

DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES
CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

THE DEMAND FOR TAX RETURN PREPARATION SERVICES

Jeffrey A. Dubin
California Institute of Technology

Michael A. Graetz
Yale Law School

Michael A. Udell
California Institute of Technology

Louis L. Wilde
California Institute of Technology



SOCIAL SCIENCE WORKING PAPER 714

December 1989
Revision May 1990

The Demand for Tax Return Preparation Services

Jeffrey A. Dubin Michael J. Graetz Michael A. Udell Louis L. Wilde*

Abstract

In this paper we focus on taxpayer choices of return preparation services. Using aggregate nested logit techniques, we find that the demand for third party preparation responds to many factors: age, education, employment status, federal auditing, and tax return characteristics. Perhaps most important among these is federal auditing. Higher federal audit rates increase the demand for tax practitioner services, but do not affect the demand for other modes of third party assistance. More generally, as the tax burden increases, or as uncertainty about true tax liability increases, the demand for all modes of third party assistance increases.

*California Institute of Technology, Yale Law School, California Institute of Technology, and California Institute of Technology, respectively. Professor Graetz thanks the National Science Foundation for research support (grant no. SES-870443) as do Professors Dubin and Wilde (grant no. SES-8701027). Professor Graetz is currently serving as Deputy Assistant Secretary (Tax Policy) of the Treasury. This article was written before Professor Graetz assumed that office and does not necessarily reflect the views of the Treasury Department.

The Demand for Tax Return Preparation Services

Jeffrey A. Dubin Michael J. Graetz Michael A. Udell Louis L. Wilde

1 Introduction

The IRS has estimated that 44.2 percent of the individual returns filed in 1979 were self prepared, and that these returns accounted for 22.8 percent of detected noncompliance. Returns prepared with third party assistance accounted for 55.8 percent of filings and 77.2 percent of the detected noncompliance. Among returns prepared with third party assistance, however, underreported tax was not uniformly distributed. Only 10.6 percent of all taxpayers used a tax practitioner (certified public accountant or attorney) yet their returns accounted for 32.5 percent of underreported tax,¹ 34.7 percent used paid preparers accounting for 40.9 percent of underreported tax,² and the remaining 10.4 percent used non-paid assistance accounting for 3.7 percent of underreported tax.³ Population frequencies of third party assistance and the associated frequencies of tax noncompliance are reported below in Table 1.

The current IRS estimate for total noncompliance from individual tax returns in 1988 is approximately \$85 billion (Internal Revenue Service, 1990). If the 1979 IRS

¹A tax practitioner is automatically able to represent clients before the IRS. Paid preparers may provide these services only after passing a written exam sponsored by the IRS, and fulfilling certain continuing education requirements.

²Paid preparers include public accountants with 6.2 percent of the returns filed, national tax services with 9.5 percent of the returns filed, local tax services with 11.6 percent of the returns filed, and miscellaneous forms of paid third party assistance with 7.4 percent of the returns filed. Public Accountants are licensed at the state level with requirements varying by state. Only four states, North Carolina, Virginia, Kansas, and Wyoming do not regulate Public Accountants. National Tax Services, which in 1979 consisted entirely of H & R Block and Beneficial Financial Services Co., and Local Tax Services often provide their own training, but do not require employees to be Public Accountants.

³These estimates are based on the Special Academic Research File of the 1979 Individual Return Taxpayer Compliance Measurement Program (TCMP). The 1979 TCMP for individual taxpayers consists of a stratified random sample of approximately 50,000 individual taxpayer returns from the population of all individual taxpayers. Each return is subjected to a detailed line by line audit, the goal of which is to uncover all taxpayer errors in reporting. The Special Academic Research File has been assembled by the Taxpayer Compliance Group of the Internal Revenue Service Research Division, and contains the 1979 Individual Return data aggregated to the IRS administrative district level. There are 58 IRS administrative districts. Each district corresponds to a state except for California, Illinois, New York, Ohio, Pennsylvania, and Texas which have multiple districts.

estimate of the proportion of non-compliance attributable to returns signed by paid preparers or practitioners is still accurate for 1988, then over \$65 billion of unpaid tax could be attributable to such returns. Although some theoretical and empirical work has begun to characterize the impact of third party assistance on tax compliance, no clear picture has yet emerged. A complete characterization would include analyses of taxpayers' choices of third party assistance (if any), the effect of third party assistance on taxpayer compliance, and the IRS posture towards returns completed in various ways. In this paper we focus on taxpayers' choices of third party assistance.^{4 5}

Mode of Preparation	Total Returns Filed		Proportion of Noncompliance
	Number	Frequency	
Self	39959000	.442	.228
<i>Non-Paid</i>			
IRS assisted ⁶	1265686	.015	.004
Other Non-Paid ⁷	8317320	.089	.033
<i>Paid Preparers</i>			
Public Accountant	5605200	.062	.102
National Tax Service	8588600	.095	.054
Local Tax Service	10488000	.116	.132
Other Paid	6690000	.074	.122
<i>Practitioners</i>			
CPA ⁸	6057200	.067	.258
Attorney and CPA	3435450	.039	.067
Total	90406000	1.000	1.000

Source: Special Academic Research File of the 1979 Individual Return TCMP

⁴With the exception of Erard (1990), who estimates both the demand for third party assistance and the effect of third party assistance on tax compliance conditional upon the mode of assistance used, the empirical literature in economics on third party assistance has focused on the effect of third party assistance on taxpayer compliance. For examples, see Mazur and Nagin (1987), Milliron and Toy (1988), Westat (1987), and Klepper, Mazur and Nagin (1988). Little is known outside the IRS about the IRS posture towards returns completed in various ways.

⁵The accounting literature that addresses the demand for accountant services has focused on the provision and quality of auditing services, and not the preparation of individual tax returns. For examples, see Benston (1985), Shockley and Holt (1983), and Dopuch and Simunic (1980).

Using aggregate nested logit techniques, we find that the demand for third party assistance responds to many factors: age, education, employment status, federal auditing, and tax return characteristics. Perhaps most important among these is federal auditing. Higher audit rates increase the demand for practitioner services, but do not affect the demand for other third party assistance. More generally we find that as the tax burden increases, *ceteris paribus*, the demand for all modes of third party assistance increase. For example, increases in the number of tax dependents, or decreases in the amount of deductions increase the demand for third party assistance. With respect to socio-economic characteristics, states with a higher percentage of the adult population having received at least a high school education show lower demand for IRS or other non-paid assistance compared to self preparation.

Among these results, it is noteworthy that increases in the amount of deductions decrease the demand for third party assistance. This is contrary to the view that greater amounts of deductions are associated with more complex tax returns and thus should increase the demand for third party assistance.

An important feature of our analysis is the use of aggregate nested logit techniques. The nested logit method is well suited to the estimation of discrete choice models when there are many alternatives and when the choices can be partitioned into groups composed of similar alternatives. The appropriate partition among modes of third party assistance reinforces two aspects of the demand for tax preparation services previously emphasized in the theoretical literature.⁹ The first concerns the service aspects of third party assistance, typically provided by tax practitioners. The second concerns the desire by taxpayers to file tax returns which accurately reflect their true tax liability. Thus we combine CPA's and attorneys into a single category labelled practitioner and we combine public accountants, national tax services, local tax services, and other paid preparers into a single category labelled paid preparers. We also combine IRS assistance with other non-paid assistance into a single category labelled non-paid assistance.¹⁰

The paper is organized as follows. Section 2 summarizes pertinent findings of the literature on third party assistance. Section 3 presents an econometric specification of the demand for tax preparation services. Section 4 describes the data we use and models we estimate. Section 5 presents our results and Section 6 contains conclusions.

⁶The IRS Assistance category consists of three services. They are IRS Advice, with a population frequency of .011; IRS Prepared, with a population frequency of .001; and IRS Reviewed, with a population frequency of .004.

⁷The other nonpaid category usually consists of a family member who helps prepare the return. We have combined in this category all VITA prepared returns. VITA is an acronym for Volunteer Income Tax Assistance, which consists of unpaid volunteers who prepare returns after receiving limited instruction, typically from the IRS.

⁸CPA is an abbreviation for Certified Public Accountant.

⁹See Reinganum and Wilde (1989), and Scotchmer (1989).

¹⁰See footnotes 5 and 6 above for a description of IRS and other non-paid assistance.

2 Literature Review

2.1 Theoretical Results

Economic theory has only recently focused on the demand for third party assistance and its role in tax compliance. Scotchmer (1989) suggests that if taxpayers are not sure of their true taxable income, they will tend to overreport their tax liability. In her models, taxpayers hire the services of knowledgeable third parties who reduce or eliminate uncertainty, and thereby lower the amount of reported tax liability.

Reinganum and Wilde (1989) propose a model that focuses exclusively on the service aspects of third party assistance. They characterize four types of equilibria, depending on whether taxpayers prefer to use tax practitioners and whether the IRS prefers them to use tax practitioners. In the empirically relevant case, which occurs when tax practitioner penalties for noncompliance are sufficiently low and the efficiency gains from using practitioners are sufficiently high, the IRS prefers taxpayers to prepare their own returns, but taxpayers prefer to use a tax practitioner. In this case, the use of a tax practitioner is associated with lower compliance and higher audit rates.

In a subsequent paper, Graetz, Reinganum and Wilde (1989) analyze a model in which taxpayers can purchase expert opinions. Favorable opinions reduce the probability of a penalty on return items for which the tax code is ambiguous. This model is similar to that of Klepper and Nagin (1989) with the exception that in the former the tax authority is modeled as an endogenous actor responding to taxpayers' decisions to purchase expert opinions.

2.2 Empirical Results

Slemrod and Sorum (1984) present data from a 1982 survey of 600 Minnesota taxpayers designed to assess the cost of compliance with the federal income tax. They estimate the marginal cost of return preparation for each respondent as a function of the willingness to work extra hours and after tax income. The main results from their analysis are that the incidence and expense of using a paid third party increases with income, age and self employment. Using the same data, Slemrod (1989), estimates the choice between using third party assistance and self preparation conditional upon the decision to itemize deductions. He finds that for taxpayers who itemize deductions, higher marginal tax rates increase the demand for third party assistance. A more recent survey undertaken by Collins, Milliron, and Toy (1988) assesses taxpayer attitudes towards filing tax returns. These authors find that nearly 70 percent of taxpayers want most to file the correct return, while 25 percent want most to minimize taxes and only a small number

want most to minimize effort. They also find that increases in income and age increase the likelihood of engaging paid third parties, while greater tax knowledge decreases the likelihood of engaging a paid third party. Hite (1987) obtains similar results, reporting that the probability of using a paid third party increases with age, complexity of return, and the probability of an audit.

Using data from the IRS 1983 Individual Master File, Long and Caudill (1987) find that increases in income, age, return complexity, number of dependents, and federal and state marginal tax rates increase the probability of using a paid third party.

Dubin, Graetz and Wilde (1989) use administrative district level data from the 1979 Taxpayer Compliance Measurement Program to estimate the probability that a paid third party will be used. They find that many forms of complex income increase the probability of using a paid third party, while higher levels of education, unemployment, and the percentage of the population between 45 and 65 years old significantly reduce the probability of using a paid third party. Unlike Long and Caudill, they find no significant relationship between the state income tax rate or self employment and the use of a paid third party.¹¹

In a recent paper, Erard (1990) studies the demand for tax preparation services using a trinomial probit of the choice to use self preparation, a tax practitioner, or any other mode of third party assistance (either paid or nonpaid). Using a random subsample of the 1979 TCMP individual records, he estimates the demand for third party assistance as a function of types of income on returns, audit history, and whether the taxpayer was over 65 years of age.¹² He finds that the probability of using a practitioner relative to self preparation increases with the presence of income from capital gains, small business or farm activity, and rents and royalties, the marginal tax rate, number of tax forms attendant the return, being over the age of 65, previous audit history, and the IRS district audit rate. He also finds that the probability of using any other form of third party assistance relative to self preparation increases with income from small business or farm activity, and rents and royalties, being married, or over the age of 65, the number of forms attendant the return, and the IRS district level audit rate, while it decreases with the marginal tax rate.¹³

¹¹One major difference between the Individual Master File data used by Long and Caudill, and the TCMP data used by Dubin, Graetz and Wilde is that the Master File data contains only amounts reported on returns while the TCMP data contain corrected amounts resulting from IRS examinations of returns. The latter study uses the collected amounts and therefore avoids potential problems with endogenous reported tax return items.

¹²Absent from his specification are socio-economic variables such as education, and tax return items such as exemptions and deductions.

¹³Feinstein (1989) has shown that it may not be possible to separate the effects of income and marginal tax rates using TCMP individual data. The significance of both income and marginal tax rates in Erard's work may result from the use of a categorical measure of income combined with a continuous measure of marginal tax rates.

The preceding review of both theoretical and empirical results suggests two hypotheses which we test in this paper. The first hypothesis is that taxpayer use of third party assistance increases with total and complex income, audit rates, tax rates, and age while it decreases with the level of education attained and the unemployment rate. The second hypothesis is that taxpayers use third party assistance to reduce uncertainty about the correct amount of tax owed and to reduce the time costs of return preparation and compliance with the tax code.

These two hypotheses are to some extent related. For example, if higher income is associated with higher opportunity costs of time, then increases in income increase the time costs associated with filing tax returns. Similarly, increases in audit rates directly increase the time costs of filing when these costs include time spent meeting the requirements of an audit. Further, more education or less income from sources other than wages, interest, or dividends may be associated with less uncertainty about the correct tax liability, and thus reduce the demand for preparation services.

3 A Model of Paid Preparer Selection

With the exception of Erard (1990), the empirical literature has focused on the choice between self preparation and all modes of third party assistance. This is an unfortunate consequence of both the data researchers have used and the models they estimate. We next describe an aggregate nested logit model which distinguishes among the modes of third party assistance.

We conceptualize the taxpayer's choice as a two stage sequential process.¹⁴ In the first stage the taxpayer chooses one of the following categories: (1) nonpaid assistance, (2) self preparation, (3) paid preparer or (4) practitioner. In the second stage the taxpayer makes a further decision regarding which mode of assistance to use within the broader category. Let $i = 1, \dots, I$ index the taxpayer choice of category, and let $j = 1, 2, \dots, J_i$ index the decision, conditional upon using category i , to choose from a set of J_i specific assistance modes. We assume that the taxpayer derives utility U_{ij} from alternative ij where

$$U_{ij} = V_{ij} + \epsilon_{ij}, \tag{1}$$

and where V_{ij} is a function of non-stochastic observed attributes, and ϵ_{ij} is a random component of unobserved attributes. The probability that the individual chooses

¹⁴Our model of individual choice falls into the class of random utility models (see McFadden, 1981).

alternative ij is

$$\begin{aligned}
P_{ij} &= P[U_{ij} \geq U_{kl}, \quad \forall kl \neq ij] \\
&= P[V_{ij} + \epsilon_{ij} \geq V_{kl} + \epsilon_{kl}, \quad \forall kl \neq ij] \\
&= P[\epsilon_{kl} - \epsilon_{ij} \leq V_{ij} - V_{kl}, \quad \forall kl \neq ij].
\end{aligned} \tag{2}$$

The strict utility component is assumed to be of the form

$$V_{ij} = \beta' X_{ij} + \alpha' Y_i \tag{3}$$

where X_{ij} is a vector of observed attributes varying with each of the j alternatives in the i^{th} category, and Y_i is a vector of observed attributes varying only with the i^{th} category. We also assume that ϵ_{ij} has a generalized extreme value distribution with cumulative distribution function

$$F(\epsilon; \theta) = \exp \left(- \sum_{i=1}^I \left(\sum_{j=1}^{J_i} e^{-\epsilon_{ij}/\theta_i} \right)^{\theta_i} \right),$$

where $\epsilon = (\epsilon_{ij})$. Under these assumptions, McFadden (1978) demonstrates that P_{ij} can be written as the product of the conditional probability $P_{j|i}$ and marginal probability P_i where each probability is in the multinomial logit form:

$$P_{j|i} = \frac{e^{\beta' X_{ij} + \alpha' Y_i}}{\sum_{j=1}^{J_i} e^{\beta' X_{ij} + \alpha' Y_i}} = \frac{e^{\beta' X_{ij}}}{\sum_{j=1}^{J_i} e^{\beta' X_{ij}}}, \tag{4}$$

$$P_i = \frac{e^{\alpha' Y_i + \theta_i' I_i}}{\sum_{i=1}^{J_i} e^{\alpha' Y_i + \theta_i' I_i}}, \tag{5}$$

where I_i denotes the expected maximum utility (or inclusive value) taxpayers derive from alternatives in the i^{th} category,

$$I_i = \log \sum_{j=1}^{J_i} e^{\beta' X_{ij}} \tag{6}$$

and θ_i is a measure of the dissimilarity of alternatives in the i^{th} category.¹⁵

¹⁵McFadden (1978) proves that if the random component of the utility function is specified with the

Our data, as already noted, has been aggregated to the IRS administrative district level. With aggregate data, we estimate $P_{j|i}$ and P_i using minimum chi-square estimation techniques.¹⁶ We form the log odds for equation (4) as

$$\log \left[\frac{P_{j|i}}{P_{1|i}} \right] = \beta'(X_{ij} - X_{i1}) \quad \forall j \neq 1. \quad (7)$$

Similarly, we form the log odds for equation (5) as

$$\log \left[\frac{P_i}{P_1} \right] = \alpha'(Y_i - Y_1) + \theta'_i(I_i - I_1) \quad \forall i \neq 1. \quad (8)$$

To estimate equations (7) and (8) requires estimates of the choice probabilities for which the natural measure is the observed population frequencies. Denote the probability of an individual in the k^{th} IRS district choosing alternative ij as P_{ij}^k . Since alternatives are exclusive, and together exhaust all possibilities, it follows that for any IRS district k , $\sum_{i=1}^I \sum_{j=1}^{J_i} P_{ij}^k = 1$. Let the number of cases in the k^{th} IRS district be N^k and denote the frequency of occurrences of alternative ij in district k as

$$F_{ij}^k \equiv \frac{1}{N^k} \sum_{h=1}^{N^k} \delta_{hij}^k \quad (9)$$

where δ_{hij}^k is an indicator equal to one if the h^{th} person in district k chooses alternative ij .¹⁷ Suppressing the superscript k , we stack the log odds of alternative ij relative to alternative $i1$ in each IRS district as

$$\log \left[\frac{F_{j|i}}{F_{1|i}} \right] = \beta'(X_{ij} - X_{i1}) + \mu_{j|i} \quad (10)$$

GEV distribution, and if θ is between 0 and 1, the resultant nested logit is consistent with random utility maximization. If $\theta = 1$, then the choice of an alternative is unaffected by the presence or absence of other alternatives, i.e., all alternatives are independent of each other. In this restrictive case, the joint probability P_{ij} is of the multinomial logit form. A value of θ in the interval $(0, 1)$ constitutes rejection of this independence, and the joint probability P_{ij} is of the nested multinomial logit form. In this case, alternatives are independent only within a category. If $\theta = 0$, the nested logit model assesses the choices among alternatives in category i as if there was a single maximal alternative.

¹⁶These techniques were developed by Berkson (1944) in the field of bioassay, but are well suited to any choice estimation problem with a limited set of alternatives and many observations for each alternative. Our development follows Amemiya (1985).

¹⁷Our observations have an average of 65 individuals per cell. Monte Carlo evidence reported in Domencich and McFadden (1975) shows that when cell sizes are large (over 30 observations) the Berkson estimator of equation (7) has very small bias, even when the cell variances are large. (see Table 5.1, page 113)

$$\text{for } i = 1, 2, 3, 4 \text{ and } j = 2, 3, \dots, J_i$$

$$\text{where } \mu_{j|i} = \log \left[\frac{F_{j|i}}{F_{1|i}} \right] - \log \left[\frac{P_{j|i}}{P_{1|i}} \right].$$

Similarly, we stack the log odds of category i relative to category 1 as

$$\log \left[\frac{F_i}{F_1} \right] = \alpha'(Y_i - Y_1) + \theta'_i(I_i - I_1) + \mu_i \quad (11)$$

$$\text{for } i = 1, 2, 3, 4$$

$$\text{where } \mu_i = \log \left[\frac{F_i}{F_1} \right] - \log \left[\frac{P_i}{P_1} \right].$$

Straightforward calculations show that

$$E(\mu_{j|i}) = 0$$

$$V(\mu_{j|i}) = \frac{1}{N} \left(\frac{1}{P_{j|i}} + \frac{1}{P_{1|i}} \right) \quad \text{for } j \neq 1, \quad (12)$$

$$\text{cov}(\mu_{j|i}, \mu_{1|i}) = \frac{1}{NP_{1|i}} \quad \text{for } j \neq 1. \quad (13)$$

$$\text{and that } E(\mu_i) = 0$$

$$V(\mu_i) = \frac{1}{N} \left(\frac{1}{P_j} + \frac{1}{P_1} \right) \quad \text{for } j \neq 1 \quad (14)$$

$$\text{cov}(\mu_j, \mu_1) = \frac{1}{NP_1} \quad \text{for } j \neq 1. \quad (15)$$

This covariance structure allows for minimum chi-square estimation of the nested logit model, correcting simultaneously for two sources of heteroskedasticity. Since P_{ij} changes with each alternative, the variance is heteroskedastic. Further, N is not constant, but varies with the size of the IRS district, adding an additional source of heteroskedasticity. Finally, induced correlation with the comparison group (the alternative in the denominator of the log odds in equations (10) and (11)) is accounted for by the covariance terms.¹⁸

¹⁸Note that sequential estimation of the simultaneous equation system given by equations (10) and (11) is fully efficient because $\text{cov}(\mu_i, \mu_{j|i}) = 0$. Amemiya (1978) shows that in general this condition does not hold for the nested logit model, and as a result, sequential estimation can entail a loss of efficiency compared with full information procedures.

4 Model Specification

4.1 Data

The Special Academic Research File of the 1979 Individual Return TCMP defines 12 modes of third party assistance. Seven cover paid preparers or practitioners and five cover nonpaid assistance. An additional mode covers self-prepared returns. To avoid districts with too few observations for an alternative, we combine all attorneys into a single category and assign the following mnemonics to the resulting practitioner alternatives:¹⁹

ATT = Attorney, or Attorney and CPA
CPA = Certified Public Accountant

The four paid-preparer alternatives are:

NTS = National Tax Service
LTS = Local Tax Service
PA = Public Accountant
OPP = Other Paid Preparer.

Finally, we combine several nonpaid assistance alternatives and assign the following mnemonics:

IRS = IRS prepared, IRS reviewed, or IRS assisted
ONP = Volunteer Income Tax Assistance, or Other Nonpaid Preparers
SELF = Self-prepared.

Our explanatory variables include the IRS audit rate, RATE, the percent of the adult population over 65 years of age, AGE65, the percent of the adult population with at least a high school diploma, EDUCATE, and the percent of the work force employed in manufacturing, PERMAN. IRS district level audit rates were constructed using data from the 1979 *Annual Report of the Commissioner of Internal Revenue*. The percent of the adult population with at least a high school education, the percent of the adult population over age 65, and the percent of the work force employed in a manufacturing industry are from the Statistical Abstract of the United States for the year 1979.

The tax return items we use reflect the TCMP adjusted levels for three basic components of the tax calculation, income, deductions, and exemptions. Objective measures of these three components of the tax return more accurately reflect the true tax situation of the taxpayer before engaging the services of a third party. EXEMPT is the

¹⁹In particular an alternative with zero observations leaves the log odds of equation (10) undefined.

total number of exemptions eligible to be claimed on returns in an IRS district net of the over 65 years of age and blindness exemption. This variable measures the number of tax dependents. DEDUCT is the total amount of allowable deductions on returns in an IRS district, and reflects the amount of downward adjustment to taxable income the taxpayer may claim. We use two measures of income; SIMPLE is the total amount of wage, interest, and dividend income reported on returns, and COMPLEX is the total amount of income from sources other than wages, interest, and dividends. Wages, interest, and dividends were, in 1979, the easiest income items to be verified by the IRS. COMPLEX income therefore represents sources that are the most difficult to trace.²⁰ We report the per return mean values for these variables in Table 2.

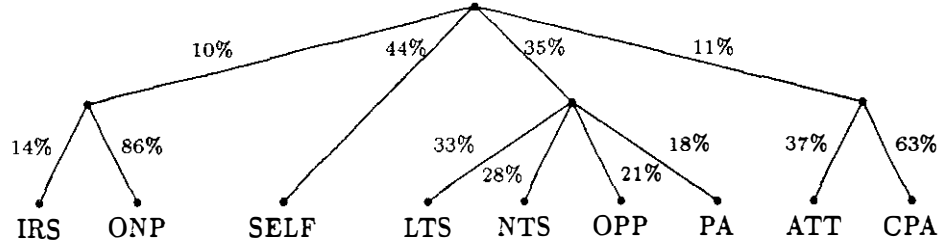
Table 2
MEAN VALUES OF EXPLANATORY VARIABLES

Category Mode Variable	Non-paid		SELF	Preparers			Practitioners		
	IRS	ONP		NTS	LTS	OPP	PA	ATT	CPA
AGE65	.153	.153	.153	.153	.153	.153	.153	.153	.153
EDUCATE	.673	.673	.673	.673	.673	.673	.673	.673	.673
PERMAN	.222	.222	.222	.222	.222	.222	.222	.222	.222
RATE	.018	.018	.018	.018	.018	.018	.018	.018	.018
DEDUCT	4,835	4,736	6,090	5,128	5,454	5,607	6,220	8,237	9,986
EXEMPT	2,098	1,793	2,122	2,532	2,700	2,490	2,662	2,642	2,811
COMPLEX	950	1,271	1,563	2,186	3,205	4,103	5,878	9,432	15,400
SIMPLE	9,716	9,456	14,417	15,508	17,306	17,314	20,112	24,980	34,304

Sources: Special Academic Research File of the 1979 Individual Return TCMP, 1979 Annual Report of the Commissioner of Internal Revenue, and Statistical Abstract of the United States.

²⁰The TCMP provides not only what the tax return reported amounts were, but what the IRS believed the true amount to be. We use the corrected amounts of deductions and exemptions as meaningful measures of the true amounts of these items. For deductions and exemptions this avoids at least two problems. The first is that many people claimed extra deductions for withholding purposes in anticipation of large deduction amounts on the tax return. The second is, prior to the Tax Reform Act of 1986, certain exemptions (such as the number of dependents) claimed on returns were difficult to verify by the IRS. This led to excessive claims of exemptions on returns. Therefore, exemptions may have been part of a tax evasion strategy, obscuring true identification of the taxpayer's situation. For example, in 1987, the first year that social security numbers were required for all exemptions claimed, the number claimed fell by over 7 million, with the bulk coming from returns that claimed multiple exemptions on 1986 returns claiming none (beyond the personal exemption) in 1987. The use of the corrected amount of exemptions eliminates these problems, and after subtracting exemptions for blindness and over 65 years of age reflects the number of persons supported by the taxpayer. Although their findings of unreported income may understate the true amount, we use the IRS corrected amounts for both income variables as proxies for the true level of income. In particular, greater amounts of nonwage reflect larger tax liabilities and increasingly more complex returns because often these sources of income require separate forms, and special accounting procedures, (as with capital gains, business and partnership income, and rental income) requiring detailed knowledge of the tax code.

Figure 1



4.2 Estimation

The choice of third party assistance is estimated in two stages, as illustrated in Figure 1. In the first stage we estimate separate choice models for specific alternatives from within the three categories defined above. The first of these categories, non-paid assistance, includes the IRS and ONP alternatives. The second, paid preparers, includes the LTS, NTS, OPP, and PA alternatives. The third, practitioners, includes the ATT and CPA alternatives. Each of these models takes the form:²¹

$$\log \left[\frac{F_{j|i}}{F_{1|i}} \right] = \beta_{0j|i} + \beta_{1j|i}AGE65 + \beta_{2j|i}EDUCATE + \beta_{3j|i}PERMAN + \beta_{4j|i}RATE \\ + \beta_{5j|i}(DEDUCT_{j|i} - DEDUCT_{1|i}) + \beta_{6j|i}(EXEMPT_{j|i} - EXEMPT_{1|i}) \\ + \beta_{7j|i}(COMPLEX_{j|i} - COMPLEX_{1|i}) \\ + \beta_{8j|i}(SIMPLE_{j|i} - SIMPLE_{1|i}). \quad (16)$$

In the second stage we estimate a choice model for the four categories, nonpaid assistance (IRS or ONP), self prepared (SELF), paid preparer (LTS, NTS, OPP, or PA), and practitioner (ATT or CPA). This model has the form:

$$\log \left[\frac{F_i}{F_1} \right] = \gamma_{0i} + \gamma_{1i}AGE65 + \gamma_{2i}EDUCATE + \gamma_{3i}PERMAN + \gamma_{4i}RATE \\ + \gamma_{5i}(DEDUCT_i - DEDUCT_1) + \gamma_{6i}(EXEMPT_i - EXEMPT_1) \\ + \gamma_{7i}(COMPLEX_i - COMPLEX_1) + \gamma_{8i}(SIMPLE_i - SIMPLE_1) \\ + \theta_i(INCV_i - INCV_1), \quad (17)$$

where $INCV_i$ is the inclusive value derived from the stage one estimates.²² In equation

²¹The normalization by alternative one in equation (16) is arbitrary. In the first stage models we use the PA, CPA, and IRS alternatives respectively as normalizations. In the second stage model we use the self-preparation category for the normalization.

²²Two issues arise with this specification. The first is that we have excluded the state tax and

(15) we use the self-prepared category for normalization as the inclusive value for alternatives in this category is zero.

5 Results

Tables 3 and 4 present our results for the sequential estimation of the aggregate nested logit model. For the socio-economic variables, AGE65, EDUCATE, and PERMAN, the coefficients are interpreted as the effect on the log odds from a 1 percent change in those rates, while DEDUCT, COMPLEX, and SIMPLE assess the effect of a one-thousand dollar per return change on the log odds. EXEMPT measures the change from an additional exemption per return on the log odds. The measure of goodness of fit we report is a modified R^2 statistic derived in Buse (1973) and discussed in Amemiya (1981).²³

Table 3 shows results from the stage one estimation. The first three columns show the estimated log odds equations for the National Tax Service, Local Tax Service, Other Paid Preparer, and Public Accountant alternatives. Considering first socio-economic variables, increases in the percent of the adult population over 65 tend to increase the likelihood of using a public accountant compared to other paid preparers. Increases in the percent of the workforce employed in manufacturing industries significantly increases the likelihood of using a public accountant compared to any other paid preparer (a result which has no immediate explanation). Finally, education has no effect on the choice between paid preparers.

A more interesting observation is that increases in the federal audit rate decrease the likelihood of using either a national or local tax service compared to the use of a public accountant, but have no effect on the choice between other paid preparers and public accountants. We surmise that this result follows from a perception by taxpayers that public accountants better shield them from audits or, in case the taxpayer does get audited, better represent the taxpayer than do tax services.

With respect to tax return items, we get a variety of significant results, but no obvious pattern is discernable.

unemployment rates. We found no effect from these variables in any of the models estimated, and so excluded them from the models presented. The second issue concerns the potential endogeneity of the IRS audit rate with the choice of tax preparation mode. Based on the test developed in Hausman (1978) we could not reject the exogeneity of the audit rate in our model.

²³The statistic is $(WSSR_r - WSSR_u) / WSSR_r$, where $WSSR_r$ is the weighted sum of squared residuals from minimum chi-square estimation of a restricted model consisting of alternative specific constants only and $WSSR_u$ is the weighted sum of squared residuals from minimum chi-square estimation of the unrestricted model.

Column four in Table 3 shows the estimated log odds for the CPA and attorney alternatives. Given that a taxpayer uses one of these practitioners, the choice between them is driven largely by exemptions and deductions. In particular, increases in exemptions increase the demand for attorneys relative to CPA's, while increases in deductions decrease the demand for attorneys relative to CPA's. Neither the federal audit rate nor income has any effect on this choice, although increases in the percent of the population over 65 tend to favor attorneys over CPA's.

The last column of Table 3 shows the estimated log odds for the IRS assistance and Volunteer Income Tax Assistance or Other Non-Paid alternatives. Even though only 14 percent of the demand for non-paid preparers accrues to the IRS (Figure 1), we observe that increases in the federal audit rate decrease the relative demand for IRS assistance. This result may follow from taxpayer mistrust of the IRS in states where auditing is high, or because the IRS perceives there to be a trade-off between return assistance and auditing; i.e., in states where the audit rate is high, less effort may be devoted by the IRS to providing taxpayer assistance.

With respect to tax return items, we observe a pattern of effects for the choice between IRS assistance and other nonpaid assistance similar to that found for the choice between attorneys and CPA's; increases in exemptions or decreases in deduction increase the demand for IRS assistance relative to other forms of nonpaid assistance.

Table 4 presents results from the stage two estimation comparing the choice of non-paid assistance, paid preparers, and practitioners, respectively, with self preparation. In these comparisons, the three socio-economic variables are largely insignificant. In fact, the choice between any form of third party assistance and self preparation is driven almost entirely by tax return items. The one exception, however, is an important one. Increases in the federal audit rate increase the demand for practitioners—CPA's or attorneys—relative to self preparation, but have no effect on the demand for nonpaid assistance or paid preparers relative to self preparation. Moreover, higher audit rates increase the demand for practitioners when the choice is between paid preparers and practitioners.

With respect to tax return items, we observe a rather consistent set of effects: increases in exemptions generally increase the demand for third party assistance, especially paid assistance; increases in deductions generally decrease the demand for third party assistance (except for practitioners), and increases in complex income generally increase the demand for third party assistance (again except for practitioners). Income from wages, interest, or dividends is largely irrelevant.

Table 3
STAGE 1 PARAMETER ESTIMATES

Log Odds Variable (<i>t</i> -statistic)	$\frac{NTS}{PA}$	$\frac{LTS}{PA}$	$\frac{OPP}{PA}$	$\frac{ATT}{CPA}$	$\frac{IRS}{ONP}$
ONE	2.514 (2.639)	3.352 (3.513)	2.101 (2.273)	-1.433 (-1.815)	-.298 (-.295)
AGE65	-.03545 (-1.630)	-.0483 (-2.291)	-.0521 (-2.036)	.0361 (1.956)	.0115 (.454)
EDUCATE	-.0065 (-.715)	-.0125 (-1.293)	-.0071 (-.796)	.0066 (.823)	.0001 (.0062)
PERMAN	-.0328 (-3.378)	-.0285 (-2.862)	-.0216 (-2.293)	.0126 (1.573)	-.0134 (-1.256)
RATE	-.2502 (-2.407)	-.2864 (-2.577)	-.0764 (-.791)	-.126 (-1.081)	-.3399 (-2.179)
EXEMPT	.174 (2.872)	-.0001 (-2.962)	.2604 (4.491)	.4535 (5.234)	.3194 (2.32)
DEDUCT	-.033 (-.833)	-.0479 (-1.548)	.0196 (.495)	-.0754 (-2.819)	-.3184 (-2.006)
COMPLEX	.0316 (1.941)	.0377 (2.091)	.0096 (1.156)	-.0026 (-.338)	.0179 (1.042)
SIMPLE	-.0007 (.0544)	.0193 (2.976)	-.0176 (-1.389)	-.0095 (-.617)	.0007 (.201)
R^2			.75	.76	.57

Table 4
STAGE 2 PARAMETER ESTIMATES OF

Log Odds Variable	<i>IRS & ONP</i> <i>Self</i>	<i>PA,LTS,NTS,OPP</i> <i>Self</i>	<i>CPA & ATT</i> <i>Self</i>
ONE	-.922 (-1.328)	-.101 (-.149)	-2.116 (-3.770)
AGE65	.0374 (1.732)	.0087 (.723)	.0151 (.967)
EDUCATE	-.0163 (-2.173)	-.0001 (-.109)	.0007 (.129)
PERMAN	-.0100 (-1.548)	-.0014 (-.251)	.0001 (.07792)
RATE	.0051 (.0729)	-.0039 (-.0691)	.1569 (2.256)
EXEMPT	.0311 (1.438)	.0482 (3.879)	.0474 (2.344)
DEDUCT	-.0297 (-2.421)	-.0376 (-4.826)	.0046 (.426)
COMPLEX	.0204 (2.278)	.0079 (2.370)	.0076 (1.646)
SIMPLE	.0001 (.115)	.0314 (1.768)	-.0063 (-.182)
<i>INC_{IRS&ONP}</i>	.621 (5.015)		
<i>INC_{PA,NTS,LTS,OPP}</i>		-.0677 (.773)	
<i>INC_{CPA&ATT}</i>			.536 (4.163)
<i>R</i> ²			.74

6 Conclusion

Our results suggest that the demand by taxpayers for third party assistance in the preparation of tax returns results from three influences, two of which have previously been emphasized in the literature, and one of which is new.

The first influence is associated with taxpayers who demand special services, such as the ability to represent the taxpayer before the IRS. Tax practitioners are uniquely qualified to provide this service, and in states where federal audit rates are higher, so too is the demand for practitioner services. The second influence is associated with taxpayers who otherwise might prepare their own returns if it were not for uncertainty about their true tax liability. For these taxpayers, greater amounts of income from sources other than wages, interest, or dividends increase the demand for third party

assistance, presumably because these types of income require additional tax forms and special accounting procedures in order to determine the correct tax liability. The third influence is associated with increases in the tax burden, as measured by the number of tax dependents or amount of deductions. In particular, we find that when the number of tax dependents increases, the demand for third party assistance increases. Furthermore, contrary to the view that more deductions imply a more complex tax return, which should increase the demand for third party assistance, we find that more deductions reduce the demand for third party assistance. These two results, which have not previously appeared in the literature, we interpret as representing the effects of changes in the tax burden on the demand for third party assistance.

We generally confirm the results of previous research, but with finer distinctions among the categories of third party assistance. For example, Erard (1990) shows that an increase in the federal audit rate increases the demand for practitioners and for all other forms of third party assistance, with the latter treated as a single group. We also find that increases in the federal audit rates increase the demand for practitioners, but not for paid preparers or nonpaid assistants, which we treat as separate groups. Similarly, while the literature generally concludes that increase in complex income increase the demand for all types of third party assistance, we find that increase in complex income increase the demand for paid and nonpaid assistants, but not for practitioners.

A logical extension of the demand for tax return preparation services presented here would be to estimate the amount of tax noncompliance attributable to returns conditional upon the mode of third party assistance used. This calculation requires estimation of two components of noncompliance. The first, presented here, is the change in demand for third party assistance, and the second is the change in noncompliance attributable to that mode of preparation. For example, an increase in the federal audit rate will increase the demand for practitioner services. If average noncompliance for this category falls by less than the proportionate increase in demand for practitioner services, then total noncompliance for this category is expected to increase. Our results have demonstrated that attendant with this increase in compliance will be a change in the composition of third party assistance. Dubin, Graetz, and Wilde (1989b), show that an increase in the federal audit rate is likely to increase total compliance. Therefore, we would expect a shift in the composition of non-compliance as individuals change their mode of tax preparation.

References

- Amemiya, T. "On a Two-step Estimation of a Multivariate Logit Model." *Journal of Econometrics* 8 (1978):13-21.
- Amemiya, T. "Qualitative Resposne Models: A Survey." *Journal of Economic Literature* 19 (1981):483-536.
- Amemiya, T. *Advanced Econometrics*. Cambridge: Harvard University Press, 1985.
- Benston, G. J. "The Market for Public Accounting Services: Demand, Supply and Regulation." *Journal of Accounting and Public Policy* 4 (1985):33-79.
- Berkson, J. "Application of the Logistic Function to Bio-Assay." *Journal of the American Statistical Association* 39 (1944):357-365.
- Buse, A. "Goodness of Fit in Generalized Least Square Estimation." *American Statistician* 27 (1973):106-8
- Collins, Julie; Milliron, Valerie; and Toy, Daniel. "Tax Preparer Usage and its Implications." Unpublished manuscript. University of North Carolina, 1988.
- Domencich, Thomas and McFadden, Daniel. *Urban Travel Demand*. New York: North-Holland Publishing Co., 1975.
- Dopuch, N., and Simunic, D. "The Nature of Comptetition in the Auditing Profession." In *Regulation and the Accounting Profession* edited by J.W. Buckley and J. F. Weston, pages 77-94. New York:Lifetime Learning, 1980.
- Dubin, Jeff; Graetz, Michael; and Wilde, Louis. "The Report of the United States to the International Fiscal Association on the Administrative and Compliance Costs of Taxation." In *Cahiers of the International Fiscal Association, 43 Congress*. Netherlands: Kluwer, 1989a, pp. 311-344.
- Dubin, Jeff; Graetz, Michael; and Wilde, Louis. "The Effect of Audit Rates on Federal Income Tax Filings and Collections, 1977-1986." Social Science Working Paper No. 711. Pasadena: California Institute of Technology, 1989b.
- Erard, Brian. "An Endogenous Switching Analysis of Preparation Mode and Tax Compliance." Unpublished Manuscript. Ann Arbor: University of Michigan, 1990.
- Feinstein, Jonathon S. "An Econometric Analysis of Income Tax Evasion and Its Detection." U.S. Internal Revenue Service. Washington D.C., 1989.
- Graetz, Michael; Reinganum, Jennifer; and Wilde, Louis. "Expert Opinions and Taxpayer Compliance: A Strategic Analysis." Social Science Working Paper No. 710. Pasadena: California Institute of Technology, 1989.

- Hausman, Jerry. "Specification Tests in Econometrics." *Econometrica* 46 (1978):1251-1271.
- Hite, Peggy. "Characteristics of Taxpayers Who Use Preparers: An Exploratory Study." Paper prepared for IRS Research Conference on The Role of Tax Practitioners in the Tax System, Washington D.C., November 16-17, 1987.
- Internal Revenue Service. *Net Tax Gap and Remittance Tax Gap Estimates*, Publication 1415. Washington: Research Division, Compliance Estimates Group, April, 1990.
- Klepper, Steven; Nazur, Mark; and Nagin, Daniel. "Expert Intermediaries and Legal Compliance: The Case of Tax Preparers." Working Paper 88-26. Pittsburgh: Carnegie Mellon University, 1988.
- Klepper, Steven and Nagin, Daniel. "The Role of Tax Practitioners in Tax Compliance." *Policy Sciences* 22 (1989):167-192.
- Long, James and Caudill, Steven. "The Usage and Benefits of Paid Tax Return Preparation." *National Tax Journal* 40 (1987):35-46.
- Mazur, Mark and Nagin, Daniel. "Tax Preparers and Tax Compliance: A Theoretical and Empirical Analysis." Unpublished manuscript. Pittsburgh: Carnegie Mellon University, 1987.
- McFadden, D. "Modeling the Choice of Residential Location." In *Spatial Interaction Theory and Planning Models*, edited by A. Karlquist et al. Amsterdam: North-Holland Publishing, 1978.
- McFadden, D. "Econometric Models of Probabilistic Choice." In *Structural Analysis of Discrete Data with Econometric Applications*, edited by Charles F. Manski and Daniel McFadden. Cambridge: MIT Press, 1981.
- Milliron, Valerie and Toy, Daniel "Tax Compliance: An Investigation of Key Features." Unpublished Manuscript. College Park: Pennsylvania State University, 1986.
- Reinganum, Jennifer and Wilde, Louis. "Equilibrium Enforcement and Compliance in the Presence of Tax Practitioners." Unpublished manuscript. Iowa City: University of Iowa, 1989.
- Scotchmer, Suzanne. "Who Profits from Taxpayer Confusion?" *Economics Letters* 29 (1989):49-55.
- Schockley, Randolph A. and Holt, Robert N. "A Behavioral Investigation of Supplier Differentiation in the Market for Audit Services" *Journal of Accounting Research* 21 (1983):545-564.

Slemrod, Joel and Sorum, Nikki. "The Compliance Cost of the U.S. Individual Income Tax System." *National Tax Journal* 37 (1984):461-474.

Slemrod, Joel. "The Return to Tax Simplification: An Econometric Analysis." *Public Finance Quarterly* 17 (1989):3-27.

Westat. "A Survey of Tax Practitioners." U.S. Internal Revenue Service. Rockville, Maryland, 1987.