

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

U·M·I

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

Order Number 9332990

**Income tax return preparation fees and tax savings of using tax
return preparers**

Lin, Suming, Ph.D.
Arizona State University, 1993

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

**INCOME TAX RETURN PREPARATION FEES AND
TAX SAVINGS OF USING TAX RETURN PREPARERS**

by

Suming Lin

**A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy**

ARIZONA STATE UNIVERSITY

August 1993

INCOME TAX RETURN PREPARATION FEES AND
TAX SAVINGS OF USING TAX RETURN PREPARERS


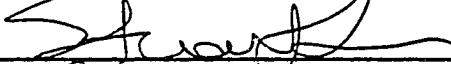
by

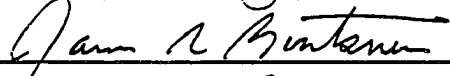
Suming Lin

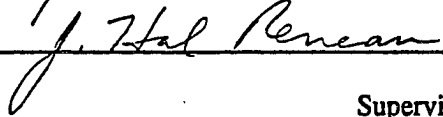
has been approved

June 1993

APPROVED:

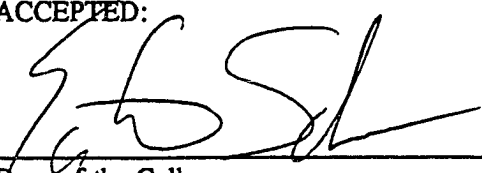

_____.Chairperson






Supervisory Committee

ACCEPTED:



Dean of the College



Dean, Graduate College

ABSTRACT

Evidence indicates that almost fifty percent of the individual federal income tax returns were prepared by tax return preparers. Using 1986-1987 income tax return data from the Statistics of Income Panel of Individual Returns, this paper examines (1) the determinants of tax return preparation fees, (2) the tax savings of using tax return preparers, and (3) the determinants of taxpayers' propensity to engage tax preparers. The principal findings are:

- (1) Tax return preparation fees are positively correlated with income during the 1986-1987 period. The presence of a supplemental schedule on a tax return generally increases preparation fees in the 1986 tax year. However, this is not true in 1987 except in the case of Schedule C attached to a sole proprietor's tax return.
- (2) The estimates of aggregate tax compliance cost expended on preparation fee are \$7.4 billion in 1986 and \$8.6 billion in 1987. The analysis appears to suggest that the Tax Reform Act of 1986 did not reverse the growth in tax compliance costs in the 1980s.
- (3) The results of endogenous switching regression indicate that, except for highly complex returns, returns prepared by tax preparers have a higher tax burden than self-prepared returns in both 1986 and 1987.
- (4) The higher the potential tax savings of using a preparer, the higher the likelihood a taxpayer is to engage a preparer.
- (5) Results from simulations indicate that the demand for preparation service is price-inelastic. This suggests that increases (decreases) in preparation price will produce increases (decreases) in tax preparers' revenue.

To the memory of my parents

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my mentor and committee chair, Dr. Charles Christian. His guidance, encouragement, and friendliness have motivated me throughout my graduate study and research. Because of his ceaseless support and patient coaching, my study in this great country has been a peaceful, pleasant, and productive academic journey. My thanks are also extended to the members of my dissertation committee. Dr. Stuart Low has been an invaluable source of help on econometric issues. Dr. James Boatsman and Dr. Hal Reneau have contributed immensely to the success of my graduate study and to the completion of this dissertation. The committee's support and encouragement can never be repaid but they will always be appreciated.

I am grateful to the Ministry of Education of the Republic of China and the Yin's Education Foundation for financial support. Also, I am grateful to the School of Accountancy at Arizona State University for providing excellent faculty and academic resources. Profound thanks are also extended to my fellow graduate students for their friendship and assistance.

Finally, I would like to thank my wife, Li-Chu, for her dedication and understanding. Her total support and willingness to sacrifice have made my dream come true.

TABLE OF CONTENTS

	Page
LIST OF TABLES	viii
I INTRODUCTION	1
1.1 Research Questions and Motivation	1
1.2 Organization of the Study	5
II LITERATURE REVIEW	6
2.1 Expenditures for Professional Tax Assistance	6
2.2 Tax Savings of Using Preparers	7
2.3 Determinants of Preparation Mode	8
III RESEARCH METHODS	13
3.1 Econometric Model	14
3.1.1 Endogenous Switching Regression Model	14
3.1.2 Estimation of Preparation Fees	18
3.2 Model's Explanatory Variables	19
3.2.1 Total Tax Burden Function	19
3.2.2 Preparation Mode Function	21
3.2.3 Preparation Fee Function	23
3.3 Sample Description	25
3.3.1 Sample Selection	25
3.3.2 Descriptive Statistics	27
IV RESULTS	29
4.1 Analysis of Tax Return Preparation Fees	29
4.1.1 Determinants of Tax Return Preparation Fees	29
4.1.2 Estimates of Aggregate Tax Compliance Cost Expended on Preparation Fees	32
4.1.3 Summary and Discussion	35
4.2 Analysis of Total Tax Burden and Tax Burden Differential between the Preparation Modes	36
4.2.1 Regression Analysis of Total Tax Burden	37

4.2.2	Comparisons of Tax Burden Differential	41
4.2.3	Summary and Discussion	43
4.3	Analysis of the Usage of Tax Preparers	45
4.3.1	Structural Form Estimates of the Preparation Mode Equation . . .	45
4.3.2	Impact of Preparation Fee Changes on Preparer Usage	47
4.3.3	An Alternative Preparation Mode Equation Estimate	48
4.3.4	Summary and Discussion	50
V	CONCLUSIONS, LIMITATIONS AND EXTENSIONS	52
	REFERENCES	57
	APPENDIX A - DESCRIPTIVE SAMPLE STATISTICS (EXCLUDING RETURNS WITH TPI IN THE UPPER ONE PERCENT RANGE)	72
	APPENDIX B - REDUCED FORM ESTIMATES OF THE PREPARATION MODE EQUATION, 1986	75
	APPENDIX C - REDUCED FORM ESTIMATES OF THE PREPARATION MODE EQUATION, 1987	77

LIST OF TABLES

TABLE	Page
1 NUMBER AND PERCENTAGE OF INDIVIDUAL FEDERAL INCOME TAX RETURNS WITH PREPARER SIGNATURE IN 1981 - 1990	60
2 SAMPLE SELECTION	61
3 DESCRIPTIVE SAMPLE STATISTICS	62
4 OLS ESTIMATES OF PREPARATION FEE EQUATIONS	64
5 PREPARATION FEE ESTIMATES FOR THE PAID-PREPARED RETURNS .	65
6 PROPORTION OF FEES TO TOTAL TAX BURDEN OF PAID-PREPARED RETURNS	66
7 ESTIMATES OF TAX BURDEN EQUATIONS, 1986	67
8 ESTIMATES OF TAX BURDEN EQUATIONS, 1987	68
9 TAX BURDEN DIFFERENTIAL BETWEEN PREPARATION MODES	69
10 STRUCTURAL FORM ESTIMATES OF THE PREPARATION MODE EQUATION	70
11 ESTIMATES OF AN ALTERNATIVE PREPARATION MODE EQUATION . .	71

I INTRODUCTION

1.1 Research Questions and Motivation

Tax return preparers have been playing a more and more important role in the individual income tax system in the United States. In 1981, forty-two percent of the 94 million individual federal income tax returns were prepared by tax return preparers. The percentage increased three percentage points to 45 percent after the Economic Recovery Tax Act of 1981. After the Tax Reform Act of 1986 (TRA86), the paid-prepared returns rose another three percentage points to account for more than 48 percent of the individual income tax returns in every year from 1987 to 1990. The trend of the preparer usage over the last decade is illustrated in Table 1, which shows more and more U.S. taxpayers rely on tax preparers to file their income tax returns in the 1980s.

Despite the increased and widespread usage of tax preparers, the question of how tax return preparation fees are determined has not been clearly answered. An investigation of the determinants of tax return preparation fees is essential to the understanding of tax compliance costs paid by the public to preparers.¹ According to Adam Smith's (1776) fourth maxim of taxation "Every tax ought to be so contrived as both to take out and to keep out of the pockets of the people as little as possible, over and above what it brings into the public treasury of the state," an ideal tax system should have minimal compliance costs, other things equal (Slemrod and Sorum 1984, 461). It is also argued that high compliance costs may act as a barrier to tax compliance because taxpayers will simply fail to file (or fail to report a particular item) because of the compliance costs involved (Roth, Scholz and Witte 1989, 118). Therefore, the understanding of tax return preparation fees, a major component of compliance costs, is

¹ The broad concept of tax compliance costs refers to all those costs incurred by taxpayers (the time spent and pecuniary costs) or by third parties (e.g., expense of employers' operating the withholding system) in complying with the requirements of the tax system, over and above the tax payments themselves (Slemrod and Sorum 1987, 461).

critical to tax policy makers in terms of evaluating the income tax system and taxpayer compliance. However, the literature on taxpayers' expenditure for professional tax assistance contains only survey studies. Survey studies have inherent limitations in that respondents' answers may not reflect their true behavior, and that respondents may not be representative of the population.

The first objective of this study is to examine the determinants of return preparation fees using a random sample of 1986-1987 tax returns from the 1979-1988 Statistics of Income Panel of Individual Returns (hereafter, the Panel). Unlike previous years, returns from 1987 and 1988 disclose a return's tax preparation fee, if itemized. According to the Internal Revenue Code §212(3) and Treasury Department Regulations §§1.212-1(l), tax preparation fees are deductible in the year paid.² That is, one can deduct fees paid in 1987 for preparing the 1986 return on the 1987 return.³ For a 1986 paid-prepared return, this paper traces forward to its 1987 return to find its preparation fee paid for the 1986 return. This procedure also is used for 1987 paid-prepared returns. After this procedure, the determinants of preparation fees can be investigated using tax return data for the 1986 and 1987 tax years, a

² IRC §212(3): In the case of an individual, there shall be allowed as a deduction all the ordinary and necessary expenses paid or incurred during the taxable year in connection with the determination, collection, or refund of any tax.

Reg. §§1.212-1(l): Expenses paid or incurred by an individual in connection with the determination, collection, or refund of any tax, whether the taxing authority be Federal, State, or municipal, and whether the tax be income, estate, gift, property, or any other tax, are deductible. Thus, expenses paid or incurred by a taxpayer for tax counsel or expenses paid or incurred in connection with the preparation of his tax returns or in connection with any proceedings involved in determining the extent of tax liability or in contesting his tax liability are deductible.

IRC §67 imposes limitations on the deductibility of preparation fees (see Section 3.1.2).

³ The Internal Revenue Service Publication 529 "Miscellaneous Deductions" (Nov. 1987 and Nov. 1988) has the same prescriptions.

period spanning TRA86. Based on the results, the aggregate preparation fee paid by the entire country can be estimated, and whether tax system changes affect the determinants of the tax preparer fees can be examined.

The second and most important objective of this study is to examine the tax savings from using tax return preparers. Long and Caudill (1987) is the only study that has examined whether professional tax assistance reduces taxpayers' tax liabilities. However, tax return preparation fees were not included in the calculation of a paid-prepared return's total tax burden because preparation fees were not available in their sample, the 1983 Individual Tax Model File. Consequently, they concluded that their finding that tax preparers contribute to tax savings was "tentative" and urged that more accurate information on the expenditures for professional tax assistance be obtained (Long and Caudill 1987, 43).⁴

Furthermore, the finding that tax preparers reduce taxpayers' tax liabilities in Long and Caudill (1987) may not be generalizable to the post TRA86 years because they used the 1983 Individual Tax Model File. There is some evidence that the benefits of using preparers have been changed. During 1988 to 1990 *Money* magazine conducted three annual tax-return tests of professional tax preparers who agreed to participate. Each year 50 preparers were asked to complete a Federal 1040 return for a hypothetical family of four with annual income of \$132,000.⁵ For the third annual test the correct amount due should have been \$12,038. However, answers for the amount of tax due ranged from \$9,806 to \$21,216 with an average of \$13,915, and no two preparers came up with the same tax due (Topolnicki 1990). The

⁴ They also acknowledged that time saved by using preparers was not included in their study. Because time saved is not a "monetary" benefit, this study excludes it from the definition of tax savings of using preparers. Instead, it is used as a separate explanatory variable of preparer usage.

⁵ The sample is not a random one because none of the "big six" accounting firms were willing to participate.

TRA86, which made vast and fundamental changes in the Internal Revenue Code, may have changed the tax benefits of using preparers.

This study uses the endogenous switching regression (ESR) model to provide new evidence on the benefit of using preparers a year before and after TRA86. The model was developed by Heckman (1976, 1979) and Lee (1976, 1978) to correct for selection bias caused by non-random sampling.⁶ In this study self-selection bias may arise because taxpayers are not randomly assigned to the self-preparation regime or the paid-preparation regime. They made their own choice based on their own characteristics and the relative advantages of each preparation mode. In the presence of self-selection, the traditional estimation procedures that ignore the non-randomness of the samples yield biased estimates of the tax burden differential (TBDIF) between paid-preparation and self-preparation (see more discussion in Section 3.1).

The ESR model includes a probit equation for the preparation mode choice, which is used to obtain selection-adjustment variables. Therefore, the third and auxiliary objective of this study is to explore the determinants of preparer usage (the determinants of preparation mode choice). Previous studies have not used tax benefits of using preparers as an explanatory variable of taxpayers' preparation mode choice. Because rational taxpayers choose either to prepare their own returns or to engage preparers based on the relative benefits or detriments of using preparers, the TBDIF of the two preparation modes should be used as an independent variable in the equation of preparation mode choice. A model of preparation mode choice without controlling for the benefits of using preparers may suffer

⁶ There are two kinds of selection biases. When the non-randomness arises from individual choices, the bias is referred to as self-selection bias. When it arises from the way the investigator designs the sample, it is referred to as sample-selection bias. Despite the differences in the terminology, the estimation methods are essentially the same (Maddala 1991, 797).

from omitted variable bias. This study provides new evidence on the determinants of preparation mode choice by introducing TBDIF into the ESR model.

1.2 Organization of the Study

The remainder of this study is organized as follows. Section II presents a review of the relevant literature. Section III discusses the research methods including the development of the econometric model, the selection of the model's explanatory variables, a description of the sample selection procedures, and descriptive sample statistics. The empirical results are reported in Section IV. Section V provides concluding remarks including a discussion of the limitations and potential extensions of this study.

II LITERATURE REVIEW

Prior research related to the research questions of this study will be briefly reviewed in the following sections. Section 2.1 contains a review of the studies on expenditures for professional tax assistance. Section 2.2 reviews literature on the tax savings of using preparers, and Section 2.3 reviews studies on the determinants of the preparation mode choice.

2.1 Expenditures for Professional Tax Assistance

Slemrod and Sorum (1984) conducted a survey of 2000 Minnesota taxpayers to investigate the compliance cost of filing 1982 federal and state individual income tax returns. They defined total compliance costs as the sum of the monetary value of the taxpayers' time devoted to tax matters and their expenditures for professional assistance and materials such as books and computer programs. They found that: (1) expenditures for professional tax assistance were higher among those age 65 or over, self-employed, unmarried and taxpayers with income greater than \$50,000; (2) education and gender were not significantly associated with expenditures for professional assistance; and (3) the total compliance cost in 1982 was between \$17 and \$27 billion, or from five to seven percent of the revenue raised by the federal and state income tax systems combined. About two billion hours of taxpayer time were spent on filing tax returns, and about \$3 billion was spent on professional tax assistance.

Blumenthal and Slemrod (1992) repeated for tax year 1989 the survey methodology employed by Slemrod and Sorum for tax year 1982. In addition to analyzing the 1989 survey data, they also merged it with the 1982 data collected by Slemrod and Sorum (1984). Their analysis suggests that: (1) monetary expenditures for professional assistance were positively correlated with income; (2) self-employed taxpayers spent about 120 more dollars on professional advice than employed taxpayers, but self-employed taxpayers on average incurred

lower preparation expenditures in 1989 than in 1982; (3) receiving rental income was associated with higher preparation expenditures in 1989 than in 1982, probably because of the complex post TRA86 rules concerning passive losses; (4) itemization raised preparation expenditures for 1982 and 1989 taxpayers, to roughly the same extent; (5) preparation expenditures were not significantly affected by differences in gender, age, marital status or education; and (6) taxpayers using professional tax assistance incurred an average cost of \$132 in 1989, compared to \$98 (1989 dollars) in 1982.

Unlike the studies that surveyed taxpayers, Swingen et al. (1992) conducted a nationwide survey of tax accountants, lawyers, enrolled agents, and unenrolled practitioners in 1989 to examine factors affecting tax return preparation fees. They found that type of tax professionals, geographic location, computer usage, staff size, number of clients and return mix significantly affected preparation fees. They also asked tax preparers what facts or methods were used to determine preparation fees. Sixty-seven percent indicated they used a fixed rate per hour, and 46 percent had a fixed rate per form.

As noted earlier, survey studies have inherent limitations in that respondents' answers may not reflect their true behavior, and that respondents may not be representative of the population. This paper will be one of the first studies on the determinants of preparation fees using archival tax return data.

2.2 Tax Savings of Using Preparers

The literature contains only one study that examined the benefits of using tax preparers. Long and Caudill (1987) employed the 1983 Individual Tax Model File, a stratified random sample of federal individual income tax returns, to investigate the determinants of tax preparer usage (reviewed in next section) and the tax benefits of using

preparers. Similar to this study, they also used an endogenous switching regression (ESR) model (described in Section 3.1 of this study). Their analysis indicates that the tax savings on paid-preparer returns ranged from \$101 to \$3805 and increased with income. Given the information of preparation fees was unavailable in their data, their definition of tax savings did not take account of preparation fees. Consequently, they acknowledged that their finding was tentative and urged that information on the expenditures for paid-preparation be included in future studies. A weakness in their ESR model is that they did not introduce tax burden differential (TBDIF) of the two preparation modes as an explanatory variable of the preparation mode choice. Because rational taxpayers choose either to prepare their own returns or to engage preparers based on the tax benefits of using preparers, the TBDIF of the two preparation modes should be used as an independent variable in the equation of preparation mode choice (Maddala 1991, 803-804). Therefore, the model in Long and Caudill (1987) might produce inaccurate estimates because of omitted variable bias.

This study will provide new evidence spanning TRA86 on the tax savings of using preparers by introducing preparation fees into the ESR model and controlling for the simultaneity between TBDIF and the preparation mode choice.

2.3 Determinants of Preparation Mode

As noted above, Long and Caudill (1987) used the 1983 Individual Tax Model File to examine the tax benefits of using preparers (reviewed in the preceding section) and to investigate the determinants of preparer usage and. Regarding the determinants of preparer usage, their logit analysis shows: (1) upper-income taxpayers, the elderly, and self-employed workers were more likely to use paid preparers; and (2) professional tax assistance was positively related to marginal tax rate and return complexity, which was proxied by the

number of specific forms and schedules included with the return. Note that they reported the determinants of preparer usage only based on the results of logit analysis, though they also conducted a probit regression in the first stage of their ESR estimation. As in this paper, the estimated probit equation was used to obtain the inverse Mills ratios that operated as the selection-adjustment variables in the second stage estimation (see Section 3.1). This paper uses probit rather than logit or both to explore factors associated with preparer usage.

Collins et al. (1990) used a mail survey of 700 households in Oklahoma and Pennsylvania to investigate factors associated with the demand for tax preparers' services. Seventy percent of the taxpayers in their sample responded that they approached the taxpaying process with the primary objective of filing the most correct return. Twenty-five percent indicated minimizing their tax liability was their primary objective. Approximately one-half of each taxpayer objective group used professional assistance. For taxpayers whose objective was to file the most correct return, preparer usage was significantly related to strong value orthodoxy, low tax knowledge, and high return complexity. For taxpayers with a tax minimization objective, preparer usage was significantly associated with high income, low social responsibility, low tax knowledge, and increased age. In addition, their results do not support the proposition that taxpayers use preparers because of fear of the IRS audit or the desire to minimize the total filing effort.

Using a random sub-sample of audited tax returns from the 1982 Taxpayer Compliance Measurement Program (TCMP), Erard (1990) found that the probability of engaging tax preparers increased with tax return complexity, but was not related to income or federal marginal tax rate. High profile occupations (chief executive, elected official), married and elderly filers also were more likely to seek professional assistance.

Recent studies have employed new econometric techniques such as bivariate probit, nested logit, and Chamberlain's conditional logit to analyze the determinants of the preparer usage. Slemrod (1989) used the same Minnesota taxpayer survey data as in Slemrod and Sorum (1984) to examine the probable compliance cost savings from tax simplification. His econometric model included a bivariate probit to estimate the determinants of the choice of whether to itemize deductions and whether to use a tax preparer. He found that itemizers' preparer usage is positively related to marginal tax rate and negatively related to years of education. Itemized returns that did not contain any dividends, interest, self-employed business income, capital gains, etc. (simple itemized returns) were less likely to be paid-prepared. However, Slemrod found no significant determinants of the non-itemizers' preparer usage.

Dubin et al. (1992) conceptualized taxpayers' choice of preparation modes as a two stage sequential process. In the first stage taxpayers choose one of the following categories: (1) non-paid assistance (IRS assisted or other non-paid), (2) self preparation, (3) paid preparer or (4) practitioner. In the second stage, taxpayers make a further decision regarding which mode of assistance to use (e.g., select CPA or lawyer from Category 4). Working with aggregate 1979 TCMP data and employing a nested logit model, they found that: (1) higher audit rates increased the demand for practitioner services but did not affect the demand for other third party assistance; (2) contrary to the view that more deductions imply a more complex tax return, which should increase the demand for third party assistance, more deductions reduced the demand for third party assistance; and (3) states with a higher percentage of the adult population having received at least a high school education showed lower demand for IRS or other non-paid assistance compared to self preparation.

Both Deere and Wolfe (1991) and Christian et al. (1992) used a balanced Panel of returns from 1982-84 and employed Chamberlain's conditional logit to investigate the determinants of preparer usage. By accounting for the heterogeneity among individuals, Chamberlain's conditional logit model may capture the unobserved yet constant over time effects of individuals' education, value orthodoxy, perception of audit probability, etc. Hence these two studies may avoid bias from omitting these variable. Deere and Wolfe (1991) and Christian et al. (1992) both found no significant correlation between preparer usage and income or marginal tax rate, which is in contrast with Long and Caudill (1987). Christian et al. (1992) indicated that when they specifically included complexity variables in the model, the significance for both income and marginal tax rates went away. They conjectured that both income and marginal tax rate might be surrogates for complexity. They also found that tax return type (i.e., Form 1040, 1040A, 1040EZ) had a significant effect on preparer usage.

Similar to this study, Christian et al. (1993) used the ESR model and the 1983 Panel to examine preparers' effect on taxpayers' income tax prepayment position. Their ESR model included a preparation mode equation in which the difference in tax prepayment between the two preparation modes (PPDIF, as they designated) was introduced as a regressor. They found that tax return complexity and taxpayer demographics were important determinants of preparer usage but the estimated coefficient of PPDIF was insignificant. The result suggests that PPDIF does not influence taxpayers' decisions on whether to engage a preparer.

In this paper, tax burden differential (TBDIF) of the two preparation modes rather than PPDIF is used as a regressor of the preparation mode choice. TBDIF is considered to be a better proxy of the monetary benefits of using tax preparers than PPDIF. In addition, the non-monetary benefits of return completion time saved by using preparers is also included as a regressor in this study. No previous study has used both monetary tax burden differential

and time saved of using preparers to examine the choice of preparation mode. This paper is expected to provide new evidence on the determinants of preparer usage.

III RESEARCH METHODS

This study follows Long and Caudill (1987) in defining preparation mode -- by the taxpayer (self-prepared) or by a tax preparer (paid-prepared) -- on the basis of the presence of a preparer's signature on the return.⁷ The preparation mode is coded DP=1, called paid-prepared, and designated as DP₁ if a preparer's signature appears on the return; otherwise, a return is coded DP=0, symbolized as DP₀, and called self-prepared. For a paid-prepared return, total tax burden is its tax liability plus the estimated fee adjusted to account for the fee's deductibility or non-deductibility on the subsequent year's tax return.⁸ For a self-prepared return, total tax burden is just its tax liability.

Section 3.1 describes the endogenous switching regression (ESR) model, which will be used to explore tax savings of using preparers. The model's explanatory variables are then described in Section 3.2. Section 3.3 explains the sample selection procedures and provides descriptive sample statistics.

⁷ It is possible that this definition is subject to measurement error because some taxpayers hire preparers only for some years but use their advice in subsequent years, some taxpayers seek advice on particular filing issues but prepare their own returns, and some paid preparers do not sign the returns even though failure to do so is illegal (Roth, Scholz, and Witte 1989, 173).

⁸ To take 1986 as an example, if a 1986 paid-prepared return reported no miscellaneous itemized deductions in its subsequent 1987 return, then its total tax burden for 1986 is its tax liability plus the whole preparation fee. If it reported miscellaneous itemized deductions in 1987 and estimated fee is less than the miscellaneous itemized deductions, then the total tax burden is tax liability plus the after-tax preparation fee, which is fee minus the tax savings of fee deduction. If it reported miscellaneous itemized deductions and estimated fee is greater than the miscellaneous itemized deductions, then the total tax burden is tax liability plus preparation fee and minus the tax savings of miscellaneous itemized deductions.

3.1 Econometric Model

3.1.1 Endogenous Switching Regression Model

To obtain consistent estimates of the tax burden differential (TBDIF) and the determinants of preparation mode choice, a two-stage endogenous switching regression (ESR) model developed Heckman (1976, 1979) and Lee (1976, 1978) is used.⁹ The estimation of the ESR model proceeds as follows. Tax burden (TB) for a paid-prepared return and for a self-prepared return are formulated as:

$$TB_{1i} = \beta_1' X_{1i} + \epsilon_{1i}, \quad (1)$$

$$TB_{0i} = \beta_0' X_{0i} + \epsilon_{0i}, \quad (2)$$

where TB = tax liability + preparation fee, if any. In the right hand side of Equations (1) and (2), X is a vector of explanatory variables determining income tax burden;¹⁰ β_1 and β_0 are the coefficient vectors; and it is assumed that the error terms

$$\epsilon_{1i} \sim N(0, \sigma_1^2), \text{ and}$$

$$\epsilon_{0i} \sim N(0, \sigma_0^2).$$

Note that the domain of both Equations (1) and (2) is the whole taxpayer population, not only those using preparers or those who prepare their own returns. In other words, the estimated β_1 and β_0 are intended to be generalizable to any taxpayer. However, neither the paid-prepared sample nor the self-prepared one is a random sample of the general population. The non-randomness arises because taxpayers make their own choice of preparation mode

⁹ The ESR model is described in Maddala (1983, Sections 8.3 and 9.7). Maddala designates it as a two-stage estimation method. However, it is a three-stage estimation in Lee (1978) and in this paper, because the probit equation is estimated twice, the reduced form estimation in the first stage and the structural form estimation in the third stage. See more discussions below.

¹⁰ It is assumed that vectors X_1 and X_0 contain the same variables in this study, though they are not necessarily the same in a ESR model.

based on their own characteristics and the relative benefits of each preparation mode. Self-selection bias will result if Equations (1) and (2) are directly estimated by the ordinary least squares (OLS) regression without adjusting for the non-randomness.

To obtain consistent estimates for Equations (1) and (2), an indicator function for preparation mode must be introduced into the model. The preparation mode function is modelled through the following equation, where the propensity to engage preparers (DP_i^*) is the dependent variable:

$$DP_i^* = \gamma_1^*(TB_{0i} - TB_{1i}) + \gamma_2 C_i + u_i, \quad (3)$$

where subscript i represents individual units ($i = 1, \dots, n$), $(TB_0 - TB_1)$ is the tax burden differential (TBDIF) between the two preparation modes, C represents a vector of other relevant explanatory variables, γ s are the coefficients, and u_i is the error term. If $DP_i^* > 0$, then a return is in the paid-prepared regime, and tax burden conditioned on the paid-prepared mode ($TB_{1i} | DP_i=1$) is observed; if $DP_i^* \leq 0$, then a return is in the self-prepared regime, and tax burden conditioned on the self-prepared mode ($TB_{0i} | DP_i=0$) is observed.

However, DP_i^* is not observable. Instead, one observes the preparation mode

$$DP_i = \begin{cases} 1, & \text{if } DP_i^* > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Plugging (1) and (2) into (3) obtains the reduced form:

$$DP_i^* = \delta' W_i + v_i, \quad (4)$$

where W includes all exogenous variables in (1), (2) and (3). The error term v_i in Equation (4) can have variance Ω^2 , but the equation can be scaled by Ω to produce a unit variance, allowing identification. The ESR model assumes that the errors $[\epsilon_{1i}, \epsilon_{0i}, v_i]$ have a trivariate normal distribution with mean vector zero and covariance matrix:

$$\begin{bmatrix} \sigma_1^2 & \sigma_{10} & \sigma_{1v} \\ & \sigma_0^2 & \sigma_{0v} \\ & & 1 \end{bmatrix}$$

In the first stage, the reduced form Equation (4) is estimated using the observed DP_i as the dependent variable with probit to estimate δ . Then the inverse Mills ratios λ_{1i} and λ_{0i} are obtained from the follow formulae:

$$\lambda_{1i} = \frac{f(\delta'W_i)}{F(\delta'W_i)},$$

$$\lambda_{0i} = \frac{-f(\delta'W_i)}{1 - F(\delta'W_i)}.$$

The $f(\cdot)$ in the λ s is the standard normal probability density function, and $F(\cdot)$ is the cumulative distribution of a standard normal function.

With the inverse Mills ratios operating as the selection-adjustment variables, Equations (1) and (2) can be estimated using the non-random selected paid-prepared returns and self-prepared returns. This is because it can be shown that:¹¹

$$E(TB_{1i} | DP_i=1) = \beta_1'X_{1i} + \sigma_{1v}*\lambda_{1i}, \text{ and}$$

¹¹ It can be demonstrated that $E(TB_{1i} | DP_i=1)$
 $=E[(\beta_1'X_{1i} + \epsilon_{1i}) | v_i > -\delta'W_i]$
 $=\beta_1'X_{1i} + E(\epsilon_{1i} | v_i > -\delta'W_i)$
 $=\beta_1'X_{1i} + E(v_i*\sigma_{1v}/\sigma_v^2 | v_i > -\delta'W_i)$
 $=\beta_1'X_{1i} + \sigma_{1v}E(v_i | v_i > -\delta'W_i)$
 $=\beta_1'X_{1i} + \sigma_{1v}*f(-\delta'W_i)/[1-F(-\delta'W_i)]$
 $=\beta_1'X_{1i} + \sigma_{1v}*f(\delta'W_i)/F(\delta'W_i)$
 $=\beta_1'X_{1i} + \theta_1*\lambda_{1i}.$

If one sets $DF_i^* = \delta'W_i - v_i$ as in Maddala (1983), then the selection is from $v_i > \delta'W_i$. The result is $\beta_1'X_{1i} - \sigma_{1v}*f(\delta'W_i)/F(\delta'W_i)$, and the two σ_{1v} s here will have different signs. Equation (6) can be proved using the same method.

$$E(TB_{0i} | DP_i=0) = \beta_0 \hat{X}_{0i} + \sigma_{0v} \lambda_{0i}$$

Therefore, in the second stage β_1 and β_0 of Equations (1) and (2) are estimated by the following two equations, which include the selection-adjustment variables λ_{1i} and λ_{0i} :

$$TB_{1i} | DP_1 = \beta_1 \hat{X}_{1i} + \theta_1 \lambda_{1i} + q_{1i}, \quad (5)$$

$$TB_{0i} | DP_0 = \beta_0 \hat{X}_{0i} + \theta_0 \lambda_{0i} + q_{0i}, \quad (6)$$

where the coefficients $\theta_1 = \sigma_{1v}$ and $\theta_0 = \sigma_{0v}$ (see Footnote 11), and q_{1i} and q_{0i} are error terms with mean vector zero.

Unless σ_{1v} and σ_{0v} are zero, i.e., v_i and ϵ_{1i} , and v_i and ϵ_{0i} are uncorrelated, using OLS to regress the observed $TB_{1i} | DP_1$ on X_{1i} and $TB_{0i} | DP_0$ on X_{0i} without including selection-adjustment variables will result in inconsistent estimates of β_1 and β_0 . This is because λ_{1i} and λ_{0i} are functions of W_i , which includes X_{1i} and X_{0i} , and hence λ_{1i} and λ_{0i} are correlated with the regressors X_{1i} and X_{0i} . In other words, the selection bias arises because the selection-adjustment variables, which are correlated with the explanatory variables, are omitted.

To estimate Equations (5) and (6) in the second stage, Lee (1978) uses OLS estimation. The β estimates obtained are unbiased and consistent, but the error terms q_{1i} and q_{0i} have been shown to be heteroscedastic (Lee 1978, 422). In the presence of heteroscedasticity the OLS estimators of the standard errors are biased and the standard tests of significance are invalid. To provide unbiased standard errors, the corrected variance-covariance matrix of β coefficients is obtained based on Greene (1981).¹²

Note that the goal of estimating Equations (5) and (6) is to obtain the β estimates of Equations (1) and (2), which are generalizable to any taxpayer. To predict a taxpayer's tax

¹² Though the corrected covariance matrix is unbiased, it is less efficient than that obtained from maximum likelihood estimation (MLE). The likelihood function of the ESR model is described in Lee and Trost (1978). The MLE in LIMDEP (Greene 1991, 653-658) did not converge, so this study relies on two-stage estimation.

burden differentials between the two preparation modes, one should compare the estimated $\hat{\beta}_1'X_{1i}$ and $\hat{\beta}_0'X_{0i}$, but the terms $\theta_1\lambda_{1i}$ and $\theta_0\lambda_{0i}$ must be excluded from the calculation of TB_{1i} and TB_{0i} .

Recall that the reduced form Equation (4) rather than Equation (3) is estimated in the first stage. After the estimation of Equations (5) and (6) in the second stage, one can substitute $(\hat{\beta}_0'X_{0i} - \hat{\beta}_1'X_{1i})$ for $(TB_{0i} - TB_{1i})$ in Equation (3) and estimate the structural equation of preparation mode choice using probit in the third stage. The structural form probit examines the determinants of preparer engagement, the third research objective of this study.

3.1.2 Estimation of Preparation Fees

As pointed out in Section I, the first objective of this study is to examine the determinants of return preparation fees. The estimation of preparation fees is needed to construct Equation (1) because the total tax burden of using preparers is defined to include the tax liability and the preparation fees. IRC §67 after TRA86 prescribes that certain miscellaneous itemized deductions, including tax return preparation fee and unreimbursed employee business expenses, are only deductible by individuals to the extent that they cumulatively exceed two percent of adjusted gross income (AGI). The deductible amount is the aggregate of such deductions reduced by two percent of AGI. Thus, the preparation fee paid for a 1986 (1987) return is not always reported on the 1987 (1988) return. In the sample there are 858 (1683) observations with preparation fees reported on the 1987 (1988) returns and these observations used preparers for their 1986 (1987) returns (see Section 3.3 for more information about the fee observations). For those paid-prepared 1986 (1987) returns with no preparation fees reported, this study regards them as having a missing value and estimates the

missing fees using a method from Slemrod and Sorum (1984) and Blumenthal and Slemrod (1992).¹³ Essentially, it estimates a regression equation using the observed data, and then the equation is used to impute a value for the unobserved preparation fee. Note that the fee estimated with this approach is probably biased because the regression uses observations only from paid-prepared returns that itemized preparation fees. Section 4.1.4 has more discussion about the potential bias.

3.2 Model's Explanatory Variables

3.2.1 Total Tax Burden Function

Equations (1) and (2) are the functions of total tax burden for paid-prepared returns and for self-prepared returns, respectively. Given income, return complexity (proxied by schedules and forms used) and demographic characteristics, these two equations would estimate a taxpayer's tax burden under the two preparation modes. The difference between the estimated tax burden of these two modes is the tax savings (or disadvantages) of using tax preparers. Drawing on Long and Caudill (1987), the tax burden of the two alternative preparation modes are both formulated as a function of total positive income (TPI), adjusted to account for the presence of main schedules included on a return (e.g., Schedules A and B), the number of other forms attached, whether taxpayer is self-employed, marital status, age

¹³ This method was referred to as a standard missing values technique by Blumenthal and Slemrod (1992, 201). In the survey study of Slemrod and Sorum (1984) and Blumenthal and Slemrod (1992) forty-six percent of the respondents of each study did not provide a usable response to the marginal wage rate question. Based on those providing a usable answer to the question, both studies estimate an equation for the wage rate as a function of a set of demographic variables. This equation was then used to impute a value for wage rate for those respondents with unusable replies to this question. Then the estimated wage rate was converted into after-tax wage rate. The after-tax wage rate was then multiplied by the number of hours a respondent spent on tax filing. The resultant monetary value of time was summed up with expenditure on fees to tax advisors to obtain the total compliance cost for a "missing values" respondent.

and the number of dependents. For the 1986 sample, Equation (1) and (2) are written as (subscript i is omitted hereafter):

$$TB_1 = \beta_{1,0} + \beta_{1,1}TPI + \beta_{1,2}SA + \beta_{1,3}SB + \beta_{1,4}SD + \beta_{1,5}SE + \beta_{1,6}SW + \beta_{1,7}FORM + \beta_{1,8}SEM + \beta_{1,9}MS + \beta_{1,10}AGE + \beta_{1,11}DEP + \epsilon_1, \quad (1)$$

$$TB_0 = \beta_{0,0} + \beta_{0,1}TPI + \beta_{0,2}SA + \beta_{0,3}SB + \beta_{0,4}SD + \beta_{0,5}SE + \beta_{0,6}SW + \beta_{0,7}FORM + \beta_{0,8}SEM + \beta_{0,9}MS + \beta_{0,10}AGE + \beta_{0,11}DEP + \epsilon_0. \quad (2)$$

For the 1987 sample, Equation (1) and (2) are as follows:

$$TB_1 = \beta_{1,0} + \beta_{1,1}TPI + \beta_{1,2}SA + \beta_{1,3}SB + \beta_{1,4}SD + \beta_{1,5}SE + \beta_{1,6}AMT + \beta_{1,7}PAL + \beta_{1,8}FORM + \beta_{1,9}SEM + \beta_{1,10}MS + \beta_{1,11}AGE + \beta_{1,13}DEP + \epsilon_1, \quad (1)$$

$$TB_0 = \beta_{0,0} + \beta_{0,1}TPI + \beta_{0,2}SA + \beta_{0,3}SB + \beta_{0,4}SD + \beta_{0,5}SE + \beta_{0,6}AMT + \beta_{0,7}PAL + \beta_{0,8}FORM + \beta_{0,9}SEM + \beta_{0,10}MS + \beta_{0,11}AGE + \beta_{0,13}DEP + \epsilon_0. \quad (2)$$

The independent variables used are defined as:

- TPI = the sum of all positive income line items reported on a return.
- SA = 1 if Schedule A (itemized deductions) present, 0 otherwise.
- SB = 1 if Schedule B (interest or dividends received greater than \$400, or all savers interest reported) present, 0 otherwise.
- SD = 1 if Schedule D (capital gains and losses) present, 0 otherwise.
- SE = 1 if Schedule E (supplement income from rents, royalties, partnerships, etc.) present, 0 otherwise.
- SW = 1 if Schedule W (two-wage earner deduction) present, 0 otherwise.
- AMT = 1 if Form 6251 (alternative minimum tax) present, 0 otherwise.
- PAL = 1 if Form 8582 (passive activity loss limitations) present, 0 otherwise.
- FORM = the number of forms contained on the return, except those already included as dummy variables.
- SEM = 1 if Schedule C or F present (taxpayer is self-employed), 0 otherwise.
- MS = 1 if married, 0 otherwise,
- AGE = 1 if either the primary or secondary taxpayer ages 65 or over, 0 otherwise,
- DEP = number of dependents,
- ϵ = the error term assumed to be normally distributed with zero mean and constant variance.

TPI rather than AGI or taxable income (TI) is used in the function because AGI and TI may be endogenous in a model of preparer usage. The recommendations and choices made by preparers are more likely to affect AGI and TI. This problem is considered less severe for TPI (Christian et al. 1992). It is also acknowledged that the presence of supplementary schedules and forms and the preparer usage also may be endogenous. Caution should be used in generalizing the results because of the potential endogeneity.

The new provisions of passive activity loss limitations and alternative minimum tax made by TRA86 are expected to have substantial influence on the tax burden of the two preparation modes. Therefore, PAL and AMT are added to the 1987 functions. Though AMT existed in 1986, it is not used as a regressor of the 1986 equation because less than one percent of the returns attached a Form 6251 in 1986 (See Table 3). To construct the variable FORM, the following forms that are disclosed in the Panel are considered: Forms 1116, 2106, 2119, 2439, 2441, 3468, 3800, 3903, 4136, 4137, 4255, 4562, 4684, 4797, 4835, 4868, 4970, 4972, 5329, 5695, 5884, 6251, 6478, 6765, 8586, 8606, 8615, and Schedule R. Note that Schedules G and W were removed after 1986, and Forms 8586, 8606, and 8615 were not used before 1987.

3.2.2 Preparation Mode Function

To specify the structural equation of the preparation mode choice (Equation 3), it is assumed that taxpayers are wealth-maximizers who compare the benefits with the costs of using a preparer in deciding which preparation mode to use. The benefits of using a preparer are defined to include:

- (1) the difference between the monetary tax burden of self-preparation (TB_0) and the tax burden of paid-preparation (TB_1), and

- (2) the time saved for a taxpayer in completing the return if a preparer were to be used.

The difference between TB_0 and TB_1 for the same taxpayer in the same year is not directly observable because a return cannot be paid-prepared and self-prepared as well in the same tax year. Therefore, the reduced form of Equation (3), i.e., Equation (4), must be estimated in the first stage (See Section 3.1.1).

The time saved for a taxpayer in completing the return is proxied by the return completion time estimated by the IRS. The IRS has estimates of the record-keeping, learning about the law and preparing time of each form and schedule. For all the forms and schedules contained on a return, the summation of its learning about the law and preparing time is used to proxy the return completion time saved if a taxpayer were to use a preparer.

The costs of using a preparer include (1) the time spent in dealing with and commuting to a preparer and (2) the additional time required by the preparer to provide documents for claiming deductions or reporting income. Because the Panel provides no data on these costs of using preparers, the effect of the costs are modelled through the constant intercept and the error term, which is assumed to be normally distributed with zero mean and constant variance. Following previous literature, demographic variables including age, marital status, self-employment status, and number of dependents are also used as explanatory variables.

To summarize, the preparation mode function is:

$$DP^* = \gamma_0 + \gamma_1*(TB_0 - TB_1) + \gamma_2CT + \gamma_3SEM + \gamma_4MS + \gamma_5AGE + \gamma_6DEP + u,$$

where CT is the time (in hours) saved for a taxpayer in completing the return if a preparer were used, u is the error term, and the remaining variables are as defined previously.

3.2.3 Preparation Fee Function

Swingen, et al. (1992) conducted a nationwide survey of tax preparers to examine factors affecting tax return preparation fees. Sixty-seven percent of the preparers indicated that they used a fixed rate per hour, and 46 percent had a fixed rate per form. Inspired by their article, this study explores how the main schedules (e.g., Schedules B, C, D and E) and return preparation time affect preparation fees. The preparation fee function for the 1986 sample is specified as:

$$\text{FEE} = \alpha_0 + \alpha_1\text{SB} + \alpha_2\text{SC} + \alpha_3\text{SD} + \alpha_4\text{SE} + \alpha_5\text{SG} + \alpha_6\text{SW} + \alpha_7\text{PTFORM} + \alpha_8\text{TPI} \\ + \phi,$$

All variables are as defined previously, except

PTFORM = the return preparation time (in hours) obtained by taking the summation of the IRS estimated preparation time of each form and schedule attached to the return, except the main schedules that are already included as dummy variables.

SG = 1 if Schedule G (income averaging) present, 0 otherwise.

ϕ = the error term assumed to be normally distributed with zero mean and constant variance.

To avoid the double counts of the presence of a main schedule and its preparation time into the function, **PTFORM** is defined as the total return preparation time (in hours) of schedules and forms included on a return excluding the preparation time of the schedules that are already included as dummy variables. Note that preparation time is used as a regressor of preparation fees, whereas return completion time (preparation time plus time of learning about the law) is a regressor of preparation mode choice in the preceding section. **TPI** is also used as a regressor because previous survey studies (Slemrod and Sorum 1984, and Blumenthal and Slemrod 1992) indicate fees are positively related to income.

Schedule A (itemized deductions) and type of return used (e.g., Form 1040, 1040A or 1040EZ) cannot be used as independent variables because the model is estimated using data only from those reported their preparation fees on Schedule A attached to Form 1040. The

presence or absence of Schedule F (farm income) is not used as a dummy variable because only 18 out of the 804 observations in the 1986 sample attached a Schedule F. Instead, its effect on preparation fee is only accounted for by including the preparation time of Schedule F in the PTFORM calculation.

The fee function of the 1987 sample is formulated as:

$$FEE = \alpha_0 + \alpha_1 SB + \alpha_2 SC + \alpha_3 SD + \alpha_4 SE + \alpha_5 AMT + \alpha_6 PTFORM + \alpha_7 TPI + \phi.$$

Unlike the 1986 function, Schedules G (income averaging) and W (two-wage earner deduction) are not in the 1987 function because they were repealed by TRA86. On the other hand, AMT is included as a dummy to represent the presence or absence of Form 6251 (alternative minimum tax calculation). TRA86 not only increased the AMT rate for individuals from 20 percent to 21 percent for the 1987 tax year, but it also greatly broadened the AMT base. As a result, many taxpayers who had never been subject to the AMT were subject to the AMT or had to make the AMT calculation. It would be informative to examine how the presence of the AMT affects preparation fee pre- and post-TRA86. However, AMT is not included in the 1986 sample because only 19 out of the 804 observations (2.4 percent) attached Form 6251 in the 1986 sample.¹⁴ The dummy variable PAL (passive activity loss limitations) is not included because the limitations did not exist in 1986, and only 2.6 percent of paid-prepared returns in 1987 attached Form 8582 (See Table 3). Also note that SG is not used in the 1986 tax burden function, but it is used in the 1986 fee function. This is because only 3.3 percent of the whole 1986 sample (n=9,019) used income averaging, but there were 9.6 percent of the 804 observations in the 1986 fee regression attached SG.

¹⁴ Among the 1,633 observations that itemized preparation fees in the 1987 sample, thirteen percent of them (212 observations) attached a Form 6251.

3.3 Sample Description

3.3.1 Sample Selection

The model parameters are estimated using 1986-87 individual income tax return data from the 1979-1988 Statistics of Income Panel of Individual Returns (the Panel), which is a part of the Ernst & Young / University of Michigan Tax Research Database.¹⁵ The Panel is a subset of the IRS Individual Tax Model File and represents a simple random sample of individual income tax returns filed each year.¹⁶

The Panel contains 10,120 individual tax returns and 193 variables in the 1986 panel, and 21,191 returns with 203 variables in the 1987 panel. From the 1986 and 1987 panels, a small number of returns were deleted for the following reasons: (1) returns not present in the subsequent year Panel;¹⁷ (2) erroneous returns;¹⁸ (3) prior year returns;¹⁹ (4) returns other

¹⁵ Other tax research databases available include: Corporate Source Book Files, the IRS Tax Practitioners Survey, and the IRS General Taxpayer Opinion Survey.

¹⁶ Imbedded within each individual year Tax Model File is a subset of returns which were chosen to be part of a panel of taxpayers whose returns would be drawn year after year. The IRS provided these returns to the Office of Tax Policy Research at the University of Michigan along with the information needed to link the subset of panel returns across years. The sampling procedures used each year are described in two annual series published by the IRS SOI Division (Individual Statistics Branch): *Individual Income Tax Returns* and *General Description Booklet for the Individual Tax Model File*.

¹⁷ For a paid-prepared return, total tax burden is its tax liability plus the estimated fee adjusted to account for the fee's deductibility or non-deductibility on the subsequent year's tax return. Therefore, only those returns with a taxpayer identification number that appears in both 1986 (1987) and 1987 (1988) Panels are retained in the 1986 (1987) sample.

¹⁸ The erroneous returns are identified by examining the SSNCODE code, which represents the sampling status. Only one out of the five SSNCODE codes were intended to be included in the 1986 panel, and two of the five codes in the 1987 panel. Therefore, returns with SSNCODE > 1 in 1986 and SSNCODE > 2 in 1987 are deleted. See Christian (1991) for a discussion of the problems that can arise if these returns are included.

¹⁹ Returns were deleted if the tax year did not match the panel file year (e.g., 1985 return that is in the 1986 Panel).

than paid-prepared or self-prepared.²⁰ The final sample consists of 9,019 returns for the 1986 analysis and 18,811 for the 1987 analysis. A reconciliation of the sample size for each year is provided in Table 2. For the fee regression, only those paid-prepared returns that itemized fees on the return of the subsequent year are retained in the sample. Returns with fees less than \$5 times the preparation time (in hours) are also excluded from the fee regression. This avoids clerical errors in data processing and exclude potential outliers with unusually low fees. The final sample size for the fee regression is 804 for the 1986 analysis and 1,633 for the 1987 analysis.

This paper uses 1986 and 1987 data for the following reasons. First, the preparation mode data are not available in the 1985 panel. Second, the preparation fee data became available in the Panel starting in 1987. Because preparation fees are deductible in the year paid, the preparation fees shown in the 1987 (1988) Panel are assumed to be the fees for preparing the 1986 (1987) return. For a 1986 paid-prepared return, this paper traces forward to its 1987 return to find the preparation fee paid for the 1986 return. This procedure also is used for 1987 paid-prepared returns. Subject to the two percent AGI rule, many paid-prepared returns did not report their preparation fees. In this case, the preparation fees are estimated (see Section 3.1.2). This enables this paper to analyze the research questions using data a year before and after the TRA86.

Note that fee may be subject to measurement error. The IRC §67 after TRA86 prescribes that certain miscellaneous itemized deductions, including tax return preparation fees and unreimbursed employee business expenses, are only deductible by individuals to the extent that they cumulatively exceed two percent of AGI. In 1992, after much criticism from

²⁰ A small number of returns in the Panel are prepared by other than the taxpayer or a paid-preparer (e.g., returns prepared by the IRS and under the Voluntary Income Tax Assistance program).

tax preparers, the IRS officially conceded "Tax return preparation expenses for part of a sole proprietor's 1040 return dealing with his trade or business (Schedule C); rents and royalties (Part I or Schedule E); farm income and expenses (Schedule F); ... , are above-the-line deductions."²¹ However, starting in 1987 some tax preparers might have allocated preparation fees to clients' Schedule C, Schedule E or Schedule F (i.e., above-the-line deductions) to avoid the two percent floor. If this is the case, measurement error in fees is inevitable because the Panel only discloses the tax return preparation fee itemized on Schedule A. In addition, preparation fees might also have been pre-billed in 1986 in order to avoid the two percent floor. Also note that in the Panel, not all returns reporting fees had miscellaneous itemized deductions greater than two percent of AGI. This is because tax return preparation fee is a line item on Schedule A. Even though the two percent floor is not met, the fee still can be reported on the return.

3.3.2 Descriptive Statistics

Table 3 reports descriptive statistics for all variables used in this study. For the year 1986, the mean tax liability of the self-prepared returns is \$2,518, less than the \$5,402 for the paid-prepared returns. The average TPI is \$22,053 for the self-prepared returns, and \$36,767 for the paid-prepared returns. The self-prepared returns also have lower average number of forms attached to tax returns, and less self-prepared returns are subject to the alternative minimum tax. A similar pattern is found in the 1987 data. For example, the average tax liability is \$2,678 for the self-prepared returns, and \$5,332 for the paid-prepared returns.

The descriptive statistics of the returns used in the fee regression are also reported in Table 3. As expected, their mean tax liability, TPI and FORM are all higher than the self-

²¹ Revenue Ruling 92-29, 1992-16 Internal Revenue Bulletin.

prepared group and the paid-prepared group. The preparation fees reported in these returns range from \$20 to \$6,227 with a mean of \$200 in 1986; and range from \$20 to \$20,730 with a mean of \$284 in 1987.

IV RESULTS

4.1 Analysis of Tax Return Preparation Fees

4.1.1 Determinants of Tax Return Preparation Fees

The first objective of this study is to explore the determinants of income tax return preparation fees. Panels A and B of Table 4 show the results of the ordinary least squares (OLS) regression analysis of preparation fees in the tax years of 1986 and 1987. In this section, the estimated coefficients of the regressors of the 1986 sample and the 1987 sample are discussed respectively, followed by comparisons of the fee determinants between the two years. Based on the OLS regression, aggregate tax compliance cost expended on return preparation fees is estimated in Section 4.1.2. A potential limitation of the fee estimation is discussed in Section 4.1.3.

For the 1986 returns, preparation fees are positively correlated with TPI (total positive income). TPI (\$1000s) has a coefficient of 1.77, which is significant ($p=0.0003$) and implies that, *ceteris paribus*, an increase of \$1000 in TPI increases the preparation fees by \$1.77.²² The presence of an additional supplementary SB, (Schedule B for interest and dividends), SC (business income), SD (capital gains and losses), SE (income from rents, royalties or partnerships) or SG (income averaging) increases preparation fees by about \$60, other things equal. However, the coefficient of SG ($p=0.3549$) is insignificant, and the coefficient of SC is only marginally significant with a p-value of 0.1001.

Surprisingly, the presence of SW (two-wage earner deduction) decreases preparation fee by about \$68 ($p=0.014$), other things equal. An examination of sample descriptive statistics reveals that about one third (32.8 percent) of paid-prepared returns include SW and about a half (48.6 percent) of paid-prepared returns that reported fee contain SW. This might

²² In this study, two-tailed p-values are used for t-statistics and one-tailed p-values for chi-square and F statistics.

indicate that preparers charge two-wage earners a lower price to retain their business. This is possible if the price elasticity of two-wage earners' demand for professional tax assistance is elastic.

PTFORM is the total return preparation time (in hours) of schedules and forms included in a return excluding the preparation time of the schedules that are already included as dummy variables. The insignificant coefficient of PTFORM ($p=0.557$) implies that the total preparation time of forms is not a significant determinant of preparation fees.

A comparison of Panels A and B of Table 4 suggests that the determinants of preparation fee for the 1987 returns are quite different from those of the 1986 returns. As shown in Panel B, the presence of SB, SD, SE or AMT (alternative minimum tax) on a 1987 return is not significantly related to the amount of preparation fees. However, the coefficient of TPI in 1987 is significant ($p=0.001$) and larger than that in 1986 ($F=69.12$, $p < .01$, Chow test). The correlations among TPI and SB, SD, SE and AMT are 0.18, 0.22, 0.25 and 0.32, respectively. These correlation coefficients are all significant at the 0.01 significance level. It is possible that returns with SB, SD, SE and/or AMT have higher TPI, and their effects on fees are intertwined with the effect of TPI.

The coefficient of SC is 174.93 ($p=0.060$), marginally significant, and larger than its 1986 counterpart ($F=67.03$, $p < .01$, Chow test). The Modified Accelerated Cost Recovery System (MACRS), the limitations of passive activity loss deductions, the new rules of alternative minimum tax and a number of other changes introduced by TRA86 instantly increased the complexity of the tax calculation of a business owner. The increased complexity might be the cause of the difference between the estimated coefficients of SC in 1986 and 1987.

As in 1986, the coefficient of PTFORM is insignificant. That is, preparation fees are not related to the preparation time of the forms attached in either 1986 or 1987.

Several other changes made by TRA86 affect the comparison basis of the determinants of preparation fees of the 1986 and 1987 returns. First, the regression uses observations only from itemized returns that reported preparation fees. However, the sample shows the percentage of itemized returns decreased from 41.7 in 1986 to 35.9 in 1987. This is because TRA86 (1) repealed the itemized deductions for state and local sales taxes, (2) applied a 7.5 percent floor to the medical expense deduction, (3) curtailed the deduction for consumer interest expense other than interest on principal or second residence, and (4) allowed miscellaneous itemized deductions only to the extent that their aggregate amount exceeds two percent of adjusted gross income.²³ Because the sampled itemized returns used in the regression analysis for 1986 and 1987 were not selected on the same basis. The comparison of the determinants of preparation fees is subject to the differential sampling basis.

A second change introduced by TRA86 that affects the determinant comparison basis is related to the TPI components in this study. TPI is defined as the sum of all positive income line items a taxpayer is required to report. The TPI of a 1986 return excluded 60 percent of long term capital gains, while TRA86 repealed the exclusion, and hence a 1987 return's TPI included the full amount of long-term capital gains. In addition, starting from the 1987 returns, taxpayers are required to disclose tax exempt interest. This also results in

²³ Prior to TRA86, taxpayers compared their zero bracket amount (ZBA) with the amount they could itemize in deciding whether itemization would be used. After TRA86, ZBA was replaced by the standard deduction. Though the ZBA of 1986 and the standard deduction amount of 1987 are of minute difference, taxpayers who used standard deduction in 1987 were not allowed to deduct their charitable deductions. Likewise, the moving expense which was deductible for the adjusted gross income (above the line) in 1986 became deductible from the adjusted gross income (below the line) and must be itemized on Schedule A. Both changes took away deductions from non-itemizers and they might have encouraged some non-itemizers in 1986 to itemize in 1987.

differences in TPI between 1986 and 1987 tax years. The data indicate that the average TPI is \$59,130 for the 804 fee-reported returns used in the regression for the 1986 tax year, and is \$67,291 for the 1633 returns used in the regression for the 1987 tax year.²⁴

Though the sampling basis is somewhat different, the goodness of fit of the fee equations is about the same between 1986 and 1987. The adjusted R^2 s are both around 0.32 and the F statistics are significant ($p < 0.01$) for both the 1986 and 1987 equations. The Breusch-Pagan test (Breusch and Pagan 1979) indicates both regressions are heteroscedastic ($p < 0.01$). Though the OLS estimators of β coefficients are unbiased and consistent, the estimated standard errors of β are biased and standard tests of significance are not correct. Without actually specifying the type of heteroscedasticity, White's heteroscedastic adjustment (White 1980) was used to provide unbiased standard errors of the β estimates.²⁵

4.1.2 Estimates of Aggregate Tax Compliance Cost Expended on Preparation Fees

The estimated equations for preparation fees in 1986 and 1987, as shown in Table 4, are used to predict the expected preparation fees of those paid-prepared returns whether reporting fees or not. Then a paid-prepared return's total monetary tax burden is obtained by adding the fee and the reported tax liability. This, in turn, allows the tax burden differential between two preparation modes to be examined in Section 4.2. In this section, the preparation fee equations are used to provide fee estimates for different return types, classified by income levels. As shown in Table 5, preparation fees increase with income in both years.

²⁴ The average TPI is \$29,280 for the 9019 returns of the whole 1986 sample, and is \$31,324 for the 18811 returns of the whole 1987 sample.

²⁵ Because the precise nature of the heteroscedasticity is unknown, neither weighted nor generalized least squares regression is employed though they may produce more efficient standard errors.

On average, preparation fees for low income returns are higher in 1986 than in 1987. The income threshold for filing increased considerably between 1986 and 1987.²⁶ There might be more low income and low complexity returns in 1987 that just filed for refunds, and tax preparers charged these returns a low price. Consequently, the average fee for low income returns decreased in 1987. On the other hand, preparation fees increase by about 48 percent (from \$324 to \$478) for those returns with TPI greater than \$50,000. Two possible reasons are provided here. First, the TRA86 changes may substantially increase return complexity for those high income returns in the 1987 tax year, compared to their 1986 counterparts. Second, the proportion of the paid-prepared returns with TPI greater than \$50,000 to the whole sample size, paid-prepared and self-prepared returns altogether, increases from 8.8 percent in 1986 to 9.9 percent in 1987. Based on these sample proportions, the number of paid-prepared returns with TPI greater than \$50,000 in the entire country increased from 4.2 million to 5.0 million between 1986 and 1987.²⁷ The increased demand may pull up the price of return preparation.

As shown in Table 5, the overall average preparation fee is \$154 for the 1986 paid-prepared returns, and is \$168 for the 1987 paid-prepared returns. The IRS Statistics of Income Bulletin reports that 48 million and 51 million returns had a preparer signature in 1986 and 1987, respectively (see Table 1). Applying the average fee payments to the number of taxpaying units, the point estimates of aggregate tax compliance cost expended on preparation fees in the 1986 and 1987 tax years are \$7.4 billion and \$8.6 billion respectively.

²⁶ For example, the filing threshold of married filing a joint return with both spouses under 65 was \$5,830 (gross income) for the 1986 tax year, and was \$7,560 for 1987. The thresholds for other status taxpayers increased by \$820 to \$2,490.

²⁷ Table 1 shows that the number of paid-prepared returns is 48 million in 1986 and 51 million in 1987.

Using survey data, Slemrod and Sorum (1984) estimate that the average fee payment of those who used professional tax assistance was \$76 (\$86.7 in 1987 constant dollars) and that about \$3 billion (\$3.4 billion in 1987 dollars) was spent on professional tax assistance in the 1982 tax year.²⁸ Employing a similar research method to that of Slemrod and Sorum (1984), Blumenthal and Slemrod (1992) estimate that the average fee charged on the 1989 paid-prepared returns was about \$132 (\$115 in 1987 dollars). This implies that the nationwide expenditure on commercial tax return preparation was about \$7.0 billion in 1989 (\$6.1 billion in 1987 dollars). Their estimates from survey studies are lower than the estimates in this paper. One possible reason is that the estimates in this paper may suffer from selection bias (see Section 4.1.3.). Another possible reason is that tax preparer industry raised the price of return preparation upon the adoption of the TRA86 in 1986 and 1987, and the price went down in 1989 when tax preparers were more adept at TRA86.²⁹ TRA86 was signed by former President Ronald Reagan on October 22, 1986. Tax preparers might have started to provide more tax planning and consultation for some of their clients along with their preparing the 1986 tax returns filed in 1987. Therefore, the return preparation fee started to increase on the 1986 tax returns. In addition, as Table 1 shows, the number of paid-prepared returns in the entire country increased from 48 million in 1986 to 51 million for 1987. The increased demand also might have caused the increases in both the average and aggregate expenditures for return preparation in 1987.

²⁸ To calculate the 1987 constant dollars of the average fee paid for the 1982 returns, the consumer price indices of 1983 and 1987 are compared. This is because the preparation fees for the 1982 returns are generally paid in 1983. This method is consistently used in this section.

²⁹ A small part of TRA86 is effective in 1986. For example, the repeal of regular investment tax credit and the imposition of new generation-skipping transfer tax.

Overall, the analysis indicates that both average and total expenditure expended on professional tax assistance is greater in 1987 than in 1986. This finding provides some confirmation of Blumenthal and Slemrod's (1992, 185) conclusion that tax reform did not reverse the growth in compliance cost (taxpayers' time and preparation fee) in the 1980s.

4.1.3 Summary and Discussion

Consistent with findings in previous survey studies (e.g., Slemrod and Sorum 1984; Blumenthal and Slemrod 1992), this study also finds tax return preparation fees are positively correlated with income in both the 1986 and 1987 tax years. However, no constant patterns of the impact of the presence of supplementary schedules on preparation fees are found during the 1986-1987 period. The analysis indicates that attaching an supplemental schedule to a tax return generally increased preparation fee in the 1986 tax year. However, this is not true in 1987 except for attaching Schedule C to a sole proprietor's tax return. Because TPI and the presence of SB, SC, SD, SE, SG, SW and AMT are correlated, the effect of income may be interwoven with the effect of these supplementary forms and schedules in producing the effects on preparation fees. Therefore, it is worth pointing out that all the dummy variables indicating the presence of a supplementary schedule have positive coefficients (though some may be insignificant), except the dummy variable of two-wage earners (SW) in 1986. The analysis also shows that the preparation time of supplementary forms is not a significant determinant of preparation fees.

A limitation exists in the fee estimation. The regression uses observations only from paid-prepared returns that itemized preparation fees, yet the results are also generalized to paid-returns that did not report fees. If there are systematic differences between these two return groups that are not accounted for by the regressors of the fee equation, then there

might be selection bias in the fee estimation. Heckman's two-stage selection adjustment model (Heckman 1976, 1979) potentially could be used to account for this bias. However, because a satisfactory specification for the indicator function for who reported preparation fee was not found, Heckman's two-stage procedure predicts negative fees for a large fraction of paid-prepared returns, especially those with low TPI. Therefore, the Heckman's model is not used. Because the estimated fees account for only three percent of the total tax burden on average, the selection bias is not expected to have a material influence on the subsequent tax burden analysis.³⁰

4.2 Analysis of Total Tax Burden and Tax Burden Differential between the Preparation Modes

The most important objective of this study is to examine the tax burden differential between paid-prepared returns and self-prepared returns. This section discusses the composition of total tax burden and then the analysis of the tax burden differential between preparation modes. For a paid-prepared return, total tax burden is its tax liability plus the estimated fee adjusted to account for the fee's deductibility or non-deductibility on the subsequent year's tax return. For a self-prepared return, total tax burden is just its tax liability. Table 6 reports the percentage of fees to the total tax burden of paid-prepared returns. Fees are estimated using the OLS regression discussed in the preceding section. For those returns with total positive income (TPI) less than \$20,000, the proportion of fees to tax burden decreases from 12.4 percent in 1986 to 7.7 percent in 1987. While the proportion remains unchanged for those with TPI between \$20,000 to \$50,000, the proportion increases from 1.6 percent to 2.5 percent for those with TPI greater than \$50,000 during this two-year

³⁰ See Table 6 and more discussions in section 4.2.

period. The overall proportion of fees to tax burden is about three percent in both the 1986 and the 1987 tax years. The data appear to suggest that the adoption of the TRA86 decreased the proportion of fees for low income taxpayers (those with TPI less than \$20,000), yet it increases the proportion of fees for high income taxpayers (those with TPI greater than \$50,000).

4.2.1 Regression Analysis of Total Tax Burden

This section presents the estimates of total tax burden equations using the endogenous switching regression (ESR) model.³¹ In the analysis, returns with TPI in the upper one percent range are deleted. The limitation is imposed for the following reasons. First, the distributions of the returns with TPI in the upper one percent are quite different between 1986 and 1987. Among the returns with upper one percent TPI in 1986, the 19 returns with highest TPI are paid-prepared; while in 1987 the two returns with highest TPI are self-prepared. The highest TPI in one regime of preparation mode choice is more than twice as large as that in the other regime in each same year. Second, in both years, the self-prepared returns account for only ten percent of the upper one percent TPI returns, far away from the about fifty-fifty distribution of each year's overall sample. The unbalanced distributions not only have a strong influence on the coefficient estimation but also affect the tax burden comparisons between these two years. For example, in estimation without excluding the upper one percent, the coefficients of the selection-adjustment variables of the paid-prepared

³¹ As noted in Section 3.1, because the maximum likelihood procedure did not converge for the parameter estimation of the ESR models, a two-stage procedure (Heckman 1976, 1979; Lee 1976, 1978) is used. The two-stage procedure provides consistent β estimates of Equations (5) and (6) but their error terms are heteroscedastic (Lee 1978, 422). To cope with the heteroscedasticity, the corrected variance-covariance matrix of β coefficients is obtained based on Greene (1981).

regime have opposite signs in 1986 (positive) and 1987 (negative). In addition, the estimation also produces a negative average tax burden for the self-prepared returns in 1986 had they been paid-prepared.

Efforts have been made to delete various combinations of outliers from the endogenous switching regression.³² Deleting different numbers of outliers influences the coefficient estimates, yet it still does not balance the distribution of paid- and self-prepared returns in the upper one percent TPI range -- perhaps because of this reason, the estimations consistently produce very large coefficients of selection adjustment variables, which result in unusual tax burden estimates. Trimming off the upper one percent of observations with the highest income is used in this study because this approach does not sacrifice too much generalizability. After trimming off the one percent of observations, the upper bound of TPI is \$146,500 in 1986, and is \$176,000 in 1987. The descriptive statistics of the sample excluding returns with TPI in the upper one percent are reported in Appendix A.

Using the trimmed 99 percent samples, the ESR estimates of tax burden equations are presented in Tables 7 and 8 for the tax years 1986 and 1987, respectively. For comparison purposes, the OLS estimates without correcting for self-selection bias are also reported in the same tables.

Comparison of ESR and OLS tax burden equations for paid-prepared returns

A comparison of the ESR estimates with those obtained from OLS reveals several interesting points. For the paid-prepared regime (Panel A in Tables 7 and 8), the signs of all the estimated coefficients are the same between the ESR and OLS models, with the exceptions

³² For example, deleting one or two outliers in the tax burden estimation in either or both preparation regimes. The potential outliers are identified from the magnitude of the studentized residual in the OLS regression.

of the constant term and the coefficients of SB and SD variables. Though the signs of the estimated coefficients are mostly the same, the magnitude of the estimated coefficients differ substantially between the two models for the SA (itemized deductions), SE (income from rents, royalties or partnerships), SEM (self-employment status), AGE (age 65 or over) and MS (marital status) variables. This indicates that self-selection bias may result if OLS is used to estimate total tax burden for the paid-prepared returns. The selection-adjustment variables have negative and significant coefficients in Panel A of both Tables 7 and 8.³³ This also indicates that returns in the paid-prepared group are selected from a population that has different characteristics from the general population consisting of both paid-prepared and self-prepared returns. Thus, correcting for self-selection bias is necessary in the estimation of total tax burden for paid-prepared returns.

Comparison of ESR and OLS tax burden equations for self-prepared returns

For the self-prepared return regime (Panel B in Tables 7 and 8) the estimated coefficient of the selection-adjustment variable is not significant in either year. It implies that the OLS estimates of the tax burden equations for self-prepared returns are not subject to significant self-selection bias. Related to this, the difference between the estimated coefficients under ESR and OLS models appears to be immaterial. Thus, correcting for the self-selection bias will not have a noticeable effect on the estimation of total tax burden for the

³³ Because the inverse Mills ratio is positive for the paid-prepared sample, the negative coefficient of the selection-adjustment variable indicates negative truncation in the ESR model. The negative truncation means that one only observes the lower sections of the tax burden distributions given fixed values of the independent variables (Lee 1978, 425). That is, if all returns with the same values of exogenous variables had been paid-prepared, those who actually used preparers would have paid less than average. However, they might have paid even less if they had prepared their own returns because the analysis below also shows that self-preparation on average incurs less tax burden than paid-preparation.

self-prepared returns. However, it is reasonable to expect that there would be some heterogeneity between the population of self-prepared returns and the population of the paid- and self-prepared returns alike. Therefore, for the following analysis of tax burden comparisons, the ESR estimates are used.

Comparison of ESR tax burden equations of paid- and self-prepared returns

Of greater concern are the comparisons of the ESR estimated coefficients between the paid-prepared regime (Equation 5; left column of Panel A in Tables 7 and 8) and the self-prepared regime (Equation 6; left column of Panel B in Tables 7 and 8). The results of the ESR estimation show that Equation (5) and Equation (6) differ markedly in both the magnitude and the sign of the estimated coefficients. This suggests that the effect of the regressors on tax burden depends on preparation mode, and hence the preparation mode choice affects taxpayers' tax burden.³⁴ The tax burden differential of a return between the two preparation modes will be discussed in next section.

There are two major differences among the ESR estimated coefficients of Equations (5) and (6). First, the coefficient of SB (interest and dividends) in Panel A, Table 7 is -475.67 ($p=0.001$), yet it is 130.22 ($p=0.281$) in Panel B. This implies that, other things equal, paid-prepared returns with SB attached in 1986 have lower tax burden than paid-prepared returns without SB attached. The same statement is not true for self-prepared returns in the same year. This might reflect that tax preparers are efficient in deducting

³⁴ This indicates that the ESR model that allows for full interactions between the regressors and the preparation modes is a better specification than a model that just uses a dummy variable to account for the effect of preparation mode. See Greene (1990, 747-748) for a discussion of the latter model.

investment expenses. The same argument can be made for the ESR estimated coefficients of SB in Table 8 (for the year 1987) if a 0.10 significance level is used.

Second, in Table 8 the ESR estimated coefficients of AMT and PAL for the paid-prepared regime are both about \$1,500 lower than those for the self-prepared regime. This may suggest that tax preparers are more effective than taxpayers in handling alternative minimum tax and passive activity loss limitations.

4.2.2 Comparisons of Tax Burden Differential

As noted in Section 4.2.1, the results of ESR estimation indicate that the preparation mode affects taxpayers' tax burden. The estimated tax burden equations in Tables 7 and 8 can be used to estimate a return's tax burden differential (TBDIF) between the two preparation modes. As indicated in Section 3.1.1, a return's tax burden of the two alternative preparation modes is estimated using the estimated $\hat{\beta}_1'X_i$ and $\hat{\beta}_0'X_i$ of Equations (5) and (6), but the selection-adjustment terms $\theta_1\lambda_i$ and $\theta_0\lambda_i$ are excluded. Tax burden differential is then estimated by $\hat{\beta}_0'X_i - \hat{\beta}_1'X_i$. If it is positive, there are tax savings from using preparers. If it is negative, then there is additional tax burden from using preparers. Table 9 reports the estimated tax burden differential between the two preparation modes for different income groups. For comparison purpose, OLS estimates are also reported along with ESR estimates. As indicated above, self-selection bias may result if OLS is used to estimate the tax burden of using tax preparers. Therefore, the subsequent discussion of tax burden differential will focus on the ESR estimates.

As Table 9 shows, the ESR model predicts that using tax preparers will on average increase tax burden for taxpayers in all income groups in both 1986 and 1987 tax years.³⁵ The estimated average tax burden increment for the whole population is about \$1,650 in both years. Overall, the tax burden differentials as estimated by the ESR model appear to be quite stable in magnitude over the 1986-87 period. In Table 9 the highest differential (\$1,962) is for the self-prepared returns with TPI less than \$20,000 in 1986, and the lowest differential (\$1,262) is for the paid-prepared returns with TPI between \$50,000 and \$176,000, the ninety-ninth percentile of the 1987 sample. If the tax burden increment is measured by the proportion of tax burden differential to total tax burden, the low income returns will have the highest tax burden increment. For example, the average proportion of burden increment for the paid-prepared returns in 1986 is 70 percent for the low income group, 25 percent for the middle income group, and 13 percent for the high income group.³⁶

Table 9 also indicates that self-prepared returns on average would have paid about \$1,800 more had they chosen to use tax preparers. Paid-prepared returns on average paid about \$1,500 more than they would have paid had they been self-prepared. The first column from the right shows the estimated fees paid by paid-prepared returns. Because preparation fees paid by paid-prepared returns on average are only \$140 in each year, far less than the tax burden differential of \$1,500, tax return preparation fees are not the only reason for the additional tax burden incurred by paid-prepared returns. For the lowest income group, the additional tax burden associated with paid-preparers is very large relative to the burden observed in Table 6. However, the additional tax burden is based on selection-corrected

³⁵ The OLS estimates also indicate that using tax preparers increases taxpayers' tax burden except for the group of TPI less than \$20,000.

³⁶ The percentage is obtained from dividing a group's average tax burden differential by the mean of the estimated tax burden of using preparers for that group.

predictions, and the observed burden reflects self-selection. Section 4.2.3 discusses some reasons why returns using tax preparers had a higher tax burden.

4.2.3 Summary and Discussion

One of the merits of this study is its use of total tax burden (tax liability plus preparation fee) rather than tax liability to better reflect taxpayers' total outlay. For paid-prepared returns with TPI less than \$20,000 the percentage of estimated fee to total tax burden was 12.4 in 1986, and 7.7 in 1987. Provided taxpayers compare their total monetary expenditure between self-preparation and paid-preparation in their preparation mode choice, a model will not fully capture low income taxpayers' preparation mode choice if it uses tax liability differential rather than tax burden differential as a regressor.

While estimating the total tax burden equation for paid-prepared returns, the ESR estimates that the selection-adjustment variable is negative and significant in both 1986 and 1987. It suggests that correcting for self-selection bias is necessary in the estimation of total tax burden for the paid-prepared returns. In addition, total tax burden equations for paid-prepared and self-prepared returns differ markedly in both the magnitude and the sign of the estimated coefficients. This provides evidence that the effect of the regressors on the tax burden depends on preparation mode, and that the preparation mode choice affects taxpayers' tax burden.

Using 1983 tax return data, Long and Caudill (1987) found that using professional tax assistance reduced taxpayers' tax liability (fee not considered). In contrast, this study provides evidence that using tax return preparers increases taxpayers' tax burden in both 1986 and 1987. On average, the additional tax burden of using tax preparers is about \$1,650. Table 9 shows that preparation fees are far less than the additional tax burden. Therefore, the

additional tax burden of using tax preparers cannot be accounted for only by the fee burden. The benefits of using professional tax assistance may have been changed after 1983 because of the enactment of TRA86 or legislative changes in preparer penalties between 1983 to 1987. Tax preparer penalties are primarily prescribed in Internal Revenue Code (IRC) §6694, §6695, §6696, §7407, §7408 and Treasury Department Circular No. 230.³⁷ Since being added to the Code in 1979, none of these IRC sections except §7408 have been changed or amended until 1989. IRC §7408 was broadened on September 4, 1982 to permit the IRS to seek an injunction against preparers who knowingly understate tax liability in violation of IRC §6701 (relating to penalties for aiding and abetting understatement of tax liability). Related to this, Treasury Department Circular No. 230 was amended in 1984 requiring an IRS practitioner to comply with new standards when providing a tax shelter opinion. Because the tax laws regarding preparer penalties were not substantially revised, it is reasonable to argue that TRA86 substantially reduced the benefits of preparer usage in 1986 and 1987. This may suggest that preparers are more careful in preparing returns during tax reform years. Consequently, paid-prepared returns, *ceteris paribus*, reported higher tax liability than self-prepared returns during tax reform years.

Two policy implications can be derived. First, if a taxpayer expects additional tax and penalties assessed by the IRS would be less than \$1,650 had he prepared his own return, then self-preparation might be the correct choice, unless the benefits of return completion time saved and the psychic cost of self-preparation is greater than monetary considerations. Second, in tax reform years self-prepared taxpayers may inadvertently underreport their tax

³⁷ §6694: Understatement of taxpayer's liability by income tax return preparer; §6695: Other assessable penalties with respect to the preparation of income tax returns for other persons; §6696: Rules applicable with respect to sections §6694 and §6695; §7407: Action to enjoin income tax return preparers; §7408: Action to enjoin promoters of abusive tax shelters, etc.

liabilities because of misunderstanding new tax laws or may underreport intentionally because of the intent to exploit the ambiguous parts of the new laws or the transitional rules. To enhance tax equity and compliance, the IRS could allocate more resources to audit the self-prepared returns filed a year before and after a major tax reform act.

A further analysis was conducted to investigate those returns that are estimated to have tax savings (positive TBDIF) from using tax preparers. The analysis reveals that these returns on average have more than twice the number of attached schedules (SCHS) and the return completion time (CT) compared to those returns having additional tax burden.³⁸ It appears to suggest that returns of high complexity are more likely to have tax savings from using professional tax assistance. This is consistent with the finding in Section 4.2.1 that tax preparers are more effective than taxpayers in handling alternative minimum tax and passive activity loss limitations.

4.3 Analysis of the Usage of Tax Preparers

4.3.1 Structural Form Estimates of the Preparation Mode Equation

The third objective of this study is to examine the determinants of the usage of tax preparers. Table 10 reports the structural form estimates of the preparation mode equation (Equation 3) for both 1986 and 1987.³⁹ As the table shows, the likelihood ratio chi-square statistics are significant ($p < 0.01$) in both Panels A and B. This suggests that the null

³⁸ In 1986 there are 161 observations estimated to have tax savings. Their average SCHS and CT are 2.89 and 55.3 hours, respectively. The other 8,768 observations in 1986 on average have 1.28 SCHS and 21.6 hours of CT. In 1987 there are 239 observations estimated to have tax savings. Their average SCHS and CT are 3.05 and 43.5 hours, respectively. The other 18,384 observations in 1987 on average have 0.93 SCHS and 21.8 hours of CT.

³⁹ The reduced form results (Equation 4) are reported in Appendix B for the 1986 sample, and Appendix C for the 1987 sample.

hypothesis that all coefficients are equal to zero should be rejected. Also, the model appears to have good prediction accuracies, with 63.6 percent of all returns correctly classified in 1986 and 66.0 percent in 1987.⁴⁰ Of particular interest is whether tax savings from using preparers have positive relationship with the propensity to engage a preparer. In Table 10 tax savings from using a preparer is represented by the variable TBDIF (tax burden differential, $\hat{\beta}_0 + X_i \hat{\beta}_1 + X_i$). The estimated coefficient of TBDIF is positive and significant ($p < 0.001$) in both Panel A and Panel B of Table 10. This suggests that the higher the tax savings from using a preparer (or the lower the unfavorable monetary tax burden of using a preparer), the higher the likelihood a taxpayer is to use a preparer. For example, if TBDIF increases by \$100, the probability of using preparers will increase by 1.2 percent in 1986 and 2.7 percent in 1987.

Consistent with prior studies (e.g., Long and Caudill 1987; Blumenthal and Slemrod 1992; and Christian et al. 1992), whether a taxpayer is self-employed is found to be a critical determinant of the preparer engagement. Evaluated at the mean values of all regressors except SEM (self-employment status), the probability that the average self-employed taxpayer engages a preparer is 15.6 percent larger than the average non-self-employed taxpayer in 1986 (22.6 percent in 1987). Other demographic variables are correlated with the usage of preparers as well. The probability that the average married taxpayer uses a preparer is about six percent higher than that of the average single taxpayer in both 1986 and 1987. On average, a taxpayer 65 or over is about eight percent in 1986 (six percent in 1987) more

⁴⁰ Using a similar ESR model, Christian et al. (1993) correctly classified 68.6 percent of the returns in the 1983 Panel.

likely to engage a preparer than a taxpayer under 65. Adding a dependent increases the usage probability by about 0.8 percent in 1986 and 1.2 percent in 1987.⁴¹

Tax return completion time (CT), which is return preparation time plus time of learning about the law, is also positively related to preparer usage ($p < .001$ in both years). The average tax return completion time is about 22 hours in both 1986 and 1987. If the completion time is decreased by ten percent (i.e., 2.2 hours), the probability of using preparers will decrease by about 0.2 percent. This suggests that though completion time is associated with preparer usage, preparer usage is primarily driven by the other factors.

4.3.2 Impact of Preparation Fee Changes on Preparer Usage

Simulations were conducted to measure the effect on the preparer usage of a ten percent decrease in preparation fee charged by preparers. Assume that a one dollar reduction in preparation fee will cause the benefits of using preparers to increase by the same amount. The simulations are conducted for three groups for each year of 1986 and 1987: the group of taxpayers having TPI less than \$20,000, the group of TPI between \$20,000 and \$50,000, and the group of TPI between \$50,000 and \$146,500 (in 1986) or \$176,000 (in 1987). The mean characteristics of each group are plugged into the estimated preparation fee equation in Table 4 to obtain the average fee of that group. Then the impact on the preparer usage of a ten percent reduction in the average fee of that group is $f(\bar{X}'\hat{\beta})\hat{\beta}_{\text{TBDIF}}*(0.1*\text{fee})$, where $f(\cdot)$ is the standard normal probability density function, $\hat{\beta}$ is the estimated coefficient vector of the structural probit, and \bar{X} is the mean vector of the independent variables for the group of interest. For all six groups, the probability increases only by 0.1 percent to 0.8 percent.

⁴¹ For the non-binary variables in Table 10, the marginal effect is calculated by $f(\bar{X}'\hat{\beta})\hat{\beta}_k$. See Judge et. al. (1985, 767).

This appears to indicate that tax preparers cannot effectively increase service volume by lowering their prices. Because the magnitude of the price elasticity of demand is found to be less than one, the demand for preparation service is price-inelastic. An even more important implication of the simulations is that increases (decreases) in preparation price will produce increases (decreases) in revenue.

4.3.3 An Alternative Preparation Mode Equation Estimate

The structural form of the preparation mode choice (Equation 3) assumes that tax burden differential ($\hat{\beta}_0 \cdot X_i - \hat{\beta}_1 \cdot X_i$) influences taxpayers' decision on the preparer engagement. However, it is also possible that only tax burden of using preparers ($\hat{\beta}_1 \cdot X_i$) or only tax burden of being self-prepared ($\hat{\beta}_0 \cdot X_i$) affects the engagement decision. If this is true, then the model using tax burden differential suffers from specification error. To test the possibility, the following alternative preparation mode equation is estimated using probit for both 1986 and 1987 samples:

$$DP^* = \gamma_0 + \gamma_1 TB_1 + \gamma_2 TB_0 + \gamma_3 CT + \gamma_4 AGE + \gamma_5 MS + \gamma_6 DEP + \gamma_7 SEM + u, \quad (10)$$

where TB_1 is the tax burden of using a preparer and TB_0 the tax burden of being self-prepared. Both TB_1 and TB_0 are estimated from the endogenous switching regression (ESR) model as described in Section 4.2. If γ_1 or γ_2 or both are not significantly different from zero, then the structural probit and hence the whole model used in this paper are misspecified.

Table 11 reports the results of the alternative preparation mode equation estimates. In both 1986 and 1987, the coefficients of TB_1 and TB_0 are significant ($p < .001$ in all four cases). This suggests that neither γ_1 nor γ_2 is zero. The likelihood ratio test also indicates

that the null hypothesis that both γ_1 and γ_2 are zero is rejected at the 0.01 level.⁴² This suggests that a model that does not include both TB_1 and TB_0 is subject to omitted variable bias.

Aside from the TB_1 , TB_0 and $TBDIF$, the estimated coefficients are quite similar between Table 10 and Table 11. There are two exceptions. First, the estimated coefficient of AGE in 1986 is positive yet insignificant ($p=0.15$) in Table 11, but is positive and significant ($p=0.001$) in Table 10. Second, the estimated coefficient of MS in Panel A, Table 11 implies that married taxpayers are less likely to retain preparers in 1986, but its counterpart estimated from the structural form predicts reversely.

In both 1986 and 1987, the coefficient of TB_1 is negative and significant ($p < .001$), and the coefficient of TB_0 is positive and significant ($p < .001$). This indicates that taxpayers are less likely to use preparers if their tax burden of using paid-preparation is increased by one dollar; and taxpayers are more likely to engage preparers if the tax burden of being self-prepared is increased by one dollar. This finding is not contradictory to the positive coefficient of $TBDIF$ (i.e., $TB_0 - TB_1$) estimated for the structural form in Table 10. Therefore, the structural form equation, which includes $TBDIF$ as a regressor, is used to model taxpayers' decision on the selection of preparation mode. The approach of using $TBDIF$ (tax burden differential) as a regressor is consistent with the models used in previous studies such as Lee (1978) (which used percentage union wage differential), Kahn and Low (1982) (which used differential between the logarithm of wage), and Christian et. al. (1993) (which used differential in the income tax prepayment associated with preparer usage).

⁴² The statistic used is likelihood ratio, $LR = -2(L_R - L_U) \sim \chi^2_m$, where L_R and L_U are the maximum of the log-likelihood function when the restrictions ($\gamma_1 = \gamma_2 = 0$) are imposed and are not imposed, respectively; and degrees of freedom m is the number of restrictions, which is two in this case (Kmenta 1986, 491). The LR is 316 ($p < 0.01$) in 1986 and is 696 ($p < 0.01$) in 1987.

4.3.4 Summary and Discussion

Though the analysis of preparer usage is but an ancillary objective of this study, several findings are noteworthy. Tax savings from using a preparer (TBDIF) and return completion time (CT) as estimated by the IRS have never been used simultaneously to model the preparation mode choice. In this study, they are found to be positively correlated with the preparer usage. However, the magnitude of their effect appears to be quite small. Prior research provides some comparable results regarding the effect of CT and TBDIF. For example, using Chamberlain's conditional logit model and the 1982-84 Panel, Christian et al. (1992) found that reducing the average time of reporting burden by 50 percent decreased the probability of preparer usage only by 2.9 percent.⁴³

In a paper similar to this study, Christian et al. (1993) used the ESR model and the 1983 Panel to examine preparers' effect on taxpayers' prepayment. In the structural form probit of that paper, the difference in individual income tax prepayment associated with the use of preparer (PPDIF, as they designate) was used as a regressor of preparation mode choice. They found that the estimated coefficient of PPDIF was not significant. This result appears to suggest that the difference in tax prepayment associated with the use of preparer (PPDIF) is not a determinant of the preparer usage. In this paper, tax savings from using a preparer (TBDIF) is found to be positively related to the preparer usage. This may reflect that TBDIF is a better proxy of the monetary benefits of using tax preparers than PPDIF. In Section 4.3.3 an alternative preparation mode equation estimate is used to compare with the

⁴³ Christian et al. (1992) defined their TIME as the sum of preparation time, the time of learning about the law and the time of sending the form to IRS for each form and schedule included in a return. In this study the time of sending the form is not included in the return completion time (CT).

structural form probit of the preparation mode choice equation (Equation 3). The results also indicate that the structural form probit is appropriate.

As reported by Long and Caudill (1987), this study also finds that the probability of using preparers is higher for self-employed taxpayers, married taxpayers, taxpayers over 64 and taxpayers with more dependents. Particularly, self-employed taxpayers have a substantially higher probability of engaging preparers than wage earners. Prior research provides two possible reasons. One is that self-employed taxpayers' returns are more complex, and tax return complexity plays a major role in the engagement decision.⁴⁴ The other is that the accountant handling the financial affairs of the business routinely prepares the owner's tax return as well (Long and Caudill 1987, 37).

A noteworthy finding from simulations is that the demand for preparation service is price-inelastic. This suggests that increases (decreases) in preparation price will produce increases (decreases) in revenue.

⁴⁴ Self-employed taxpayers have to complete Schedule C for business income or Schedule F for farm income and Schedule SE for self-employment taxes.

V CONCLUSIONS, LIMITATIONS AND EXTENSIONS

This paper presents empirical evidence that, *ceteris paribus*, returns prepared by tax preparers generally have a higher tax burden (tax liability plus fee) than self-prepared returns in 1986 and 1987, except for returns of high complexity. Because preparation fees are far less than the additional tax burden, the evidence suggests that paid-prepared returns report higher tax liabilities. This contrasts with previous evidence that tax preparers reduce tax liabilities (Long and Caudill 1987). The unfavorable tax burden of paid-prepared returns in 1986 and 1987 is more likely associated with the enactment of the Tax Reform Act of 1986 (TRA86) than associated with the minor tax law changes in preparer penalties between 1983 and 1987. Because this study uses data spanning TRA86, the evidence should be interpreted as paid-prepared returns report higher tax burden during tax reform years. Whether the tax benefits of using preparers are persistently reduced after 1983 or by TRA86 is a question that cannot be answered without using tax return data of other years.

Because tax return data in the Panel are unaudited, the finding that paid-prepared returns report higher tax burden does not necessarily mean they have higher tax burden after IRS examination. To decide whether using preparers is preferable, one must also take account of additional tax and penalties assessed by the IRS, the non-monetary benefits of return completion time saved, and the relative psychic cost of self-preparation. The evidence that paid-prepared returns report higher tax liabilities may suggest self-prepared returns are more likely to underreport tax, intentionally or inadvertently, during tax reform years. To enhance tax equity, the IRS could allocate more resources to audit self-prepared returns filed during tax reform years.

Subject to potential selection bias, this study also provides evidence that higher income and the presence of Schedules B, C, D or Schedule E in 1986 and Schedule C in 1987 are associated with higher return preparation fees. Based on the fee equation, the aggregate

tax compliance cost expended on commercial tax return preparation is estimated to be \$8.6 billion in 1987 (post TRA86), higher than the \$7.4 billion predicted for 1986, and the \$3 billion estimated for 1982 by Slemrod and Sorum (1987). This finding provides evidence consistent with Blumenthal and Slemrod's (1992, 185) conclusion that tax reform did not reverse the growth in compliance cost in the 1980s. Tax policy makers should be concerned about the increased expenditure for return preparation. It has been argued that high compliance costs may act as a barrier to compliance because taxpayers will simply fail to file (or fail to report a particular item) because of the compliance costs involved (Roth, Scholz and Witte 1989, 118).

A unique merit of this paper is the inclusion of tax burden differential (TBDIF) between self-preparation and paid-preparation as an explanatory variable in modelling the propensity to engage preparers. The analysis shows that TBDIF is positively related to the choice of being paid-prepared. This indicates that it is necessary to control for the simultaneity between TBDIF and the choice of preparation mode in the endogenous switching regression model.

In addition to using TBDIF to proxy the monetary benefits of using preparers, this study also uses return completion time (CT) to represent the non-monetary benefits of time saved by using preparers. TBDIF and CT are both found to be positively related to preparer usage in both 1986 and 1987. Though there are many studies on the determinants of preparer usage, no one has ever used these two variables together as predictors. This study shows that increases in the benefits of TBDIF and time saved increase the preparer usage.

Demographic variables including self-employment status, marital status, age and number of dependents are shown to be important determinants of preparer usage. The

estimated coefficients also suggest that the determinants of preparation mode choice are not changed remarkably between 1986 and 1987.

This paper also conducts simulations to measure the effect on preparer usage of a ten percent decrease in the preparation fee charged by preparers. The results indicate that the preparers cannot effectively increase service volume by lowering their service charges. The simulations also show that the demand for preparation service is price-inelastic. This suggests that increases (decreases) in preparation price will produce increases (decreases) in revenue.

This study is subject to several limitations. First, the tax liability investigated is based on the reported tax liability on the tax return. Ideally, a taxpayer's total individual income tax burden should be measured by the tax liability as examined by the IRS plus the preparation fee if a preparer were used. Though the IRS Taxpayer Compliance Measurement Program (TCMP) data contain both the reported and examined tax liability, access to the TCMP data is highly limited, and tax return preparation fees are not included in TCMP data.

Second, the literature has provided evidence that different factors are associated with the engagement of different types of tax preparers (Dubin et al. 1992) and that type of tax preparer significantly affects preparation fees (Swingen et al. 1992). It is possible that type of tax preparer also affects the tax savings of paid-prepared returns. However, this study cannot examine this question because the Panel does not reveal whether a preparer is a CPA, attorney, enrolled agent, or unenrolled practitioner.

Third, though the relevant tax laws and the IRS publications instruct taxpayers to deduct their preparation fees paid in 1987 (1988) for the 1986 (1987) returns in the 1987 (1988) tax year, it is possible that some taxpayers may deduct their preparation fees paid for their 1986 (1987) returns on the 1986 (1987) returns. This causes measurement error.

However, it is assumed that most preparers are aware of the correct treatment of the deductibility of preparation fees, and that the measurement error is not significant.

Fourth, state income tax liability is not included in the calculation of total tax burden. In 1986 and 1987 all states except Alaska, Florida, Nevada, South Dakota, Texas, Washington, and Wyoming imposed a state income tax on individuals.⁴⁵ One can argue that a taxpayer's total income tax burden should be measured by the federal income tax, state income tax and preparation fee. However, state income tax is not included in the tax burden comparison between the two preparation modes for the following reasons. Though a taxpayer may engage a preparer to prepare the federal tax return and state tax return simultaneously, it is also possible that a taxpayer only purchases professional assistance for the federal tax return but self-prepares the state tax return. Because data on the preparation mode of state income tax returns are not available, this paper cannot analyze the tax burden differential defined to include the state income tax liability.

A natural extension of this study would examine the persistence of the reduction in the monetary benefits of using tax preparers after TRA86 by examining more recent data. When the 1989 and 1990 tax return data are added to the Panel, tax burden including fees could be estimated for 1988 and 1989. Data from the extended period 1986-1989 could be used to directly address this question. TCMP could be used in another extension, if fee data are added. Because a return's additional tax and penalties assessed by the IRS are included in the TCMP data, the tax burden differential estimated with TCMP reflects the genuine monetary benefit of using preparers. In addition, whether there are different tax savings of using attorneys, CPAs, enrolled agents, and unenrolled practitioners can be examined using the

⁴⁵ In Connecticut, New Hampshire, and Tennessee, individuals were taxed only on certain interest and dividend income in 1986 and 1987.

TCMP data because it discloses type of preparer. However, the latest TCMP data (the 1988 TCMP) does not include preparation fees, so only tax liability differential rather than total tax burden differential can be examined.

REFERENCES

- Blumenthal, M., and J. Slemrod. 1992. The compliance cost of the U.S. individual income tax system: A second look after tax reform. *National Tax Journal* (June): 185-202.
- Breusch, T. S., and A. R. Pagan. 1979. A simple test for heteroscedasticity and random coefficient variation. *Econometrica* 47 (September): 1287-94.
- Christian, C. W. 1991. Endogenous sampling in the Statistics of Income Panel of individual returns. *Journal of Economic and Social Measurement* 16: 199-219.
- _____, S. Gupta, and S. Lin. 1992. Determinants of tax preparer usage: Evidence from panel data. Working paper, Arizona State University, Tempe, AZ.
- _____, S. Gupta, G. J. Weber, and E. Willis. 1993. An Empirical Investigation of Tax Preparer Effects on Taxpayers' Prepayment Position. Working paper, Arizona State University, Tempe, AZ.
- Collins, J. H., V. C. Milliron, and D. R. Toy. 1990. Factors associated with household demand for tax preparers. *The Journal of the American Taxation Association* (Fall): 9-25.
- Deere, D., and C. Wolfe. 1991. Tax rates and the usage of paid tax return preparers. Working paper, Texas A & M University, College Station, TX.
- Dubin, J., M. Graetz, M. Udell, and L. Wilde. 1992. The demand for tax return preparation services. *The Review of Economics and Statistics* (February): 75-82.
- Erard, B. 1990. Tax practitioners and tax compliance: A microeconomic analysis of the decision to engage a tax preparer and its consequences. Unpublished manuscript, University of Michigan.
- Greene, W. H. 1981. Sample selection bias as a specification error: comment. *Econometrica* 49 (May): 795-98.
- _____. 1990. *Econometric Analysis*. New York, NY: Macmillan.
- _____. 1991. *LIMDEP 6.0*. Bellport, NY: Econometric Software.
- Heckman, J. J. 1976. The common structure of statistical models of truncation, sample selection, and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement* 5 (Fall): 475-92.
- _____. 1979. The sample selection bias as a specification error. *Econometrica* 47 (July): 153-62.
- Internal Revenue Service. 1987. Miscellaneous Deductions. *Publication 529* (November). Washington, D.C.: U.S. Government Printing Office.

- Judge, G. G., W. E. Griffiths, R. C. Hill, H. Lutkepohl, and T. C. Lee. 1985. *The Theory and Practice of Econometrics*. 2d ed. New York, NY: John Wiley & Sons.
- Kahn, L. M., and S. A. Low, 1982. The relative effects of employed and unemployed job search. *The Review of Economics and Statistics* 64(2): 234-41.
- Kmenta, J. 1986. *Elements of Econometrics*. 2d Ed. New York, NY: Macmillan.
- Lee, L. F. 1976. Estimation of limited dependent variable models by two-stage methods. Ph. D. dissertation, University of Rochester, NY.
- _____. 1978. Unionism and wage rates: A simultaneous equation model with qualitative and limited dependent variables. *International Economic Review* 19 (June): 415-33.
- _____, and R. P. Trost. 1978. Estimation of some limited dependent variable models with application to housing demand. *Journal of Econometrics* 8 (December): 357-82.
- Long, J. E., and S. B. Caudill. 1987. The usage and benefits of paid tax return preparation. *National Tax Journal* 40(1): 35-46.
- Maddala, G. S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. Econometric Society Monographs in Quantitative Economics. New York: Cambridge University Press.
- _____, 1991. A perspective on the use of limited-dependent and qualitative variables models in accounting research. *The Accounting Review* 66 (October): 788-807.
- Roth, J. A., J. T. Scholz, and D. Y. Witte, eds. 1989. *Taxpayer Compliance, Volume 1: An Agenda for Research*. Philadelphia: University of Pennsylvania Press.
- Shehata, M. 1988. The impact of SFAS No. 2 and the Economic Recovery Tax Act of 1981 on firms' R and D spending behavior. Ph.D. dissertation, University of Florida, FL.
- _____, 1991. Self-selection bias and the economic consequences of accounting regulation: An application of two-stage switching regression to SFAS No. 2. *The Accounting Review* 66 (October): 768-87.
- Slemrod, J. 1989. The return to tax simplification: An econometric analysis. *Public Finance Quarterly* 17 (1): 3-27.
- _____, and N. Sorum. 1984. The compliance cost of the U. S. individual income tax system. *National Tax Journal* (December): 461-74.
- Smith A., 1776. *The Wealth of Nations*, Book V, 311 (London: Strahan and Prefton, 1805, Volume III, 270-71).

Swingen, J., P. Greene, and S. Long. 1992. Taxpayer compliance burden: An examination of factors affecting tax return preparation fees. Paper presented at the SWFAD Conference, San Antonio, Texas.

Topolnicki, D. M. 1992. The pros flub our third annual tax-return test. *Money* 19(3): 90-98.

White, H. 1980. A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica* 48 (May): 817-38.

TABLE 1

NUMBER AND PERCENTAGE OF INDIVIDUAL FEDERAL INCOME
TAX RETURNS WITH PREPARER SIGNATURE IN 1981 - 1990

<i>Tax Year</i>	<i>No. of Returns Filed</i>	<i>No. of Returns with Preparer Signature</i>	<i>% of Returns with Preparer Signature</i>
1981	94,157,000	39,534,000	42.0
1982	95,574,000	42,116,000	44.1
1983	95,699,000	43,258,000	45.2
1984	96,652,000	45,220,000	46.8
1985	99,704,000	46,685,000	46.8
1986	102,652,000	48,049,000	46.8
1987	103,462,000	50,991,000	49.3
1988	107,259,000	51,507,000	48.0
1989	110,129,000	52,808,000	48.0
1990	112,596,000	54,499,000	48.4

Source: Compilation from various tables in "Internal Revenue Service, *Statistics of Income Bulletin*," Fall 1983, Spring 1985, Spring 1987, Fall 1989, Fall 1990 and Summer 1992.

TABLE 2
SAMPLE SELECTION

	<u>1986</u>	<u>1987</u>
Observations in panel	10,120	21,191
Returns not in subsequent year panel ^a	(926)	(1,967)
Erroneous returns ^a	(6)	(0)
Prior year returns ^a	(102)	(251)
Not paid or self-prepared ^a	<u>(67)</u>	<u>(162)</u>
Final sample size	<u>9,019</u>	<u>18,811</u>
Self-prepared returns	4,589	9,602
Paid-prepared returns	<u>4,430</u>	<u>9,209</u>
Returns reporting fee ^a	858	1,683
Fee < \$5*preparation time (in hours) ^a	<u>(54)</u>	<u>(50)</u>
Sample size for fee regression	<u>804</u>	<u>1,633</u>

^a See Section 3.3 for descriptions.

TABLE 3
DESCRIPTIVE SAMPLE STATISTICS^a

	1986 SAMPLE				1987 SAMPLE			
	Full Sample n=9019	Self- Prepared n=4589	Paid- Prepared n=4430	Fee Reported n=804	Full Sample n=18811	Self- Prepared n=9602	Paid- Prepared n=9209	Fee Reported n=1633
TLIA	3934	2518	5402	10273	3977	2678	5332	9805
(Tax Liability)	0	0	0	0	0	0	0	0
	602200	98250	602200	602200	1580000	1580000	794100	597700
	4572	3040	6040	10288	4656	3245	6029	10019
		0		200		0		284
FEE	n/a ^b	0	n/a ^b	20	n/a ^b	0	n/a ^b	20
		0		6227		0		20730
TB	3951	2518	5435	10452	3999	2678	5377	10056
(Total Burden)	0	0	0	28	0	0	0	25
	603175	98250	603175	603175	1580000	1580000	794100	618430
	4590	3040	6076	10457	4681	3245	6083	10264
DP (Preparer)	.491	0.000	1.000	1.000	0.490	0.000	1.000	1.000
TPI (\$1000)	29.281	22.053	36.767	59.130	31.324	23.543	39.438	67.291
	0.000	0.000	0.000	2.332	0.000	0.000	0.000	1.035
	1317.550	465.650	1317.550	1292.370	7441.560	7441.560	2897.460	2897.460
	32.951	25.594	40.007	59.307	35.372	27.286	43.236	67.844
CT (Completion time)	22.809	11.249	34.784	27.863	22.427	10.780	34.571	27.234
	1.090	1.090	1.090	5.070	1.060	1.060	1.060	5.050
	376.190	336.770	376.190	376.190	368.340	359.040	368.340	368.340
PT (Preparing Time)	5.454	3.841	7.124	7.640	4.797	3.246	6.415	6.847
	0.730	0.730	0.730	3.020	0.660	0.660	0.660	2.750
	39.820	31.830	39.820	39.820	40.150	33.900	40.150	40.150
	0.343	0.212	0.478	0.617	0.325	0.183	0.473	0.657
FORM	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	6.000	4.000	6.000	6.000	7.000	5.000	7.000	7.000
PTFORM	4.689	3.429	5.994	6.299	4.073	2.903	5.293	5.483
	0.730	0.730	0.730	3.020	0.660	0.660	0.660	2.750
	34.480	27.170	34.480	34.480	34.250	31.720	34.250	34.250
	0.743	0.687	0.802	0.945	0.685	0.629	0.742	0.940
DEP	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	11.000	8.000	11.000	6.000	14.000	9.000	14.000	7.000
AGE	0.123	0.083	0.164	0.103	0.115	0.068	0.164	0.106
MS	0.495	0.398	0.594	0.716	0.500	0.377	0.589	0.724
SEM	0.141	0.068	0.217	0.172	0.138	0.065	0.215	0.165
SA	0.417	0.306	0.531	0.937	0.359	0.251	0.471	0.934
SB	0.338	0.239	0.440	0.525	0.333	0.222	0.449	0.533

(Continued)

	1986 SAMPLE				1987 SAMPLE			
	Full Sample n=9019	Self- Prepared n=4589	Paid- Prepared n=4430	Fee Reported n=804	Full Sample n=18811	Self- Prepared n=9602	Paid- Prepared n=9209	Fee Reported n=1633
SC	0.122	0.063	0.183	0.153	0.121	0.059	0.186	0.184
SD	0.128	0.065	0.193	0.300	0.152	0.079	0.229	0.353
SE	0.147	0.061	0.235	0.302	0.137	0.055	0.223	0.303
SF	0.025	0.007	0.043	0.022	0.022	0.008	0.038	0.022
SG	0.033	0.011	0.056	0.096	n/a ^c	n/a ^c	n/a ^c	n/a ^c
SW	0.272	0.218	0.328	0.483	n/a ^c	n/a ^c	n/a ^c	n/a ^c
AMT	0.007	0.001	0.012	0.024	0.034	0.004	0.065	0.130
PAL ^d (F8582)	n/a ^d	n/a ^d	n/a ^d	n/a ^d	0.017	0.009	0.026	0.039

^a In each cell the number in the first row is the mean, in the second row is the minimum, and in the third row is the maximum. For TLIA, TB and TPI, the mean values conditioned on TLIA > 0 are also reported in the fourth row.

^b Not applicable because not all returns reported tax return preparation fee.

^c Schedules G and W are no longer used after 1986.

^d Form 8582 (passive activity loss limitation) was not used before 1987.

TABLE 4
OLS ESTIMATES OF PREPARATION FEE EQUATIONS

PANEL A. ESTIMATES FOR 1986 PAID-PREPARED RETURNS:

Variable	Coefficient	Std. Error	t-ratio	Prob t >x
Constant	22.094	41.93	0.527	0.59824
SB	46.641	14.73	3.166	0.00155
SC	64.857	39.44	1.644	0.10013
SD	72.376	24.04	3.011	0.00260
SE	60.049	25.05	2.397	0.01653
SG	60.340	65.23	0.925	0.35493
SW	-67.676	27.53	-2.458	0.01397
PTFORM	4.0585	6.914	0.587	0.55721
TPI	1.7748	0.494	3.593	0.00033

Sample consists of 804 returns that
itemized fee \geq \$5*preparation time (hours).
 $R^2=0.323$, adjusted $R^2=0.316$,
 $F[8, 795]=47.34$, $p<0.01$.
Breusch-Pagan $\chi^2=4690.29$, $df=8$, $p<0.01$,
White's heteroscedastic adjustment used.

PANEL B. ESTIMATES FOR 1987 PAID-PREPARED RETURNS:

Variable	Coefficient	Std. Error	t-ratio	Prob t >x
Constant	-22.662	48.02	-0.472	0.63695
SB	32.052	53.50	0.599	0.54910
SC	174.93	93.07	1.880	0.06017
SD	25.129	64.23	0.391	0.69561
SE	66.283	67.38	0.984	0.32523
AMT	-5.7273	130.6	-0.044	0.96501
PTFORM	-11.314	8.476	-1.335	0.18194
TPI	4.4061	1.340	3.288	0.00101

Sample consists of 1633 returns that
itemized fee \geq \$5*preparation time (hours).
 $R^2=0.321$, adjusted $R^2=0.319$,
 $F[7, 1625]=110.0$, $p<0.01$.
Breusch-Pagan $\chi^2=10799.8$, $df=7$, $p<0.01$,
White's heteroscedastic adjustment used.

TABLE 5

PREPARATION FEE ESTIMATES FOR THE PAID-PREPARED RETURNS

<i>TPI</i> <i>(\$1000s)</i>	1986			1987		
	<i>Standard Deduction</i>	<i>Itemized Deduction</i>	<i>Overall</i>	<i>Standard Deduction</i>	<i>Itemized Deduction</i>	<i>Overall</i>
0-20	\$87	\$125	\$95	\$44	\$67	\$47
20-50	131	139	137	120	134	128
Over 50	344	322	324	388	489	478
Overall	\$108	\$194	\$154	\$86	\$261	\$168

TABLE 6
 PROPORTION OF FEES TO TOTAL TAX BURDEN OF PAID-PREPARED
 RETURNS

<i>TPI</i> <i>(\$1000s)</i>	1986			1987		
	<i>Fee</i>	<i>Tax</i> <i>Burden</i>	<i>Percent</i> <i>-age</i>	<i>Fee</i>	<i>Tax</i> <i>Burden</i>	<i>Percent</i> <i>-age</i>
0-20	\$95	\$767	12.4	\$47	\$614	7.7
20-50	137	3608	3.8	128	3333	3.8
Over 50	324	20534	1.6	478	19135	2.5
Overall	\$154	\$5551	2.8	\$168	\$5495	3.1

TABLE 7
ESTIMATES OF TAX BURDEN EQUATIONS, 1986

Panel A. Paid-prepared sample:

Variable	<u>ESR Estimate</u>		<u>OLS Estimate</u>	
	Coefficient	t-ratio	Coefficient	t-ratio
Constant	1348.5	1.289	-1006.6	-13.779 ^{***}
TPI	218.65	85.599 ^{***}	216.41	35.797 ^{***}
SA	-1318.3	-6.164 ^{***}	-889.16	-9.428 ^{***}
SB	-475.67	-3.227 ^{***}	-228.11	-2.407 [*]
SD	-110.88	-0.735	94.150	0.688
SE	-997.74	-3.858 ^{***}	-469.77	-4.056 ^{***}
SW	-477.43	-4.012 ^{***}	-395.09	-2.910 ^{***}
FORM	-441.43	-4.283 ^{***}	-260.71	-2.510 [*]
SEM	89.689	0.330	650.95	5.580 ^{***}
MS	-1434.6	-9.921 ^{***}	-1219.5	-10.756 ^{***}
AGE	-1100.4	-4.617 ^{***}	-647.84	-5.313 ^{***}
DEP	-251.16	-6.457 ^{***}	-248.85	-6.372 ^{***}
LAMBDA	-2076.3	-2.264 [*]		

Adjusted R²=0.791
Observations=4348
Heteroscedastic adjustment used.

Adjusted R²=0.791
Observations=4348
Breusch-Pagan $\chi^2=15173$,
White's hetero. adj. used.

Panel B. Self-prepared sample:

Variable	<u>ESR Estimate</u>		<u>OLS Estimate</u>	
	Coefficient	t-ratio	Coefficient	t-ratio
Constant	-735.33	-1.678	-864.29	-19.686 ^{***}
TPI	197.09	112.383 ^{***}	197.38	44.008 ^{***}
SA	-658.14	-3.383 ^{***}	-713.37	-8.421 ^{***}
SB	130.22	1.077	98.243	1.565
SD	-434.60	-2.791 ^{***}	-473.36	-3.083 ^{***}
SE	-429.19	-1.260	-526.57	-3.245 ^{***}
SW	-161.28	-2.287 [*]	-170.37	-1.659
FORM	-498.41	-4.262 ^{***}	-530.43	-7.305 ^{***}
SEM	442.89	1.338	347.90	2.909 ^{***}
MS	-768.20	-7.381 ^{***}	-793.79	-10.228 ^{***}
AGE	-436.83	-1.913	-500.15	-4.870 ^{***}
DEP	-311.22	-16.010 ^{***}	-310.75	-13.837 ^{***}
LAMBDA	270.09	0.295		

Adjusted R²=0.872
Observations=4581
Heteroscedastic adjustment used.

Adjusted R²=0.872
Observations=4581
Breusch-Pagan $\chi^2=16121$,
White's hetero. adj. used.

Note: * significant at the .05 level, *** significant at the .01 level.

TABLE 8
ESTIMATES OF TAX BURDEN EQUATIONS, 1987

Panel A. Paid-prepared sample:

Variable	<u>ESR Estimate</u>		<u>OLS Estimate</u>	
	Coefficient	t-ratio	Coefficient	t-ratio
Constant	844.20	1.540	-936.81	-19.367 ^{***}
TPI	195.64	126.733 ^{***}	193.53	54.688 ^{***}
SA	-1160.5	-8.948 ^{***}	-790.49	-10.250 ^{***}
SB	-153.77	-1.667	71.089	1.264
SD	-482.16	-5.469 ^{***}	-309.16	-3.902 ^{***}
SE	-785.85	-6.139 ^{***}	-437.35	-5.700 ^{***}
AMT	298.49	1.751	663.53	2.827 ^{***}
PAL	3454.5	17.810 ^{***}	3591.7	7.926 ^{***}
FORM	-265.00	-5.259 ^{***}	-186.99	-3.292 ^{***}
SEM	254.28	1.661	699.66	8.789 ^{***}
MS	-1402.9	-15.176 ^{***}	-1173.8	-16.650 ^{***}
AGE	-1483.8	-9.641 ^{***}	-1053.1	-12.875 ^{***}
DEP	-378.48	-14.810 ^{***}	-378.21	-15.245 ^{***}
LAMBDA	-1564.6	-3.269 ^{***}		

Adjusted R²=0.831
Observations=9041
Heteroscedastic adjustment used.

Adjusted R²=0.831
Observations=9041
Breusch-Pagan $\chi^2=27421$,
White's hetero. adj. used.

Panel B. Self-prepared sample:

Variable	<u>ESR Estimate</u>		<u>OLS Estimate</u>	
	Coefficient	t-ratio	Coefficient	t-ratio
Constant	-1071.6	-3.515 ^{***}	-896.82	-21.810 ^{***}
TPI	187.09	121.683 ^{***}	186.59	48.674 ^{***}
SA	-790.23	-4.944 ^{***}	-703.00	-9.270 ^{***}
SB	98.469	0.966	150.32	3.289 ^{***}
SD	-127.04	-1.043	-69.263	-0.708
SE	-692.88	-3.153 ^{***}	-574.76	-2.359 [*]
AMT	1732.5	3.634 ^{***}	1959.8	0.728
PAL	5048.4	25.578 ^{***}	5074.3	3.688 ^{***}
FORM	-479.83	-8.060 ^{***}	-455.69	-5.074 ^{***}
SEM	114.69	0.460	252.26	2.383 [*]
MS	-1058.4	-10.933 ^{***}	-1008.9	-20.067 ^{***}
AGE	-956.35	-4.502 ^{***}	-841.47	-11.310 ^{***}
DEP	-380.79	-20.775 ^{***}	-381.61	-21.253 ^{***}
LAMBDA	-367.67	-0.575		

Adjusted R²=0.822
Observations=9582
Heteroscedastic adjustment used.

Adjusted R²=0.822
Observations=9582
Breusch-Pagan $\chi^2=363639$
White's hetero. adj. used.

Note: * significant at the .05 level, *** significant at the .01 level.

TABLE 9
TAX BURDEN DIFFERENTIAL BETWEEN PREPARATION MODES

PANEL A. THE 1986 RETURNS:

TPI RANGE	ALL RETURNS		SELF-PREPARED		PAID-PREPARED		
	ESR	OLS	ESR	OLS	ESR	OLS	Fee
0 - 20000	(1802)	50	(1962)	53	(1549)	47	(95)
20000 - 50000	(1455)	(141)	(1625)	(119)	(1324)	(159)	(137)
50000 - 146500	(1678)	(891)	(1727)	(701)	(1650)	(998)	(261)
0 - 146500	(1654)	(141)	(1832)	(69)	(1466)	(215)	(141)

PANEL B. THE 1987 RETURNS:

TPI RANGE	ALL RETURNS		SELF-PREPARED		PAID-PREPARED		
	ESR	OLS	ESR	OLS	ESR	OLS	Fee
0 - 20000	(1772)	7	(1856)	5	(1633)	10	(47)
20000 - 50000	(1622)	(97)	(1724)	(96)	(1543)	(97)	(128)
50000 - 176000	(1359)	(102)	(1546)	(182)	(1262)	(60)	(339)
0 - 176000	(1659)	(46)	(1786)	(45)	(1525)	(50)	(137)

Note: Tax burden differential = $\beta_0 \bar{x} - \beta_1 \bar{x}$. If it is positive, there are tax savings from using preparers. If it is negative, then there is additional tax burden from using preparers.

TABLE 10
STRUCTURAL FORM ESTIMATES OF THE PREPARATION MODE EQUATION

PANEL A. THE 1986 SAMPLE:

Variable	Coefficient	Std. Error	t-ratio	Prob t >x
Constant	0.27644	0.6812E-01	4.058	0.00005
TBDIF	0.31274E-03	0.3068E-04	10.195	0.00000
CT	0.18667E-02	0.3173E-03	5.884	0.00000
SEM	0.39723	0.5364E-01	7.406	0.00000
MS	0.14225	0.3706E-01	3.839	0.00012
AGE	0.20291	0.5207E-01	3.897	0.00010
DEP	0.30963E-01	0.1374E-01	2.254	0.02420

Log-Likelihood..... -5734.5
 Restricted (Slopes=0) Log-L. -6186.1
 $\chi^2=903.05$, $df=6$, $p<0.01$
 Returns correctly classified as self-prepared 3322 (n=4581)
 Returns correctly classified as paid-prepared 2359 (n=4348)
 The overall classification accuracy is 63.6%

PANEL B. THE 1987 SAMPLE:

Variable	Coefficient	Std. Error	t-ratio	Prob t >x
Constant	0.88160	0.5716E-01	15.425	0.00000
TBDIF	0.68487E-03	0.2936E-04	23.326	0.00000
CT	0.19090E-02	0.2179E-03	8.760	0.00000
SEM	0.58745	0.3736E-01	15.723	0.00000
MS	0.16729	0.2365E-01	7.075	0.00000
AGE	0.14637	0.3725E-01	3.929	0.00009
DEP	0.20258E-01	0.1005E-01	2.016	0.04376

Log-Likelihood..... -11632
 Restricted (Slopes=0) Log-L. -12901
 $\chi^2=2538.1$, $df=6$, $p<0.01$
 Returns correctly classified as self-prepared 7374 (n=9582)
 Returns correctly classified as paid-prepared 4912 (n=9041)
 The overall classification accuracy is 66.0%

TABLE 11
ESTIMATES OF AN ALTERNATIVE PREPARATION MODE EQUATION

PANEL A. THE 1986 SAMPLE:

Variable	Coefficient	Std. Error	t-ratio	Prob t >x
Constant	0.54523	0.7143E-01	7.633	0.00000
TB1	-0.49950E-03	0.3361E-04	-14.861	0.00000
TB0	0.56107E-03	0.3544E-04	15.830	0.00000
CT	0.15047E-02	0.3179E-03	4.733	0.00000
SEM	0.33361	0.5419E-01	6.157	0.00000
MS	-0.16638	0.4300E-01	-3.869	0.00011
AGE	0.76500E-01	0.5315E-01	1.439	0.15008
DEP	0.60572E-01	0.1397E-01	4.337	0.00001

Log-Likelihood..... -5628.6

Restricted (Slopes=0) Log-L. -6186.1

$\chi^2 = 1114.9$, $df = 7$, $p < 0.01$

Returns correctly classified as self-prepared 3204 (n=4581)

Returns correctly classified as paid-prepared 2598 (n=4348)

The overall classification accuracy is 65.0%

PANEL B. THE 1987 SAMPLE:

Variable	Coefficient	Std. Error	t-ratio	Prob t >x
Constant	0.72765	0.6004E-01	12.119	0.00000
TB1	-0.62459E-03	0.3039E-04	-20.549	0.00000
TB0	0.64991E-03	0.2999E-04	21.672	0.00000
CT	0.18345E-02	0.2181E-03	8.412	0.00000
SEM	0.58737	0.3744E-01	15.688	0.00000
MS	0.12022	0.2429E-01	4.950	0.00000
AGE	0.21059	0.3806E-01	5.533	0.00000
DEP	0.24103E-01	0.1008E-01	2.392	0.01676

Log-Likelihood..... -11587

Restricted (Slopes=0) Log-L. -12901

$\chi^2 = 2627.5$, $df = 6$, $p < 0.01$

Returns correctly classified as self-prepared 7150 (n=9582)

Returns correctly classified as paid-prepared 5189 (n=9041)

The overall classification accuracy is 66.3%

APPENDIX A
DESCRIPTIVE SAMPLE STATISTICS
(EXCLUDING RETURNS WITH TPI IN THE UPPER ONE PERCENT RANGE)

	1986 SAMPLE ^a			1987 SAMPLE ^a		
	99% Sample n=8929	Self- Prepared n=4581	Paid- Prepared n=4348	99% Sample n=18623	Self- Prepared n=9582	Paid- Prepared n=9041
TLIA	3084	2467	3734	3022	2317	3770
(Tax	0	0	0	0	0	0
Liability)	53220	34380	53220	123800	123800	56840
	3589	2980	4184	3544	2810	4308
		0			0	
FEE	n/a ^b	0	n/a ^b	n/a ^b	0	n/a ^b
		0			0	
TB	3097	2467	3761	3038	2317	3802
(Total	0	0	0	0	0	0
Burden)	53220	34380	53220	123800	123800	56840
	3603	2980	4212	3562	2810	4308
DP	.487	0.000	1.000	0.485	0.000	1.000
(Preparer)						
TPI	26.122	21.701	30.780	27.005	21.872	32.444
(\$1000)	0.000	0.000	0.000	0.000	0.000	0.000
	146.400	136.900	146.400	175.800	171.500	175.800
	29.399	25.293	33.413	30.365	25.294	35.389
CT	22.246	11.238	33.844	22.060	10.690	34.111
(Completion	1.090	1.090	1.090	1.060	1.060	1.060
time)	355.400	336.770	355.400	362.400	359.000	362.400
PT	5.376	3.835	7.000	4.725	3.231	6.308
(Preparing	0.730	0.730	0.730	0.660	0.660	0.660
Time)	34.440	31.830	34.440	36.000	31.720	36.000
	0.333	0.211	0.460	0.307	0.181	0.442
FORM	0.000	0.000	0.000	0.000	0.000	0.000
	6.000	4.000	6.000	6.000	5.000	6.000
FFFORM	4.636	3.426	5.911	4.028	2.895	5.230
	0.730	0.730	0.730	0.660	0.660	0.660
	30.430	27.170	30.430	32.580	31.720	32.580
	0.740	0.686	0.796	0.681	0.628	0.738
DEP	0.000	0.000	0.000	0.000	0.000	0.000
	11.000	8.000	11.000	14.000	9.000	14.000
AGE	0.123	0.083	0.165	0.115	0.068	0.166
MS	0.491	0.398	0.589	0.477	0.376	0.584
SEM	0.139	0.068	0.214	0.137	0.064	0.213
SA	0.411	0.305	0.523	0.353	0.250	0.462
SB	0.332	0.239	0.431	0.327	0.220	0.440
SC	0.120	0.063	0.180	0.119	0.058	0.184
SD	0.121	0.064	0.181	0.146	0.077	0.219

(Continued)

	1986 SAMPLE ^a			1987 SAMPLE ^a		
	99% Sample n=8929	Self- Prepared n=4581	Paid- Prepared n=4348	99% Sample n=18623	Self- Prepared n=9582	Paid- Prepared n=9041
SE	0.139	0.060	0.222	0.130	0.053	0.211
SF	0.024	0.007	0.043	0.022	0.008	0.037
SG	0.030	0.010	0.051	n/a ^c	n/a ^c	n/a ^c
SW	0.271	0.218	0.327	n/a ^c	n/a ^c	n/a ^c
AMT	0.005	0.001	0.009	0.028	0.004	0.054
PAL ^d (F8582)	n/a ^d	n/a ^d	n/a ^d	0.014	0.008	0.021

- ^a In each cell the number in the first row is the mean, in the second row is the minimum, and in the third row is the maximum. For TLIA, TB and TPI, the mean values conditioned on TLIA > 0 are also reported in the fourth row.
- ^b Not applicable because not all returns reported tax return preparation fee.
- ^c Schedules G and W are no longer used after 1986.
- ^d Form 8582 (passive activity loss limitation) was not used before 1987.

APPENDIX B

REDUCED FORM ESTIMATES OF THE PREPARATION MODE EQUATION,
1986

<i>Variable</i>	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>Prob t >x</i>
Constant	-0.54890	0.2442E-01	-22.475	0.00000
CT	0.20226E-03	0.3446E-03	0.587	0.55728
SEM	0.49290	0.5357E-01	9.201	0.00000
MS	0.16437	0.3972E-01	4.138	0.00004
AGE	0.37332	0.4958E-01	7.530	0.00000
DEP	-0.22905E-03	0.1396E-01	-0.016	0.98691
TPI	-0.17830E-02	0.9443E-03	-1.888	0.05900
FORM	0.17443	0.2864E-01	6.091	0.00000
SA	0.33942	0.3546E-01	9.572	0.00000
SB	0.19655	0.3623E-01	5.426	0.00000
SD	0.20508	0.4970E-01	4.126	0.00004
SE	0.50313	0.4681E-01	10.747	0.00000
SW	0.60142E-01	0.4285E-01	1.404	0.16045

Log-Likelihood..... -5524.7

Restricted (Slopes=0) Log-L. -6186.1

$\chi^2=1322.8$, $df=12$, $p<0.01$

Returns correctly classified as self-prepared 3403 (n=4581)

Returns correctly classified as paid-prepared 2424 (n=4348)

The overall classification accuracy is 65.3%

APPENDIX C

REDUCED FORM ESTIMATES OF THE PREPARATION MODE EQUATION,
1987

Variable	Coefficient	Std. Error	t-ratio	Prob t >x
Constant	-0.55716	0.1661E-01	-33.546	0.00000
CT	0.64721E-03	0.2382E-03	2.717	0.00659
SEM	0.48874	0.3787E-01	12.906	0.00000
MS	0.23288	0.2433E-01	9.571	0.00000
AGE	0.48672	0.3506E-01	13.884	0.00000
DEP	-0.59568E-03	0.1035E-01	-0.058	0.95410
TPI	-0.21879E-02	0.6457E-03	-3.388	0.00070
FORM	0.82366E-01	0.2251E-01	3.659	0.00025
SA	0.39167	0.2622E-01	14.938	0.00000
SB	0.23443	0.2524E-01	9.288	0.00000
SD	0.22483	0.3237E-01	6.947	0.00000
SE	0.44644	0.3407E-01	13.104	0.00000
AMT	0.74819	0.8893E-01	8.413	0.00000
PAL	0.10801	0.9141E-01	1.182	0.23736

Log-Likelihood..... -11327.

Restricted (Slopes=0) Log-L. -12901.

$\chi^2=3146.4$, df=13, p<0.01

Returns correctly classified as self-prepared 7144 (n=9582)

Returns correctly classified as paid-prepared 5278 (n=9041)

The overall classification accuracy is 66.7%
