NBER WORKING PAPER SERIES

A STRUCTURAL EQUATION MODEL FOR TAX COMPLIANCE AND AUDITING

Kurt Beron

Helen V. Tauchen

Ann Dryden Witte

Working Paper No. 2556

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 1988

The work reported in this paper was made possible by a grant from the National Science Foundation and the financial support of the University of North Carolina's Research Council. Their support of our work in no way indicated support for our approach or conclusions. We thank Dennis Cox and Bill Lefbom of the U.S. Internal Revenue Service for helping us obtain the data. The research reported here is part of the NBER's research program in Taxation. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research. Support from The Lynde and Harry Bradley Foundation is gratefully acknowledged.

NBER Working Paper #2556
April 1988

A Structural Equation Model for Tax Compliance and Auditing

ABSTRACT

In this paper, we estimate a three equation model for taxpayers' reported income and tax liability and for the probability of an audit. Our work differs from previous studies in that our dependent variables in the compliance equations are taxpayer reports rather than a variable related to auditor estimates of noncompliance and in that we estimate a structural equation for audits.

We find that audits stimulate compliance although the effect is not large and is not statistically significant for all groups. Audits are more effective at inducing accurate reporting of subtractions from income than of income. Reduced-form results suggest that IRS activities other than audits have significant compliance effects.

Results for the sociodemographic variables are interesting and help to explain some seemingly incongruous findings in the literature. We find compliance to be higher, if anything, in areas with less educated and older taxpayers, a large proportion of households headed by females, and a mostly native born population.

Kurt Beron
University of Texas
Dallas, TX

Helen V. Tauchen
University of North Carolina
Chapel Hill, NC

Ann Dryden Witte
Wellesley College and the NBER
Wellesley, MA

1. Introduction

The study of tax compliance promises important scientific and policy insights. Form a scientific point of view, the reporting of tax liabilities provides an arena in which we can observe the compliance behavior of most of the adult population. This contrasts markedly with economists' study of most criminal behavior (e.g., murder, property offenses) where a small, and possible aberrant, portion of the population is at risk. At this time of large budget deficits, understanding compliance behavior also offers the possibility of "painlessly" reducing the deficit.

The potential richness of the tax arena has not escaped the research community and studies of tax compliance behavior have burgeoned during the 1980s. However, progress to date has been somewhat disappointing. Researchers report seemingly contradictory results and research on tax compliance has not advanced beyond research on the deterrence of general criminal offenses. We believe that it is time to step back, assess current research findings, and try other approaches. A critical approach to tax compliance research seems particularly valuable at this time because of the extensive research that is underway on this topic.

In this paper, we provide an analysis of a 1969 data set that combines information from the IRS's Statistics of Income (SOI) program, the Census, internal IRS documents and a special data set compiled by the IRS in the mid 1970s. This last data set has been used by a number of researchers in recent empirical studies of tax compliance.

In carrying out the analysis, it is important to take into account a number of issues that are central to compliance research. First, consider the dependent variable used in previous empirical studies, namely an estimate of

compliance based on Internal Revenue Service (IRS) auditors' findings. Given the difficulty of IRS auditors in detecting unreported in income, the use of such a dependent variable provides results that are very difficult to interpret. Instead we analyze the actual reporting decisions of taxpayers. As is described in detail in Section 4, the use of these types of dependent variables offers a number of advantages. Primarily, it makes estimation of compliance models less complex and interpretation of results more straightforward. Second, we specify a structural model that allows us to infer the factors affecting taxpayers' income reports and reports of subtractions (e.g., deductions and exemptions) on the tax return. Other compliance research considers only a single compliance decision.[1] We know, however, that the probability of detection is much higher for subtractions than income reports and we suspect that taxpayer behavior differs for these two quite different types of compliance activities. Noncompliance from underreports of income is often an act of omission while noncompliance from overstatement of subtractions requires actual misstatements. Third, the work reported here represents the first time, as far as we are aware, that empirical research has specified and estimated a structural equation for audits. Yet, theory suggests that the proper specification of compliance equations requires careful consideration of the audit process.

To summarize our results briefly, we find that taxpayers significantly underreport adjusted gross income (AGI) and that increasing the probability of audit increases reported income and tax liability. We also find that audits are more effective in inducing accurate reporting of subtractions from income than of income. In addition, there is evidence that detectability of noncompliance through audits is a determinant of taxpayers' reporting behavior.

The results that audits affect compliance are rather comforting and support other findings. However, both our results and the results of previous studies vary considerably across taxpayer groups and are not always significant. Previous studies using different models all report that, in general, audits deter tax noncompliance but find significant deterrent effects of audits for different taxpayer groups. To us this is disturbing and suggests that economists' understanding of the effects of audits on compliance is not very firm. In the conclusions we suggest reasons why this might be the case.

Comparison of our structural and reduced form results suggests that IRS enforcement actions other than audits have significant effects on compliance. To date research, including our own, has focused on the effect of audits and largely ignored other IRS actions. It appears that enhancing our understanding of the effect of tax administration on compliance will require economists to broaden their perspectives.

The results for sociodemographic variables are also of interest in that they help to explain seemingly incongruous results in the literature and to provide important guidance for future research. We find in general that compliance is higher in areas with less educated and older populations and in which a large percent of the population is native born or in female-headed households. The estimated coefficients on these variables in ours and others' work depend critically though on the dependent variable used and the vector of explanatory variables included in the model. Consistent implications emerge, however, when the coefficients on these variables are interpreted carefully on the basis of the model specified.

The paper is organized as follows. In the next section, we present our conceptual framework. The two sections that follow describe the data and our

empirical model. Section 5 and 6 contain a discussion of the empirical results for our structural and the reduced form equations respectively. The final section contains our conclusions.

2. Conceptual Framework

Our model of tax compliance and enforcement is a simultaneous equation model with equations for taxpayers' reports and for the probability of an audit. Specifically, to reflect the reporting behavior of taxpayers we include equations for reported adjusted gross income (AGI) and the reported total tax liability. We use separate equations for reported AGI and tax liability rather than a single reporting equation because of the differences in the magnitude and possible causes of income underreporting and the overstatement of subtractions (i.e.,adjustments, exemptions, deductions and credits). The results for the reported AGI equation allow us to determine the factors affecting income reporting while comparison of the results for the reported AGI and tax liability equations allows us to infer the factors associated with the overstatement of subtractions.[2]

To specify the equations for reported AGI and total tax liability, we surveyed the theoretical literature on tax compliance.[3] This literature suggests that taxpayer's reports of income and tax liability depend on the enforcement policies of the tax authorities, income, tax rates, and tastes and preferences. Although not explicitly mentioned in the literature, the theoretical models of tax compliance also imply that taxpayers' reports of income and tax liability depend on provisions of the tax code regarding income exclusions and subtractions.

We chose audits as the measure of enforcement policy since this is the enforcement action that has been of most interest in the literature on tax compliance and since the IRS believes that audits are its most effective instrument for stimulating accurate taxpayer reports (Controller General of the United States, 1976, p.1). To specify the audit equation, we read the available documents regarding the way in which the Internal Revenue Service (IRS) selects returns for audit[4] and discussed the selection process with IRS personnel at the national and district level. The IRS selects the largest proportion of returns for audit using carefully developed computerized formulas, known as the DIF formulas. The parameters of the DIF formulas are estimated using data from the IRS's Tax Compliance Measurement Program (TCMP). Under this program, the IRS every three or four years selects approximately 50,000 returns to be examined by its best auditors. Taxpayers are asked to substantiate every line item on their return, and auditors record both what the taxpayer filed and what they believe to be the correct report. These data are used to estimate equations that relate the extent of underpayment to the characteristics of the filed return. In subsequent years, the IRS uses the parameters estimates from these equations to score each return filed in regard to its audit potential. The returns with the highest scores are targeted for audit. The actual number of returns audited through the DIF program depends on the reporting behavior of taxpayers, the type and number of audit personnel available, and the amount of audit resources involved in special audit programs.[5] To reflect this audit selection process, we model the probability of an audit as determined by taxpayers' reports which are endogenous to the model, by taxpayer characteristics that the IRS both observes and may legally use for audit selection, and by the level of IRS resources.

3. The Data

We obtain our data on reported AGI and total tax liability and on the number of returns filed from the Statistics of Income (SOI) files and our information on audits from the IRS's Project 778 data base. Our measures of income and tastes and preferences come from the 1970 Census of Population and Housing.[6] Information on IRS audit resources was obtained from an IRS internal document (IRS, 1969).

All data except that on IRS work load (returns per full-time-equivalent employee) are aggregated to the three digit zip code level since data on audits are not available at the individual level. Information on IRS work load is aggregated to the IRS District Office level since this is the administrative unit responsible for conducting audits. IRS Districts are coterminous with state boundaries except in the most populous states (e.g., California) where multiple Districts are established. Tax return data are for 1969 returns filed in 1970, and audit data relate to audits performed in 1969.

Information on audits is available for seven distinct groups defined on the basis of AGI and income source (i.e., whether or not some income is received from a business or farm.)[7] The IRS separates the returns into these audit classes because the it believes that compliance behavior varies across groups and develops separate audit selection formulas for each group.

Both the IRS and academic researchers (Witte and Woodbury, 1985, and Dubin and Wilde, 1988) have tested for the homogeneity of taxpayer behavior across the audit classes and concluded that aggregation is not appropriate. We therefore estimate separate models for the different audit classes. Since we are able to obtain income estimates for only the five low (incomes below $10,000 in 1969 dollars) and middle (incomes between $10,000 and $30,000 for

proprietors and between $10,000 and $50,000 for wage and salary workers in 1969 dollars) income audit classes, we present results for only these audit classes. The definitions of the audit classes are given in Table 1. There are no good estimates of income for the two high income groups from the Census or other sources.[8] These high income returns account for less than one percent of the returns filed in 1969. Results for a specification that omits the income variable for these two audit classes are available on request.

4. The Empirical Model

Our model is a three equation model for reported AGI (AGI^r), reported total tax liability (T^r), and the log odds of an audit (A). We discuss our specification for the two reporting equations first and then turn our attention to the specification of the audit equation.

We see reported AGI and report total tax liability as determined by the log odds of an audit (our measure of enforcement actions which is endogenous),[9] the degree to which noncompliance can be detected (D) in an audit (our measure of the penalties), the level of income (I), a vector of variables that reflect important aspects of the tax code regarding filing requirements, income exclusions and subtractions from income (TC), and a vector of sociodemographic variables that control for differences in tastes and preferences (S^r). Formally, our two equations for taxpayer reports are

$$AGI^r = a_0 + a_1 A + a_2 D + a_3 I + a_4 TC + a_5 S^r + e_g$$

and

$$T^r = b_0 + b_1 A + b_2 D + b_3 I + b_4 TC + b_5 S^r + e_t$$

where the e's are random error terms and the a's and b's are the parameters to be estimated.

Our measure of the log odds of an audit is based on the audit coverage in the geographical area. We do not measure penalties directly but rather seek to measure them indirectly through our measures of detectability. The vast majority of audits result in some change, generally an increase, in tax liability.[10] However, the taxpayer generally receives no formal penalty but rather is assessed the taxes due plus interest for the period of nonpayment.[11] Thus, for most taxpayers it is what the auditor uncovers that is the important factor in determining the costs of noncompliance.[12]

We seek to control for the degree of detectability in two ways. First, we estimate separate equations for reported AGI and tax liability. As the IRS readily acknowledges (e.g., US Department of the Treasury, 1983), it is far more able to uncover overstatements of subtractions from income than understatements of income. Thus, we would expect audits to have a stronger effect on reported tax liability where overstatements of subtractions as well as underreporting of income are reflected than on reported AGI. Second, we include two variables related to information reporting to the IRS and other types of record keeping. In 1969, the major paper trails used by the IRS were the W-2 forms.[13] To reflect the relative availability W-2s we include the percent of employment in manufacturing and in the service industries. Income from manufacturing firms is generally covered by W-2 reporting and any underreporting of income is likely to be detected by an audit. In contrast, income earned in the service industries is more likely to be in a form (e.g.,

cash, commissions) less subject to detection. Even the proprietor audit classes in areas where much employment is in manufacturing will be likely to have "paper trails" for receipts because of the extensive record keeping activities of large manufacturing enterprises and because a significant portion of the income for individuals in the proprietor audit classes is from wages and salary.

Our measure of income is obtained from the Census and varies according to the income category included in the audit class. We chose Census income as our measure of "true" income rather than an estimate of income from TCMP audit results because TCMP auditors fail to uncover a substantial fraction of unreported income.[14] We have not explicitly introduced a measure of the marginal tax rate as an explanatory variable because for our data set the tax rate has no variation that is independent of income.[15] The coefficient on our income variable should therefore be interpreted as measuring the effect of changes in income on compliance when the tax rates change with income in accordance with the tax code.

To reflect relevant features of the tax code, we include variables for the percent of the population over 65, average family size, the percent of the population owning their own home, and the percent unemployed. These variables control for legitimate nonfilings by very low income individuals (i.e., in 1969, generally those with incomes below $600), the special allowances for senior citizens, the exemptions for dependents, the tax deductability of mortgage interest and property taxes, and the exclusion of unemployment benefits from income. Some of these variables may also be related to tastes and preference regarding tax compliance. For instance, there has been some speculation that in areas with high unemployment there might be more resentment

of the government and less willingness to pay the mandated taxes. The two variables measuring employment structure (the percent of the population employed in manufacturing and in services) may serve to control for the most important adjustment to income, the adjustment for employee business expenses, as well as detectability.

The theoretical work on tax compliance provides relatively little direction in selecting the variables that should be included to reflect tastes and preferences. The limited empirical work on tax compliance and the survey research do, however, provide some valuable guidance. See Witte and Woodbury (1983b), Kinsey (1984a), and Jackson and Milliron (1986) for surveys. The variables that are consistently related to tax compliance behavior include sex, race, education, age, and place of birth. We include measures of each in our model.

Our specification of the taxpayer compliance equations differs from the literature in that our dependent variables are reported AGI and tax liability rather than measures derived from tax auditor estimates of noncompliance. We believe that using variables related to the taxpayers' reports offer several important advantages for estimating the compliance model and interpreting the results. In addition, the use of reported AGI and tax liability as the equations for the taxpayers' behavior is implied by much of the theoretical literature on tax compliance and recommended by the National Academy of Sciences' Panel on Research on Taxpayer Compliance (Scholz, Roth and Witte, 1988).

With reported AGI and total tax liability as the dependent variables in the equations for taxpayers' behavior it is reasonable to assume a normal distribution for the error terms in the equations and to use standard

techniques for estimating the equations. Further, the error term can be interpreted as arising from true stochasticness in taxpayers' behavior or from omitted variables related to that behavior (e.g., knowledge of the tax code). When the dependent variable is affected by tax auditors' estimate of unreported income, these assumptions and interpretations are not valid. In this case, the error term reflects any elements of the auditors' and taxpayers' behaviors that are not captured by the model. Further, the error term has some unusual features that make it difficult to handle statistically, particularly in a simultaneous equation setting. First, the error term is not continuously distributed since taxpayers' reports and auditors' findings often coincide. For example, with unreported income as the dependent variable there may be an atom at zero. Second, the density of the error term may be distributed asymmetrically about the atom since underreports of taxable income are far more common than over reports. While there has recently been some very good progress on methods to handle these difficulties (Schmidt, 1986), the issue is not resolved.

The advantages of using reported income as a dependent variable, rather than a variable such as unreported income which is determined in part by tax auditors' findings, are, perhaps, even more important for the interpretation of the estimated coefficients. With reported income as the dependent variable, the effects of explanatory variables can arise either from taxpayers' reporting behavior or from specific provisions of the tax code. With a dependent variable that is influenced by tax auditors' findings, the effects of explanatory variables might also reflect the auditors' behavior. Sorting out even the separate effects of the tax code and compliance behavior may not be easy, but at least the tax code is known and thus one should be able

to make some reasonable interpretation of the coefficients for the reported income equations. By way of contrast, very little is known about tax auditors' behavior other than that they miss a substantial amount of underreporting and that they are far better at catching overstatements of deductions and exemptions than at uncovering underreports of income (U.S. Department of the Treasury, 1983). This lack of information compounds considerably the task of interpreting the coefficients in an equation for unreported income (or for percentage of noncompliance) and is one of the primary reasons that we use taxpayer reports as the dependent variables in the compliance equations.

The third equation in our model is the equation for the log odds of an audit. We see the log odds of an audit as determined by reported AGI and reported total tax liability which are endogenous to the model, by a vector of variables that are observed by the IRS and that may be legally used to select returns for audit (S^a) and by the level of IRS resources (R). Formally,

$$A = c_0 + c_1 AGI^r + c_2 Tr + c_3 S^a + c_4 R + e^a$$

where e^a is the random error and the c's are the parameters to be estimated.

We choose the log odds of an audit as the dependent variable for the audit equation for two reasons. First, this functional form ensures that the probability of an audit is between zero and one for all values of the explanatory variables. Second, the specification nests the random audit strategy, assumed in much theoretical work, as a special case.[16] A random audit strategy would be implied if only measures of IRS resources were significantly related to the log odds of an audit.

The variables that are available from the filed tax return and that the IRS may potentially use for audit selection are the measures of employment structure, the unemployment rate, the percent of the population over 65,

average family size and the percent of the population who own their own homes. Our measure of IRS resources is the number of returns per full-time-equivalent district employees.[17]

Although all federal tax returns in any audit class are scored by the same DIF formula for audit potential, the chances of a tax return being audited, ceteris paribus, differ considerably across IRS districts. The reason for the differences in rules across districts is related to resource availability. Based on the DIF scores, the IRS develops a master plan annually for the number of return examinations to be conducted by audit class in each district. However, the IRS has not been able to locate its staff in a fashion to carry out its plan. In some regions it has been able to examine substantially fewer returns than indicated by the optimal base plan whereas in other regions more audits were conducted. Since much of the reason for the differences in the probability that tax returns are selected for audit, ceteris paribus, is related to resource constraints, we use a measure of IRS audit resources to control for differences in the audit rules across taxpayers.

Note that it is the variation in the completion of the optimal plan across districts that is the important in this respect rather than just the difference in resources across regions. A difference in resources relative to the number of returns is not by itself necessarily evidence that the audit selection rule varies across taxpayers. If the IRS targets its resources towards areas that are believed to be particularly noncompliant, then higher than average resources might just reflect higher concentration of noncompliant taxpayers. There is considerable evidence in government reports (GAO, 1976 and Wilt, 1986) that the IRS has not been able to distribute its resources among districts in away to achieve best its revenue and compliance goals given its

budget. In addition, Long (1985) concludes that the "introduction of TCMP compliance data did not bring about any dramatic restructuring in audit coverage--even when it disclosed regions or return classes with much lower compliance levels which were receiving less audit attention than more compliant groups (p. 29)."

We achieve identification of our model by use of exclusion restrictions which are justified on the basis of institutional, legal and informational constraints. Specifically, we identify the two reporting equations by excluding our measure of IRS resources from these equations. The level of IRS resources affects reporting behavior only indirectly through its effect on the likelihood that a taxpayer is audited. Identification of the audit equation is obtained by the exclusion of economic and socio-demographic variables that are not available to the IRS on the tax return or information documents (e.g., whether or not an individual was foreign born) or, if available, cannot legally be used to select returns for audit (e.g, sex).

We estimate the model by two stage least squares. We use a single equation method rather than a systems method of estimation (e.g., three stage least squares) because we believe that some of our equations (e.g., the equation for reported AGI) are more completely specified than others (e.g, the audit equation) and we are concerned about spreading any omitted variable bias across equations. (See Intriligator (1978) for a discussion.)

5. Empirical Results for the Structural Model

Tables 2 thru 4 contain the results of the two-stage least squares estimation for the reported AGI and tax liability and for the log odds of an audit equations respectively. Note first that the model explains a significant

proportion of the variation in reported AGI and tax liability and in the log odds of an audit for each of the five audit classes. As we had anticipated, we model reporting behavior more successfully than auditing behavior. For reported AGI, the R^2s range from .14 for low income wage and salary workers with itemized deductions to .79 for low income proprietors while for tax liability the range of the R^2s is from .13 for middle income wage and salary workers to .79 for low income proprietors. Our model explains between 6 (for low income wage and salary workers with itemized deductions) and 31 percent (for low income proprietors) of the variation in the log odds of an audit.

As might be expected, the coefficients on income are by far the most significant of the estimated coefficients in both the reported AGI and tax liability equations. The coefficients on income in the reported AGI equations range from .27 for low income taxpayers taking standard deductions to .97 for middle income wage and salary workers. Taxpayers in all audit classes reported significantly less than an additional dollar of AGI for each extra dollar of income.

There are legal exclusions of certain types of income from taxation and legal adjustments to income that account for some differences between income and reported AGI. Further, in 1969, some individuals (i.e.,those with total income below $600 unless they had net earnings from self employment above $400) in the low income groups (i.e., AGIs below $10,000) were not required to file. However, recall that we include variables to control for the major income exclusions (the percent over 65 to control for the exclusion of Social Security benefits and the unemployment rate to control for the exclusion of unemployment compensation), adjustments (the two employment structure variables to control for employee business expenses, the largest adjustment to income in 1969), and

some of the major causes of extremely low income (the unemployment rate and the percent of the population over 65). We do not control for some exclusions (e.g., income on municipal bonds), adjustments (e.g., contributions to retirement plans by the self employed), and causes of incomes below the filing level (e.g., dependent teenagers). However, it seems unlikely that such exclusions, adjustments (total adjustments were less than .1 percent of AGI in 1969) and legitimate nonfilings could explain the low rate at which income shows up in reported AGI.

The coefficients on income in the tax liability equations provide an estimate of the marginal tax rate on total income (income before exclusions and adjustments by the tax code). The marginal tax rates implied by these coefficients do not seem unreasonable. For the low income groups the estimated marginal rate is between 7 and 10 percent. As expected, middle income groups had higher marginal rates (22 percent for middle income proprietors and 27 percent for middle income wage and salary workers).

For all audit classes, we find that increasing the odds of an audit increases reported AGI and tax liability. However, for reported AGI the effect is significant for only two groups and for reported tax liability it is significant for three groups. The magnitudes of the effects of audits on tax compliance are modest with the elasticities for reported tax liability with respect to the audit variable being between .19 and .31 when significant. These elasticities imply that a one percent increase in audit coverage would lead to approximately a $1.4 billion increase in tax revenue in current dollars for the five audit classes considered. This is an increase in 1969 tax revenue from these groups of approximately .69 percent.

High audit rates appear to be more effective in eliciting accurate reporting of subtractions from income than accurate reporting of income. The elasticity of reported AGI with respect to the audit variable is smaller than the elasticity for reported tax liability for all five audit classes and the difference inelasticities is more pronounced for the middle than the low income audit classes. These findings support IRS's belief that its auditors are better able to find overreported subtractions than underreported income. The findings are also consistent with survey findings (Westat, 1980) that relatively more of the noncompliance for blue-collar communities is from underreported income whereas for white-collar communities relatively more of the noncompliance is from overstated subtractions from income.

Recall that we have no direct measure of penalties, but rather estimate separate equations for reported AGI and tax liability and include two measures of detectability (the percent of employment in manufacturing and in services) as admittedly crude proxies for penalties. Our results provide some limited support for greater detectability being associated with compliance. Specifically, the finding that reported tax liability is more responsive to audit rates than is reported AGI supports the hypothesis that the effectiveness of high audit rates in eliciting honest reporting on the tax return is related to the detectability of noncompliance. It appears that much of the effect of audits operates through subtractions from income which must be substantiated by the taxpayer rather than through income reporting.

Also, we find higher level of employment in manufacturing (where most income may be reported to the IRS by the employer) to be associated with higher levels of reported tax liabilities for all audit classes and higher levels of employment in services[18] (where a substantial portion of all income may not be

covered by W-2 or other reports) generally to be associated with lower levels of reported tax liability. The results are significant, however, for only two of the low income groups.

Results for sociodemographic variables are interesting and help to explain some seemingly incongruous findings that appear in the literature. As in other studies that include measures of both income and education, we find that, if anything, taxpayers in areas with a more educated population report lower AGI and pay less in taxes. In their summary of the factors related to tax compliance, Jackson and Milliron (1986) suggest that the more educated may be less compliant because they better understand the opportunities for evasion and are more willing to play the audit lottery than are the less educated.

Studies that omit income from the specification (e.g., Dubin and Wilde, 1988) find that the educated are, if anything, more compliant with the tax laws. Indeed, if we omit income from our specification, the coefficients on the education variable becomes positive. Income and education are, however, highly correlated. With income omitted from the specification, the education variable serves partially as an income proxy and the estimated coefficient on education must be interpreted accordingly.

The percent of the population that is nonwhite is not significantly related to income reporting behavior for any of the groups we consider although reported tax liability is significantly lower for nonwhites than for whites in two of the low income audit classes. The findings for race in other studies are mixed. The results obtained depend to some extent on the dependent variable used and on whether or not income is included as an explanatory variable. When income is excluded from the specification and an estimate of the compliance rate that is indirectly related to auditors' findings is used as

the dependent variable (Witte and Woodbury, 1985, Dubin and Wilde, 1988), the coefficient on the variable for the percent of the population that is nonwhite is significant for most audit classes and when significant generally, but not always, indicates that whites are more compliant than other racial groups. In contrast, our results which control for income and use measures of taxpayers' reports as dependent variables show no evidence that whites are significantly more accurate in their income reports and only weak evidence that whites are significantly more accurate in their reports of subtractions from income. Clearly, economists' understanding of the relationship between race and compliance is not very strong. This mirrors the results for this variable in the general crime literature. See Schmidt and Witte (1984), Wilson and Herrnstein (1985) or Blumstein, et al. (1986) for a discussion.

The results for the percent foreign born and the percent of households female headed generally support previous work. Except for low-income wage and salary workers, we find, as do previous studies, that areas with a relatively large percent of families headed by females and a relatively small percent of the population foreign born tend to report higher AGI and tax liability. The better compliance for areas with a larger percentage of families headed by women appears to result predominantly from higher income reports.

The variables related to the taxpayers' ages are the percentage of the population older than 65 and the average age of the population between 18 and 65. The coefficients on these variables may reflect any age or cohort effects related to compliance. In addition, the variable for the percentage of the population over 65 is used to control for the exclusion of Social Security benefits from taxable income and the extra exemption allowed individuals over 65. The coefficients on this variable will be related to both age and tax code

effects and would be negative if being over 65 affected reporting only through the special provisions of the tax code for older taxpayers.

As in other studies using information related to 1969 returns (Clotfelter, 1983, Witte and Woodbury, 1985, and Dubin and Wilde, 1988), our results provide some, but not overwhelming, support for the common perception that older taxpayers are more compliant than younger ones (see Jackson and Milliron, 1986). For two audit classes (low and middle income proprietors) we find a significant and positive relationship between reported tax liability and average age. For middle income proprietors we also find significantly higher reported income and tax liability in areas with a relatively high percent of the population over 65. For low income proprietors the relationship is reversed. We believe that the finding for the low income audit class is a result of the regressive structure of the tax code benefits afforded individuals over 65. At low income levels the tax code effects outweigh any additional compliance behavior of older individuals whereas at higher income levels the compliance effect is predominant.

The remaining three variables were included to control for exclusions, adjustment and other subtractions from income allowed by the tax code. When significant the results for average family size and percent of housing owner occupied are as anticipated: negative and significant in the tax liability equation reflecting the exemptions for children and the deductions for mortgage interest payments and property taxes.

The findings for the unemployment variable do not seem to reflect the exclusions of unemployment compensation from income reporting. Indeed, the coefficients on the unemployment rate are generally positive in the reported AGI equation and significantly positive for two of the low income groups. It

may be that income reporting increases with the unemployment rate because of increased filings associated with attempts to secure tax refunds.

The coefficients on the unemployment rate are consistently negative in the tax liability equation, but are only significant for low income proprietors. Our results suggest that unemployment has only a very weak association with compliance and any tendency for areas with high rates of unemployment to have lower levels of compliance stems from overstatements of subtractions not understatements of income. Thus, it does not appear that the effect of unemployment can be explained by operation in the underground economy. The effects may be due to antigovernment attitudes or to a simple attempt to overcome financial stress by obtaining a bigger tax refund. However, the effects are not strong.

Studies that use an estimated compliance rate based on auditor findings as the dependent variable for compliance equations (Witte and Woodbury, 1985 and Dubin and Wilde, 1988) generally find increased unemployment to be associated with decreased compliance. We suspect that these results may reflect in large part the auditors' ability to detect overstated deductions.

As noted earlier, our results for the audit equation although significant are not very strong. See Table 4. We find the probability of an audit to be significantly associated with information available on taxpayers' returns for all audit classes. That is, we find no support for the IRS pursuing a random audit strategy for any audit class. For four of the five audit classes we consider, auditing is significantly associated with the endogenous taxpayers' reports (i.e., by reported AGI and total tax liability). For the remaining audit class (middle income wage and salary workers) the probability of audit is not significantly affected by the endogenous taxpayer

reports but is significantly affected by variables that are related to information available on taxpayers' returns.[19] Specifically for middle income wage and salary workers we find that the probability of audit is higher in areas where a larger percentage of individuals is employed in services and other nonmanufacturing jobs and where there is a higher percentage of the population over 65. We suspect that these variables are proxying specific aspects of taxpayers' reports that we do not observe although they might be mirroring IRS targeting of specific professions or income sources.

6. Reduced Form Results

Economic studies of taxpayer compliance have focused on audits as the enforcement activity of the IRS. This is perhaps as natural choice given the theoretical and empirical models for participation in other illegal activities and the IRS's emphasis on the importance of audits. As reported in the previous section, however, both our work and the previous literature find relatively weak deterrent effects from audits. Admittedly, this empirical work is based on aggregate data and given the relatively low audit coverage this is a potential problem. In addition, the emphasis on audits may be misplaced. The IRS undertakes many other enforcement and service activities such as matching the income claims on the return against 1099 and W-2 documents, checking math errors, providing information and education programs, pursuing delinquent accounts, instituting detailed investigations and recommending civil and criminal cases against tax evaders. The reduced form estimates for reported AGI and tax liability reveal that these other activities may also have significant compliance effects.

While estimates of the structural equations suggest that compliance behavior varies markedly across taxpayer groups and that audits have limited deterrence effects, the reduced from equations reveal a consistent pattern of compliance behavior and show significant effects of IRS activities. Consider the reduced form results for reported tax liability reported in Table 5. These results indicate that increases in IRS work load (returns per full-time -equivalent employee) consistently and significantly decrease reported tax liability.[20] This result when coupled with our much weaker structural findings regarding the effects of audits suggests that IRS activities other than audits may have significant impacts on compliance.

The reduced form results for the sociodemographic and tax code related variables also are also far more significant and consistent than in the structural model.[21] This suggests that the variability of the structural results across taxpayer classes may be due at least in part to the fact that IRS activities other than audits affect compliance and perhaps that academic researchers have not been very successful in modeling the audit process.

7. Conclusions

At this point it seems worthwhile to step back and consider what we can conclude regarding compliance behavior not only from the work reported here but also from other work that has used data related to 1969 tax returns (e.g., Clotfelter, 1983, Witte and Woodbury, 1985, Dubin and Wilde, 1988). There are two important reasons for giving thought to what has been learned in the course of this work. First, there are now a number of studies that use data related to 1969 returns and there is some controversy surrounding the reported results. Second, the 1969 data sets contain information that is only available through

the IRS and the recent work represents the first time that personnel outside the IRS have been allowed to analyze such data and publish results. It is important that we learn as much as possible from the 1969 results because the IRS is now making more recent data available and a number of researchers are developing proposals based on the new data sets.

All authors who consider the issue (Witte and Woodbury, 1985, Dubin and Wilde, 1988 and the results reported here) conclude that for most audit classes audits have positive effects on compliance but that the magnitudes of the effects are small. Further, different models suggest that audits deter noncompliance for different groups. Witte and Woodbury find deterrent effects (mainly lagged effects) for all audit classes except the low income class that takes standard deductions. Dubin and Wilde find a deterrent effect for four of the seven audit classes and a significant deterrent effect for low income wage and salary workers who take standard and who take itemized deductions. (For high income wage and salary workers they find that a higher probability of being audited is associated with less compliance.) We find a deterrent effect of audits for the five low and middle income audit classes that we consider and a significant effect for low income returns with standard deductions and for low and middle income proprietors. Obviously the results depend critically on model specification. This is quite worrisome and suggests that economists' understanding of the effects of audits on tax compliance is not very thorough.

For some audit classes, part of the problem may be that there is very little noncompliance to explain. None of the empirical studies have found a deterrent effect of audits for middle income wage and salary workers who were the most compliant audit class in 1969. Interestingly, the audit class for which our model provides the best the description of compliance behavior is

also the least compliant audit class, namely low income proprietors.

Given the attention that has been focused on how audit coverage affects tax compliance, it is rather disappointing that empirical researchers have not made more progress. We believe that the major reasons for the state of affairs relate to the nature of the data utilized and the secrecy surrounding the IRS's audit selection formulas. All studies that have considered the effect of audits on compliance have used data that is aggregated to at least the three digit zip code level. The use of such data to estimate the audit selection rule is potentially a severe problem. With audit coverage in the range of one to five percent, the audit selection system is primarily an "outliers" system. To be more specific, the IRS's DIF formulas are primarily selecting returns for audit that are outliers in terms of reporting behavior. The averaging implicit in aggregate data washes out most information on outliers and, hence, may only very poorly reflect the audit selection rule.

In order to maintain the integrity of its audit selection system, the IRS must release data in a form that precludes academic researchers motivated by intellectual curiosity or private greed from estimating an audit selection rule that closely duplicates the IRS's formulas. This means that researchers may have great difficulty estimating meaningful structural equations with audits. Indeed, one can question the ethics of doing so. Yet, as noted earlier, different types of audit rules have quite different implications for the specification of compliance equations. This is obviously a major issue that the IRS and compliance researchers need to consider very carefully.

An additional reason why empirical studies may not have found consistently significant deterrent effects is that the economic research thus far has concentrated on the effects of only one IRS enforcement activity,

namely audits. Other enforcement and service activities of the IRS (e.g., taxpayer services and document matching for W-2 and 1099 forms) may also have significant compliance effects. Part of the reason why the reduced form model shows a more consistent deterrent effect than do the structural models for audits may be related to how other IRS activities affect compliance. Very little is known about the effects of IRS enforcement efforts other than audits or about the effects of service activities. Some work has been done though on how programs directed at return preparers might affect reporting behavior (Scholz, Roth, and Witte, 1988).

While results from the 1969 data sets allow no definitive conclusions regarding the effects of audits on compliance, they provide very valuable and quite consistent guidance regarding the type of sociodemographic variables to include to control for"tastes and preferences". This is quite important since theory provides little direction in this area. Specifically, the results from the study of the 1969 data base indicate that the age, sex, and race of the taxpayer, the taxpayer's education, and taxpayer's country of birth affect compliance. The results of work with the 1969 data base also suggest that it is important to consider the specifics of the tax code when constructing and interpreting the coefficients on sociodemographic variables. Variables such as the age of the taxpayer may be related both to attitudes towards compliance and to the special allowances in the tax code for older individuals.

Careful comparison of existing work also indicates that when interpreting the coefficients of compliance models it is important to consider the model specification very carefully. The results one obtains appear to be quite sensitive to the dependent variable utilized and the vector of explanatory variables included. All previous research that has used 1969 data

bases has estimated a single compliance equation with a dependent variable which is either directly (Clotfelter, 1983) or indirectly (Witte and Woodbury, 1985, and Dubin and Wilde, 1988) linked to auditors' findings.[22] The results of this research needs to be interpreted with care because coefficients reflect auditors' as well as taxpayers' behavior. For example, our results suggest that previous findings (Witte and Woodbury, 1985; Dubin and Wilde, 1988) that increased unemployment is generally associated with decreased compliance may be due to tax auditors' abilities to uncover overstated deductions.

In addition, the seemingly mixed evidence regarding the relationship between education and compliance results from differences is the explanatory factors included in the models and is quite understandable when one considers the specifications of the models. If income is included in the model, then we find that the more educated are less compliant. If, however, the specification of the compliance equation does not include a measure of income, then the educated may appear to be more compliant (e.g., Dubin and Wilde, 1988). Without an income variable in the model, however, the education variable reflects the change in compliance associated jointly with an increase in education and the average change in income that accompanies the increased education.

Endnotes

[1]Klepper and Nagin (1987) use the 1982 TCMP data to examine how factors related to individual line items on the return affect compliance.

[2]Adjustments have, of course, been subtracted from income to obtain AGI. Thus, when considering results for our AGI equation some subtractions will be included. However, adjustments to income in 1969 were very small (less than $700 thousand or less than .1 percent of reported AGI) and, thus, our results should not be affected much by adjustments. Reported AGI reflects mainly the income reporting decision. Our data contain no measures of income before adjustments.

[3]See Witte and Woodbury (1983b), Kinsey (1984b), Tauchen and Witte (1986), and Scholz, Roth and Witte (1988) for surveys.

[4]See, for example, Controller General of the United States (1976), Weddick (1983), or Wilt (1986).

[5]See Tauchen and Witte (1986b) for additional details regarding the audit selection process.

[6]We obtained some Census information from the Project 778 files and other variables directly from the Census tapes.

[7]See Witte and Woodbury (1983a, 1985) or Dubin and Wilde (1988) for additional description of the Project 778 data.

[8]For example, Census data contains only open-ended income categories for high income individuals.

[9]Note that if the taxpayer report affects the probability of audit that it is the parameters on these variables not the simple probability of audit that will affect taxpayer compliance. We consider the approach taken in this paper a first approximation to properly modeling of compliance behavior when the taxpayer's reports affect the probability of an audit. See Tauchen and Witte (1986a and 1986b) for a discussion of a way to model compliance when the audit rule depends on taxpayer behavior.

[10]In 1969, seventy percent of all returns audited resulted in auditors' findings that adjustments were necessary. See Commissioner of Internal Revenue (1970).

[11]For example, in 1969 when two and a half million returns were audited, only 25,110 were received by the appelate division and 2,293 were received for full-scale criminal investigation.

[12]Civil penalties are a set rate on the amount of taxes the auditor estimates to have been underpaid. Thus, detectability is central to tax penalties when civil penalties are imposed as well. Only criminal penalties are not directly related to the auditors' estimates of the amount of underpayment. There were only 526 criminal convictions for tax fraud in 1969.

Criminal tax cases generally relate to such IRS special programs as the organized crime and drug enforcement programs or to notorious cases of evasion such as the recent case against the Reverend Moon.

[13]There was information reporting on dividend and interest income in 1969, but the IRS had not perfected its matching programs and, hence, did not make very effective use of this type of paper trail information.

[14]The IRS estimates that auditors are able to uncover only $1 in every $3 that is reported on 1099 documents (US Department of the Treasury, 1983). Until 1979 the TCMP auditors were not provided with 1099 forms.

[15]We developed a marginal tax rate variable that reflected differences in state tax codes and differences in subtractions related to differences in costs of living and other factors. Perhaps because of the aggregate nature of our data, the marginal tax variable was essentially an income proxy.

[16]See Reinganum and Wilde (1985) or Tauchen and Witte (1986a).

[17]Our count of the number of IRS full-time-equivalent employees excludes those at the regional service centers since they do not conduct audits and the district employees who handle alcohol, tobacco and firearms enforcement.

[18]Remember that these employment structure variables are also included to control for legal subtractions from income. We would expect individuals in the service industries to have higher levels of employee business expenses and contributions to self employment pension plans. These adjustments would lead to negative coefficients on the percent employed in services in the reported AGI equation. However, the effect should not be very large since total adjustments were only .1 percent of AGI in 1969.

[19]The reduced form results for the reported AGI equation are also available upon request. As for the structural equations, the reduced form results show that the IRS activities have a greater effect upon reported subtractions than on reported income.

[20]As in the structural model, we believe that the significant and negative coefficients on the variables of the percent of the population over 65 for the low income classes result from the nature of the income exclusions and exemptions allowance given to older taxpayers. At low income levels, these tax code effects swamp the compliance effects associated with age.

[21]The dependent variable used by Witte and Woodbury (1985a) and Dubin and Wilde (1988) particularly troublesome. The measure of compliance is created from the DIF scores. Results using this dependent will reflect the DIF formulas as well as taxpayer and auditor behavior.

References

Beron, K., H. Tauchen and A. D. Witte, (1987) "A Simultaneous Equations Model of Tax Reporting and IRS Auditing," Working Paper, Department of Economics, Wellesley College, Wellesley, MA.

Blumstein, A., J. Cohen, J. A. Roth and C. A. Visher, eds. (1988) Criminal Careers and "Career Criminals", 2 vols. Washington D.C.: National Academy Press.

Clotfelter, C. T., (1983) "Tax Evasion and Tax Rates: An Analysis of Individual Returns," Review of Economics and Statistics, 47, 363-373.

Commissioner of Internal Revenue (1970) 1970 Annual Report, Washington, D.C.: Department of the Treasury.

Controller General of the United States (1976) How the Internal Revenue Service Selects Individual Income Tax Returns for Audit, Washington, D.C.: General Accounting Office.

Controller General of the United States (1979) Who's Not Filing Income Tax Returns?: IRS Needs Better Ways to Find Them and Collect Their Taxes, Washington D.C.: General Accounting Office.

Dubin, J. A., M. J. Graetz and L. L. Wilde (1987) "Are We a Nation of Tax Cheaters? New Econometric Evidence on Tax Compliance," American Economic Review, 77, 240-245.

Dubin, J. A. and L. L. Wilde (1988) "An Empirical Analysis of Federal Income Tax Auditing and Compliance," National Tax Journal, forthcoming.

Greenberg, J. (1984) "Avoiding Tax Avoidance: A (Repeated) Game Theoretic Approach," Journal of Economic Theory, 32, 1-13.

Internal Revenue Service (1969) IRS Office Location, Washington D.C.: Internal Revenue Service.

Intriligator, M. D., (1978) Econometric Models and Applications, Englewood Cliffs, N.J.: Prentice Hall, Inc.

Jackson, B. R. and V. C. Milliron (1986) "Tax Compliance Research, Findings, Problems and Prospects," Journal of Accounting Literature, 5, 125-165.

Kinsey, K. A. (1984a) "Survey Data on Tax Compliance: A Compendium and Review, Working Paper 84-1, American Bar Foundation, Chicago.

Kinsey, K. A. (1984b) "Theories and Models of Tax Evasion," Working Paper 84-2, American Bar Foundation, Chicago.

Klepper, S. and D. Nagin (1987), "The Anatomy of Tax Evasion," Working Paper, Carnegie-Mellon University.

Landsberger, M. and I. Meilijson (1982) "Incentive Generating State Dependent Penalty System," Journal of Public Economics, 19, 333-352.

Long, S. (1985) "The Impact of Information on Law Enforcement," Working Paper, School of Business Administration, Syracuse University.

Poterba, J. A. (1987) "Tax Evasion and Capital Gain Taxation," American Economic Review, 77, 234-239.

Reinganum, J. F. and L. L. Wilde (1985) "Income Tax Compliance in a Principal-Agent Framework," Journal of Public Economics, 26, 1-34.

Schmidt, P. (1986) "Statistical Issues in Modelling Taxpayer Compliance," paper presented at the National Academy of Sciences' Symposium on Taxpayer Compliance, San Padre Island, Texas, January 15-17.

Schmidt, P. and A. D. Witte (1984) An Economic Analysis of Crime and Justice, Orlando, Fla.: Academic Press.

Scholz, J., J. Roth and A. D. Witte, eds. (1988), Report of the NAS Panel on Taxpayer Compliance Research, Washington, D.C.: National Academy Press, forthcoming.

Tauchen, H. V. and A. D. Witte (1986a) "Economic Models of How Audit Policies Affect Voluntary Tax Compliance" in Proceedings of the Seventy-eighth Annual Conference of the National Tax Association-Tax Institute of America, Columbus, OH: National Tax Association, 39-45.

Tauchen, H. V. and A. D. Witte (1986b) "The Impact of Law and Administrative Rules and Compliance: The Case of the Federal Income Tax", proposal to the National Science Foundation, available from the National Bureau of Economic Research.

US Department of the Treasury, Internal Revenue Service (1983) Income Tax Compliance Research: Estimates for 1973-1981, Washington, D.C.: Office of the Assistant Commissioner for Planning, Finance and Research.

Weddick, J. L., Jr. (1983) "Looking for a Needle in a Haystack--How the IRS Selects Returns for Audit," The Tax Advisor, November.

Westat Inc. (1980) Individual Income Tax Compliance Factors Study Qualitative Research, Prepared for the Internal Revenue Service.

Wilson, J. Q. and R. J. Herrnstein (1985) Crime & Human Nature, New York: Simon and Schuster.

Wilt, D. E. (1986) "Trends in DIF and DIF Related Source Examinations" in U.S. Department of the Treasury, Trend Analysis and Related Statistics, Washington, D.C.: Internal Revenue Service.

Witte, A. D. and D. F. Woodbury (1983a) "The Effect of Tax Laws and Tax Administration on Tax Compliance," Working Paper 83-1, Department of Economics, University of North Carolina, Chapel Hill.

Witte, A. D. and D. F. Woodbury (1983b) "What We Know About Factors Affecting Compliance with the Tax Laws" in P. Sawicki, ed. Income Tax Compliance Research: A Report of the ABA Section on Taxation, Washington, D.C.: American Bar Association.

Witte, A. D. and D. F. Woodbury, (1985) "The Effect of Tax Laws and Tax Administration on Tax Compliance: The Case of the U.S. Individual Income Tax," National Tax Journal, 38, 1-13.

Table 1: Definitions of the IRS Audit Classes

Acronymn	Definitions of Audit Class
LI-SD	Low-income tax returns taking the standard deduction (AGI < $10,000 with no Schedule C or F)
LI-ID	Low-income tax returns taking itemized deductions (AGI < $10,000 with no Schedule C or F)
LI-C&F	Low-income propreitor tax returns (AGI below $10,000 with Schedule C or F)
MI-W&S	Middle-income wage and salary workers (AGI between $10,000 and $50,000 with no Schedule C or F)
MI-C&F	Middle-income propreitor tax returns (AGI between $10,000 and $30,000 with Schedule C or F)

Table 2: Empirical Results for Reported Adjusted Gross Income: 2SLS
(t-values in parentheses)

Variable	Audit Class LI-SD	LI-ID	LI-C&F	MI-W&S	MI-C&F
Log odds of an audit -- endogenous	648.13 (2.87)	1693.37 (1.05)	450.75 (3.20)	2491.81 (0.72)	1072.27 (1.07)
Income	.27 (6.47)	.43 (4.96)	.60 (14.46)	.97 (8.72)	.90 (15.78)
Detectability:					
% Employed in manufacturing	8.77 (3.52)	-.34 (-0.10)	2.67 (2.21)	15.18 (0.67)	4.00 (0.74)
% Employed in Services	6.30 (1.55)	-20.37 (-1.40)	-8.23 (-2.84)	-8.87 (-0.26)	11.42 (0.65)
Sociodemographics:					
% High School Education	3.52 (.90)	2.91 (0.42)	-8.38 (-5.05)	-15.89 (-2.71)	-22.23 (-3.11)
% Nonwhite	.11 (.11)	-.97 (-0.40)	.42 (0.49)	5.13 (0.45)	1.71 (0.35)
% Female-headed Households	-29.52 (-1.90)	-20.83 (-1.05)	25.98 (4.76)	60.09 (2.12)	52.45 (2.17)
% Foreign Born	11.87 (2.70)	10.52 (0.55)	-7.24 (-2.48)	-52.80 (-1.79)	-49.04 (-2.67)
Average Age	-6.80 (-.50)	-70.51 (-0.88)	-6.68 (-0.78)	23.20 (0.53)	26.12 (0.70)
Tax Code Variables:					
% Over 65 Years	14.63 (1.24)	65.70 (0.93)	-13.58 (-2.46)	35.69 (1.29)	56.51 (2.51)
Average Family Size	39.95 (.44)	278.49 (0.96)	-70.99 (-1.40)	-274.90 (-1.05)	-389.80 (-1.58)
% Housing Owner Occupied	1.55 (.76)	.30 (0.08)	2.44 (1.55)	-3.51 (-0.37)	1.61 (0.23)
% Unemployed	39.04 (3.32)	67.97 (3.10)	13.55 (1.41)	-36.47 (-0.77)	18.00 (0.39)
Constant	4442.44 (4.25)	12238.68 (1.59)	3958.82 (3.72)	7785.44 (0.58)	3319.67 (0.72)
R^2	.42	.14	.79	.52	.58
F	47.30	10.26	248.66	66.96	83.74
N	858	852	858	827	801

Table 3: Empirical Results for Reported Total Tax Liability: 2SLS
(t-values in parentheses)

Variable	Audit Class				
	LI-SD	LI-ID	LI-C&F	MI-W&S	MI-C&F
Log odds of an audit -- endogenous	109.60	290.17	71.50	2035.37	648.83
	(2.66)	(1.04)	(3.21)	(0.89)	(1.72)
Income	0.07	0.10	0.09	0.27	0.22
	(9.17)	(6.38)	(14.42)	(3.68)	(10.01)
Detectability:					
% Employed in manufacturing	1.42	0.65	0.32	12.77	0.49
	(3.14)	(0.11)	(1.66)	(0.86)	(0.24)
% Employed in Services	.46	-4.15	-2.03	-16.77	-3.90
	(0.62)	(-1.65)	(-4.43)	(-0.76)	(-0.59)
Sociodemographics:					
% High School Education	0.61	0.23	-1.45	-7.87	-10.15
	(0.85)	(0.19)	(-5.53)	(-2.04)	(3.78)
% Nonwhite	-0.31	-0.53	-0.36	2.43	-1.59
	(-1.66)	(-1.27)	(-2.61)	(0.33)	(-0.88)
% Female-headed Households	-4.81	-3.66	2.49	16.33	14.22
	(-1.69)	(-1.06)	(2.89)	(0.88)	(1.57)
% Foreign Born	1.75	1.92	-1.48	-30.73	-24.65
	(2.19)	(0.58)	(-3.20)	(1.59)	(-3.58)
Average Age	2.89	-4.79	3.96	20.96	24.95
	(1.16)	(-0.35)	(2.94)	(0.73)	(1.78)
Tax Code Variables:					
% Over 65 Years	1.78	10.74	-2.35	4.26	14.21
	(0.82)	(0.88)	(-2.68)	(0.24)	(1.68)
Average Family Size	-11.02	5.07	-35.74	-256.13	-280.03
	(-0.67)	(0.10)	(-4.44)	(-1.49)	(-3.02)
% Housing Owner Occupied	-.32	-1.06	-.32	-5.82	-.63
	(-0.85)	(-1.67)	(-1.26)	(-0.94)	(-0.24)
% Unemployed	-0.12	-1.78	-5.38	-46.40	-16.92
	(-0.05)	(-0.47)	(-3.54)	(-1.49)	(-0.99)
Constant	381.16	1449.62	257.94	5966.08	1270.73
	(1.99)	(1.09)	(1.53)	(0.68)	(0.73)
R^2	.56	.22	.78	.13	.35
F	81.77	18.50	231.84	9.34	33.03
N	858	852	858	827	801

Table 4: Empirical Results for the Log Odds of An Audit (x 1000): 2SLS
(t-values in parentheses)

Variable	Audit Class				
	LI-SD	LI-ID	LI-C&F	MI-W&S	MI-C&F
Reported AGI ($1000s) -- endogenous	6.16	-0.82	.48	-.03	.10
	(7.79)	(-2.02)	(1.68)	(-0.50)	(1.34)
Reported Tax Liability ($100s) -- endogenous	-23.91	1.88	.06	-0.02	-0.60
	(-8.39)	(1.25)	(0.04)	(-0.07)	(-1.72)
IRS Resources:					
Returns per FTE	32.68	-31.75	-95.62	-40.35	-104.16
	(0.62)	(-1.27)	(-3.12)	(-1.49)	(-2.90)
Detectability:					
% Employed in manufacturing	-15.35	-.03	-3.44	4.17	7.41
	(-3.51)	(-0.02)	(-1.57)	(-3.03)	(3.81)
% Employed in Services	-22.72	3.94	20.79	13.31	23.28
	(-4.60)	(1.72)	(5.54)	(4.47)	(5.67)
Tax Code Variables:					
% Over 65 Years	25.06	-26.29	.13	10.74	13.23
	(2.80)	(-3.43)	(0.02)	(1.77)	(1.63)
Average Family Size	-69.83	81.90	121.76	38.70	100.36
	(-0.70)	(1.50)	(2.01)	(0.79)	(1.47)
% Housing Owner Occupied	-16.95	2.31	-1.87	1.65	4.63
	(-5.46)	(0.84)	(-1.03)	(0.97)	(-2.06)
% Unemployed	-210.55	4.25	30.48	12.49	-26.51
	(-6.43)	(1.51)	(1.62)	(0.79)	(-1.23)
Constant	-14509	-121	-6998	-3721	-4445
	(-8.53)	(0.79)	(-10.50)	(-10.63)	(-9.45)
R^2	.16	.07	.31	.13	.09
F	17.71	6.51	42.26	13.65	8.52
N	858	852	858	827	801

Table 5: Empirical Results for Reported Total Tax Liability: Reduced Form
(t-values in parentheses)

Variable	Audit Class LI-SD	LI-ID	LI-C&F	MI-W&S	MI-C&F
Returns per FTE (1000s)	-9.16 (-4.74)	-7.13 (-3.07)	-8.24 (-4.58)	-47.37 (-2.71)	-45.83 (-2.53)
Income	0.08 (26.31)	0.09 (22.37)	0.11 (37.67)	0.21 (19.26)	0.19 (16.64)
Detectability:					
% Employed in manufacturing	0.50 (3.51)	-0.19 (-1.08)	0.27 (2.08)	0.33 (.28)	2.25 (1.78)
% Employed in Services	-1.14 (-4.65)	-1.70 (-5.60)	-.99 (-4.33)	2.06 (0.90)	5.53 (2.26)
Sociodemographics:					
% High School Education	-1.19 (-8.23)	-0.92 (-5.31)	-0.88 (-6.60)	-6.38 (-5.48)	-6.61 (-5.37)
% Nonwhite	-0.21 (-2.08)	-0.32 (-2.54)	-0.31 (-3.29)	-3.47 (-3.42)	-2.73 (-2.42)
% Female-headed Households	2.43 (4.42)	-.60 (-0.89)	4.10 (8.00)	26.35 (5.20)	22.19 (4.04)
% Foreign Born	0.15 (0.45)	-1.40 (-3.50)	-2.07 (-6.68)	-16.25 (-5.18)	-16.72 (-5.08)
Average Age	7.11 (7.45)	8.70 (7.59)	4.96 (5.58)	30.57 (3.64)	31.02 (3.45)
Tax Code Variables:					
Over 65 Years	-2.81 (-4.44)	-1.48 (-1.95)	-2.88 (-4.88)	12.83 (2.40)	13.55 (2.37)
Average Family Size	-46.46 (-8.89	-43.22 (-6.87)	-21.96 (-4.51)	-169.48 (-3.59)	-183.14 (-3.62)
% Housing Owner Occupied	-.76 (-4.29)	-.86 (-4.05)	-.53 (-3.21)	-1.86 (-1.18)	-1.93 (-1.16)
% Unemployed	-3.04 (-2.90)	-2.64 (-2.09)	-3.53 (-3.61)	-37.28 (-3.93)	-34.08 (-3.45)
Constant	-80.23 (2.84)	85.03 (2.49)	-243.22 (9.24)	-1655.81 (-6.10)	-1477.21 (-5.25)
R^2	.80	.71	.88	.57	.53
F	259.74	160.68	470.63	85.97	71.46
N	857	851	857	826	800