hired), (6b) is irrelevant and (6a) collapses to

$$H_i = \alpha_I' X_{Hi} + u_{HIi}. \qquad [6a']$$

Similarly, when J_{Ni} is zero, (6d) is irrelevant and (6c) reduces to

$$H_i = \alpha_N' X_{Hi} + u_{HNi}. \qquad [6c']$$

It is worth noting that the system outlined here is more general than a "tobit"-type system in which one structural equation determines both the probability that a household pays for professional assistance and the amount spent, conditional on spending any positive amount. In this more general structure, the two decisions are allowed to respond differently to exogenous variables. This differentiation is appropriate if, for example, there are fixed costs involved in purchasing professional tax advice. The existence of fixed costs is in fact reasonable, as the preparer must become familiar with the return before the return can be completed. For complicated and idiosyncratic returns with large fixed costs associated with preparation, it is plausible that a

taxpayer may be unlikely to hire professional advice, although—if any assistance is purchased—a substantial amount will be required.

It is well known that direct estimation of (6a) – (6d) by ordinary least squares will yield biased coefficient estimates because the expected value of the error term, conditional on the household being observed in a particular itemization and professional assistance regime, is generally nonzero and correlated with the explanatory variables. Our strategy in this event is to utilize a two-stage estimation procedure. The first stage is to estimate the bivariate discrete choice model outlined above of whether to itemize deductions and whether to hire professional tax assistance, conditional on the itemization decision. In the second stage, the estimates of the discrete choice model are utilized to control for the self-selection biases in the estimation of the simultaneous system of (6a) – (6d).

The first stage of the estimation begins with our stated assumptions that both ϵ_i and ν_{Ii} and also ϵ_i and ν_{Ni} have a joint normal distribution, and that each term has unit variance. Next we divide the observations into four regimes as follows:

$$R_1\colon\ I_i^* > 0,\ J_{Ii}^* > 0 \qquad \text{(itemize, hire assistance)} \qquad [7]$$

$$R_2\colon\ I_i^* > 0,\ J_{Ii}^* \leqslant 0 \qquad \text{(itemize, don't hire assistance)}$$

$$R_3\colon\ I_i^* \leqslant 0,\ J_{Ni}^* > 0 \qquad \text{(don't itemize, hire assistance)}$$

$$R_4\colon\ I_i^* \leqslant 0,\ J_{Ni}^* \leqslant 0 \qquad \text{(don't itemize, don't hire assistance)}$$

If we denote the bivariate distribution function of ϵ and ν_I as $f_I(.,., \rho_I)$ and the joint density of ϵ and ν_N as $f_N(.,.,\rho_N)$, the likelihood function is

$$\prod_1 \int_{\delta_I' Z_i}^{\infty} \int_{-\gamma' Z_i}^{\infty} f_I(\epsilon, \nu_I, \rho_I) d\epsilon_i du_{Ii} \prod_2 \int_{\infty}^{-\delta_I' Z_i} \int_{\gamma' Z_i}^{\infty} f_I(\epsilon, \nu_I, \rho_I) d\epsilon_i du_{Ii} \qquad [8]$$

$$\prod_3 \int_{-\delta_N' Z_i}^{\infty} \int_{\infty}^{-\gamma' Z_i} f_N(\epsilon, \nu_N, \rho_N) d\epsilon_i du_{Ni} \prod_4 \int_{-\infty}^{-\delta_N' Z_i} \int_{-\infty}^{-\gamma' Z_i} f_N(\epsilon, \nu_N, \rho_N) d\epsilon_i du_{Ni}$$

We employ the method of maximum likelihood to estimate the parameters γ', δ'_I, δ'_N, ρ_I, and ρ_N.

As mentioned above, the results of the first-stage estimation are to be used in the estimation of the parameters of (6a) – (6d). To see how this works we first write the reduced forms of these equations as

Regime 1 $(I_i = 1,\ J_{Ii} = 1)$

$$H_i = k_I[(\alpha'_I + \alpha_I\beta'_I)X_i + u^o_{HIi}] \qquad [9a]$$

$$B_i = k_I[(\beta'_I + \beta_I\alpha'_I)X_i + u^o_{BIi}] \qquad [9b]$$

Regime 2 $(I_i = 1,\ J_{Ii} = 0)$

$$H_i = \alpha'_I X_i + u_{HIi} \qquad [9c]$$

Regime 3 $(I_i = 0,\ J_{Ni} = 1)$

$$H_i = k_N[(\alpha'_N + \alpha_N\beta'_N)X_i + u^o_{HNi}] \qquad [9d]$$

$$B_i = k_N[(\beta'_N + \beta_N\alpha'_N)X_i + u^o_{BNi}] \qquad [9e]$$

Regime 4 $(I_i = 0,\ J_{Ni} = 0)$

$$H_i = \alpha'_N X_i + u_{HNi} \qquad [9f]$$

where $k_I = 1/(1 - \alpha_I\beta_I)$,

$$k_N = 1/(1 - \alpha_N\beta_N)$$

$$u^o_{HIi} = u_{HIi} + \alpha_I u_{BIi}$$

$$u^o_{BIi} = u_{BIi} + \beta_I u_{HIi}$$

$$u^o_{HNi} = u_{HNi} + \alpha_N u_{BNi}$$

$$u^o_{BNi} = u_{BNi} + \beta_N u_{HNi}$$

X_i is now the set of all exogenous variables in (6a) – 6d), and the vectors α'_I and β'_I are expanded to be compatible with X_i.

The problem with estimating (9a) – (9f) by ordinary least squares is that the error terms are not necessarily uncorrelated with the explanatory variables, given the self-selection rules. Fortunately, the estimates obtained from the bivariate probit model allow us to form a consistent estimator of $E(u^o_{HIi} \mid I_i = 1, J_{Ii} = 1)$ and of the other error terms in (9b) – (9f), conditional on regime. Thus, in order to estimate the parameters of (9a), for example, we utilize an ordinary least squares procedure for the observations in Regime 1 that not only includes as explanatory variables the union of X_H and X_B but also includes two other variables that represent the conditional expected value of ϵ_i and γ_{Ii}. Lee et al. (1980) have shown, in a similar model, that the estimated standard errors from the ordinary least squares equations will underestimate the true standard errors, as their calculation ignores the fact that $\hat{\gamma}$, $\hat{\delta}_I$, and $\hat{\rho}_I$ are themselves estimated.

4. DATA

The data for this study are drawn from a mail survey of Minnesota households' tax-filing behavior in 1982 sent to a random sample of 2,000 Minnesota residents. Of the 653 questionnaires returned, 79 were eliminated from the sample because of incompleteness of response or because the household did not file 1982 income tax returns, leaving a total of 574 usable responses. A detailed description of the data and survey procedures is presented in Slemrod and Sorum (1984).

The questionnaire first requests some demographic information, in particular the respondents' sex, age, level of education completed, income, employment status, occupation, and wage rate or reservation wage. (In assessing this information, it is important to realize that the cover letter pointedly asks that the addressee refer the questionnaire to the "person in [the] household most familiar with filing [the] income tax returns.") It then solicits

information about the household's income tax return itself. The taxpayer is asked which, if any, of the three federal tax returns and which, if either, of the two Minnesota state tax forms was filed. In addition, responses are sought concerning whether the return featured itemized deductions, whether it was a joint return, and which of several sources of income were received. The remainder of the questionnaire is devoted to collecting information about the household's cost of filing tax returns. This section asks how many hours were spent during the year and requests a breakdown of the hours into various categories. In addition, any money spent on tax assistance or otherwise spent in filing returns is solicited.

A subset of the information collected in the survey was used in this study. The precise definition of the variables considered is as follows:

ENDOGENOUS VARIABLES

H:	total hours spent on tax matters
B:	total dollar expenditure on professional tax assistance

EXOGENOUS VARIABLES

AGE:	age of respondent
AGESQ:	age squared
EDU:	years of education completed
MAR:	dummy variable equal to one if respondent is married; equal to zero otherwise
EASY:	dummy variable equal to one if tax return did not contain any dividends, interest, self-employed business income, capital gains, rental income, pension, annuity, or other income; equal to zero otherwise
INT:	dummy variable for presence of interest income
DIV:	dummy variable for presence of dividend income
SEBUS:	dummy variable for presence of self-employed business income
CAPGL:	dummy variable for presence of capital gains or losses
RENT:	dummy variable for presence of rental income

PENANN:	dummy variable for presence of pension or annuity income
OTHER:	dummy variable for presence of "other" income
TIMEVAL:	value of time (per hour), measured as after-tax wage rate or reservation wage
TAX:	marginal tax rate applicable to deductible expenses, expressed as a fraction

The interpretation of most of these variables is clear; some, however, merit further comment. Responses to the age and education question were in ranges. Each taxpayer was assigned the midpoint of the indicated range. Those indicating "over 65" were assigned an age of 70, and those indicating graduate-level education were assigned 18 years. The marginal tax rate was calculated from information about income reported by the taxpayer. The calculation takes into account both the federal and state income tax and the possible deductibility of one tax in the calculation of taxable income for the other level of government. Because a usable wage rate was not supplied on 46% of the returned questionnaires, a wage equation was estimated for those who did supply a usable answer and used to impute a wage rate from those who did not.[2] The exogenous variables in this equation were the income level, employment status, occupation, age, level of education completed, marital status, and sex. The marginal tax rate applied to the gross wage is the same as that described above plus the social security tax.

All of the exogenous variables listed above are included in the vectors Z and X_H. The vector X_B includes all of the above except TIMEVAL.

5. RESULTS

The results of the ordinary least squares estimations are presented in Table 1. Two separate reduced-form equations were estimated for both own time spent and for professional assistance purchased. One version (denoted A) includes all the exogenous variables plus a dummy variable for itemization status (ITEM).

The second version (denoted B) includes all these variables plus an interactive dummy term that is equal to the product of TAX and ITEM, called TAXITEM. This particular interactive term is included because of the fact that, for itemizers only, the tax rate affects not only the return to reducing taxable income by one dollar but also directly affects the effective price of professional assistance.

The results in Table 1 indicate that households with more complicated returns (as measured by the presence of various sources of income) spend more time and money on tax matters.[3] Particularly strong associations exist between compliance cost and the presence of self-employed business income and capital gains or losses. Households with both sources of income spend approximately 32 hours and $65 more than households with neither source of income. The value of time variable is estimated to have a significant positive association with monetary expense and a positive but not significant effect on own hours spent. The theory would suggest that a high value of time should induce substitution away from own time toward professional advice. The fact that time spent is positively correlated with the after-tax rate may be due to the fact that it is a better measure of competency in tax matters than educational attainment, so that higher-wage-rate people have both a higher opportunity cost of time and a higher return to investing their time working on their tax return.

In both Version A and B of the model, the estimated effect of the marginal tax rate on hours spent is positive, though not statistically significant. On the other hand, the regression equations do pick up a strong positive relationship between the tax rate and the amount of professional assistance purchased. This is apparent in Version A, where an increase in the tax rate of 0.1 is associated with $16 more spent. In Version B, this association is shown to be present only for itemizers, but is very large ($21 more with an increase of 0.1) and significant for that group.

Finally, we note that, in Version A, the dummy variable for itemization is positive and at least marginally significant. The magnitude of the implied effect is large. Itemizers are estimated to spend 8.8 more hours on tax matters and $24.30 more. In Version B, the estimated impact of itemization depends on the marg

TABLE 1
Parameter Estimates: Ordinary Least Squares Using All Observations

Independent Variable	Hours		Professional Assistance	
	A	B	A	B
Constant	1.15	-3.36	1.93	62.86
	(19.96)	(23.17)	(36.58)	(42.17)
AGE	-0.52	-0.53	-1.93	-1.83
	(0.85)	(0.86)	(1.56)	(1.56)
AGESQ	0.0074	0.0076	0.030*	0.029*
	(0.0094)	(0.0094)	(0.017)	(0.017)
EDU	0.89	0.91	-0.68	-0.86
	(0.66)	(0.66)	(1.21)	(1.20)
MAR	6.96*	7.02*	-11.67	-12.42*
	(4.44)	(4.44)	(8.13)	(8.08)
EASY	-9.60	-9.47	-18.63	-20.37
	(7.90)	(7.92)	(14.49)	(14.41)
INT	-3.75	-3.61	-22.26*	-24.06**
	(6.36)	(6.38)	(11.66)	(11.61)
DIV	-6.13	-6.05	-5.00	-6.11
	(4.42)	(4.43)	(8.11)	(8.07)
SEBUS	23.42**	23.39**	36.09**	36.52**
	(4.76)	(4.77)	(8.73)	(8.68)
CAPGL	8.32*	8.30*	29.40**	29.72**
	(5.08)	(5.09)	(9.32)	(9.26)
RENT	2.05	1.93	16.78*	18.47*
	(5.73)	(5.74)	(10.50)	(10.44)
PENANN	-2.42	-2.53	-2.10	-0.61
	(5.97)	(5.98)	(10.95)	(10.89)
OTHER	4.15	4.21	14.23	13.37
	(6.42)	(6.43)	(11.77)	(11.70)
TIMEVAL	0.41	0.42	2.16**	2.00**
	(0.28)	(0.28)	(0.52)	(0.52)
TAX	11.67	26.87	160.87**	-44.77
	(21.29)	(45.01)	(39.02)	(81.92)
ITEM	8.76*	14.23	24.30**	-49.79*
	(4.93)	(15.11)	(9.04)	(27.50)
TAXITEM		-19.07		257.94**
		(19.73)		(90.52)

(continued)

TABLE 1 Continued

Independent Variable	Hours		Professional Assistance	
	A	B	A	B
Mean of Dependent Variable	26.9	26.9	51.4	51.4
Sample Size	574	574	574	574
R^2	0.107	0.107	0.215	0.226
Standard Error of Regression	42.3	42.3	77.5	77.0

NOTE: Standard errors are in parentheses.

*Significant at 90% confidence level; **significant at 95% confidence level.

tax rate. For a tax rate of 0.3, the estimated impact is approximately the same as in Version A.

Table 2 contains the results of estimating the more general model outlined in detail in Section 3.[4] In the first three columns are the results of estimating the bivariate probit model of the choices of whether to itemize and whether to hire professional assistance, conditional on itemization status. The results of within-regime, ordinary least squares estimates with sample selection bias correction terms are presented in the last six columns.

The estimates of Table 2 indicate that the presence of a very simple return (EASY = 1) has a significant negative association with the probability of hiring professional help for itemizers and, for itemizers who pay for help, a negative association with the amount of own time spent on tax filing. Few other relationships approach statistical significance. Having self-employment income (SEBUS = 1) is positively associated with hiring professional help for itemizers and has a strong positive effect on own time spent for itemizers who pay for help and on the amount of professional assistance purchased for all who purchase any help at all. Having capital gains or losses (CAPGL = 1) has a strong positive relationship only with the amount of professional assistance purchased by itemizers.

The results also indicate that while TIMEVAL does not significantly affect the probability of purchasing assistance, it is strongly positively associated with the amount purchased for itemizers and positively, but not significantly, related to the amount purchased for nonitemizers. The effect of TIMEVAL on own time spent is mixed. A higher value of TIMEVAL increases the likelihood of being an itemizer, which is associated with higher H. Given the regime, TIMEVAL has an insignificant positive association with H, except for the nonitemizing, zero B, households, for whom there is a strong negative relationship.

A higher marginal tax rate increases the likelihood of itemizing, and increases the likelihood of purchasing professional assistance, though for itemizers only. In addition, for itemizers, the amount of professional assistance purchased is positively associated with the marginal tax rate.

The bottom of Table 2 reports the coefficients of the two selectivity variables for each equation. In each case the first listed selectivity variable is the expected value of the error term in the itemization probit equation (ϵ_i), conditional on the regime chosen. The second listed selectivity variable is the expected value of the error term in the professional assistance probit equation (ν_{Ii} for itemizers, and ν_{Ni} for nonitemizers), conditional on the regime chosen. The estimated coefficients on the selectivity variables are large relative to the unadjusted standard errors only in the H equation for households in Regime 4 (nonitemizing, not paying for professional assistance). The negative signs of the estimated coefficients indicate that there will be a tendency to underestimate H when either the probability of itemization or the probability of paying for professional assistance is overestimated.

The interpretation of the coefficients of each of the two selectivity variables is the estimated covariance between the error in the two relevant probit equations and the error in the ordinary least squares equation. The inner product of the two coefficient estimates and the expected value of the error terms is the estimate of the expected unobservable component. For each of the six equations, this value is very large in absolute value compared to the observed mean of the dependent variable. This implies that

TABLE 2

Parameter Estimates: Bivariate Probit Model and Ordinary Least Squares with Sample Selection Correction

Independent Variable	Bivariate Probit			OLS on Regime 1		OLS on Regime 2	OLS on Regime 3		OLS on Regime 4
	I^*	J^*_I	J^*_N	H	B	H	H	B	H
Constant	-1.08 (0.81)	1.67 (1.26)	-0.752 (2.23)	34.7 (49.7)	2.83 (81.3)	-18.1 (246.3)	-73.9 (61.1)	-121.8 (142.2)	-1.48 (30.5)
AGE	0.055 (0.034)	-0.035 (0.040)	0.020 (0.259)	-1.66 (1.92)	-2.56 (3.14)	-3.61 (3.34)	0.43 (4.31)	7.39 (10.0)	-8.48 (5.08)
AGESQ	-0.00065 (0.00037)	0.00052 (0.00048)	0.00006 (0.0025)	0.0258 (0.024)	0.041 (0.039)	0.035 (0.049)	0.014 (0.048)	-0.065 (0.11)	0.085 (0.059)
EDU	-0.014 (0.030)	-0.079** (0.026)	-0.101 (0.257)	-1.06 (2.49)	0.73 (4.07)	6.38 (8.33)	0.19 (2.75)	-5.08 (6.40)	4.86** (2.15)
MAR	0.375** (0.169)	0.027 (0.187)	-0.291 (0.937)	4.81 (9.33)	-26.4* (15.3)	-17.9 (14.0)	-5.9 (22.6)	33.0 (52.8)	-33.8 (24.3)
EASY	-0.131 (0.331)	-0.646** (0.302)	-0.089 (1.30)	-37.5 (24.3)	-6.41 (39.7)	52.6 (68.4)	-27.4 (20.8)	-28.6 (48.4)	8.84 (19.5)
INT	-0.163 (0.280)	-0.628** (0.259)	0.341 (0.746)	-20.5 (21.1)	-13.8 (34.6)	53.5 (66.3)	-9.9 (16.5)	18.7 (38.3)	-8.62 (21.4)
DIV	0.514** (0.198)	0.101 (0.174)	-0.944 (0.633)	-10.8 (9.5)	-13.1 (15.5)	-26.5 (20.6)	-32.9 (37.3)	-22.0 (86.8)	-34.3 (48.4)
SEBUS	0.182 (0.248)	0.291 (0.163)	0.677 (2.27)	30.5** (12.4)	43.6** (20.3)	-9.9 (32.4)	39.3 (32.0)	137.8 (74.5)	-16.0 (24.8)
CAPGL	0.284 (0.262)	0.125 (0.167)	0.119 (1.59)	13.3 (9.3)	38.4** (15.2)	-16.5 (17.6)	1.42 (30.1)	70.8 (70.0)	-37.9 (29.0)

RENT	-0.116 (0.300)	0.111 (0.203)	1.48 (2.19)	1.79 (9.64)	16.7 (15.8)	16.9 (14.3)	63.1 (31.6)	92.2 (73.6)	-69.0** (29.7)
PENANN	-0.157 (0.233)	-0.277 (0.203)	-1.13 (2.74)	-17.1 (12.9)	14.7 (21.0)	29.6 (33.7)	-47.5 (30.0)	-62.2 (69.7)	76.7** (21.6)
OTHER	0.064 (0.307)	0.061 (0.210)	-0.18 (0.72)	4.37 (10.9)	40.4** (17.8)	9.6 (10.8)	-30.3 (24.5)	13.0 (57.1)	-17.3 (8.7)
TIMEVAL	0.0219 (0.0155)	-0.0087 (0.0091)	0.040 (0.19)	0.205 (0.574)	3.39** (0.94)	0.37 (0.76)	3.36 (2.73)	7.02 (6.37)	-4.50** (2.16)
TAX	2.27** (0.90)	1.73** (0.80)	-0.699 (7.44)	15.9 (64.0)	242.5** (104.8)	-138.3 (203.3)	-96.1 (155.3)	27.3 (361.3)	-218.1 (217.7)
Correlation Coefficient	-	0.69 (0.43)	-0.48 (4.67)	-	-	-	-	-	-
Selectivity Variable #1				-113.3 (90.6)	-145.0 (148.1)	-73.1 (52.7)	36.6 (123.2)	204.1 (286.7)	-223.3 (126.2)
Selectivity Variable #2				134.8 (105.6)	11.6 (172.7)	-53.3 (153.6)	64.2 (54.2)	151.3 (126.1)	-128.8** (38.0)
Means of Dependent Variable	-	-	-	30.7	91.9	27.4	16.2	55.3	14.1
Sample size				302	302	172	32	32	68
R^2				0.10	0.34	0.16	0.65	0.57	0.67
Standard Error of Regression				51.7	84.5	32.8	20.0	46.6	14.9

NOTE: Standard errors are in parentheses. These statistics are computed without adjusting for the fact that the selectivity variables are estimated.

**Coefficient estimate is at least twice the unadjusted estimated standard error.

the unobservable component of tax-filing behavior, for both time and money spent, is very large compared to the component explained by the independent variables. In several cases these coefficients imply implausible predicted behavior of a taxpayer in a regime other than the one in which he or she is actually observed.

The large standard errors of the coefficients of the sample selection bias correction terms suggest that these implausible predictions are due to the multicollinearity of the estimate of the unobserved component of the choice of regime decision and the determinants of the conditional continuous choice of H and B. Remember that the same set of exogenous variables explains both the choice of regime and the compliance behavior conditional on regime. That the two relationships may be identified is entirely due to the nonlinear relationship of the sample selection bias correction terms with the explanatory variables. In sum, the estimation technique is unable to distinguish precisely between the effect of the unobservable influences that determine the choice of regime and the effect of the explanatory variables that affect behavior within a regime.

What conclusions can be drawn from the regression results presented in Table 1 and 2? Although the results from the two estimation approaches are not in all cases consistent, some clear findings do emerge. First, the presence of certain sources of income causes greater expenditure of time and money on tax compliance. This is especially true for self-employment income, and is also observed for capital gains income. A higher value of time, which we expect to be associated with a substitution away from own time to the use of professional assistance, is positively associated with the use of professional assistance in the simple OLS equations of Table 1, but when the regimes are separated, this effect is found only for itemizers. The estimated effect of the marginal tax rate on compliance behavior is somewhat mixed. Neither of the techniques indicates a strong positive association between the tax rate and the amount of one's own time spent on tax matters. The regressions do reveal, however, a positive association between the marginal tax rate and the expenditure on

professional tax assistance. The probit estimates imply that, at least for itemizers, a higher tax rate increases the likelihood of purchasing assistance. Furthermore, for itemizers who do purchase assistance, the amount of assistance purchased is positively associated with the marginal tax rate.

6. THE EFFECT OF TAX SIMPLIFICATION ON COMPLIANCE COST

Slemrod (1985) uses the preceding econometric results to simulate the likely impact on compliance cost of (1) eliminating the system of itemizing deductions and (2) installing a "flat-rate" income tax system. Here the basic simulation findings are reported, and the reader is referred to the other article for details.

Simulations based on both empirical models indicate that eliminating the system of itemized deductions would be accompanied by a substantial reduction in expenditures on professional assistance, ranging from 28% to 39%.[5] The percentage of taxpayers who use professional assistance is predicted to decline by 12.6 percentage points. The models have different predictions about the impact on taxpayer time spent. While the OLS models of Table 1 predict a decline of per-taxpayer hours of slightly more than 3, the regime-by-regime estimates with endogenous regime selection predict a slight increase of 0.5 hours per taxpayer. Both models predict that the total resource cost of compliance would decline, although the decline is large only using the OLS estimates (15% to 20%). In sum, not itemizing deductions would apparently save some hours of record keeping and also eliminate the primary reason that many people seek professional help. In addition, it increases the net price of professional tax assistance from $(1-t)P_B$ to P_B for former itemizers. Because current itemizers account for 64% of payments for assistance in the weighted sample, the substantial increase in price that accompanies the elimination of itemization apparently would have a large aggregate impact.

The simulations indicate that although moving to a flat-rate[6] tax system without itemizing would entail less resource cost than the current system, the move to a flat rate would not by itself

reduce the aggregate cost of compliance by very much, if at all. Although the OLS results indicate a small increase in both aggregate hours spent and professional assistance purchased if the tax schedule is flattened, the more general model predicts a small decline in both time and money spent when the rate schedule is changed. Compared to the policy of eliminating only itemization, the change in the resource cost of compliance is estimated at between a 2% saving and a 6% additional cost.

These results are relevant to understanding the impact of the Tax Reform Act of 1986, because the new law changed the income tax system in the direction of both of these simulated reforms. As a result of raising the standard deduction, eliminating the deductibility of sales tax and (eventually) personal interest payments, and subjecting miscellaneous deductions to a floor, the fraction of taxpayers who itemize will fall by about 7%. Furthermore, the structure of rates was significantly flattened, with the top rate reduced from 50% to 28%. Because these two aspects of the new law are partial moves toward eliminating itemization and instituting a flat tax, the simulation results suggest that, in the long run, the Tax Reform Act of 1986 will reduce the use of tax professionals below what it otherwise would be, although no significant reduction in the amount of taxpayer time spent on tax matters can be predicted with confidence. In the short run, however, both the reliance on tax professionals and taxpayer time will likely rise as the public adjusts to the new income tax system.

Other aspects of the new law have implications for the compliance cost, some simplifying (e.g., eliminating the two-earner credit and income averaging) and some complicating (e.g., limiting IRA and interest deductions), and the net effect on compliance cost is not clear. Including expenses on professional tax assistance in miscellaneous deductions that can be deducted only to the extent they exceed 2% of AGI will, by increasing the marginal cost of tax advice for some taxpayers, reinforce the conclusion that the use of professional assistance will tend to decline in the long run.

There is one potentially significant reason that these results may underestimate the resource savings to be derived from a change to a flatter-rate schedule. It is that the analysis assumes

that the taxpayer's sources of income do not change when the marginal tax rate changes. To the extent that the current rate structure induces individuals to engage in income-earning activities that require a relatively high cost of compliance, the estimates presented here will underestimate the resource savings.

7. CONCLUSIONS

Microeconomic data of the kind analyzed in this article are potentially a rich source of information about tax-compliance behavior. This analysis of tax simplification suggests that significant resource savings can be expected from eliminating the system of itemized deductions, although no saving from changing to a single-rate tax structure can be confidently predicted.

Information about the likely resource savings of a particular tax simplification scheme is properly seen as one input of many that should be considered by policymakers. The allocational and distributional impact of the changes must be weighed as well. It is hoped that this research can begin the task of enriching the debate about tax simplification by bringing quantitative evidence to bear on its benefits and costs.

NOTES

1. For models of tax evasion, see Allingham and Sandmo (1972) and Cross and Shaw (1981, 1982).

2. Although there is a substantial literature concerned with imputing wage rates to nonworking individuals based on their characteristics and the wage rate and characteristics of working individuals, direct application of that methodology to the problem at hand seems inappropriate, because there is no clear distinction between those for whom a wage rate is available and those for whom it is not available. (Examples of unusable replies to the wage-rate question are "time-and-a-half," "retired," or "variable." Others left the answer space blank.) Thus it is invalid to claim that the reservation wage of those with no wage rate available (because they are not working) must be greater than the wage rate that could be earned. The fact that TIMEVAL is estimated for some taxpayers is not taken into account in the calculation of the standard error of the regression estimates.

3. In interpreting the coefficients of the sources of income variables, it is important to bear in mind that EASY takes on a value of one only when none of the sources of income is

present. Thus, for example, the estimated impact of having dividends, compared to having no nonwage income, is found by subtracting the coefficient of EASY from the coefficient of DIV.

4. The standard errors reported in the last six columns of Table 2 are not adjusted for the fact that the selectivity variables are estimated. Thus claims of statistical significance cannot be supported. In the subsequent text, I refer to an estimate as being statistically insignificant if the estimated coefficient is less than its unadjusted standard error. It is unlikely that the corrected standard errors would alter this assessment.

5. Because the results of Table 2 indicate implausible predictions of behavior of taxpayers who change regimes, the simulation results are based on the bivariate probit estimates of Table 2 and OLS estimates of within-regime behavior estimated without the selectivity terms. See Slemrod (1985) for details.

6. The policy simulated institutes a flat federal tax rate of 20% (which is approximately equal yield) and does not change the state tax system. Note that a switch to this system would not reduce all taxpayers' marginal tax rate. On the contrary, many low-income individuals would face a higher rate.

REFERENCES

ALLINGHAM, M. G. and A. SANDMO (1972) "Income tax evasion: a theoretical analysis." J. of Public Economics 1: 323-338.

BRANNON, G. M. (1979) "Simplification and other objectives," pp. 191-204 in C. H. Gustafson (ed.) Federal Income Tax Simplification. Philadelphia: American Law Institute.

CROSS, R. and G. K. SHAW (1981) "The evasion-avoidance choice: a suggested approach." National Tax J. 34: 489-491.

CROSS, R. and G. K. SHAW (1982) "On the economics of tax aversion." Public Finance 37: 36-47.

LEE, L. F., G. S. MADDALA, and R. P. TROST (1980) "Asymptotic covariance matrices of two-stage probit and two-stage tobit methods for simultaneous equations models with selectivity." Econometrica 43: 491-503.

LEE, L. F. and R. P. TROST (1978) "Estimation of some limited dependent variable models with applications to housing demand." J. of Econometrics 8: 357-382.

MADDALA, G. S. (1983) Limited-Dependent and Qualitative Variables in Econometrics. Cambridge: Cambridge Univ. Press.

SANDFORD, C. (1973) The Hidden Costs of Taxation. London: Institute for Fiscal Studies.

SICKLES, R. C. and P. SCHMIDT (1978) "Simultaneous equations models with truncated dependent variables: a simultaneous tobit model." J. of Economics and Business 31: 11-21.

SLEMROD, J. (1983) "Optimal tax simplification: toward a framework for analysis." Proceedings of the Seventy-Sixth Annual Conference of the National Tax Association—Tax Institute of America.

SLEMROD, J. (1985) The Return to Tax Simplification: An Econometric Analysis. Working Paper no. 1756. Cambridge, MA: National Bureau of Economic Research.

SLEMROD, J. and N. SORUM (1984) "The compliance cost of the U.S. individual income tax system." National Tax J. 37: 361-370.

Joel Slemrod is Associate Professor of Business Economics, Economics, and Public Policy at the University of Michigan and is Director of the Office of Tax Policy Research there. In addition to the cost of administering tax systems, his research interests include tax compliance and international aspects of taxation.