The Determinants of Individual Income Tax Compliance

Estimating The Impacts of
Tax Policy, Enforcement, and IRS Responsiveness

Department of the Treasury
Internal Revenue Service
Publication 1916 (Rev. 11-96)
Catalog Number 22555A

Department
of the
Treasury
**Internal
Revenue
Service**

The Determinants of Individual Income Tax Compliance

Estimating the Impacts of Tax Policy, Enforcement, and IRS Responsiveness

November 1996

Alan H. Plumley
IRS Research Division

This is a copy of a Ph.D dissertation prepared for Harvard University, slightly modified for distribution by the Internal Revenue Service. The estimates presented herein are subject to change. In fact, this revision is the second IRS printing, and includes corrections of a number of typographical errors contained in the August 1996 printing. Other refinements are being made continually. Comments and questions are welcome, and may be directed to Alan Plumley as follows:

Mail: IRS Research Division CP:R:R:AR:E
1111 Constitution Ave. NW, Washington, DC 20224

Phone/VoiceMail: (202) 874-0508

Fax: (202) 874-0634

E-mail:AHPLUM50@M1.IRS.GOV

Suggested Citation
Internal Revenue Service
The Determinants of Individual Income Tax Compliance: Estimating The Impacts of Tax Policy, Enforcement, and IRS Responsiveness
Publication 1916 (Rev. 11-96)
Washington, DC: 1996

ABSTRACT

This paper presents an econometric analysis of the impact of a wide variety of potential determinants of voluntary compliance with individual income tax filing and reporting obligations. Based on perhaps the richest dataset yet compiled (by state and year, from 1982 through 1991), including data on taxpayer behavior, IRS actions, and other factors, the analysis finds significant compliance effects attributable to many tax policy and tax administration parameters, including: audits; the matching of third-party information documents; the issuance of targeted nonfiler notices; criminal tax convictions; marginal tax rates; the burden associated with completing the myriad tax forms and schedules; and the preparation of returns by the IRS Taxpayer Service function.

The Determinants of Individual Income Tax Compliance:

Estimating The Impacts of Tax Policy, Enforcement, and IRS Responsiveness

Contents

The Determinants of Individual Income Tax Compliance:
Estimating The Impacts of Tax Policy, Enforcement, and IRS Responsiveness

1. Introduction

For many years, the Internal Revenue Service (IRS) has presumed that its activities promote better income tax compliance in the general population—both through deterrence and taxpayer service—but it has never been able to quantify this impact, or even verify that it exists. This paper does both, providing the first empirical basis for choosing the best combination of major IRS activities to improve voluntary filing and reporting compliance among individuals—a capability that is increasingly needed in this climate of declining budgets. It also estimates the compliance impact of important tax policy parameters—most notably marginal tax rates.

The success in estimating these compliance effects is not the result of sophisticated, new statistical techniques, but rather the creative application of straightforward techniques to the right data, which took eight years to compile from IRS reports and databases, as well as numerous other sources. The analysis uses panel data over a ten-year period (1982-91) aggregated to the state level to estimate one filing compliance equation and three separate reporting compliance equations. Among the findings: the deterrent effect of audits in the general population is about 11 times as large as the adjustments proposed by the audits themselves, but nonfiler notices, information document matching, and return preparation assistance are more cost-effective in boosting revenue.

1.1 Background

The federal income tax system operates on a self-assessment basis. That is, the government expects taxpayers to determine their own tax obligations and to pay voluntarily whatever is due—both regularly (through withholding from wages and through estimated tax payments, if necessary) and at year end (by filing tax returns and paying any additional balances due). By placing the onus on taxpayers, the government avoids the costly alternative of determining each individual's tax liability and doing whatever it must to collect it.

However, one cost of relying so heavily on the voluntary compliance of taxpayers is that not all tax is voluntarily paid. The IRS estimates that the gross individual income tax gap (the difference between what taxpayers should pay and what they actually do pay voluntarily and timely) was about $94 billion for Tax Year 1992—about 20 percent of total individual income tax receipts and over half the size of the budget deficit![1]

Congress has taken several steps to strengthen voluntary compliance with the income tax laws. These actions fall into two major categories: requirements (e.g., requiring withholding of tax at the source of income and requiring third party information reporting) and deterrents (e.g., giving the IRS certain enforcement powers, and stipulating the penalties that those who are caught through this enforcement must pay).

Probably the most widely known example of an IRS deterrent is the "audit" of an individual's tax return. In Fiscal Year 1992, for example, IRS completed examinations of just over 1 million tax returns of individuals. Although these returns were for several tax years of liability, this is approximately the same number of examinations conducted on Tax Year 1992

[1] IRS (1990).

returns, which numbered about 115 million—implying an audit coverage rate of roughly 0.9 percent. As a result of the examinations completed in FY 1992, IRS recommended that the taxpayers in question pay additional tax and penalties totaling $6.0 billion.[2] The additional revenue that the government will collect from this "yield" (following all appeals, litigation, and collection efforts) is the *direct* revenue effect of those 1 million examinations. However, it is quite likely that the audits also had an *indirect* revenue effect—inducing some amount of voluntary compliance in the population at large through the general deterrent effect of the examinations (referred to at IRS as the "ripple effect" of the examinations), and (perhaps) by influencing the voluntary compliance of the contacted taxpayers in subsequent years (referred to as the "subsequent year effect"). Indeed, one of the purposes of IRS enforcement is establishing a credible deterrent to noncompliance. It has generally been believed, for example, that many taxpayers would perceive increased auditing by IRS as an increase in their chances of being audited, and that they would improve their voluntary compliance as a result.

In addition to the various requirements and deterrents that Congress has implemented specifically to improve voluntary compliance, other actions or laws may have influenced compliance indirectly. Examples of this phenomenon may include a change in compliance behavior resulting from a change in tax policy (e.g., the marginal tax rate structure), or from some change in the public's attitude toward the IRS (which may arise from changes in IRS's responsiveness to taxpayers' needs).

The focus of this paper is the indirect behavioral response of taxpayers (as measured by changes in their voluntary filing of required income tax returns, and their reporting of income and offsets to income on those returns) to changes in IRS enforcement, IRS's responsiveness, and basic tax policies.[3] Quantifying these responses could help to shape tax policy and tax administration for the foreseeable future—especially given the need to reduce the budget deficits and to make the best use of government resources.

1.2 Previous Research

The indirect revenue effect of audits is beginning to receive attention from researchers, but little—if any—empirical research has been done to quantify the separate compliance effects of enforcement, tax policy and IRS responsiveness. What makes such research challenging, of course, is that the compliance impact of government actions is never observed in isolation; it can only be estimated. Voluntary income tax compliance is probably determined by a wide variety of factors that interact differently for each individual. Although many such factors have been suggested, and several studies have focused on some of them, nothing has emerged to guide policy-makers concerning the relative merits of alternative approaches to improving compliance. This may be due in part to the fact that no comprehensive theory exists that explains the compliance behavior of taxpayers. It may also be because very little data are available to test such theories.

One of the earliest attempts to model taxpayer compliance was done by Allingham and Sandmo (1972), which applied the utility-maximization approach of Becker (1967) to tax compliance. Allingham and Sandmo's simple model predicts the intuitive result that taxpayers will voluntarily report more income in response to either an increase in the probability of being detected, or an increase in the penalty imposed on those who are caught. However, the model is inconclusive in predicting the response to an increase in the tax rate; the net response is the sum of two terms in their model—one negative (suggesting a decrease in income reported as the tax rate increases), and another, which is most probably positive, assuming that taxpayers' risk aversion

2 IRS (1992).

3 Strictly speaking, my analysis *quantifies* such behavioral responses, but it does not address *why* taxpayers behave the way they do. Although traditional deterrence mechanisms may be responsible, this study cannot prove it.

decreases with income. Allingham and Sandmo liken these two terms to a positive income effect and a negative substitution effect. They reason that the substitution effect means that an increase in the tax rate makes it more profitable to underreport income at the margin (i.e., the higher the tax rate, the more money is retained when one underreports a dollar of income). The income effect, however, is most likely positive because an increase in the tax rate reduces net income, and assuming decreasing absolute risk aversion, the taxpayer is less willing to underreport income than before. Which of these two effects is stronger depends in part on the taxpayer's degree of risk aversion.

This early model—and virtually everything that has followed—says nothing about IRS responsiveness to taxpayers, or about nonfiling. Roth, Scholz, and Witte (1989) gives an excellent overview of the theoretical and empirical work on tax compliance through the 1980's, but several empirical studies are worth mentioning here. At least three studies (Clotfelter (1983), Cox (1984), and Dubin, Graetz and Wilde (1990)) have explored the possibility that the actual marginal tax rate to which individuals respond is the sum of the federal and state marginal rates, and that since the federal income tax law is applied uniformly across states, using the state marginal tax rate alone has equivalent explanatory power. The primary reason for making this assumption arises from econometric pragmatism: to gain cross-sectional variation in the marginal tax rate variable. However, for panel studies such as this one (i.e., a time series of cross sections), the difficulties inherent in creating a state marginal tax rate variable, and the opportunities associated with the variation in federal tax rates over time, caused me to explore the role of federal tax rates alone.

Four econometric studies that attempted to estimate the indirect effect of audits on compliance are worth noting. Erard (1992) focused solely on the "subsequent-year effect" of audits on the reporting compliance of those who were audited, and reports inconclusive results. Three other studies examined the general deterrent effect of audits.[4] Tauchen, Witte, and Beron (1989) was a cross-sectional study based on the 1979 Taxpayer Compliance Measurement Program (TCMP) micro database, and concluded that the impact of audits is weak at best. The other two studies were aggregate analyses. Dubin, Graetz and Wilde (1990) used panel data (aggregated at the state level, 1977-1986) to estimate both reported tax per return and returns filed per capita, and finds a very large and significant deterrent effect of audits. Beron, Tauchen, and Witte (1992) uses 1969 data aggregated at the 3-digit ZIP Code level to estimate average Adjusted Gross Income and average tax reported, and finds a weak indirect effect of audits that is limited to certain taxpayer groups.

1.3 Advances Made By This Research

This study improves upon all of the earlier work in this area. It is based on one of the most comprehensive datasets ever compiled on the potential determinants of voluntary compliance, and provides an empirical basis for choosing the best mix of strategies for improving and maintaining voluntary compliance.

Before describing the details of my model (section 2) and the results (section 3), I provide in this section an overview of the advances made by this research. These advances are almost entirely in the realm of improved data and an improved econometric specification; no new estimating procedures are developed. Since my approach is most similar to that of Dubin, Graetz and Wilde (1990) (which I will hereafter refer to as DGW), I will describe the advances with respect to that important work. There are several obvious similarities between this study and

[4] Actually, these studies examined the combined indirect effect of audits on the compliance of those who were not audited (the "ripple" effect) as well as of those who were audited (the "subsequent-year" effect). But since the methodologies could not distinguish between these effects, and since the "ripple" effect is presumed to dominate, the studies can be thought to examine general deterrence.

DGW. Each attempts to estimate the impact of various factors (including audits and tax rates) on the voluntary income tax filing and reporting compliance of individuals using a 10-year panel of data aggregated to the state level. Each accounts for the endogeneity of audit rates. However, there are many important differences between these studies. I describe these differences—and explain why they represent improvements—below.

1.3.1 Dependent Variables

DGW estimates three equations: one for reported tax per return (a measure of reporting compliance), one for returns filed per capita (reflecting filing compliance), and one for "assessed liability" (reported tax plus additional tax and penalties proposed by audits) per return. These dependent variables control for cross-sectional and time variations by dividing by the number of returns filed or by population. My dependent variables use as denominators exogenous surrogates for reporting and filing obligations, giving them a meaning closer to traditional measures of voluntary compliance. My estimation of filing obligations by state from Census data, for example, creates a very powerful measure of the filing rate—obviating the need for such DGW variables as the number of households per capita, and the percent of households on welfare. Moreover, using reported tax per return filed as a dependent variable makes it difficult to interpret the results since the explanatory variables could conceivably influence both the numerator and the denominator. For example, a positive coefficient could indicate that the explanatory variable increases reporting compliance, but it is also possible that it increases average tax reported by *decreasing* filing compliance among low-income taxpayers. The implications of those two possibilities are dramatically different. In fact, one may initially view the latter possibility as the correct interpretation of DGW's positive coefficient on audit rate in the reporting equation and its negative coefficient on audits in the filing equation—especially given DGW's assessment that "one way to escape audits has been simply not to file." DGW reports, however, that the magnitudes of these coefficients is such that the net effect of audits is to increase dollars reported—but extraneous side calculations are necessary to conclude this, illustrating the cumbersome nature of the specification.

The numerators are also quite different in the two studies. DGW uses the *tax* reported on returns as its measure of voluntary reporting compliance. I avoid tax as a measure of reporting compliance because several of the potential determinants of voluntary compliance (e.g., marginal tax rates, filing thresholds, marital status, and allowable child exemptions) also have a direct role in the calculation of tax from gross income, making it difficult to separate their impact on compliance. I use three more useful measures instead: total income reported, total offsets reported, and net income reported (income minus offsets). These three equations also have the advantage of providing insight into the major forms of noncompliance (underreporting income vs. overstating offsets to income or to tax), and they allow consistency comparisons across equations (since income minus offsets equals net income). I have also controlled for the extent to which the tax rules have changed concerning the amount of income that must be reported and the amount that may be claimed as offsets. I have done this largely by defining income and offsets in three different ways, and by estimating separate equations for each definition.

My data also reflect two important qualitative improvements. First, DGW uses dollars of tax reported on returns as tabulated from IRS's Statistics of Income (SOI) samples. I use the same source, but since these samples are not designed to be accurate at the state level, I have adjusted the sample weights (by state and tax form type) so that they conform to actual return filings by state. Second, I have restricted the number of returns filed and the dollars reported on those returns to include only those returns that were required to be filed—excluding those with no tax liability, but were filed to claim a refund of withheld tax or to claim the Earned Income Tax Credit.

Finally, I do not estimate anything like "Assessed Liability per Return" (ALR) because it is too misleading. Not only does it suffer from all of the disadvantages of reported tax as a dependent variable, but the proposed audit adjustments component of ALR is an incorrect measure of the direct revenue effect of audits. This is because the proposed audit adjustments ("recommended additional tax and penalties"), which DGW compiled from IRS's Commissioner's Annual Reports, are *not* the amounts assessed through enforcement; a large portion of these "recommended" adjustments is never assessed, due to successful taxpayer appeals and litigation, and the rate at which these recommendations are ultimately assessed has varied greatly over time and across states. Moreover, even the proposed adjustments are endogenous with audit rates; since IRS allocates its audit resources so as to audit only those returns it perceives to be most noncompliant, the average audit "yield" declines with audit rate. Even though DGW recognizes that ALR is endogenous with audit rates, and that as voluntary compliance increases (say, due to increased audit rates) average audit yield is likely to decrease proportionately, the paper does not recognize that audit yield decreases with audit rates by design. This complicates the choice of instruments for audit rate, since an effective instrument now has to be unrelated to both taxpayer compliance *and* overall resource levels. A final source of confusion introduced by the inclusion of proposed audit adjustments with the tax voluntarily reported arises from the fact that they relate to very different time frames, which could be important in a longitudinal study such as this. Recommended audit adjustments are typically made more than a year after the return is filed; the larger the adjustment, the greater the likelihood that the time lag is more than one year. Therefore, the impact of variables whose effect on tax obligations varies over time may be difficult to estimate reliably with this dependent variable.

1.3.2 Independent Variables

A more obvious improvement in this study is the inclusion of a much richer set of explanatory variables—richer both in quantity and in quality. This allowed the specification of each equation to be uniquely suited to the differences in the dependent variables; DGW employed the same specification for each equation. I have grouped my variables into the following categories: Tax Policy, Burden/Opportunity, IRS Enforcement, IRS Responsiveness, and Demographics/Economics. I have constructed *federal* marginal tax rate variables in lieu of the average *state* tax rate used in DGW, and have included a variety of other tax policy variables, as well: the filing threshold, a state amnesty indicator, allowed child exemptions, and state and local taxes that are deductible federally. I include the prevalence of sole proprietors rather than of farmers, since they are a more prevalent (and, arguably, more important) indicator of the opportunity to avoid (and, perhaps, to evade) taxes. Other Burden/Opportunity variables I have included are the burden (in hours) needed to complete and file all required returns and schedules, and the percentage of returns prepared by a paid practitioner.

The only IRS activity DGW includes is the audit rate. This study also includes the audit rate, but it is the first to use the audit *start* rate instead of the audit *closure* rate. Since audits are typically closed several years after the returns are filed, and closures in any given year relate to many different prior tax years, the start rate better represents the percentage of returns filed in a given year that are audited. I have also included variables for four additional enforcement activities: the information return matching program, nonfiler notices, refund offsets, and criminal tax convictions. This is also the first study of its kind to include variables related to IRS's non-enforcement activities. Two included variables relate to IRS's Taxpayer Service telephone assistance and return preparation services. Variables considered, but not found to have a significant impact on compliance include other Taxpayer Service activities (correspondence and educational outreaches), and the speed with which refunds are processed and sent to taxpayers.

1.3.3 Estimation Technique

DGW actually employs a more sophisticated econometric procedure, but I do not believe that it is appropriate in this context. That study uses a "random effects" model to control for unobservable state-specific factors that do not vary over time, whereas I employ a "fixed effects" model to account for this common peculiarity of panel data. (I explain in section 2.2.1 why this is a preferable approach in this context.)

1.3.4 Identification

One of the most critical features of any study of this kind is the choice of instrument(s) for audit rates. DGW uses two: Budget per Return (BPR), and information documents (other than W-2s) filed divided by the number of tax returns filed. Although these variables reflect a very creative use of available data, each has serious drawbacks as an instrument for the audit rate. (The reasons for this are given in section 2.2.2.) This paper introduces for the first time two instruments that help to explain the audit rate, but which are unrelated to compliance: the percentage of auditor time directly devoted to audits (Direct Examination Time, or DET), and the average DET per audit. As productivity-related measures, these variables *cause* the audit rate to increase or decrease quite independent of taxpayer compliance. Strong evidence that my reporting equations are identified is given in section 3.2.1.

1.3.5 Functional Form

The DGW specification is strictly linear. In contrast, I use the logarithms of those independent variables that are likely to have a non-linear effect on compliance. This is especially important among the enforcement variables, which almost certainly achieve diminishing indirect marginal returns to effort, much like their direct revenue effects.

2. The Model

It is tempting for economists to develop theoretical models of individual (micro) tax compliance behavior. However, there are two significant reasons—one theoretical and one practical—why these models may be inadequate. The theoretical reason is that much of that behavior is governed by what is called in the literature "general deterrence."[5] As Nagin (1978) correctly observes, "general deterrence is inherently an aggregate phenomenon since it is reflected in the behavior of the entire population."[6] That is why analyses of criminal sanctions have generally been aggregate studies. Many studies of tax compliance have been aggregate, also, but this has typically been due to a lack of access to micro-level compliance data, such as IRS develops in its Taxpayer Compliance Measurement Program (TCMP). The aggregate studies, therefore, have tended to start with micro-level economic models, and attempted to estimate these using the lowest level of aggregation possible—such as at the 3-digit, or 5-digit ZIP Code level. (See, for example, Beron, Tauchen, and Witte (1992)). The lower the level of aggregation, however, the less realistic is the model. That is because such models implicitly assume that the general deterrent operates only within the strict confines of each unit of observation (e.g., a ZIP Code boundary),

[5] We may think of general deterrence as including both negative influences, such as IRS enforcement actions, as well as potentially positive influences, such as IRS responsiveness to taxpayers' needs. Although the latter may not intuitively be considered a deterrent, it undoubtedly influences the general population just like a deterrent. For example, many of those who have good or bad experiences with IRS efforts to help them presumably share their experiences and perceptions with their friends, who may change their own compliance accordingly. This is completely analogous to the way in which perceptions about IRS enforcement are developed in the general population.

[6] Nagin (1978), p. 99.

and it seems obvious that people will develop their compliance perceptions and propensities based on the information they get from a wide variety of sources from many locations. In fact, many people today interact more with people from outside their ZIP Code (such as at work) than they do with others in the immediate vicinity of their residence.

The practical reason why micro models may be inadequate is that it is virtually impossible to quantify deterrence-type activities (like audit rates and Taxpayer Service phone calls) in any meaningful way for each individual observation. Inevitably, these variables are aggregated in some way, then imputed to individual observations; they are therefore subject to the same limitations as the low-level aggregate studies, and are poor substitutes for the individual perceptions called for in the theoretical models.

In order to avoid or minimize these problems, I have aggregated all data to the state level. It would be nice to have been able to aggregate to the IRS district level,[7] but it would have been extremely difficult to have aggregated most of the non-IRS data in this way. Fortunately, the IRS data could easily be combined to derive state-level aggregations. Such an aggregate analysis suits the IRS from the perspective of its usefulness, as well; IRS is not as interested in individual behavior as it is in aggregate behavior. Although the aggregate behavior is certainly the sum of the behaviors of individuals, IRS is interested in the bottom line: how could we allocate resources differently to improve voluntary compliance, and hence, net revenues? Aggregation bias cuts both ways: it makes it hard to use aggregate results to estimate a micro model of individual behavior, but by the same token, it is equally problematic to estimate a micro model using micro data, and then to generalize the results to make aggregate calculations. It is best to make aggregate calculations based on aggregate data. Although it may not have the sophistication of some micro models, and it cannot model individual motivations, it seems appropriate in this context.

2.1 The Data

The data collected for this study form one of the most comprehensive datasets ever compiled of potential determinants of voluntary filing and reporting compliance, and took over eight years to assemble in usable form. Many of the IRS variables had never before been assembled for any study, and were available only on paper or microfiche tables—often in a form requiring some manipulation to derive the desired concept. Some IRS and external data were available as representative samples of individuals; certain IRS variables were available at the district level; and certain external data were available at the state level. All variables were aggregated to the state level,[8] and were compiled for a ten-year period: 1982-1991. The panel nature of the data increased the number of observations, and also captured important variations in both compliance and in its determinants over time. Appendix A contains a detailed summary of the sources and derivations of the raw data used to create the variables included in this study.

2.1.1 Measures of Voluntary Compliance

The IRS recognizes three types of voluntary compliance: *filing compliance* (the timely filing of any required return); *reportingcompliance* (the accurate reporting of income and of tax liability); and *payment compliance* (the timely payment of all tax obligations). This study focuses on both filing compliance and reporting compliance (both income reporting and offset reporting—subtractions such as deductions, exemptions, adjustments, and credits). The most basic measure

7 Until just recently, IRS had 63 districts, each of which was a single state or a portion of a single state.

8 The District of Columbia (DC) is included in Maryland both because of its small size and because most IRS data are not available for DC separately, since IRS's Baltimore District includes DC with all of Maryland. Moreover, Alaska is excluded from the data for reasons explained below.

of voluntary compliance is what the taxpayers actually did: how many returns were filed, and how much income and offsets they reported on those returns. However, while analyzing the possible determinants of this taxpaying behavior, one clearly has to control for the corresponding "true" obligations—returns required to be filed, income required to be reported, and offsets allowed to be claimed. One could control for these true obligations either by including them among the explanatory variables, or by dividing the basic voluntary measures by them to derive appropriate compliance ratios. I have adopted the ratio alternative. Either way, these true obligations are never observed; they can only be estimated.

It is tempting to use TCMP's compliance data—thorough audits on a representative sample of taxpayers—to estimate true obligations, but even these data suffer greatly from a general lack of good information on the true filing and reporting obligations of the individuals in the samples.[9] Moreover, TCMP data on individual income tax reporting compliance are currently available for only four years (1979, 1982, 1985, and 1988), and only one survey on filing compliance is available (1988), making it difficult to construct a useful panel of data from the TCMP. I therefore controlled for true filing and reporting obligations using non-TCMP data.[10]

Filing Compliance

My measure of filing compliance [*FilingRate*] is the ratio of the number of required returns actually filed to the total number required to be filed, expressed as a percentage. The number of required returns actually filed was aggregated by state from IRS's Statistics of Income (SOI) samples of individual returns for each year.[11] All three basic compliance measures (the number of returns, and the amount of income and offsets reported on those returns) correspond to returns *required* to be filed. This includes all returns having a positive tax liability or net losses. This definition excludes returns filed "unnecessarily," as well as those having no tax liability, but are filed to claim a refund of any withholding, or solely to claim the refundable Earned Income Tax Credit (EITC). I have excluded returns filed solely to claim the EITC since EITC noncompliance tends to *increase* filing, whereas usual filing noncompliance *decreases* filing.

I have defined the denominator—the number of returns required to be filed—in the same way, and have estimated it by state for each year from the micro data files of the Current Population Survey (CPS) compiled by the Bureau of the Census.[12] The CPS is the most comprehensive annual U.S. census of individuals, and recognizes families and households, as well. In order to structure the data to reflect potential tax returns instead of individuals, I first combined the information about spouses into combined records, approximating a jointly-filed tax

---

9 Since IRS auditors are not omniscient, TCMP inevitably falls short of identifying all unreported income and the corresponding tax liability. (Alexander and Feinstein (1987) attempted to estimate the undetected unreported income using a sophisticated detection-controlled econometric technique, but while this would be useful for making aggregate adjustments to estimates of noncompliance, using it to impute specific amounts of undetected unreported income to the individuals in the sample is likely to make inferences about taxpayer behavior very sensitive to any errors introduced by the imputation.) Estimating true filing obligations is even more troublesome at the micro level, since nonfilers, by definition, tend not to be identified. (Erard and Ho (1995), for example, implemented a sophisticated analysis of the recent nonfiler TCMP micro database to estimate both the number of nonfilers and the tax gap associated with them.)

10 It is somewhat ironic that I have chosen not to use TCMP data—even though I have access to them—while others have used aggregate data only because they had no access to TCMP. This seems to be the right choice for this study, however.

11 State-by-state aggregations of filing data are available from IRS's Individual Masterfile (IMF) of all returns, but comprehensive reporting data are much too difficult to produce from the IMF—making it impossible to restrict our counts to required returns. The SOI samples are the only alternative—both for filing data and for reported data. However, even though the SOI samples of 100,000 or so individuals are relatively easy to access and manipulate, they are not designed to be representative of state populations. Therefore, I adjusted the SOI sample weights for each year to make these files conform to the actual number of returns filed in each state.

12 The "March Supplement" of the CPS is compiled each year, and typically has a sample size of 80,000 to 90,000 records. Each file reflects income received in the prior calendar year, so I compared the SOI estimates for a given tax year with CPS estimates based on data compiled the following March.

return.[13] I then estimated which records corresponded to dependents in the tax sense, since special filing rules apply to them.[14] Next, I estimated which individuals might have been eligible to claim the Head of Household filing status, making it a simple matter, then, to divide the remaining records into the Single and Married-Filing-Jointly statuses. Finally, I calculated the correct standard deduction and exemption values based on the estimated filing status and the tax law for the year in question. Using the amounts of various types of income reported on the CPS, I was then able to estimate whether the "return" had income over the filing threshold for that year. Any record that did, or which showed negative income, I counted as a required return. Appendix B contains detailed information about the logic of this estimation process, as well as a summary of the CPS variables used and the relevant tax parameters for each year. Figure 1 illustrates the resulting national trend in the *FilingRate* variable.

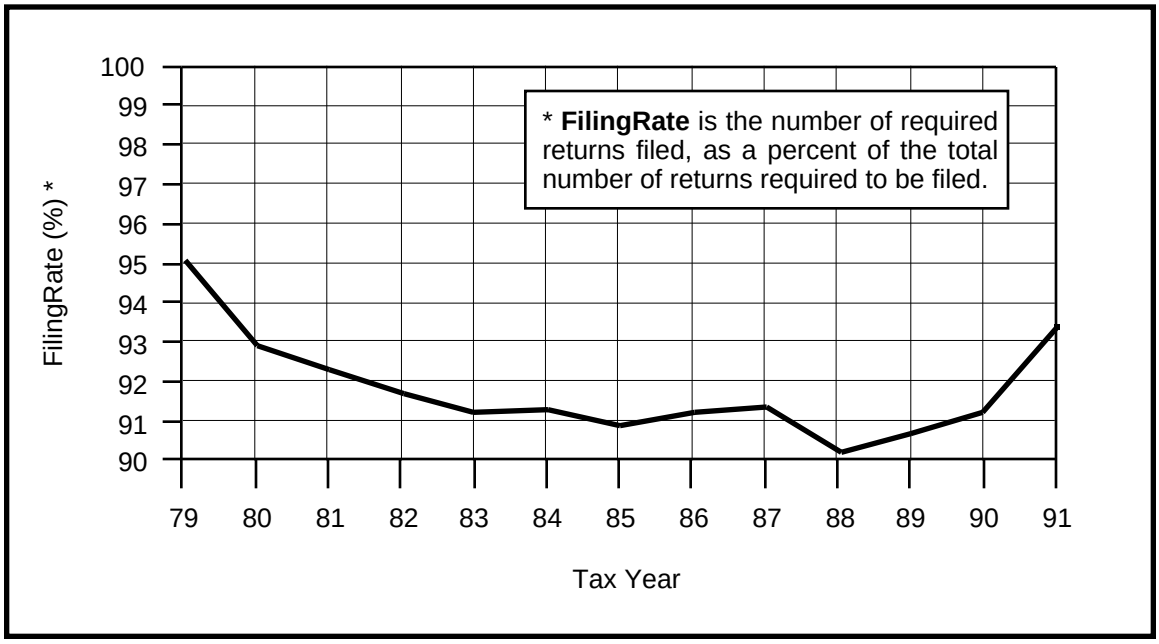


Figure 1. Filing Compliance Measure: The number of required individual income tax returns filed (per SOI), as a percent of the total number of returns required (estimated from CPS), Tax Years 1979-1991

As Figure 1 illustrates, the national *FilingRate* has remained in the low 90's, with a general downward trend until 1988, and a significant improvement thereafter. (Average *FilingRates* by state are tabulated in Appendix E.)

Reporting Compliance

Reporting *non*compliance takes two forms: *under*reporting income and *over*stating offsets to income (i.e., exemptions, adjustments, and deductions) or to tax (i.e., credits). Since

[13] I had to assume that all couples filed jointly, since there is no conclusive information on the CPS to suggest otherwise. This is a fairly small approximation, however, since very few couples file separately in reality.

[14] Unfortunately, the CPS files compiled for me excluded records for children under the age of 15. This resulted in a slight underestimation of filing obligations—and a corresponding overestimation of the FilingRate variable—for all states, but this was especially fatal for Alaska. The state of Alaska pays from its oil holdings a "Permanent Fund Dividend" on the order of $1,000 each year to every permanent state resident—including children—and this dividend is taxable federally. Virtually every child in Alaska, therefore, has an obligation to file a federal income tax return—especially after the Tax Reform Act of 1986, when the filing threshold for dependent children was drastically reduced in most cases. Because of this, the number of Alaska returns filed doubled for tax year 1987. Since my CPS files could not reflect this, I chose to drop Alaska from my analysis, leaving me with 49 states rather than 50.

the various determinants of compliance presumably influence income reporting and offset reporting differently, I have measured reporting compliance by means of three amounts that taxpayers voluntarily report on filed returns: the total amount of income that they report, the amount of offsets[15] they report, and the corresponding net income—income minus offsets. It is tempting to monitor the tax reported on the returns, but the impact of tax policy parameters (e.g., marginal tax rates and filing thresholds) on the tax voluntarily reported is twofold—influencing the tax calculation directly, and potentially influencing compliance propensities, as well. By focusing on income and offsets separately—before the tax calculation—we can isolate the impact of tax policy parameters on reporting compliance. I estimate the net income equation both because it is close to a tax concept, and because it is a useful check on the results from the separate income and offsets equations.

The amounts of income and offsets reported were aggregated by state from the SOI files for each year. However, during the 1982-1991 time period, the rules that govern what income must be reported and what amount of offsets may be claimed changed significantly for many items (see Appendix C for complete details). Unless we can control for these rule changes by modifying the data to reflect constant-law amounts, or by including adequate explanatory variables, we may misinterpret the effects of tax policy variables (like marginal tax rates) that are highly correlated with these other tax law changes.[16] I handled this problem by defining the three reporting compliance measures (income, offsets, and net income) in three separate ways: (A) excluding all components whose reporting rules changed during the period (except for ones that could be controlled for by creating constant-law data or by including appropriate explanatory variables; this definition included on the order of 97 percent of total income, 30 to 60 percent of adjustments, 94 percent of itemized deductions, and 30 to 60 percent of credits); (B) making all of those adjustments, but including income and offset components whose rules were changed only by the Tax Reform Act of 1986 (TRA86); and (C) including all income and offset items regardless of rule changes. Comparing the results using these three sets of definitions gives us insight into the impact of changes in the rules governing what must (or can) be reported.

As with filing compliance, each of the reporting concepts was expressed as a ratio. However, unlike for filing compliance, it is not possible to develop comprehensive exogenous estimates of the amount of income or offsets that *should be* reported. Therefore, to control for true reporting obligations, each of these reported amounts was divided by the amount of Personal Income estimated for the National Accounts by the Bureau of Economic Analysis (BEA) for the appropriate state and year. Although Personal Income does not correspond exactly to the amount of "true" income that must be reported on federal income tax returns (it includes the income of all those who do not need to file returns, for example; see Appendix D for a detailed comparison), and it is certainly not the same as "true" offsets allowable, it is a very effective control for these concepts. This is because it is probably the most comprehensive individual income variable available annually at the state level, and because it is derived substantially independent of tax return data. That is, it is reasonably exogenous to income tax compliance and income tax administration decisions. Table 1 summarizes the data sources for each of the major components of Personal Income. As the table shows, four major components are based (at least in part) on individual income tax returns. However, BEA adjusts these amounts to account for underreporting using data from the IRS Taxpayer Compliance Measurement Program, the Information Returns Program, audits, and other data (see Parker (1984)).

---

15 All of these offsets are subtractions from income, except credits, which are subtractions from tax. In order to combine these amounts into a single offsets concept, I converted the credit amounts on a given return to the equivalent income offset amounts by dividing them by the marginal tax rate faced by the return.

16 I am indebted to Brian Erard for making this observation.

Table 1. Data Sources for Major Components of Personal Income

Component of Personal Income	Data Source
Wages and salaries	State Unemployment Insurance data (Form ES 202)*
Other labor income	Forms 5500 submitted by employers and plan managers
Non-farm proprietor income	Forms 1040*
Farm proprietor income	U.S. Department of Agriculture
Rental income of persons	Census of Housing, Bureau of the Census
Royalty income of persons	Forms 1040*
Personal dividend income	Forms 1120 submitted by payers of dividends
Personal interest income	Forms 1120 submitted by payers of interest*
Transfer payments	Federal budget, Social Security Administration, etc.

Source: Thae Park, Bureau of Economic Analysis, Department of Commerce; see also BEA (1994).

* These data are adjusted by TCMP, IRP, audit , and other data; see Parker (1984).

The resulting ratios [*IncomePct*, *OffsetsPct*, and *NetIncomePct*] must not be interpreted as compliance rates in the same sense that the *FilingRate* ratios can be, but in the context of the regression analysis described below, dividing by Personal Income (and including additional control variables among the explanatory variables) is an effective way to control for variations in true obligations over time and across states.

The national trends in the income and offset reporting variables are illustrated in Figures 2 and 3, respectively. (Average values by state are tabulated in Appendix E.) One of the most striking things about Figure 2 is to see that reported income as a percent of Personal Income generally declined through 1986, and then rose sharply. A decomposition of total reported income into the separate types of income reveals that most of the increase after 1986 (presumably in response to TRA86), can be attributed to Schedule E income (consisting mainly of rents, royalties, and income from partnerships, S-Corporations, estates and trusts), and to wage and salary income. Of the Schedule E income types, S-Corporation income is not included in Personal income, and yet increased after TRA86, so this may have contributed to the overall rise in *IncomePct*. However, excluding all Schedule E income from *IncomePct* does not change the results appreciably; *IncomePct* still rises after 1986, and the econometric parameter estimates are similar. The reason for the rise in wage and salary reporting is less clear. The biggest difference in the three definitions of income reported is that the unadjusted measure of total income reported declines after 1988. This is due primarily to the decline in capital gains reported during that period. Given that capital gains are not included in Personal Income, and that the timing of gains realizations is often influenced by tax law changes, I excluded it from income definitions A and B.

The trend of offsets under definition A follows a pattern quite similar to that of income reported. The main reason for the sharp rise after TRA86, however, is the large increases in standard deductions and exemptions. Although these increases imply larger filing thresholds (resulting in fewer required returns), their net effect is to increase offsets reported as a percent of Personal Income. I control for the increase in standard deductions and exemptions by including several explanatory variables. The large differences between the amount of offsets claimed under

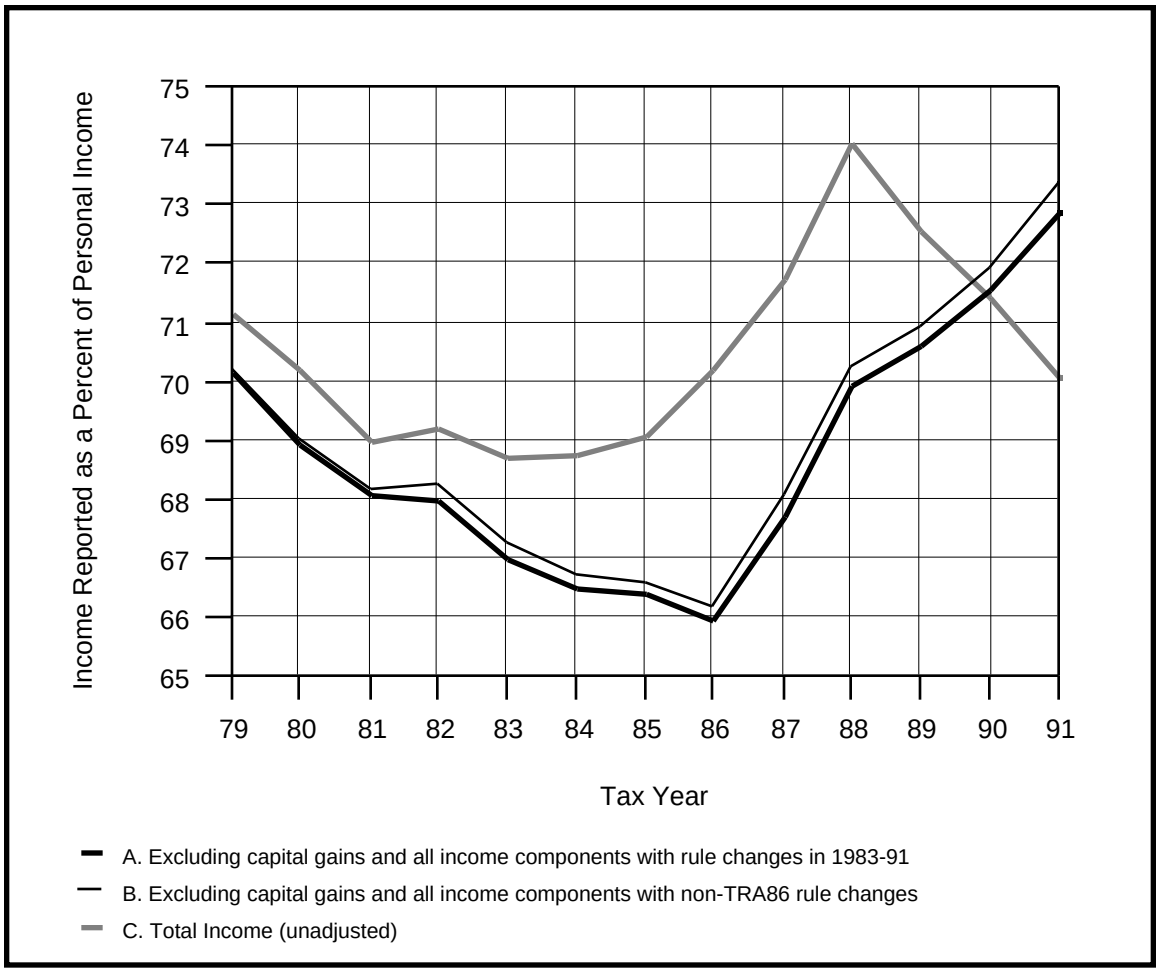


Figure 2. Income Reporting Compliance Measures: The amount of income reported on returns (per SOI), as a percent of the amount of Personal Income (per BEA), for the three definitions of income, Tax Years 1979-1991

definitions B and C compared with definition A reflect the large number of offset items excluded from definition A because the rules for what can and cannot be claimed on these items changed during the period—especially due to TRA86.

2.1.2 Potential Determinants of Voluntary Compliance

I have included over twenty variables likely to influence the four measures of voluntary compliance described above—one filing equation and three reporting equations (income, offsets, and net income). Apart from *FilingRate*, which I include in the reporting equations (since the more returns that are filed, the more income and offsets will be reported overall), these explanatory variables fall into five categories: Tax Policy, Burden/Opportunity, IRS Enforcement, IRS Responsiveness, and Demographics/Economics. These variables are defined in Table 2, and are described in greater detail below—including a discussion of my *a priori* judgment of the direction (sign) of each one's impact on compliance. Descriptive statistics (and units) for all the variables are provided in Appendix E.

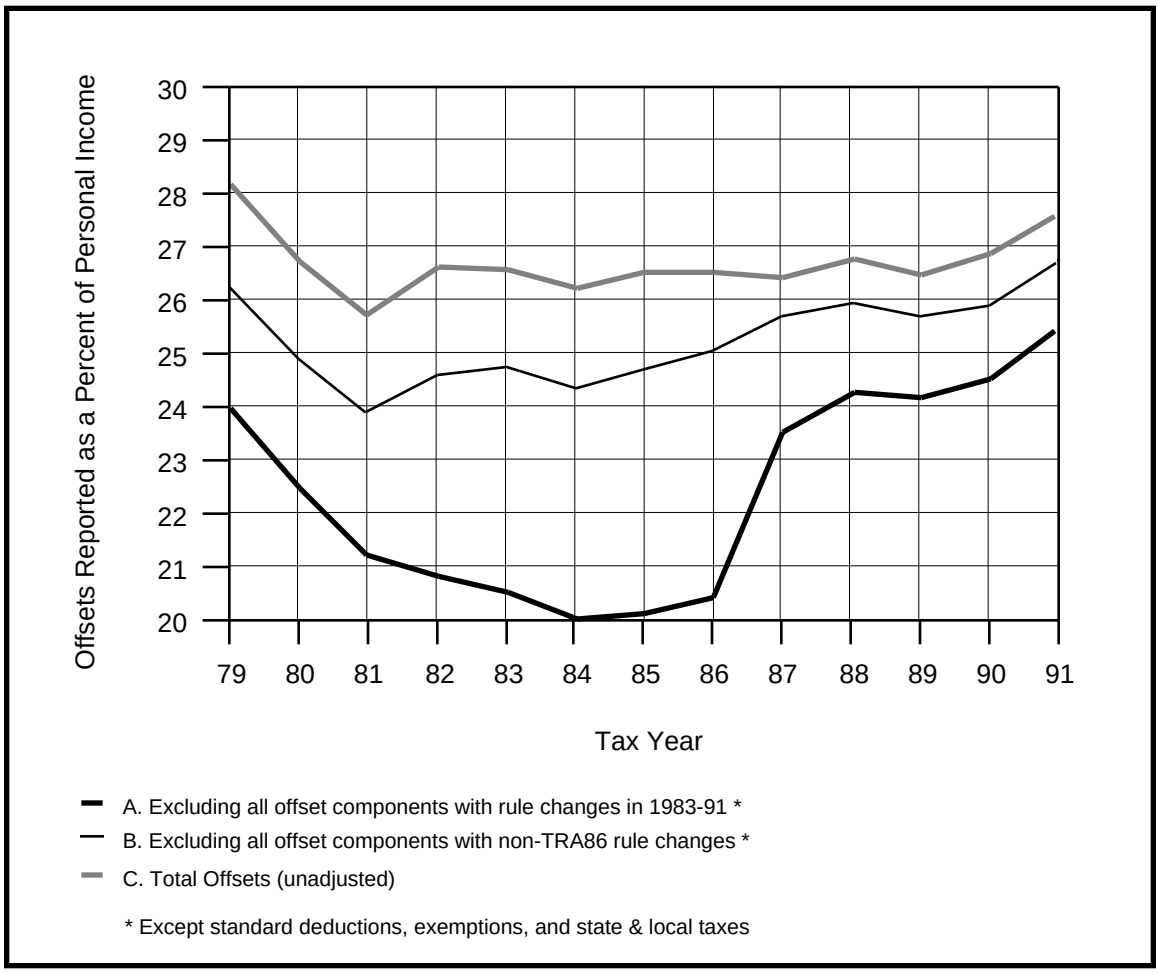


Figure 3. Offsets Reporting Compliance Measures: The amount of offsets reported on returns (per SOI), as a percent of the amount of Personal Income (per BEA), for the three definitions of offsets, Tax Years 1979-1991

Tax Policy

I control for changes in the values of standard deductions and exemptions through the use of two variables: *FThresholdPct*, which is the aggregate filing threshold value[17] among all required returns (estimated from the CPS), expressed as a percentage of Personal Income; and *ChildExemptsPct*, which is the aggregate exemption value for all children among all required returns (estimated from the CPS), also expressed as a percentage of Personal Income. Increased filing thresholds will decrease the *number* of returns filed, but it is not clear whether increased thresholds would affect the *FilingRate* one way or the other; it may depend on whether those who no longer need to file have a higher filing rate than those who are still required to file. If filing thresholds and dependent exemptions have any impact on how much income is reported (controlling for the filing rate), that impact is likely to be positive, since the larger these offsets are, the more income that can safely be reported without paying more tax. *FThresholdPct* and *ChildExemptsPct* should both have a positive impact on *OffsetsPct*, since these are two of the most important income offsets.

[17] The filing threshold for a given return is the amount of total income below which the taxpayer generally does not need to file a return. In most cases, this is defined as the sum of the standard deduction applicable to the taxpayer plus the value of personal exemptions to which he is entitled. Personal exemptions include those for the taxpayer and spouse (if filing jointly), and exclude exemptions claimed for dependents; claiming dependent exemptions requires filing a return.

Table 2. Definitions of Explanatory Variables

Variable	Numerator	Denominator
Tax Policy		
FThresholdPct	Filing threshold on required returns (CPS)	Personal Income (BEA)
Amnesty5	Dummy indicating whether state has had an amnesty in the last 5 years	
MargTaxRate@$15K	Marginal tax rate at $15K taxable income (weighted for married-single mix by state and year)	
MargTaxRate@$57K	Marginal tax rate at $57K taxable income (weighted for married-single mix by state and year)	
ChildExemptsPct	Value of exemptions for children (CPS)	Personal Income (BEA)
StateTaxPct	State income, property & sales tax revenues	Personal Income (BEA)
Burden / Opportunity		
AvgBurden	Total tax form burden on required returns (CPS)	Number of potential returns (CPS)
SoleProps	Number of sole proprietors (CPS)	Number of potential returns (CPS)
SolePropTFS	SoleProps x percentage of non-farm employment in Trade, Finance & Service sectors	
PaidPrep	No. of returns prepared by paid practitioner (SOI)	Number of returns filed (SOI)
IRS Enforcement		
AuditRate	Number of district audits started in fiscal year (AIMS)	Returns filed in prior tax year (SOI)
IRP_DocRate	No. of IRP documents matched against returns	Number of potential returns (CPS)
TDI_TotRate	Total number of TDI notices issued	Number of potential returns (CPS)
RefOffRate	Number of refunds offset for outstanding debts	Number of refunds
CID_ConvRate	Criminal convictions	Population, in millions (Census)
IRS Responsiveness		
TPS_CallsPC	Number of telephone calls handled by TPS	Population, in thousands (Census)
TPS_RetPrepPC	Number of returns prepared by TPS	Population, in thousands (Census)
Demographics / Economics		
Singles	Number of singles among potential returns (CPS)	Number of potential returns (CPS)
Under30	Number of potential returns under age 30 (CPS)	Number of potential returns (CPS)
Over64	Number of potential returns over age 64 (CPS)	Number of potential returns (CPS)
PCBirths	Number of births (HHS)	Population, in thousands (Census)
AvgPI	Personal Income (BEA)	Number of potential returns (CPS)
AvgPIgrowth	Annual growth in AvgPI	
ExclIncomePct	Income on potential returns that is not taxable	Personal Income (BEA)
UnemplRate	Unemployment rate (among those 16 and older)	

Abbreviations:

AIMS	Audit Information Management System (IRS Examination function)
BEA	Bureau of Economic Analysis, national accounts (Commerce Department)
CID	Criminal Investigation Division (IRS)
CPS	Current Population Survey (Census Bureau)
HHS	U.S. Department of Health & Human Services
IRP	Information Returns Program, document matching (IRS)
SOI	Statistics of Income (IRS)
TDI	Taxpayer Delinquency Investigation, nonfiler program (IRS Collection function)
TPS	Taxpayer Service function (IRS)

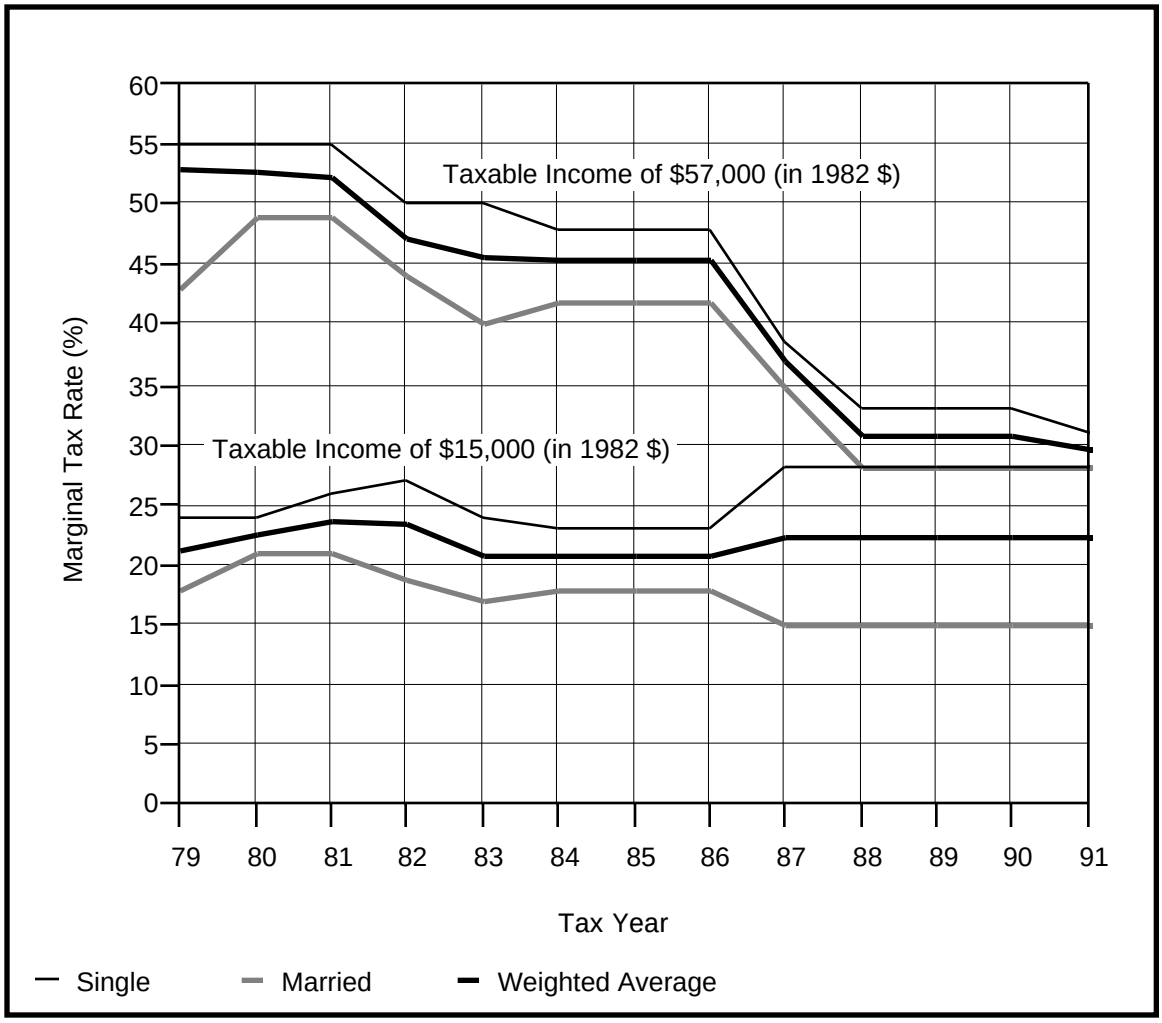


Figure 4. Marginal Tax Rates: The national trend for two taxable income levels, Singles and Marrieds Filing Jointly, and the weighted average of the two, Tax Years 1979-1991

Marginal tax rates are difficult to reflect in an aggregate analysis—especially since the tax rate schedules apply to everyone equally in a given year. Also, since the average marginal tax rate for each state population does not by itself reflect the progressivity of the tax rate schedules, or the possibility that marginal rates have different impacts on the reporting compliance of taxpayers at different income levels, and since any average marginal rate calculated from filed returns is endogenous, I developed an alternative way to estimate the impact of marginal rates. This involved determining the marginal tax rate for each year from the tax rate schedules (included in the Form 1040 tax package) at two separate levels of taxable income: $15,000 and $57,000, expressed in constant 1982 dollars.[18] These marginal rates saw significant changes over time (see Figure 4), and varied across states each year due to the widely variable mix of single and married taxpayers from state to state.[19] I constructed a composite marginal tax rate variable for each income level by weighting the single and married rates by the relative mix of singles and marrieds among potential

[18] $15,000 was chosen as a fairly modest income, and $57,000 was chosen because it is the highest level of taxable income (in 1982 dollars) for which the Single and Married-Joint marginal tax rates were different for each year. If the marginal tax rate were the same for these two rate schedules in any year, then we would not observe any variation in the variable across states for that year.

[19] The ratio of marrieds to singles in a given state typically ranged from 0.6 to 1.2.

returns (which I estimated from the CPS). (I also included the percent of singles in the state population among the demographic explanatory variables to ensure that these marginal tax rate variables are not simply surrogates for the married-single mix.) As Figure 4 shows, the marginal tax rate at $15,000 of taxable income remained relatively stable, whereas the rate at $57,000 of taxable income dropped significantly over the 1982-1991 period of my analysis. It is not completely clear *a priori* whether the effect of marginal tax rates is positive or negative; it most likely depends on how risk-averse the typical taxpayer is relative to the magnitude of other important tax parameters. Since low-income taxpayers are most likely more risk averse than higher income taxpayers, however, marginal rates are likely to have more of a positive effect on those with low income.

Two additional tax policy variables relate to state taxes. *StateTaxPct* is the amount of state income, property, and sales tax revenues that were deductible federally, expressed as a percentage of Personal Income. We should expect this to have a positive impact on *OffsetsPct. Amnesty5* is a dummy variable that indicates whether the state has had a tax amnesty[20] within the five years up to and including the year in question. Since state and federal tax administration is linked in many important ways, it is conceivable that the federal government realizes some of the short-term gains that the states have enjoyed following their amnesties. However, it is also possible that taxpayers reduce their compliance following an amnesty—either because they expect another one will follow, or because they feel they deserve to receive some of the government's leniency from which their less compliant neighbors have benefitted. Therefore, it is not clear *a priori* whether state amnesties have a positive or a negative impact on federal compliance.

Burden / Opportunity

It seems reasonable to presume that the more complex the tax system becomes, and the harder it becomes to comply with one's tax obligations, the more likely people will become noncompliant—either unintentionally due to confusion, or willfully out of frustration. On the other hand, the fewer opportunities there are to be noncompliant (e.g., with the introduction of requirements that payers of certain types of income report this information both to the recipient and to the IRS), the less noncompliance we should expect. These two factors are actually much the same: the complexity of the tax system (and therefore the burden associated with complying with it) arises to a large extent from the various opportunities left open for noncompliance (e.g., underreporting business income), and the many mechanisms in the tax system to minimize those opportunities (e.g., detailed forms and schedules, complicated rules, and lengthy instructions). Moreover, what is complexity to one (e.g., itemized deductions) may be opportunity to another.

Recognizing that business income among individuals presents both tax-paying complexities and opportunities, I have included two variables: *SoleProps,* the percent of potential returns having non-farm sole proprietor income (per the CPS); and *SolePropTFS*, an interaction term between *SoleProps* and the percent of nonfarm employment in the Trade, Finance, and Services sectors. Sole proprietors generally keep a fairly visible public profile, and thus may find it difficult to hide from the IRS's notice. However, dealing largely in cash and "moonlighting" may provide many entrepreneurs—especially in the Trade, Finance, and Services sectors—the opportunities to evade taxes by not filing or by not reporting all of their income. It is not clear whether *SoleProps* should have a positive or negative impact on compliance, but we should expect *SolePropTFS* to have a less positive, or even negative, impact compared with *SoleProps.*

---

[20] By 1991, 33 states had conducted some form of amnesty. Virtually all of these waived some or all penalties associated with nonfiling if a delinquent return were filed within a specific time period. Many amnesties also focused on accounts receivable.

The extent to which taxpayers pay others to prepare their returns has been related to compliance before, and it seems appropriate to control for it. It is not obvious, however, whether paid tax practitioners improve reporting compliance (say, by asking their clients if they had any extra income from easy-to-forget sources, or by avoiding misunderstanding as to what offsets are available to their clients, and how to calculate them), or whether they reduce the amount of net income reported (say, by pointing out legal tax avoidance strategies open to their clients).

A variable that has not been related to compliance before is the burden associated with getting, learning how to use, completing, and filing the various required tax forms and schedules. This "burden" was estimated by an IRS study conducted in response to the Paperwork Reduction Act. Estimates of the burden (measured in hours) associated with each form and schedule have been published in the Form 1040 tax package in the most recent years, and I have applied the same methodology to develop estimates for the earlier years. (These estimates are tabulated in Appendix F.) I then multiplied these burden estimates for each form and schedule by the number of corresponding forms and schedules required to be filed, which I estimated from the CPS (see Appendix F for details on the logic employed). I then derived *AvgBurden* by dividing this total burden estimate for the state and year by the corresponding number of potential returns indicated by the CPS. (I use *potential* returns instead of *filed* or *required* returns because an increase in the filing threshold could easily increase the average burden among the remaining required returns, but should be considered a burden reducer.) We should expect this burden to decrease filing compliance, as well as the reporting of offsets, since the hassle associated with maintaining the correct records and completing the paperwork can only diminish taxpayers' willingness to file returns and to claim offsets to which they might be entitled. It is not clear *a priori* what impact burden might have on income reporting compliance, though the "hassle factor" could have a negative effect.

IRS Enforcement

The primary enforcement variable thought to influence voluntary compliance is the audit coverage rate—the percentage of returns audited. Presumably, if taxpayers respond to the deterrent effect of audits, they (subjectively, at least) try to estimate their chances of being audited. This probability depends both on what they report on their returns (making the audit rate endogenous with compliance) and on the prevailing level of audit resources in their area. Traditionally (in IRS reports, and therefore in academic research), the latter concept has been expressed as an average audit coverage rate, which is usually defined as the number of audits closed in a given fiscal year divided by the corresponding number of returns filed in the prior calendar year. Since the length of audits varies widely, and is often longer than one year, the traditional coverage rate concept does not accurately reflect the average percentage of returns filed in a given year that are eventually audited. That is better captured by the percentage of audits *started* in a given year, which has never been used in an analysis of taxpayer compliance. It is possible, however, that audits send different signals to the general population when they are started compared with when they are closed. For example, the message that gets "rippled" to friends and neighbors when an audit begins presumably focuses on the *fact* of the audit, and may shape their perception of their own likelihood of getting audited. In contrast, the message communicated when an audit ends probably has more to do with the *quality* of the audit, and may shape others' perceptions more of the *consequences* of the audit (good or bad) than of its *likelihood.* Including *both* audit measures—which are obviously highly correlated—however, introduces the problem of multicollinearity into the analysis, so I have included just the audit start rate alone, since it displayed the greater predictive power.[21]

[21] My definition includes only the person-to-person audits conducted by the district offices; it excludes the simple correspondence audits conducted by the service centers.

Many have observed that over the last two decades voluntary compliance seems to have fallen concurrent with a decline in the audit coverage rate. However, as shown in Figure 5, much of the decline in conventional, labor-intensive audits has been accompanied by a very significant rise in IRS's ability to detect noncompliance through the use of automated matching of third-party information documents with tax returns in its Information Returns Program (IRP). We ought to control for this shifting of enforcement resources, since taxpayers may not perceive much of a difference between getting caught by a person and getting caught by a computer. For this reason, some recent analysis[22] has included the number of computer-generated notices to taxpayers (called CP-2000's) arising from such mismatches. I do not think that that is the correct variable, however. The deterrent effect of the IRP document matching program is achieved when taxpayers believe that virtually every mismatch will be detected and pursued. It is not the number of mismatches *found* that reflects the level of enforcement, but rather the number of documents *actuallymatched*. In fact, in recent years, the reporting of wage, salary, interest and dividend income has steadily improved while the number of mismatches has generally declined. Compliance has improved in these areas because taxpayers have increasingly understood that virtually all mismatches will be detected (a "coverage rate" approaching 100 percent); the number of mismatches has fallen as a result. Therefore, I have included the average number of IRP documents processed per potential return [*IRP_DocRate*] as an explanatory variable. (As with *AvgBurden*, this is divided by the number of *potential* returns; since IRP documents are submitted even for those who have no filing obligation, we should not compare the number of IRP documents with only the required returns.) Since one of the uses of these documents is to identify nonfilers, the more IRP documents that are processed, the greater the likelihood that a potential nonfiler will choose to file. Although the overall impact of IRP has been to improve the reporting of income, as well, it is not clear what impact new types of IRP documents have had in the recent past. Their positive deterrent may be mitigated somewhat by a "What the IRS doesn't know won't hurt them" type of mentality. That is, taxpayers could improve their reporting of IRP-covered income types, but reason that income not reported to IRS is easy to conceal. The more that taxpayers are aware of the limits of IRS knowledge, the more opportunity they have to underreport their income.

I am not aware of any other study that has attempted to measure the impact of three other IRS enforcement programs. The first, the Taxpayer Delinquency Investigations (TDI) program, is specifically targeted toward nonfilers. Based on either the presence of sufficient IRP-detected income without a return having been filed (known as the "IRP-Nonfiler" program), or on the absence of a return from someone who had filed the previous year (known as the "Stopfiler" program), IRS issues up to four TDI notices to potential nonfilers. If the notices do not yield the required returns, a more-intensive investigation may be conducted by the IRS Collection function. *TDI_TotRate* includes all such notices issued by state and year, and we should expect that this has a positive effect on the *FilingRate*.

Most of the effort of IRS's Criminal Investigation Division (CID) is intended to improve voluntary compliance by catching and prosecuting tax fraud cases. This can improve compliance in two ways: either as a deterrent among those tempted to defraud the government, or as an encouragement to the general population (to the extent that they don't want to see criminals go scott-free). It is possible, however, that the latter effect primarily works in reverse; that is, criminal convictions might not improve compliance, but the lack of them might erode compliance. Since one of the primary mechanisms for influencing the general population by CID activities is the publicity surrounding the cases—especially if convictions result—I have included *CID_ConvRate*, which I have defined as the number of criminal convictions obtained per million people in the population.

---

22 For example, an early draft of Dubin, Graetz and Wilde (1990).

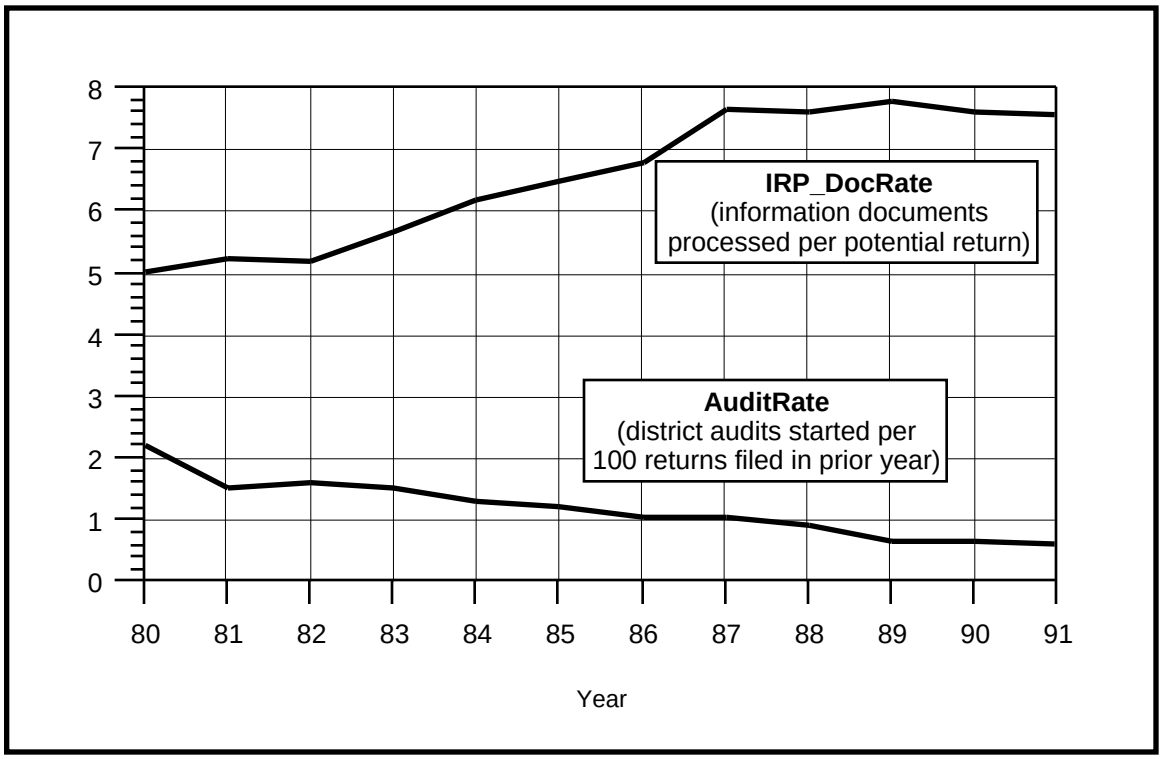


Figure 5. IRP and Audit Coverage Trends: The number of IRP documents processed per potential return, and the audit start rate, 1980-1991

The third IRS enforcement activity unique to this study involves not paying some or all of a refund to a taxpayer in order to satisfy delinquent child support payments, or some other debt. I have included *RefOffRate*, which is the percentage of refunds offset in this way. It seems likely that if a taxpayer has such a debt, he is likely to find ways to avoid being offset in the future (such as by adjusting his withholding, not filing altogether, or both) if he has been offset once, or if he hears of others getting smaller refunds than they had claimed.

IRS Responsiveness

One of the most significant contributions of this study is a better understanding of the extent to which IRS responsiveness to taxpayer needs influences voluntary compliance. To my knowledge, this has not been studied before. Unfortunately, reliable and consistent measures of such responsiveness have generally not been maintained by IRS for very long. In most cases, only one or two years worth of data are available. However, I have compiled data for two important types of Taxpayer Service (TPS) activities: the number of telephone calls handled per thousand people in the population [*TPS_CallsPC*], and the number of returns TPS helps to prepare, also per thousand people in the population [*TPS_RetPrepPC*]. We should expect *TPS_RetPrepPC* to contribute positively to both filing and reporting compliance. The impact of the telephone calls is somewhat ambiguous, however. Since these calls are almost always initiated by taxpayers, the ones who choose to call are are probably not representative of the overall population. Taxpayers call for two major reasons: seeking information about the administrative progress or status of their account, or seeking clarity on some substantive tax law issue that they face. Generally, a pleasant experience with the IRS (e.g., getting the correct answers in a reasonable amount of time) ought to contribute to higher voluntary compliance, but since the

quantity of phone calls is not always related to their *quality*, the likely impact of the calls is not straightforward. There are two additional factors that add uncertainty: (1) the people who ask the IRS substantive tax law questions are likely to call because they are generally compliant, and will tend to seek the correct answer rather than not call and take a risk, and the answers to their questions might either be in their favor or not; and (2) the number of calls answered does not always relate to the number attempted, but not answered, or to the number of times a taxpayer had to try before being served. Including both of these TPS variables, however, could shed important light on the impact that these activities have.

Demographics and Economics

It is probable that a variety of demographic and economic factors help to shape an individual's tax compliance behavior. Other factors help to control for fluctuations in the dependent variables (specifically Personal Income in the denominator of the reporting compliance measures) that are not related to compliance.

The demographic variables include the prevalence of singles as percent of potential returns [*Singles*], the percentage of potential returns in young and old age categories [*Under30* and *Over64*], and the number of births per thousand in the population [*PCBirths*]. Marital status has important tax implications, and is likely to be negatively correlated with compliance. The two age categories included tend to have less income, but it is not obvious whether they contribute to higher filing compliance. The per capita birthrate may be related to people's present satisfaction level and to their optimism about the future. If so, it may be related to better compliance.

The economic variables include: the level and rate of growth of Personal Income [*AvgPI* and *AvgPIgrowth*, respectively]; a measure of the income included in Personal Income, but excluded from taxable income [*ExclIncomePct*]; and the unemployment rate among those age 16 and older [*UnemplRate*]. *AvgPI* and *AvgPIgrowth* control for income differences across states, which must be done through the inclusion of independent variables in the *FilingRate* equation. (Recall that the three reporting compliance measures are divided by Personal Income, so including Personal Income on the right-hand side of these equations would introduce non-linearity with respect to income, which may or may not be realistic.) *ExclIncomePct* controls for fluctuations in Personal Income (and therefore in the income reporting compliance measures) related to sources of income not reported on income tax returns—such as the income of those not required to file, veterans' benefits, and child support payments—and should therefore have a negative sign. Finally, a high unemployment rate is likely to cause taxpayers to become less compliant; to the extent that they have less disposable income, they are more likely to cut corners on their taxes—or maybe not even file at all.

2.2 Estimation Approach

Three important complexities had to be addressed in order to estimate successfully the impact of these various potential determinants on voluntary compliance. The first arose from the panel structure of the data (a time series of cross sections). The second dealt with the endogeneity of audit rates. The third involved the problem of accounting for changes in the rules as to what income should be reported and what offsets could be claimed.

2.2.1 Panel Data

In the standard regression model, the unexplained error term is assumed to vary randomly across all observations. With a cross-sectional and time-series panel structure, however, certain left out (perhaps unmeasurable) variables may vary cross-sectionally (i.e., across states in this study), but not over time (such as cultural factors), while other variables may vary over time, but not across states (such as features of the tax law). The effects of these left-out variables are captured in the error term, which is no longer random across all observations. The accepted econometric approach in such cases is to assume that the error term is made up of components: one varying only across individuals (i.e., states), reflecting the "individual effect;" sometimes another component that varies over time, but not across individuals, reflecting a "time effect;" and then the usual error term, varying randomly across all observations. Pioneering work applying this technique to economic problems include Kuh (1959) in estimating investment equations, Mundlak (1961) and Hoch (1962) estimating production functions, and Balestra and Nerlove (1966) on demand functions.

The existence of such individual and time effects can be determined from a simple analysis of variance examination of the residuals obtained from the standard regression that assumes no error components. Such a test on my data reveals a very significant presence of both individual (i.e., state) and time effects.

There are two principal ways to model these error components. One is to view them as "fixed effects" for each individual (state) and time period (year). Under this assumption, the effects are not only unique to each state or time period, but they are constant, or "fixed." If a separate regression of the same equation were estimated for each state on its own, all of the estimated parameters would be identical for each state, except for the constant term; the "individual effect" would appear in combination with the overall constant term, making the constant vary across states. The other possibility is that the individual effect is random rather than constant. Under the "random effects" assumption, all components of the error term (the individual effect, the time effect, and the remaining disturbance) are distributed with a mean of zero and a constant variance. However, the variance of the individual effect is unique to each individual (state), and the variance of the time effect is unique to each time period (year), while the variance of the remaining disturbance is common across all individuals and time periods.

Fixed Effects vs. Random Effects

Which is the best way to model the error term? Two considerations have led me to choose the fixed effects model. First, the random effects approach is best suited when the panel consists of individuals tracked over time who are representative of a much larger population. But if the individuals *are* the population of interest, then the individual effects would most appropriately be fixed.[23] In this problem, the "individuals" are states, and *are* the population of interest; they are not a sample drawn from a larger population, so the fixed effects model would seem the most appropriate. Second, the random effects model may lead to biased parameter estimates if the individual effects include unobserved factors that are correlated with the included explanatory variables.[24] The usual example of such an unobserved factor is an individual's inherent "ability" (or intelligence) which is undoubtedly correlated with education (say, in an equation to explain income differences). In this study, the unobservable state effects undoubtedly include factors having to do with people's attitudes toward and perceptions about the federal government generally, and tax compliance specifically. If so, then these effects are almost by definition

[23] See Hsiao (1986), pp. 42-43, and Baltagi (1995), p. 13.

[24] See Mundlak (1978), and Hsiao (1986), pp. 43-46.

correlated with the included explanatory variables, since these variables are included for the very reason that they are thought to influence taxpayer attitudes and perceptions. For example, if IRS is not responsive to taxpayer needs (e.g., by answering telephone enquiries correctly and efficiently), then those taxpayers—and perhaps others influenced by them—may develop unobserved attitudes toward taxpaying (built on frustration, or perceived unfairness, for example) that cause them to be less compliant next time around. So, by both considerations, the fixed effects model would seem to be the most appropriate specification.

Estimation Procedure

The most straightforward way to estimate a fixed effects model is to include an overall constant term as well as dummy variables for all but one of the states and for all but one of the years. This captures the fixed effects explicitly, leaving only one error component: the standard random disturbance. I employed this procedure, which is sometimes called a Least Squares-Dummy Variable (LSDV) approach (or 2SLSDV for Two-Stage Least Squares), since it gives insight into the magnitude and distribution of the effects.

2.2.2 Endogeneity of Audit Rates

By carefully constructing my dependent and explanatory variables (for example, using the CPS to estimate filing obligations and certain economic and demographic characteristics of the potential filing population, and using marginal tax rates taken directly from the tax rate schedules), I have avoided a number of potential endogeneity problems. However, the problem of endogenous audit rates remains, and is one of the most significant estimating challenges in any comprehensive study of IRS's impact on voluntary compliance. The problem arises because not only do audit rates presumably influence taxpayers' perceptions of their chances of getting audited—and, therefore, their compliance decisions—but IRS allocates its audit resources based on its perceptions of taxpayers' noncompliance. That is, audit rates and voluntary compliance are jointly (simultaneously) determined. Since these audit resources are allocated on a district-by-district basis (where most districts encompass an entire state), this endogeneity can be expected to manifest itself in state-level data. All of the most recent attempts to estimate the determinants of voluntary compliance have included audit rates, and have recognized this endogeneity. However, I am not confident that these earlier studies have successfully identified the compliance equations in which audit rates appear. Dubin, Graetz, and Wilde (1990) employs an instrumental variables approach in which IRS budget per return filed and the the number of information returns per tax return filed are used as instruments for audit rate. Since district budgets vary in large part due to IRS perceptions of the variations in taxpayer compliance, since information documents have their own impact on voluntary compliance (particularly on filing compliance), since the number of returns filed (the denominator for each instrument) is itself endogenous, and since data on information documents were not compiled by state, these would seem to be poor instruments for the audit rate. Beron, Tauchen, and Witte (1992) uses a similar approach, including as an instrument the number of returns filed per IRS employee in a given district—in essence, the inverse of the budget per return instrument described above. This is intended to reflect the fact that district audit staffing levels are often dictated by constraints unrelated to compliance, but this variable includes *all* types of district staffing, and it is unclear whether the constraints on optimal resource allocation are significant enough to say that this variable is an effective instrument for audit rates. Since the paper does not include, for the sake of comparison, results using the endogenous audit rate in the reporting equations, we can only speculate whether the predicted audit rates based on this instrument corrected the bias of the parameter estimate, or exacerbated it.

Direction of Bias

To understand the presence and direction of this bias, consider the following generic income reporting and audit rate equations:

$$\begin{aligned} \text{IncomePct} &= \alpha_0 + \alpha_1 \text{AuditRate} + \varepsilon \\ \text{AuditRate} &= \beta_0 - \beta_1 \text{IncomePct} + \nu \end{aligned}$$

We would expect the impact of *AuditRate* on *IncomePct* (i.e., α_1) to be positive, and the impact of *IncomePct* on *AuditRate* (β_1) to be negative. If the endogeneity is not controlled for, what is the relationship between the desired parameter estimate (α_1) and the error term? A positive ε implies a larger value of *IncomePct*. The second equation, then, because the coefficient on *IncomePct* is negative, suggests that the larger value of *IncomePct* is associated with a lower value of *AuditRate*. So, a positive ε is associated with a smaller *AuditRate*. This means that the coefficient on *AuditRate* (α_1) would be biased *downward* if we estimated the equation using the endogenous *AuditRate* directly, with no correction for the endogeneity. This negative bias can be seen in state level data; many states with high audit coverage also have *low* compliance. This does not mean that audits *reduce* compliance; rather, it simply reflects that IRS intentionally allocates its audit resources where they are needed most. Using similar logic, one would expect a *positive* bias of the *AuditRate* parameter in the *OffsetsPct* equation—on the assumption that *OffsetsPct* has a positive impact on *AuditRate*. I examine these biases, and the extent to which they are corrected, in the discussion of results, section 3.2.1.

Instruments for AuditRate

The ideal instrument for *AuditRate* would be something that helps to predict *AuditRate*, but is not related to compliance. I have chosen two audit productivity-related measures to fulfill this purpose. Presumably, the more productive the auditors are in a given year, the more audits they can start in that year. And, to a lesser extent, perhaps, the more time applied per audit last year, the greater the likelihood that those audits will be completed (and new audits started) this year. Audit productivity is routinely measured in several ways, but the best definition for this purpose is probably the percent of all examiner time available that is applied to the direct examination of returns. Direct Examination Time (DET) is generally on the order of 50 percent of total time, and varies widely across districts and by year. Non-direct activities include vacations, training, travel, and an assortment of administrative duties. I have used as instruments both the DET percent and the one-year lag of the average DET per audit.

2.2.3 Alternate Definitions of Income and Offsets

As discussed earlier, the amount of income that taxpayers report is often influenced by changing rules that dictate what must be reported. For example, prior to the Tax Reform Act of 1986 (TRA86), taxpayers could exclude from total income the first $100 of dividends received ($200, on married-joint returns). After 1986, there was no such dividend exclusion. Unless this is controlled for, we might interpret the increase in reported income to be an improvement in voluntary compliance instead of a straightforward response to the new rule. This would be especially detrimental if the elimination of the dividend exclusion were correlated with other policy parameters included as explanatory variables (like marginal tax rates). Many more such rule changes have occurred over the 1982-1991 timeframe of this study—especially in the rules governing offsets.

No single approach to this problem is ideal, so I have employed a combination of techniques. Where possible, I have included explanatory variables that control for certain rule changes—such as *FThresholdPct*, *ChildExemptsPct*, and *StateTaxPct*. In the case of dividends, I was able to construct a constant-law dividend variable by applying the dividend exclusion to the post TRA86 micro data before aggregating to the state level. But more often than not, neither of these two approaches could control for specific rule changes. Therefore, I have constructed the reporting compliance variables according to three different definitions to test the sensitivity of my results to these rule changes. The basic approach (Definition A) was to exclude from total income or total offsets *any* component whose rules changed during the period, and neither of the other forms of controlling for the changes was available. The other extreme (Definition C) included all income and offsets components regardless of rule changes. A hybrid of these two (Definition B) excluded only those components of total income or total offsets whose rules changed in years other than 1986. Definition A has the advantage of being purged of such rule changes, allowing a straightforward interpretation of the parameter estimates. However, it has the drawback that it is incomplete; it says nothing about the influence of the explanatory variables on the excluded components of income and offsets. By comparing the results from all three alternatives, however, it should be possible to conclude something more definitive about the sensitivity of the results to unaccounted-for tax rule changes.

2.3 Model Specification

The model consists of four compliance equations (*FilingRate*, *IncomePct*, *OffsetsPct*, and *NetIncomePct*), and one first-stage *AuditRate* equation. They are all estimated using single-equation procedures (LSDV and 2SLSDV) to avoid the likelihood of introducing omitted variable bias across equations. The equations are estimated from the panel of 49 states (i.e., excluding Alaska, as discussed earlier) over ten years (1982-1991), giving 490 observations for each variable.

2.3.1 Functional Form

Many of the explanatory variables can be expected to have a non-linear effect on compliance, reflecting diminishing returns to IRS effort, for example. When this is plausible, I have expressed the independent variables as logarithms.[25] For such variables that frequently take on values between zero and one, or values near one, I have used the logarithm of one plus the variable. Otherwise, all variables are modeled linearly.

2.3.2 First-Stage AuditRate Equation

It is both difficult and unnecessary to estimate a structural equation for *AuditRate*. We need only estimate a first-stage *AuditRate* equation using appropriate instruments, and then use the predicted *AuditRate* in the structural compliance equations. My specification for the *AuditRate* equation, therefore, was as follows:

$$\text{AuditRate} = \delta_0 + \Sigma\sigma_{1i}\text{State}_i + \Sigma\tau_{1t}\text{Year}_t + \delta_1\text{DET\%} + \delta_2\text{Ln}(\text{AvgDET}_{-1}+1) + \varepsilon_1$$

As with all of the specifications that follow, there are 48 State dummies (i = 1 to 48) and 9 Year dummies (t = 1 to 9), and all of the state and time subscripts on the other variables are omitted for simplicity.

---

[25] I have used *natural* logarithms. However, since the instruction for natural logs in my econometric software is simply "log," I inadvertently thought I had been using base 10 logs in prior versions of this report. Although this did not affect the econometric results, it does change any calculations based on those results. This version of the report includes the corrected nomenclature and calculations. For example, see Figure 6, Table 5, and Appendices G, H and I.

2.3.3 FilingRate Equation

Many of the determinants of reporting compliance are not relevant to filing compliance, and vice-versa. I used the following specification for the *FilingRate* equation:

$$\begin{aligned}\text{FilingRate} = \; & \varphi_0 + \Sigma\sigma_{2i}\text{State}_i + \Sigma\tau_{2t}\text{Year}_t + \varphi_1\text{FThresholdPct} + \varphi_2\text{Amnesty5} \\ & + \varphi_3\text{Ln(AvgBurden)} + \varphi_4\text{SoleProps} + \varphi_5\text{SolePropTFS} \\ & + \varphi_6\text{IRP\_DocRate} + \varphi_7\text{Ln(TDI+1)} + \varphi_8\text{Ln(RefOffRate+1)} \\ & + \varphi_9\text{TPS\_RetPrepPC} \\ & + \varphi_{10}\text{Singles} + \varphi_{11}\text{Under30} + \varphi_{12}\text{Over64} \\ & + \varphi_{13}\text{AvgPI} + \varphi_{14}\text{AvgPIgrowth} + \varphi_{15}\text{UnemplRate} + \varepsilon_2\end{aligned}$$

The explanatory variables unique to the *FilingRate* equation are: *Amnesty5*, since the common characteristic of all the various state amnesties was that they targeted nonfilers; Ln(*TDI*+1), since the TDI program is only targeted to nonfilers; Ln(*RefOffRate*+1), since the incidence of refund offsets is likely to induce some nonfiling, but is likely to manifest itself in reporting compliance (if at all) through smaller amounts of tax withheld, which is not reflected in any of my reporting compliance variables; and *AvgPI* and *AvgPIgrowth*, in order to control for income variations across states and over time (this is handled in the reporting compliance equations by dividing the dependent variables by Personal Income). In addition, *IRP_DocRate* is included as a linear variable here, whereas it enters the income reporting equations in logarithm form. This is because the filing decision is an either-or choice; more IRP documents would induce more people to file—not greater filing compliance among those who already file.

The other variables of primary interest in this equation are: *FThresholdPct*, which should indicate whether raising the filing threshold increases or decreases filing compliance; Ln(*AvgBurden*), which is likely to diminish filing compliance; and *TPS_RetPrepPC*, which we would expect to improve filing compliance.

2.3.4 IncomePct Equation

The income and offset reporting equations include *FilingRate* as an explanatory variable, but this is not endogenous. This is because, although we would expect that greater filing compliance (i.e., more people filing) would increase the aggregate amounts of income and of offsets reported, the amounts reported on filed returns do not affect whether or not people file returns. The specification of the *IncomePct* equation, then, is as follows:

$$\begin{aligned}\text{IncomePct} = \; & \alpha_0 + \Sigma\sigma_{3i}\text{State}_i + \Sigma\tau_{3t}\text{Year}_t + \alpha_1\text{FilingRate} \\ & + \alpha_2\text{FThresholdPct} + \alpha_3\text{MargTaxRate@\$15K} + \alpha_4\text{MargTaxRate@\$57K} \\ & + \alpha_5\text{ChildExemptsPct} \\ & + \alpha_6\text{Ln(AvgBurden)} + \alpha_7\text{SoleProps} + \alpha_8\text{SolePropTFS} + \alpha_9\text{PaidPrep} \\ & + \alpha_{10}\text{Ln(pAuditRate+1)} + \alpha_{11}\text{Ln(IRP+1)} + \alpha_{12}\text{Ln(CID+1)} \\ & + \alpha_{13}\text{TPS\_CallsPC} + \alpha_{14}\text{TPS\_RetPrepPC} \\ & + \alpha_{15}\text{Singles} + \alpha_{16}\text{Under30} + \alpha_{17}\text{Over64} \\ & + \alpha_{18}\text{PCBirths} + \alpha_{19}\text{ExclIncomePct} + \alpha_{20}\text{UnemplRate} + \varepsilon_3\end{aligned}$$

Like the other two reporting compliance equations, this equation includes several variables—in addition to the *FilingRate* variable itself—that are not in the *FilingRate* equation. Additional enforcement variables include Ln(*pAuditRate*+1) (the predicted results from the first-stage *AuditRate* equation), and Ln(*CID*+1); additional tax policy variables include *MargTaxRate@$15K*, *MargTaxRate@$57K*, and *ChildExemptsPct* (which controls for the changing value of dependent exemptions); *PaidPrep* is an additional Burden/Opportunity variable, which is relevant only to filers; and *TPS_CallsPC* is an additional IRS Responsiveness variable, which seems to be mostly relevant to filers. The reporting compliance equations also include *PCBirths*, which does not appear to affect *FilingRate* significantly. Finally, as pointed out earlier, Ln(*IRP*+1) is used for the income reporting equations instead of *IRP_DocRate* in linear form.

The *IncomePct* equation has two variables not included in the *OffsetsPct* equation: Ln(*IRP*+1), since virtually all information documents report income rather than offsets; and *ExclIncomePct*, which accounts for types of income included in Personal Income that do not need to be reported on tax returns—either because the income falls below the filing threshold (and no return is required), or the type of income is not taxable.

The explanatory variables of primary interest in the *IncomePct* equation are: Ln(*pAuditRate*+1), *FThresholdPct*, the two *MargTaxRate* variables, Ln(*AvgBurden*), Ln(*IRP*+1), Ln(*CID*+1), and the two *TPS* variables. These represent the most important tax policy and tax administration variables that could conceivably be manipulated so as to foster better voluntary compliance.

2.3.5 OffsetsPct Equation

This equation is much like the income reporting equation, but with some important differences. The specification is as follows:

$$
\begin{aligned}
\text{OffsetsPct} = \; & \beta_0 + \Sigma\sigma_{4i}\text{State}_i + \Sigma\tau_{4t}\text{Year}_t + \beta_1\text{FilingRate} \\
& + \beta_2\text{FThresholdPct} + \beta_3\text{MargTaxRate@\$15K} + \beta_4\text{MargTaxRate@\$57K} \\
& + \beta_5\text{ChildExemptsPct} + \beta_6\text{StateTaxPct} \\
& + \beta_7\text{Ln(AvgBurden)} + \beta_8\text{SoleProps} + \beta_9\text{SolePropTFS} + \beta_{10}\text{PaidPrep} \\
& + \beta_{11}\text{Ln(pAuditRate+1)} + \beta_{12}\text{Ln(CID+1)} \\
& + \beta_{13}\text{TPS\_CallsPC} + \beta_{14}\text{TPS\_RetPrepPC} \\
& + \beta_{15}\text{Singles} + \beta_{16}\text{Under30} + \beta_{17}\text{Over64} \\
& + \beta_{18}\text{PCBirths} + \beta_{19}\text{UnemplRate} + \varepsilon_4
\end{aligned}
$$

As discussed above, the *OffsetsPct* equation does not include two of the variables included in the *IncomePct* equation: Ln(*IRP*+1), and *ExclIncomePct*. However, it includes one other variable not in the *IncomePct* equation: *StateTaxPct*, which controls for variations in the amount of state and local taxes that are deductible federally (and therefore contribute to offsets).

The same tax policy and tax administration variables are of primary interest in the *OffsetsPct* equation (to the extent that they are included). However, some of these variables are likely to affect offsets and income in opposite ways, while others have the same kind of impact.

2.3.6 NetIncomePct Equation

The final reporting equation represents income minus offsets, and so includes all of the explanatory variables included in either of the two separate reporting equations, as follows:

$$\begin{aligned}\text{NetIncomePct} = {} & \gamma_0 + \Sigma\sigma_{5i}\text{State}_i + \Sigma\tau_{5t}\text{Year}_t + \gamma_1\text{FilingRate} \\ & + \gamma_2\text{FThresholdPct} + \gamma_3\text{MargTaxRate@\$15K} + \gamma_4\text{MargTaxRate@\$57K} \\ & + \gamma_5\text{ChildExemptsPct} + \gamma_6\text{StateTaxPct} \\ & + \gamma_7\text{Ln(AvgBurden)} + \gamma_8\text{SoleProps} + \gamma_9\text{SolePropTFS} + \gamma_{10}\text{PaidPrep} \\ & + \gamma_{11}\text{Ln(pAuditRate+1)} + \gamma_{12}\text{Ln(IRP+1)} + \gamma_{13}\text{Ln(CID+1)} \\ & + \gamma_{14}\text{TPS\_CallsPC} + \gamma_{15}\text{TPS\_RetPrepPC} \\ & + \gamma_{16}\text{Singles} + \gamma_{17}\text{Under30} + \gamma_{18}\text{Over64} \\ & + \gamma_{19}\text{PCBirths} + \gamma_{20}\text{ExclIncomePct} + \gamma_{21}\text{UnemplRate} + \varepsilon_5\end{aligned}$$

This equation represents the "bottom line" of reporting compliance; it differs from *tax* reporting compliance only because the amount of tax reported reflects the application of the tax rate schedule to net income. (Recall that I have defined offsets—and therefore net income—to include the income-offset value of credits, which are tax offsets. Therefore, this *net* income is not synonymous with *taxable* income, as reported on tax returns.)

Even though the *NetIncomePct* equation alone cannot provide the insight into the method of noncompliance (i.e., underreporting income or overstating offsets) that the other two equations provide, it does serve two useful purposes. First, when we are interested in *tax* compliance, *NetIncomePct* is the most direct way to estimate it. Second, it provides a useful check on the other two reporting equations; we should expect that when a variable is included in all three equations, the coefficient in the *NetIncomePct* equation should be roughly equal to the coefficient in the *IncomePct* equation minus the coefficient in the *OffsetsPct* equation. This should especially be true when the coefficients are all significant, and can help to evaluate the results when one of the coefficients is not significant.

3. Results

The principal results for all four compliance equations are summarized in Table 3. These results pertain to the most restricted definition of the dependent variables *IncomePct*, *OffsetsPct*, and *NetIncomePct* (Definition A), which excludes all income and offset components for which the rules changed during the 1982-1991 period as to what should be reported—unless the rule change could be reflected in the data (as with the dividend exclusion), or the change could be controlled for with explanatory variables (as with the value of standard deductions and exemptions, and the deductibility of state and local income taxes). These results are discussed below, first with respect to filing compliance, and then with respect to reporting compliance. In the case of reporting compliance, all three equations are discussed concurrently. For convenience, the explanatory variables in Table 3 are grouped in the five major categories (Tax Policy, Burden/Opportunity, Enforcement, IRS Responsiveness, and Demographics/Economics) used in the discussion.

Following this discussion of the principal results is a section that compares the results for all three definitions of the reporting compliance variables (Definitions A, B, and C), focusing on the sensitivity of the results to rule changes in the 1982-1991 time period. The fourth section describes a number of additional variables that were tested as potential determinants of voluntary

compliance, but were found not to have any significant impact. Finally, section 3.5 discusses the estimated relative merits of expanding the five IRS activities found to have a positive impact on voluntary compliance.

3.1 Filing Compliance

3.1.1 Impact of Tax Policy Parameters

Of the two policy variables included, only the filing threshold (the sum of one's standard deduction and personal exemptions) significantly influences the *FilingRate*. The impact is strongly negative, which probably reflects two phenomena: first, as the filing threshold increases, some people who are still required to file stop doing so (perhaps out of confusion); and second, low-income people (who are most affected by a change in the filing threshold) may exhibit higher filing compliance than those with higher incomes—raising the overall *FilingRate* while they are required to file, but causing a drop in that rate when they no longer need to file. The latter possibility is consistent with the strongly negative impact of *AvgPI* on *FilingRate*. State tax amnesties apparently have no significant impact on federal filing compliance, but the results suggest that they may have a weak positive influence.

3.1.2 Impact of Burden / Opportunity

All of the included variables significantly affect filing compliance—at least at the 10 percent level of significance. The burden (in hours) associated with the various tax forms and schedules seems to reduce the *FilingRate*, as one might expect. This seems to confirm IRS concerns that as the forms get more numerous and complex, requiring more time to maintain records and to complete the paperwork, more people decide to forget about filing altogether. The coefficients on the two sole proprietor variables suggest that sole proprietors, generally, improve the *FilingRate*—except for those within the Trade, Finance, and Service sectors, who have a strong negative impact on filing compliance.[26] This may reflect the fact that businesses typically have multiple "paper trails," making it hard for them to hide from the IRS, but that these three special sectors tend to be associated with the "underground economy," including cash-based businesses and "moonlighting."

3.1.3 Impact of Enforcement Activities

Both the matching of third-party information documents (*IRPDocRate*) and the issuance of TDI nonfiler notices (Ln(*TDI*+1)) provide a fairly strong deterrent against nonfiling, as we would expect. The more information that is provided to the IRS by the payers of income, the more people will file required returns, since it is harder for them to hide. Furthermore, when IRS uses this information (and prior filing patterns) to issue TDI notices to presumed nonfilers, the general population seems to respond with a higher *FilingRate* than it would otherwise. The two programs apparently complement each other nicely to promote the filing of required returns.

Refund offsets seem to have a negative impact on filing compliance, but it is too weak to be considered significant. This suggests that some taxpayers may stop filing in order to avoid paying the debts being addressed by these offsets, but the most likely response—if any—seems to be to adjust their withholding to minimize their refunds (a phenomenon that I cannot verify with a model focused solely on filing and reporting compliance).

---

[26] Note that *SolePropTFS* is actually an interaction term, the product of *SoleProps* (the number of proprietors as a percentage of all potential returns) times *TFSEmplPct* (Trade, Finance, and Service employment as a percent of total nonfarm employment). It seems reasonable to view the interaction term, though, as representing the relative concentration of proprietors in these three sectors.

Table 3. Determinants of Voluntary Filing and Reporting Compliance, Definition A *

	Equation			
Explanatory Variables	FilingRate	IncomePct	OffsetsPct	NetIncomePct
FilingRate		0.345586 (7.64)	0.137683 (8.72)	0.207853 (5.57)
FThresholdPct	-3.569438 (-8.31)	1.182627 (4.04)	0.857427 (8.33)	0.339758 (1.40)
Amnesty5	0.207335 (0.67)			
MargTaxRate@$15K		1.221297 (1.17)	-0.663976 (-1.80)	1.921272 (2.22)
MargTaxRate@$57K		-1.978458 (-1.00)	0.530545 (0.76)	-2.442911 (-1.50)
ChildExemptsPct		1.475395 (1.86)	0.457696 (1.63)	1.000080 (1.52)
StateTaxPct			0.145114 (1.99)	-0.101522 (-0.59)
Ln(AvgBurden)	-11.929189 (-1.78)	3.383676 (0.55)	-3.550471 (-1.77)	4.888900 (0.96)
SoleProps	1.953925 (2.44)	1.428688 (1.84)	0.274123 (0.98)	1.169128 (1.77)
SolePropTFS	-3.414896 (-2.30)	-2.925128 (-2.02)	-0.527449 (-1.01)	-2.399908 (-1.95)
PaidPrep		-0.166282 (-4.81)	-0.014858 (-1.23)	-0.153009 (-5.36)
Ln(pAuditRate+1)		16.158539 (3.37)	3.313904 (2.00)	13.892113 (3.46)
IRP_DocRate	1.565057 (2.71)			
Ln(IRP+1)		-1.121633 (-0.23)		0.675205 (0.16)
Ln(TDI+1)	3.850765 (1.81)			
Ln(RefOffRate+1)	-0.873704 (-1.13)			
Ln(CID+1)		0.932191 (3.08)	0.314909 (2.96)	0.593380 (2.37)
TPS_CallsPC		-0.003994 (-1.34)	-0.000742 (-0.71)	-0.003378 (-1.36)
TPS_RetPrepPC	0.146118 (2.18)	0.130914 (2.07)	-0.007756 (-0.35)	0.136453 (2.61)
Singles	-0.551763 (-5.77)	0.266954 (1.46)	0.072872 (1.14)	0.190927 (1.26)
Under30	0.186049 (1.95)	-0.098600 (-1.10)	-0.008269 (-0.26)	-0.091372 (-1.23)
Over64	0.242260 (2.40)	-0.075873 (-0.76)	-0.042744 (-1.23)	-0.022223 (-0.27)
PCBirths		0.991262 (4.55)	0.253864 (3.58)	0.734990 (4.02)
AvgPI	-0.920930 (-4.32)			
AvgPIgrowth	0.192294 (3.68)			
ExclIncomePct		-0.642278 (-1.60)		-0.917979 (-2.77)
UnemplRate	-0.200099 (-1.50)	-0.370384 (-2.97)	0.078118 (1.84)	-0.428625 (-4.14)
Adj. R-Squared	0.627007	0.757573	0.919223	0.801260

* 2SLSDV estimates (just LSDV for the FilingRate equation) from state-level panel data for 1982-1991;
t-statistics in parentheses; variables in **bold** are the primary tax policy and tax administration parameters of interest.

3.1.4 Impact of IRS Responsiveness

The effort of IRS's Taxpayer Service (TPS) function to help taxpayers prepare their returns (*TPS_RetPrepPC*) seems to have a significant, positive impact on the overall *FilingRate*. Specifically, we can interpret the estimated coefficient to mean that for every additional seven returns prepared by TPS per thousand of population,[27] the *FilingRate* will increase by one percentage point. This suggests that some of the induced returns may be ones prepared by TPS, but also that some of the new filers are influenced indirectly by the TPS outreach.

3.1.5 Impact of Demographic and Economic Variables

Five of the six demographic and economic control variables included have significant and intuitive effects on the *FilingRate*. A greater concentration of singles within a state is strongly associated with *lower FilingRates*. This may reflect less of a requirement to file among singles, a misunderstanding as to when dependents need to file their own returns (especially now with relatively low filing thresholds for dependents), or—like auto insurance companies have learned—that singles are typically less careful than their married counterparts. However, higher concentrations of the population in the under-30 and over-64 age categories each seem to improve filing compliance. This may reflect fewer opportunities among the young and the old to avoid IRS notice.

As noted earlier, income seems to be strongly and negatively associated with filing compliance. Specifically, every additional thousand (1992) dollars of average Personal Income (*AvgPI*) is associated with a drop in the *FilingRate* of almost one percentage point. This suggests that those who are able to hide from IRS entirely can conceal significant amounts of income—and as long as it works, they'll find ways to make more such income—while those who contribute little to aggregate Personal Income either have no requirement to file, or derive most of their income from wages and interest, which are hard to hide from IRS.

Of the two variables reflecting the state of the economy, only the rate of real income growth (*AvgPIgrowth*) has a strongly positive impact on *FilingRate*, which we would expect. As real income increases in the general population, fewer people are tempted to cut corners by not filing a tax return. By the same token, as the unemployment rate increases, filing compliance seems to decline. While this result is intuitive, however, it does not appear to be significant.

3.2 Reporting Compliance

Before discussing the impact that specific variables have on voluntary reporting compliance, it is necessary to ensure that these equations are appropriately identified, given the endogeneity of audit rates.

3.2.1 Identification

As mentioned earlier, the first-stage *AuditRate* equation is a function of two audit variables related to productivity: *DET%* and *AvgDET*, where DET (Direct Examination Time) is the time that auditors apply directly to the examination of returns, as opposed to the time they spend on leave, in training, or performing various administrative duties. The specification for this equation is as follows:

---

[27] Note that the coefficient (0.146) is approximately one-seventh.

$$\text{AuditRate} = \delta_0 + \Sigma\sigma_{1i}\text{State}_i + \Sigma\tau_{1t}\text{Year}_t + \delta_1\text{DET\%} + \delta_2\text{Ln}(\text{AvgDET}_{-1}+1) + \varepsilon_1$$

In words, the percentage of returns subject to a *started* audit this year is a function of the percentage of their time that auditors apply directly to such audits (*DET_Pct*) and the logarithm of one plus the average amount of time spent directly on audits last year. The more time that auditors are able to apply to examinations this year ought to increase the number of examinations conducted, suggesting that δ_1 ought to be positive. Similarly, if more time was applied to the average audit last year, more audits should be closer to being completed—and new audits started—this year, suggesting that δ_2 also ought to be positive. The actual results for this regression (with t-statistics in parentheses) are as follows:

$$\underset{(1.84)}{\delta_1 = 0.007570}, \qquad \underset{(0.58)}{\delta_2 = 0.057984}, \qquad \text{Adj. } R^2 = 0.766978$$

Except for the fact that the impact of the prior year's *AvgDET* is not statistically significant, these results seem encouraging.

There are two indications that the reporting equations estimated using *AuditRate* predicted from this first-stage regression (*pAuditRate*) are identified. First, when we compare the coefficient on the predicted (and presumed exogenous) Ln(*pAuditRate*+1) from these reporting equations with what is obtained using same specifications, but using the actual (endogenous) *AuditRate* (i.e., Ln(*AuditRate*+1)), we see that the two-stage approach corrects the downward bias anticipated in the coefficients in both the *IncomePct* equation and the *NetIncomePct* equation (see section 2.2.2 above). These comparisons are given in Table 4 for each of the reporting

Table 4. Comparison of AuditRate Coefficients in the Three Reporting Compliance Equations, Estimated Using the Endogenous and Exogenous AuditRate Variables

	Definition A	Definition B	Definition C
IncomePct Equation			
Ln(AuditRate+1) [endogenous]	-0.188470 (-0.17)	-0.234193 (-0.21)	0.747762 (0.54)
Ln(pAuditRate+1) [exogenous]	16.158539 (3.37)	15.823423 (3.30)	17.100229 (2.79)
OffsetsPct Equation			
Ln(AuditRate+1) [endogenous]	-0.014032 (-0.04)	0.400095 (0.91)	0.468282 (0.93)
Ln(pAuditRate+1) [exogenous]	3.313904 (2.00)	0.838515 (0.44)	-0.525384 (-0.24)
NetIncomePct Equation			
Ln(AuditRate+1) [endogenous]	-0.164486 (-0.18)	-0.600534 (-0.69)	0.505934 (0.38)
Ln(pAuditRate+1) [exogenous]	13.892113 (3.46)	15.751865 (4.11)	17.449324 (2.97)

t-statistics in parentheses. All equations specified as in Table 3. For a complete tabulation of these results, see Appendix C.

compliance equations for all three definitions of the dependent variables. For both the *IncomePct* and *NetIncomePct* equations, assuming that the *AuditRate* is exogenous yields parameter estimates that are very small and very insignificant. Accounting for the endogeneity of *AuditRate*, however, yields parameter estimates that are larger (as anticipated) and highly significant.

Notice, however, that the same pattern is true with the *OffsetsPct* equation—even though we anticipated that the coefficient on *AuditRate* would be biased *upward*, and that the corrected coefficient would be negative. Understanding this counter-intuitive result first requires making two observations: first, the estimated impact of Ln(*pAuditRate*+1) on *OffsetsPct* is much less than its impact on *IncomePct*; and second, the estimated impact of Ln(*pAuditRate*+1) on *NetIncomePct* (shown as about 13.89 in Table 4) is very significant, and is the logical combination of the separately-estimated coefficients in the *IncomePct* and *OffsetsPct* equations (shown as 16.16 and 3.31, respectively in Table 4—their difference being 12.85). All the results are significant and internally consistent, but why the unanticipated sign on the *AuditRate* parameter in the *OffsetsPct* equation? The logical explanation seems to lie in the fact that the claiming of offsets is not a simple matter. For example, some offsets, such as medical expenses and miscellaneous deductions, are specifically limited by the amount of Adjusted Gross Income reported on the return; as more income is reported, we should expect more to be claimed for these types of offsets. Likewise, most credits are not refundable; they are limited by the amount of tax due. If more income (and, therefore, more tax) is reported, we should expect more to be claimed as credits. It may also be true that if taxpayers feel compelled to report more income (e.g., in response to an increased *AuditRate*), they may seek to find additional offsets to reduce the bite somewhat. This may be especially true if (as one might expect) taxpayers perceive that they may become audit targets if the offsets they claim seem out of line to the IRS with respect to their income; so, if they do not report all of their income, they may consciously avoid claiming all of their potential offsets. The fact that the impact of *AuditRate* on *OffsetsPct* is much smaller than its impact on *IncomePct* may also mean that audits have some negative impact on the amount of offsets claimed, after all, but that this effect is more than compensated for by the phenomena just described. In any event, if the net impact of *OffsetsPct* on the *AuditRate* is actually negative (as is the impact of *IncomePct*), then the coefficient on *AuditRate* in the *OffsetsPct* equation is biased downward, rather than upward, and the 2SLS estimates correct this bias. This reasoning suggests that *OffsetsPct* is dependent, in part, on *IncomePct*—a possibility not controlled for in this specification. Not controlling for this does not affect the results of either the *IncomePct* equation or the *NetIncomePct* equation (which represents the "bottom line" results); in fact, estimating the three separate reporting equations allows us to identify such dependencies. As long as we are able to interpret the results for the *OffsetsPct* equation with this dependency in mind, we can make the correct inferences.

The second fact consistent with the proposition that the reporting equations are identified is the result of specification tests for endogeneity based on Hausman (1978). This procedure tests for statistically significant differences between a regression that assumes endogeneity and an otherwise identical one that does not. The results of this test are given for each pair of such specifications in the tables in Appendix C. In every important case,[28] the test statistic is strongly significant, which is consistent with notion that *AuditRate* is endogenous, and that using *pAuditRate* corrects for that endogeneity. Moreover, experimentation with alternate specifications for the first-stage *AuditRate* equation yielded no alternatives that produced Hausman test statistics as strongly significant as these. This, together with the observed bias correction, suggests that the reporting equations are identified.

---

[28] The only exceptions to this are the OffsetsPct equation under Definitions B and C. This is consistent with the drawbacks of these two definitions.

3.2.2 Impact of Tax Policy Parameters

Now let us turn to discussing the the impact of the various determinants of reporting compliance, as summarized in Table 3. As expected, *FilingRate* positively and significantly influences the reporting of both income and offsets; as more people file returns, they obviously report some of each.

More interesting are the effects of the five tax policy parameters included in the three reporting equations. First, having controlled for the *FilingRate*, an increase in the filing threshold seems to increase the reporting of both income and offsets significantly. However, since *FThresholdPct* has a strongly negative impact on *FilingRate*, which is included in the reporting equations, the net impact of *FThresholdPct* must take this into account. For example, a one percentage point increase in *FThresholdPct* decreases the *FilingRate* by roughly 3.57 percentage points, which, in turn, decreases *IncomePct* by about 1.23 percentage points (-3.57 times 0.35, the coefficient on *FilingRate* in the *IncomePct* equation). So, the increase in *FThresholdPct* has virtually *no* net impact on *IncomePct*; the apparent increase of 1.18 is just offset by the reduction of 1.23 due to a decline in the *FilingRate*. Similar arithmetic reveals that the *net* impact of a one percentage point increase in *FThresholdPct* is an increase in *OffsetsPct* of 0.37 and a decrease in *NetIncomePct* of 0.40. (The fact that these two numbers are opposite and roughly equal is consistent with the net impact on *IncomePct* being zero. That is: $0.00 - 0.37 \approx -0.40$). So, an increase in the filing threshold—controlling for its impact on *FilingRate*—increases the amount of offsets claimed (as we would expect), but has no effect on income reported.

The influence of marginal tax rates on the voluntary reporting of income and offsets is even more interesting. An increase in the marginal tax rate for those with $15,000 of taxable income increases *IncomePct* by 1.22 and decreases *OffsetsPct* by 0.66, resulting in an increase in *NetIncomePct* of 1.92, which is strongly significant. In other words, low-income taxpayers seem to respond to an increase in marginal rates by becoming more compliant. This is consistent with the Allingham and Sandmo (1972) prediction when the positive "income effect" is stronger than the negative "substitution effect" (see section 1.2 above). This is most likely to be true for low-income people, since they are typically more risk-averse than those with higher incomes, and since the penalties for underreporting their income would represent a much larger share of their net income. The impact of marginal tax rates on those with $57,000 of taxable income appears to be in the opposite direction (i.e., decreasing the reporting of income and net income, and increasing the amount of offsets claimed), but the effect is only marginally significant. A significant impact would have meant that the positive "income effect" is weaker than the negative "substitution effect," meaning that these taxpayers would, on balance, increase the amount of income they underreport if the marginal tax rate increased, since each dollar of underreported income is now worth more (in terms of tax savings) than before. The fact that the results are not very significant suggests that the "income" and "substitution" effects almost evenly balance each other at this income level. It would be nice to be able to estimate the impact of marginal rates at some larger income (hypothesizing that they would have a significant negative impact on reporting compliance), but the changes in marginal rates over the 1982-1991 time period and the panel structure of my data preclude testing this hypothesis at larger incomes.[29] However, even without

[29] Recall that the two *MargTaxRate* variables were constructed as weighted averages of the Married and Single tax rates for the income level in question, where the weights are the percentage of marrieds and singles in the potential filing population by state and year. This only works if the Married and Single tax rates are different; if they are the same in any year, then there will be no variation in the tax rate variable across states for that year, making interpretation of the results much more difficult. Unfortunately (for this purpose, at least), the Tax Reform Act of 1986 drastically reduced the number of marginal rate brackets, so that incomes over the $57,000 (in 1982 dollars) used here typically face the same marginal rate regardless of marital status; a marginal tax rate variable defined in this way at a much larger income level would exhibit virtually no variation after 1986, and would essentially be a dummy variable for all those post-TRA86 years. Since I already include dummy variables for each of those years separately to account for the fixed time effects, it would be difficult to interpret the results of such a specification.

being able to test that hypothesis fully, these results suggest that the impact of marginal tax rates is different at different income levels, and that reporting compliance might be improved with a flatter rate schedule.

Both of the variables included to control for changes over time in the value of important offsets seem to affect *OffsetsPct*, as expected. The impact of *ChildExemptsPct* is larger than that of *StateTaxPct*, but is not as significant. The smaller coefficient on *StateTaxPct* makes sense, since not all state taxes are claimed as offsets; only taxpayers who itemize deductions claim them. Somewhat surprising, however, is the finding that *ChildExemptsPct* significantly affects *IncomePct* also.[30] This may simply suggest that taxpayers with children tend to have more income, but it may also mean that they report more of their income. In any case, the coefficient of 1.00 on *ChildExemptsPct* in the *NetIncomePct* equation suggests that the net effect of dependent exemptions is to increase the reporting of *net* income dollar for dollar.

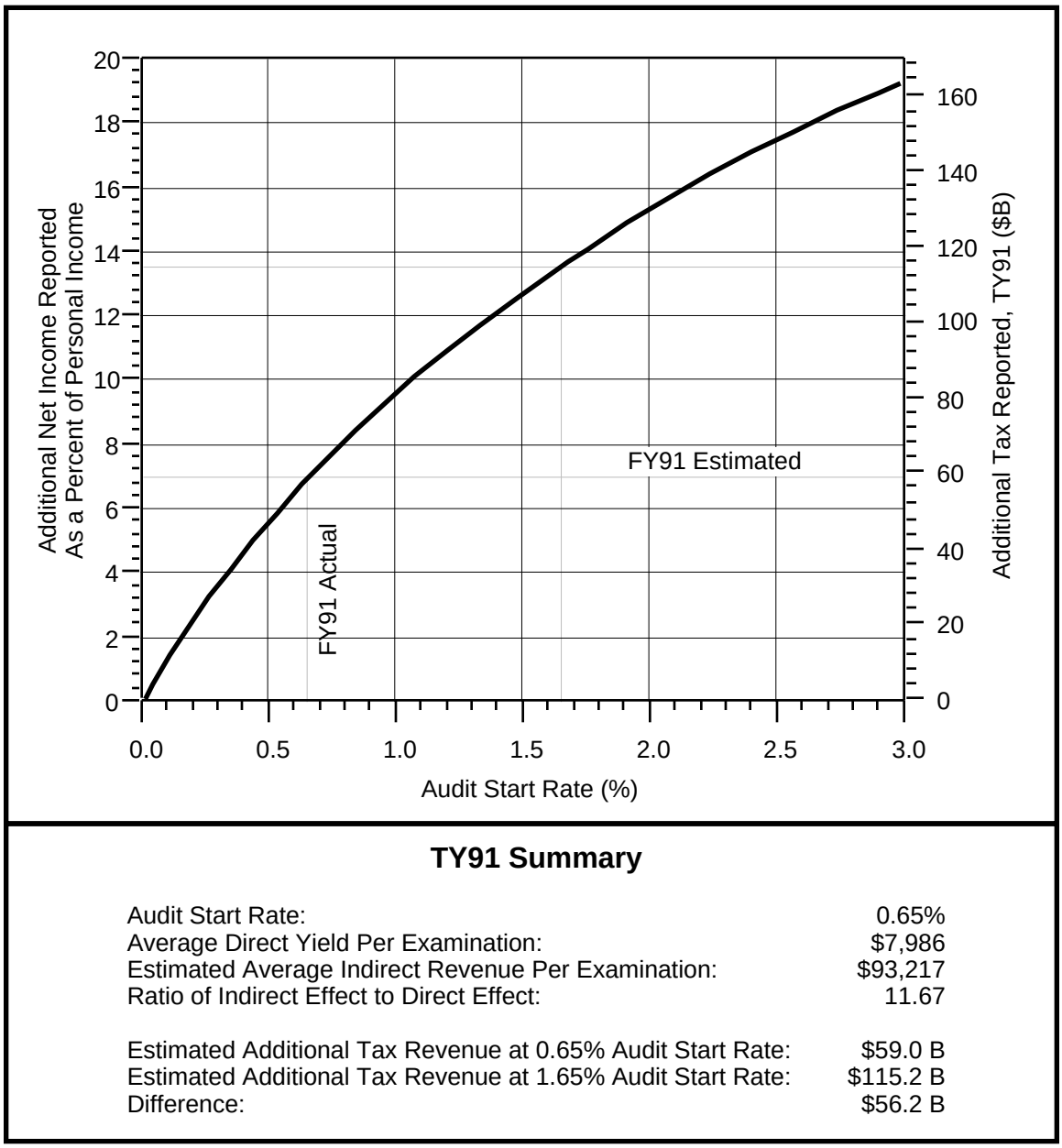
3.2.3 Impact of Burden / Opportunity

As with its impact on filing compliance, *AvgBurden* seems to be related to lower amounts of offsets being reported, although this impact is not strongly significant. Apparently, taxpayers do not claim some offsets because they view it as too burdensome (in terms of the time required) to do so. The impact of the forms burden on income reporting—even *NetIncomePct*—does not appear to be significant, though it tends to be positively related. This may be because higher-income taxpayers—especially those with business income—typically have to fill out more forms and schedules, and thus face a higher burden.

The three opportunity variables all seem to influence income reporting significantly, but not offset reporting. The impact of *SoleProps* on *IncomePct* is much like its impact on the *FilingRate*, probably for much the same reasons; a higher concentration of proprietors generally is associated with more income reported, but that effect is offset by the negative impact of higher concentrations of employment in the Trade, Finance, and Service sectors. The opportunity variable with the most significant impact on income reporting is *PaidPrep*, the percentage of returns prepared with the help of a paid practitioner. Interestingly, controlling for other important determinants of income reporting, preparers seem to *reduce* the amount of income reported. Clearly, many people seek the help of professional tax preparers for the express purpose of reducing their tax liability. But one would suppose that the primary method for doing this would be by finding additional offsets that can be claimed. The results in Table 3, however, indicate that, if anything, the impact of paid preparers is to *reduce* the amount claimed as offsets (although this result is not significant). The very strong negative impact of preparers on income reporting is probably a reflection of their preparation of returns with primarily business-source income, which is net of business expenses; it seems reasonable to assume that professional tax preparers routinely reduce the amount of business income needing to be reported on tax returns by maximizing the business expenses that are claimed.

---

30 I had not intended to include *ChildExemptsPct* in the IncomePct equation, but concluded that it belonged when I noticed that its coefficient in the NetIncomePct equation was fairly significant, but not close to zero minus the coefficient in the OffsetsPct equation (in fact, it also has a positive sign)—suggesting that *ChildExemptsPct* also influenced income reporting.



TY91 Summary

Audit Start Rate:	0.65%
Average Direct Yield Per Examination:	$7,986
Estimated Average Indirect Revenue Per Examination:	$93,217
Ratio of Indirect Effect to Direct Effect:	11.67
Estimated Additional Tax Revenue at 0.65% Audit Start Rate:	$59.0 B
Estimated Additional Tax Revenue at 1.65% Audit Start Rate:	$115.2 B
Difference:	$56.2 B

Figure 6. The Indirect Effect of the Audit Start Rate: NetIncomePct and additional tax reported (in TY91 dollars) as a function of the audit start rate, based on Table 3 results (i.e., Definition A).

3.2.4 Impact of Enforcement Activities

One of the most important findings of this analysis is that audits have a strong, positive impact on reporting compliance. As discussed above in the context of identification issues, the impact of audits seems to be to increase the reporting of income and, to a lesser extent, offsets. The estimated impact of *AuditRate* on *NetIncomePct* is illustrated in Figure 6, which also expresses the magnitude of this "ripple effect" in terms of the additional dollars of tax induced as a function of *AuditRate* (using Tax Year 1991 as an example). As the details at the bottom of Figure 6 indicate, the average indirect effect of the audits started in 1991 was over 11.6 times as large as the average adjustment directly proposed by audits closed that year. Moreover, if the

AuditRate were to have been 1.65 percent in 1991 instead of the actual 0.65 percent, an additional $56 billion of additional tax would have been reported voluntarily.[31] Similarly, these results suggest that if the *AuditRate* had remained constant at its 1982 level of 1.62 percent, the cumulative impact through 1991 would have been that an additional $257 billion of tax would have been reported voluntarily (see Appendix G for the detailed calculations). This is strong evidence that audits are a potent tool to foster voluntary compliance. In fact, since the effect is significantly larger than the direct revenue effect of the audits, these results suggest that the allocation of audit resources (which is currently based almost solely on their direct revenue potential) ought to be modified to give more weight to this indirect effect on voluntary compliance.[32]

Perhaps one of the most surprising findings is that IRP document matching does not have any significant impact on reporting compliance. This seems inconsistent with the familiar observation (and common intuition) that those types of income covered by third-party information reporting (such as wages, interest, and dividends) exhibit a much greater degree of reporting compliance than income types not subject to IRP (see, for example, IRS (1988)). There are two important reasons for this counter-intuitive finding. First, by the time of the ten-year period studied in this analysis (1982-1991), most of the IRP-generated improvements in voluntary reporting compliance had already been realized. Although it was not until the mid-1980's that IRS actually expanded its matching program to encompass virtually all IRP documents received, most taxpayers apparently assumed that the matching was always in place, judging by the earlier compliance statistics. Therefore, there was little more compliance improvement to be realized by 1982. The second reason for this finding is that one of the effects of the IRP program is to make it clearer to taxpayers how much IRS knows about them. The more people are aware of what information is given to IRS and what is not, the more they may be tempted to hide some of what is not reported. In other words, they might take the attitude, "What the IRS doesn't know won't hurt them!" The fact that the estimated coefficients are very insignificant suggests that this kind of response cannot be very strong or widespread, which may be because of the strong deterrent effect of audits.

The final enforcement variable in the reporting equations—criminal convictions arising from the work of the IRS Criminal Investigation Division (CID)—has a highly significant and positive impact on income reporting, and an equally significant, but smaller, positive impact on offsets reporting. These convictions seem to have the same kind of impact on reporting as do audits (increasing both income and offsets), but perhaps for entirely different reasons. As mentioned earlier, the most important influence that these convictions have on the general population may be to satisfy the typical taxpayer that criminals are not going scott-free, thus encouraging him to pay his "fair share." When the opposite perception prevails (i.e., that scofflaws get away with flagrant tax violations), it could very well cause some otherwise law-abiding taxpayers to underreport some of their own tax obligation in protest. It is reasonable to assume that the smaller, but positive, impact on *OffsetsPct* is attributable to the same phenomena causing the positive impact of audits on *OffsetsPct*—such as the fact that many offsets are made possible by larger reported income, and the likelihood that many taxpayers try to keep their offsets fairly proportional to their income so as to avoid an audit.

---

31 These estimated tax effects are based on the average marginal tax rate in 1991.

32 Since this specification uses the overall average audit rate for all individual income tax returns, however, these results do not help us to allocate audit resources cost-effectively to the different classes of individual returns (e.g., business vs. non-business, or low-income vs. high income). If (as is likely) the indirect effect of audits varies according to which classes of returns are audited, then we should allocate audit resources at the margin according to the combination of the direct plus indirect revenue-to-cost ratio. The best way to estimate the extent to which the indirect effect varies across audit classes is to include *AuditRate* data for the separate classes. Lacking such data (for now), I tested interactions of *AuditRate* with both *SoleProps* and *AvgPI* (see Appendix H). The results were inconclusive, but that may be because the mix of returns audited is more important than the mix of returns filed.

3.2.5 Impact of IRS Responsiveness

The impact of Taxpayer Service's (TPS's) return preparation efforts on income reporting compliance is much like its impact on filing compliance—significant and positive, as one might expect. This type of activity seems to have no impact on offset reporting compliance, however—perhaps indicating that the typical return prepared by TPS claims few offsets, or that the TPS assistance results in some taxpayers claiming more in offsets than they otherwise would have, and other taxpayers claiming less, with little significant net impact on the population.

The telephone calls that TPS handles seem to have a weakly significant *negative* impact on income reporting, however. This may indicate that many taxpayers who call IRS find out that they do not need to report certain income (which they perhaps thought might be the case, but wanted confirmation). It seems reasonable to assume that those who believe there is a good chance that their questionable income must be reported do not even call. However, it is quite possible that this negative impact of TPS telephone calls is due more to the *quality* of service that taxpayers receive. If they have a hard time getting through, or need to spend a long time getting a response, or if they are treated rudely or incompetently, the taxpayers who call—and, perhaps others influenced by them—may underreport some of their income in protest. Whatever the reason, although the effect is only marginally significant, it does suggest the need for further study, and perhaps some improvement in the way these calls are handled.

3.2.6 Impact of Demographic and Economic Variables

In contrast with their impact on filing compliance, the concentration in the state population of singles, those under 30 years old, and those over 64 years old do not seem to have a significant impact on either income reporting or offset reporting. However, the signs on these coefficients in all three reporting equations are opposite their signs in the *FilingRate* equation. That is, a higher concentration of singles seems to be associated with lower filing compliance, but higher levels of income and offset reporting, while higher concentrations of both young and old taxpayers tend to strengthen filing compliance, but diminish income and offset reporting. This may mean that singles who file tax returns are marginally more compliant in reporting their income than are married taxpayers, and that the extreme age groupings have less income (and offsets) to report.

Per capita births has a strongly significant, positive impact on both income and offset reporting. This may confirm the hypothesis raised earlier—that the birthrate reflects the general level of optimism and, therefore, taxpayers' willingness to comply with their tax obligations—but we cannot be certain. If the value of dependent exemptions were not already controlled for, we might view *PCBirths* as a surrogate for *ChildExemptsPct*, but this does not seem probable under the circumstances.

The *ExclIncomePct* variable controls for income included in Personal Income that did not need to be reported on tax returns—either because certain individuals were not required to file returns, or because of the types of income that are not taxable. As expected, this has a negative impact on income reporting; the more that does not need to be reported on returns, the less will be reported on returns.

Finally, the unemployment rate appears to have a strongly significant, negative impact on reporting compliance—much like its impact on filing compliance. For example, at higher unemployment rates, more people might save money by cutting back on their taxes—both by reducing the income they report and by increasing the offsets that they claim. However, these results might not reflect a deterioration of reporting compliance at all; they may simply reflect the

fact that unemployment reduces the amount of income that needs to be reported on tax returns (due to fairly high filing thresholds) more than it reduces Personal Income (the denominator of *IncomePct*). On the other hand, unemployment may not influence the reporting of offsets (if the unemployed tended to have few offsets anyway) as much as it decreases Personal Income—resulting in an increase in *OffsetsPct.* Although this is not a variable that can be manipulated by the government for tax purposes, these highly significant results suggest that the role of unemployment on reporting behavior must not be underestimated.

3.3 Alternate Definitions of Income and Offsets

All of the preceding results pertain to income and offsets defined by Definition A. That is, the dependent variables were constructed to omit many components of income and offsets—those subject to changing rules during the 1982-1991 period. Although this definition included on the order of 97 percent of total income, 30 to 60 percent of adjustments, 94 percent of itemized deductions, and 30 to 60 percent of credits each year, and although using this definition seems necessary in order to control for the many rule changes, it leaves open the question of what impact these various determinants have had on the income and offset components excluded entirely from Definition A. It may well be that some of these components are the *avenue* of noncompliance, perhaps *because of* the changing rules.

To check this possibility, I estimated the same equations using Definition B (which includes components whose rules changed only in the Tax Reform Act of 1986) and Definition C (which includes all components, regardless of rule changes). The results for the three definitions are compared in Appendix C for all three reporting equations. Generally, the results are similar for all three definitions (with the greatest differences typically observed among the tax policy variables), suggesting that most of the various determinants have much the same impact on the excluded components of income and offsets that they do on those included in Definition A, or that the aggregate effect of any different impacts is not large. This allows us to generalize the results for Definition A more confidently.

3.4 Other Potential Determinants of Voluntary Compliance

A number of other variables were excluded from the final analysis because they were found to have no significant impact on compliance. Since that is a useful finding in and of itself, it seems appropriate to mention these variables briefly here.

One of the big surprises had to do with a tax policy variable. Starting in 1987, taxpayers have been required to supply the Social Security Number (SSN) of any dependents they claim as exemptions who are above some specified age. That age has progressively declined in the years that followed. I was able to use detailed Census population data to quantify by state and year the percent of the under-18 population for whom SSNs were required, but found that this had no significant impact—even on *OffsetsPct.* One would assume it would have been possible to detect the seemingly significant reduction in the number of dependent exemptions claimed shortly after this rule went into effect. It may have been captured already in the fixed year effects, however, making the SSN variable somewhat collinear with the year dummies—especially since it was zero for all observations prior to 1987.

The only additional Burden/Opportunity variable tested represented the concentration of farmers, and was also insignificant. There are too many potential reasons for this result to speculate which one(s) might be responsible.

Three additional enforcement variables were tested—all of them related to other variables that have been included. In addition to the number of CID convictions obtained, I tested the number of CID investigations started; in addition to the percentage of refunds that were offset, I checked the average dollar amount of the refund offsets; and instead of the total number of TDI nonfiler notices issued, I tried just the number of TDI first notices (out of a series of four possible notices). None of these additional variables or alternatives was useful, but this is not too surprising. Since the most common and most effective means of communicating CID's activities to the general population is the media coverage of arrests and convictions, it seems reasonable to find that the start of new investigations is not an important determinant. It also makes sense that the dollar value of refund offsets has little impact on compliance; the decision of whether or not to file is more likely influenced by the prevalence of refund offsets than on the amount, and the response of filers to the prospect of getting a partial (or zero) refund is most likely to be to change their withholding, which is not reflected in my reporting compliance variables. Finally, since nonfilers are often so hard to find, it is not surprising that the first notice alone is not as important a determinant of filing compliance as the combined effect of all of the TDI notices.

I also tested three additional measures of IRS Responsiveness: the speed in issuing refunds; the volume of Taxpayer Service (TPS) correspondence with taxpayers; and TPS educational efforts. Refund speed (measured as the percent of refunds taking longer than 35 or 45 days to process) could conceivably influence taxpayer attitudes about the IRS, and, therefore, subsequent compliance, but I found no evidence for this—even though the variable has exhibited much variation cross-sectionally and over time. The two additional TPS variables represent somewhat less important activities than the phone calls handled and returns prepared, and it is not surprising that they did not significantly influence compliance.[33]

Several demographic/economic variables did not exhibit a significant impact on compliance, as well. These include the concentration of males (among potential Single and Head of Household returns), the percentage of the population with at least some college education, population density, the concentration of employment in the Trade/Finance/Service sectors (although this was significant when included as an interaction term with SoleProps), and the concentration of employment in the Construction/Mining/Manufacturing sectors. It may be that these (or other) demographic/economic variables might exhibit more significance in slightly different specifications, but it is likely that many of the included variables have picked up whatever influence might be attributed to these.

3.5 Implications for Resource Allocation

Given that several IRS activities seem to have a positive impact on voluntary filing and/or reporting compliance, it is natural to want to compare these activities with respect to their relative effectiveness in improving compliance. Such an awareness could be used to improve the allocation of current IRS resources (e.g., by expanding some activities at the expense of others), and to identify which activities to concentrate on in any expansion of IRS resources intended specifically to increase revenue or to promote voluntary compliance. These results give IRS this kind of insight for the first time—at a time when shrinking budgets are forcing tough decisions.

A useful way to compare IRS activities is according to how well they help achieve the objective of maximizing net revenue—total revenue net of costs. Maximum net revenue is achieved—at any constrained budget level—when the ratio of marginal revenue to marginal cost is the same for all activities. If this condition is not true, then additional revenue can be obtained with the same budget by reallocating resources from the low revenue-to-cost programs to the high

[33] I have been able to compile data on the accuracy of the phone calls, as well, but since these quality checks began only in 1984, I have not included this variable in the analysis.

revenue-to-cost programs.

For many IRS activities (e.g., audits), "revenue" may include both the amounts paid directly by the taxpayers who are contacted by the activity as well as amounts paid by the general population as an indirect response to the deterrent effect of the activity. Other activities (e.g., offsetting refunds to help settle non-tax debts) may have just an indirect effect on tax compliance. In every case, the marginal revenue must reflect only the dollars *collected* (i.e., not the amounts that are proposed or assessed, but are never collected), and the marginal cost must include all additional overhead and support costs necessary to collect the marginal revenue.

Table 5 compares the cost-effectiveness in 1991 of the five IRS activities found in this study to have a significant and positive impact on voluntary compliance. Although these comparisons include only the indirect effect of the activities on compliance (since their marginal direct effects are not available), the results are very informative. Marginal indirect revenue functions were derived for each activity based on the *FilingRate* and *NetIncomePct* equations estimated for Definition A (see Appendix I). The first panel summarizes the current level of activity in 1991, and shows that TDI notices are by far the most cost-effective at the margin in producing revenue (recall that they do this by promoting better filing compliance). The second panel shows how much each activity would have had to be expanded (and the corresponding cost) in 1991 in order for that activity to induce the voluntary reporting of an additional $10 billion of tax. Notice that an increase in the number of TDI notices would again be the cheapest alternative (if it were feasible to issue over 460 percent more of these notices), followed by an increase in the number of returns prepared by TPS. The problem with the second panel is that it may not be feasible to expand some of these activities to the rate required to induce an additional $10 billion of tax. The third panel of Table 5 illustrates a more realistic expansion of each activity; it assumes that the nationwide level of the activity is increased to the largest rate observed within any state during the 1982-1991 period. In this scenario, audits produce the most revenue by far, followed by the return preparation assistance of TPS.

In each panel, the *least* cost-effective activity is criminal convictions. However, these results support the belief within CID that their principal role is to foster voluntary compliance. In fact, the marginal indirect revenue-to-cost ratio of 16.3 at their current level of activity is far greater than the average direct effect of audits. Also, Panel C of Table 5 highlights the fact that a realistic expansion of CID activities may produce more indirect revenue than the largest realistic expansion of TDI notices—even though TDI notices are the *most* cost-effective in producing indirect revenue. These results emphasize that using estimates of indirect effects to guide IRS resource allocations must take into account the practical constraints imposed on the expansion of most IRS activities. However, they also illustrate that the potential benefits for resource allocation and for revenue generation are immense.

Table 5. Indirect Revenue-to-Cost Comparisons for Five IRS Activities, 1991 [a]

A. Actual Level of Activity

IRS Activity	Rate	Number of units	Cost [b] per unit ($)	Marginal indirect Revenue/cost Ratio
Audit start rate (%)	0.647	632,819	1,298	54.6
TPS- Returns prepared/thousand	3.33	840,126	13.74	395.9
IRP documents/potential return	7.56	978,512,924	0.031	668.0
TDI notices/potential return (%)	3.43	4,436,942	0.305	3,766.1
CID convictions/million	10.51	2,650	103,064	16.3

B. Rate Required to Induce an Additional $10 Billion of Tax

IRS Activity	Rate	Increase in units	% Increase	Cost ($M)
Audit start rate (%)	0.797	147,469	23.3%	191.4
TPS- Returns prepared/thousand	7.29	998,414	118.8%	13.7
IRP documents/potential return	11.30	483,860,602	49.4%	15.0
TDI notices/potential return (%)	19.25	20,469,781	461.3%	6.2
CID convictions/million	88.41	19,640	741.1%	2,024.2

C. Nationwide Rate Increased to Highest Rate Observed Within Any State

IRS Activity	Rate	Increase in units	Additional Indirect tax Revenue ($M)	Marginal indirect Revenue/cost Ratio
Audit start rate (%)	3.510	2,802,462	115,072	19.9
TPS- Returns prepared/thousand	23.02	4,963,934	27,001	395.9
IRP documents/potential return	10.51	381,705,783	7,889	668.0
TDI notices/potential return (%)	9.29	7,582,359	5,545	1,621.4
CID convictions/million	51.52	10,340	7,405	3.6

[a] Revenues exclude amounts collected directly from the taxpayers contacted; the appropriate marginal revenue-to-cost comparison includes both the direct and the indirect effects. Audits, returns prepared by TPS, and CID convictions typically result in direct revenue at no additional cost, while IRP matching and TDI notices typically require additional contact with the taxpayers (at additional cost) to generate direct enforcement revenue (in fact, some level of such direct enforcement contacts are probably necessary to ensure that the matching and notices are credible deterrents).

[b] Source: IRS Compliance Planning & Finance: Budget & Resource Allocation Group; includes all appropriate overhead, support, and follow-on costs.

4. Conclusions

This study breaks new ground in answering the important question of what the government can do to foster better voluntary compliance with the individual income tax. The quality and breadth of the data, together with a simple, yet powerful, econometric specification, have provided new insights into several important—yet previously unexplored—determinants of filing and reporting compliance, and have yielded perhaps the most believable estimates yet of the compliance effects of factors that have been studied before.

Although at least one previous study of this sort has been based on state-level panel data, this analysis has broken from the usual preference for cross-sectional micro data, arguing that a state-level panel is the *mostappropriate* data structure with which to quantify what are inherently aggregate phenomena. (Micro models may provide insight into the *mechanisms* of taxpayer behavior, however.) Data innovations include: the estimation of the required filing population using the Current Population Survey; the use of Personal Income from the National Accounts in the denominator of the reporting compliance measures (together with additional explanatory variables to control for variations in Personal Income not related to changes in actual tax obligations); the use of two productivity-related variables to estimate audit rates; the use of the audit start rate, instead of the widely-used audit closure rate; the use of two exogenous federal marginal tax rate variables, capturing the variation of marginal rates at both low and moderate incomes separately; and the introduction of several determinants of compliance that had received little or no quantitative attention until now, including the burden associated with the time to complete the myriad tax forms and schedules, three new enforcement variables (TDI nonfiler notices, refund offsets, and CID convictions), and several measures of IRS responsiveness to taxpayers (telephone calls handled and returns prepared by Taxpayer Service, and the speed with which refunds are issued).

Most previous studies of income tax reporting compliance have estimated equations for tax reporting; some have also estimated an equation for some income concept (e.g., Adjusted Gross Income). This analysis avoids the interpretive problems associated with estimating a tax equation (in which the compliance effect of important tax policy parameters cannot be separated from the direct role of these parameters in calculating tax from reported income), and estimates separate equations for the amount of total income reported and total offsets claimed (including the income-offset value of tax credits). Such a specification provides direct insight into the separate problems of underreported income and overstated offsets. It also allows the estimation of a third reporting compliance equation—net income, which is total income minus total offsets. This third equation is not only the "bottom line" with respect to reporting compliance, but it is a useful check on the other two reporting equations.

Another major innovation of this study is the use of three different definitions of income and offsets to account for the fact that several income and offset components have been subject to changing rules as to what should (or can) be reported in any given year. Estimating separate equations for each definition has prevented inadvertently attributing to some of the explanatory variables changes in reporting caused entirely by the changed rules.

Perhaps the greatest need suggested by this research follows from the large, strongly significant, and positive deterrent effect of audits on the general population. If the indirect effect of audits is, on average, about eleven times as large as their direct effect on revenue, then we should seek to understand this "ripple effect" in much greater detail. For example, it is reasonable to assume that this indirect effect is different among different groups of taxpayers (e.g., business vs.

non-business, or high-income vs. low-income). If so, then, given the relative magnitude of the indirect effect, an understanding of these differences is probably much more crucial to the optimal allocation of audit resources than our current understanding of the different *direct* yields to be expected within these groups.

This study provides the first quantitative basis for allocating IRS resources optimally across its various enforcement and non-enforcement activities. Perhaps it can also lead to a method for evaluating the effectiveness of those activities, since it controls for a wide variety of other determinants of voluntary compliance as well.

Bibliography

Alexander, C. and J. Feinstein (1987). "A Micro-Econometric Analysis of Income Tax Evasion," Massachusetts Institute of Technology, mimeo.

Allingham, M.G., and A. Sandmo (1972). "Income Tax Evasion: A Theoretical Analysis," *Journal of Public Economics* 1:323-338.

Baltagi, B. (1995). *Econometric Analysis of Panel Data*, (John Wiley & Sons, Chichester, England).

Balestra, P., and M. Nerlove (1966). "Pooling Cross-Section and Time Series Data in the Estimation of a Dynamic Model: The Demand for Natural Gas," *Econometrica* 34:585-612.

Becker, G. (1967). "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, 78(2): 526-36.

Beron, K.J., H.V. Tauchen, and A.D. Witte (1992). "The Effect of Audits and Socioeconomic Variables on Compliance," in *Why People Pay Taxes: Tax Compliance and Enforcement*, Joel Slemrod, ed., (The University of Michigan Press, Ann Arbor), p. 67-89.

Bureau of Economic Analysis (BEA) (1986). "Relationship Between Personal Income and Adjusted Gross Income: Revised Estimates, 1947-83," *Survey of Current Business*, May 1986.

_____ (1994). "Annual Revision of the U.S. National Income and Product Accounts," *Survey of Current Business*, July 1994, p. 7-48.

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

Cox, D. (1984). "Raising Revenue in the Underground Economy," *National Tax Journal* 39:283-288.

Dubin, J., M. Graetz and L. Wilde (1990). "The Effect of Audit Rates on the Federal Individual Income Tax, 1977-1986," *National Tax Journal* 43:395-409.

Erard, B. (1992). "The Influence of Tax Audits on Reporting Behavior," in *Why People Pay Taxes: Tax Compliance and Enforcement,* Joel Slemrod, ed., (The University of Michigan Press, Ann Arbor), p. 95-114.

_____, and C. Ho (1995). "Searching for Ghosts: Who Are the Nonfilers and How Much Tax Do They Owe?," paper presented at the March 1995 Tax Research Symposium at the University of Michigan.

Hausman, J.A. (1978). "Specification Tests in Econometrics," *Econometrica* 46:1251-1271.

Hoch, I. (1962). "Estimation of Production Function Parameters Combining Time-Series and Cross-Section Data," *Econometrica* 30:34-53.

Hsiao, C. (1986). *Analysis of Panel Data*, Econometric Society Monographs No. 11, (Cambridge University Press, Cambridge, England).

Internal Revenue Service (IRS) (1988). *Income Tax Compliance Research*, Publication 7285 (3-88), Washington, D.C.

_____ (1990). *Income Tax Compliance Research: Net Tax Gap and Remittance Gap Estimates (Supplement to Publication 7285)*, Publication 1415 (4-90), Washington, D.C.

_____ (1992). *Commissioner's Annual Report*, Washington, D.C.

Kuh, E. (1959). "The Validity of Cross Sectionally Estimated Behavior Equations in Time Series Applications," *Econometrica* 27:197-214.

Mundlak, Y. (1961). "Empirical Production Function Free of Management Bias," *Journal of Farm Economics* 43:44-56.

_____ (1978). "On the Pooling of Time Series and Cross Section Data," *Econometrica* 46:69-85.

Nagin, D. (1978). "General Deterrence: A Review of the Empirical Evidence," in *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates*, Alfred Blumstein, Jacquelin Cohen, and Daniel Nagin, eds., (Panel on Research on Deterrent and Incapacitation Effects, National Academy of Sciences, Washington, DC), p. 95-139.

Park, T.S., and J.S. Reeb (1989). "Personal Income and Adjusted Gross Income, 1984-1986," *SOI Bulletin*, Vol. 8, No. 3, (Internal Revenue Service, Statistics of Income), p. 71-78.

Parker, R.P. (1984). "Improved Adjustments for Misreporting of Tax Return Information Used to Estimate the National Income and Product Accounts, 1977" *Survey of Current Business*, June 1984, p. 17-25.

Roth, J.A., J.T. Scholz, and A.D. Witte (1989), eds. *Taxpayer Compliance, Volume 1: An Agenda for Research*, (Panel on Taxpayer Compliance Research, Committee on Research on Law Enforcement and the Administration of Justice, Commission on Behavioral and Social Sciences and Education, National Research Council, University of Pennsylvania Press, Philadelphia).

Tauchen, H.V., A.D. Witte , and K.J. Beron (1989). "Tax Compliance: An Investigation Using Individual TCMP Data," NBER Working Paper No. 3078, (National Bureau of Economic Research, Inc., Cambridge, MA).

Appendix A

Data Sources and Derivations

The following information describes where the raw data used in this study came from, and what had to be done to derive the components (typically the numerators and denominators) necessary to form the variables used in the analysis. The 30 variables actually used in the analysis (and several more that were not included in the final specification) were derived from the 45 data elements described below. The descriptions of the variables and data elements are grouped in the major categories used in the report: Dependent Variables, Tax Policy, Burden/Opportunity, Enforcement, IRS Responsiveness, and Demographics/Economics.

A. Dependent Variables

FilingRate = [A1]/[A2] *100
IncomePct = [A3]/[A5] *100
OffsetsPct = [A4]/[A5] *100
NetIncomePct = {[A3]-[A4]}/[A5] *100

[A1] Number of Returns Filed

Sources:

a. IRS Statistics of Income (SOI) Individual/Sole Proprietor Data File—a representative sample of filed returns containing detailed income, offset, tax, and other information from each return, which can be weighted to project to the entire filing population. Available for each tax year.

b. IRS Individual Masterfile (IMF)—the accounting system that records all tax transactions of individuals. This study used IMF return filings summarized by state by tax year by type of tax form (1040, 1040A, or 1040EZ).

Derivation:

a. Determine which filed returns on the SOI file were required to file:
   - Those with Gross Income > Filing Threshold; OR
   - Those with Gross Income < 0; OR
   - Those that filed a Schedule SE; OR
   - those with Schedule C Net Income < 0 OR Schedule F Net Income < 0.

   Note: this definition excludes returns with no tax liability that are filed solely to claim a refund of prepaid taxes or the Earned Income Tax Credit.

b. Aggregate the records of required returns by form type, by state, by year, weighted to project to the entire population.

c. Adjust the aggregations using the ratio of IMF to SOI filings by form type, by state, by year. This re-weighting corrects for the SOI sampling error inherent in state-level aggregations of the SOI data, which is due to the fact that the SOI sample is not designed to be representative at the state level. (This same adjustment is applied to all SOI data, including data elements A3, A4, C4, and D5.)

d. Add the adjusted filings across form type to get total filings by state by year.

[A2] Number of Returns Required

Source: Current Population Survey, Bureau of the Census

Derivation: See Appendix B.

[A3] Income Reported

Source: Statistics of Income (SOI)

Derivation: Amounts aggregated according to three alternative definitions. (See Appendix C.) The amounts were adjusted by the IMF/SOI ratio (see data element [A1]).

[A4] Offsets Reported

Source: Statistics of Income (SOI)

Derivation: Amounts aggregated according to three alternative definitions. (See Appendix C.) The amounts were adjusted by the IMF/SOI ratio (see data element [A1]).

[A5] Personal Income

Source: Regional Analysis Division, Bureau of Economic Analysis (BEA), Department of Commerce.

Derivation: Available in digital format, at the state level.

B. Tax Policy

FThresholdPct = [B1]/[A5] *100
Amnesty5 = [B2]
MargTaxRate@$15K = {[B3]*[F2] + [B4]*[F3]}/{F2]+[F3]}
MargTaxRate@$57K = {[B3]*[F2] + [B4]*[F3]}/{F2]+[F3]}
ChildExemptsPct = [B5]/[A5] *100
StateTaxPct = [B6]/[A5] *100
TINsForTots = [B7]/[F2] *100

[B1] Aggregate Filing Threshold on Required Returns

Source: Current Population Survey, Bureau of the Census

Derivation: See Appendix B.

[B2] State Tax Amnesty

Source: IRS Research Division study of state amnesties.

Derivation: Amnesty5 is a dummy variable taking the value of 1 for the year in which the state had an amnesty (if any), and for the next 4 years as well; it otherwise had the value of zero.

[B3] Marginal Tax Rate Schedule: Singles

Source: Tax rate schedule for Single Filers included in the Form 1040 tax package each year.

Derivation: Marginal tax rate bracket associated with the specified amount of taxable income ($15,000 or $57,000 in constant 1982 dollars per the GDP deflator).

[B4] Marginal Tax Rate Schedule: Married-Joint

Source: Tax rate schedule for Married-Joint filers included in the Form 1040 tax package each year.

Derivation: Marginal tax rate bracket associated with the specified amount of taxable income ($15,000 or $57,000 in constant 1982 dollars per the GDP deflator).

[B5] Aggregate Value of Allowed Dependent Child Exemptions

Source: Current Population Survey, Bureau of the Census

Derivation: The exemption value for the year in question multiplied by the number of children under 18 (never married) associated with the primary taxpayer on required returns. (See Appendix B.)

[B6] Aggregate Value of Income, Property, and Sales Taxes Collected by State

Source: Census of Governments, Bureau of the Census

Derivation: The sum of income, property, and sales tax receipts by the state government by year. (Since sales tax were no longer deductible on federal returns after 1986, sales tax receipts were excluded from the sum after 1986.)

[B7] Number of Children Under 18 Needing an SSN on the Return

Source: Population Division, Bureau of the Census

Derivation: Population estimates are available on diskette by state by year by age. The age at which a Taxpayer Identification Number (TIN) has been required in order to claim the dependent exemption has changed since it was first required on 1987 returns, as shown in Table A-1. Since TINs were not required before 1987, this variable is zero for all states from 1982 through 1986.

Table A-1

Tax Years	Age
1987-88	5
1989-90	2
1991	1

C. Burden/Opportunity

AvgBurden = [C1]/[C2]
SoleProps = [C3]/[C2] *100
SolePropTFS = [C3]/[C2] * [F9]/[F10] /100
PaidPrep = [C4]/[A1] *100

[C1] Aggregate Burden on Required Returns

Source:
a. Current Population Survey, Bureau of the Census
b. Tax forms and instructions
c. IRS Tax Forms Burden Estimation Model developed by Arthur D. Little, Inc.

Derivation:

a. Estimate which forms and schedules are required with required returns each year, and estimate the total number of each that is required in the population. (See Appendix F.)
b. Estimate the hours of burden per form and schedule by year. (See Appendix F.)
c. Aggregate Burden is the sum (across all form and schedule types) of the burden associated with a particular form or schedule multiplied by the number of returns requiring that form or schedule.

[C2] Number of Potential Returns

Source: Current Population Survey, Bureau of the Census

Derivation: The number of "returns" that would be required if the filing threshold were zero. This combines any married records into the appropriate "Married-Joint" return (i.e., spouses on the CPS are not both counted as separate potential returns). (See Appendix B.)

[C3] Number of Sole Proprietors in the Population

Source: Current Population Survey, Bureau of the Census

Derivation: The number of people with self-employment income, per the CPS (variable I51B; see Appendix B for a listing of the CPS variables used in this study).

[C4] Number of Returns Prepared by a Paid Practitioner

Source: Statistics of Income (SOI)

Derivation: The number of filed required returns that were prepared by a paid practitioner. The counts were adjusted by the IMF/SOI ratio (see data element [A1]).

D. Enforcement

AuditRate = [D1]/[A1]$_{-1}$ *100
IRP_DocRate = [D2]/[C2] *100
TDI_TotRate = [D3]/[C2] *100
RefOffRate = [D4]/[D5] *100
CID_ConvRate = [D6]/[F1] *1,000,000
DET_Pct = [D7]/[D8] *100
AvgDET = [D7]/[D9] *100

[D1] Number of Audits Started

Sources:

a. 1982-1985: IRS Audit Information Management System (AIMS) Table 1.0. Annual (fiscal year) data available by district on microfiched tables.

b. 1986-1988: IRS Audit Base Inventory System (ABIS) Table 1.1. Monthly data available by district on microfiched tables.

c. 1989-1991: IRS Audit Information Management System (AIMS) Table 37. Monthly data available by district on microfiched tables.

Derivation: These are the audits conducted by district office Revenue Agents and Tax Auditors only; they do not include audits conducted in the service centers (typically by Tax Examiners).

a. 1982-1985:

Table A-2. Data Required to Derive the Number of Audits Started From AIMS Table 1.0

Item*	Description	Derivation
(2)	Audits completed	
(4)	Total Inventory of returns in audit	
(5)	Returns awaiting classification	
(6)	Local Definition	
(7)	RPM Suspense	
(8)	Unassigned, selected for audit	
(10)	Assigned, no time applied	
A	Beginning inventory (audit status ≥ 12)	Prior year ending inventory
B	Ending inventory (audit status ≥ 12)	(4)-(5)-(6)-(7)-(8)-(10)
C	Audits started	B-A+(2)

* Numbered items are columns on Table 1.0; lettered items are derived.

b. 1986-1988:

- Total number of audits started = (DIF-source starts) + (Other, actual starts)
- Annual starts derived as the sum of the twelve monthly figures
- For 1986, tables were available only for the second half of the fiscal year (April 86 - September 86). Analysis of other years, however, revealed that the last 6 months accounted for very nearly 50 percent of all starts, on average, for all districts. So, the 1986 counts were derived by doubling the number of audits started in the last 6 months of that fiscal year.

c. 1989-1991:
- Total number of audits started = (Revenue Agent starts) + (Tax Auditor starts)
- Annual starts derived as the sum of the twelve monthly figures

---

[D2] Number of IRP Documents Matched Against Returns

Sources:

a. 1985-1991: IRS Information Return Masterfile (IRMF). Number of documents available by type of document by state by year—but only for documents received by IRS on magnetic media (the state field is not transcribed from paper documents). (The share of documents received on magnetic media grew from almost 70% in 1985 to over 80% in 1991.)

b. 1982-1991: IRMF data compiled by IRS Research Division's Projections & Forecasting Group. Total number of documents processed—whether received on magnetic media or on paper—by type of document by year, but not by state (national totals only).

Derivation:

a. For each year, 1985-91, compute each state's share of all IRP documents that have a state indicator on the IRMF.

b. For each state, compute the 1985-91 average of its annual share of all IRP documents that have a state indicator on the IRMF.

c. Adjust the 1985-91 cross-sectional distribution slightly to force the state average shares to sum to 100%. The adjustment is proportional to the average shares in step (b).

d. The imputed number of IRP documents for each state is the product of the national total number of IRP documents (from source b) times the corresponding average state share given in step (c) above. Thus, this series has a constant distribution across states over time, and varies over time for a state according to the trend in the national totals.

[D3] Number of Nonfiler (Taxpayer Delinquency Investigation) Notices Issued

Sources:

a. 1988-1991: IRS Collection function, COINS database (available on diskette by district by year).

b. 1985-1987: IRS Collection function, TDI Cumulative Report NO-5000-4, Part 3, available on microfiche by district by year.

c. 1982-1984: IRS Collection function, TDI Cumulative Report C-4, available on microfiche by district by year.

Derivation:

a. Derive the total number of notices issued as the sum of all first, second, third, and fourth notices issued, for both IMF Nonfiler notices and IMF Stop-Filer notices (tabulated separately on the reports).

b. Account for the fact that all districts served by the Memphis Service Center (Indianapolis, IN; Richmond, VA; Parkersburg, WV; Greensboro, NC; Louisville, KY; and Nashville, TN) started using the new Report NO-5000-4 in August 1984, while the remaining districts started using the new report in November of 1984. This required transcribing the appropriate monthly totals for FY84 and FY85, then adding the correct figures from the two reports to get annual totals by district.

[D4] Number of Refunds Offset for Outstanding Debts

Source:

a. 1982-1983: Virtually no refunds were offset for other debts.

b. 1984-1991: IRS Debtor Masterfile (DMF). Data on the total number and amount of offsets imposed nationally available by year.

c. 1987, 1989-1991: IRS Debtor Masterfile (DMF). Data on the number and amount of offsets imposed available by state by year.

d. 1985-1991: IRS Debtor Masterfile (DMF). Data on the total number and amount of offsets imposed available by service center by year.

Derivation:

a. 1982-1983: Zero offsets.

b. 1984:

- Service center totals estimated by applying 1985 distribution across service centers to 1984 national total.
- State totals estimated by applying 1987 distribution of service center totals across states to the estimated 1984 service center totals (from the step above).

c. 1985-1986: State totals estimated by applying 1987 distribution of service center totals across states to the known 1985 and 1986 service center totals.

d. 1988: State totals estimated by applying the average of the 1987 and 1989 distribution of service center totals across states to the known 1988 service center totals.

e. 1987, 1989-1991: Data already available by state.

[D5] Number of Refund Returns

Source: Statistics of Income (SOI)

Derivation: The number of filed returns that claimed a refund. The counts were adjusted by the IMF/SOI ratio (see data element [A1]).

[D6] Number of Criminal Tax Convictions

Source: IRS Criminal Investigation Division (CID), Ad Hoc Report 48. Data available by district by year from paper reports.

Derivation: The number of convictions was already available by state, but for some years, the total number of convictions was derived as the sum of convictions arising from the General Enforcement Program (GEP) and the number arising from the Special Enforcement Program (SEP).

[D7] Direct Examination Time (DET)

Source:

a. 1982-1987: IRS Audit Time Tracking Report (ATTR) Table Z-19. Annual (fiscal year) data on DET and DET% available only by grade by audit class by district on microfiched tables. Data are for Revenue Agents only.

b. 1988-1991: IRS Audit Information Management System (AIMS) Table 37. Annual (fiscal year) data available by audit class by district on microfiched tables. Data are for Revenue Agents only.

Derivation:

a. 1982-1987: Total DET is the sum of the separate times applied by Revenue Agents in three grade categories: grade 11 and under; grade 12; and grade 13.

b. 1988-1991: DET was directly available from the reports for all grades of agents combined.

[D8] Total Examination Time

Source:

a. 1982-1987: IRS Audit Time Tracking Report (ATTR) Table Z-19. Annual (fiscal year) data on DET% available by grade by district on microfiched tables. Data are for Revenue Agents only.

b. 1988-1991: IRS Audit Information Management System (AIMS) Table 37. Annual (fiscal year) data available for all grades of agents combined by district on microfiched tables. Data are for Revenue Agents only.

Derivation:

a. 1982-1987:
   - Total time for each grade category is DET for the grade divided by the DET% for the grade.
   - Total time over all is the sum of total time across grades.

[D9] Number of Audits Closed

Source: IRS Commissioner's Annual Report. Data are available on the number of audits closed by district by year. The data compiled for this study exclude service center audits.

Derivation: The number of audits closed is available directly from this report.

E. IRS Responsiveness

TPS_CallsPC = [E1]/[F1] *1,000
TPS_RetPrepPC = [E2]/[F1] *1,000
TPS_CorrespPC = [E3]/[F1] *1,000
TPS_EdnPC = [E4]/[F1] *1,000
Refund45 = [E5]/[D5] *100

[E1] Number of Telephone Calls Handled by Taxpayer Service

Source: IRS Taxpayer Service RMIS reports. The number of telephone calls answered (toll-free and non-toll-free) is available on paper reports by district by year. However, not all districts have toll-free call sites, and the districts that do handle toll-free calls serve taxpayers in several states.

Derivation:

a. Toll-free calls are imputed to the states served by each toll-free call site in proportion to their 1988 filing populations. For example, the Jacksonville call site serves Florida and South Carolina. Since Florida accounted for 79.6 percent of the returns filed by Florida and South Carolina taxpayers in 1988, 79.6 percent of the toll-free calls handled by Jacksonville are imputed to Florida. The states served by each call site are listed in Table A-3.

b. Total calls handled by a district is the sum of the number of toll-free calls imputed by the method described above plus the number of non-toll-free calls handled by the district.

[E2] Number of Returns Prepared by Taxpayer Service

Source: IRS Taxpayer Service RMIS reports.

Derivation: This is available on the paper reports.

[E3] Number of Units of Correspondence Handled by Taxpayer Service

Source: IRS Taxpayer Service RMIS reports.

Derivation: This is available on the paper reports.

Table A-3. States Served by the Taxpayer Service Toll-Free Call Sites

Call Site Location	States Served					
North Atlantic Region						
Boston	MA	ME	NH	10%NY	RI	VT
Brooklyn	35%NY					
Buffalo	55%NY	CT				
Mid-Atlantic Region						
Newark	NJ					
Philadelphia	60%PA					
Pittsburgh	40%PA	DE	65%MD			
Baltimore	35%MD					
Richmond	VA					
Southeast Region						
Atlanta	GA	AL	LA	MS		
Jacksonville	FL	SC				
Nashville	TN	AR	NC			
Central Region						
Cincinnati	50%OH	WV				
Cleveland	50%OH					
Indianapolis	IN	KY				
Detroit	MI					
Midwest Region						
Chicago	IL					
Milwaukee	WI					
St Paul	MN	MT	ND	SD		
Des Moines	IA					
St Louis	MO					
Omaha	NE					
Southwest Region						
Dallas	60%TX	KS	NM	OK		
Houston	40%TX					
Denver	CO	UT	WY			
Phoenix	AZ					
Western Region						
Laguna Niguel	10%CA					
Sacramento	10%CA					
San Jose	10%CA					
Seattle	WA	10%CA				
Anchorage	AK					
Portland	OR	10%CA	ID			
San Francisco	20%CA	NV				
Los Angeles	30%CA					
Honolulu	HI					

[E4] Number of Educational Efforts Conducted by Taxpayer Service

Source: IRS Taxpayer Service RMIS reports.

Derivation: This is available on the paper reports.

[E5] Number of Refunds Taking More Than 45 Days to Process

Source: Statistics of Income (SOI).

Derivation:

a. Select only refund returns from the SOI micro data file.

b. Use the Document Locator Number (DLN) Julian date as the return received date (which, for refund returns, is a pretty good approximation).

c. Derive the refund Julian date from the IMF posting cycle using Table A-4.

d. Calculate the number of days elapsed between the two Julian dates. If the difference is negative, add 365 to the difference. (These generally occur when the filing date is late in the year and the refund is issued early in the next year)

e. Count the weighted number of refunds for each state (based on the taxpayers' state of residence) for which the elapsed time is over 45 days.

Table A-4. Refund Julian Dates Associated With Posting Cycles, 1980-1992

	Tax Year (SOI)												
	79	80	81	82	83	84	85	86	87	88	89	90	91
Cycle	Processing Year (XX)												
Posted	80	81	82	83	84	85	86	87	88	89	90	91	92
XX01	18	16	15	14	13	11	10	16	15	13	12	11	10
XX02	25	23	22	21	20	18	17	23	22	20	19	18	17
XX03	32	30	29	28	27	25	24	30	29	27	26	25	24
XX04	39	37	36	35	34	32	31	37	36	34	33	32	31
XX05	46	44	43	42	41	39	38	44	43	41	40	39	38
XX06	53	51	50	49	48	46	45	51	50	48	47	46	45
XX07	60	58	57	56	55	53	52	58	57	55	54	53	52
XX08	67	65	64	63	62	60	59	65	64	62	61	60	59
XX09	74	72	71	70	69	67	66	72	71	69	68	67	66
XX10	81	79	78	77	76	74	73	79	78	76	75	74	73
XX11	88	86	85	84	83	81	80	86	85	83	82	81	80
XX12	95	93	92	91	90	88	87	93	92	90	89	88	87
XX13	102	100	99	98	97	95	94	100	99	97	96	95	94
XX14	109	107	106	105	104	102	101	107	106	104	103	102	101
XX15	116	114	113	112	111	109	108	114	113	111	110	109	108
XX16	123	121	120	119	118	116	115	121	120	118	117	116	115
XX17	130	128	127	126	125	123	122	128	127	125	124	123	122
XX18	137	135	134	133	132	130	129	135	134	132	131	130	129
XX19	144	142	141	140	139	137	136	142	141	139	138	137	136
XX20	151	149	148	147	146	144	143	149	148	146	145	144	143
XX21	151	150	149	148	151	150	150	150	149	150	150	148	150
XX22	158	163	162	161	160	150	150	163	162	160	150	148	150
XX23	165	170	169	168	167	165	164	170	169	167	166	165	164
XX24	172	177	176	175	174	172	171	177	176	174	173	172	171
XX25	179	184	183	182	181	179	178	184	183	181	180	179	178
XX26	186	191	190	189	188	186	185	191	190	188	187	186	185
XX27	193	198	197	196	195	193	192	198	197	195	194	193	192
XX28	200	205	204	203	202	200	199	205	204	202	201	200	199
XX29	207	212	211	210	209	207	206	212	211	209	208	207	206
XX30	214	219	218	217	216	214	213	219	218	216	215	214	213
XX31	221	226	225	224	223	221	220	226	225	223	222	221	220
XX32	228	233	232	231	230	228	227	233	232	230	229	228	227
XX33	235	240	239	238	237	235	234	240	239	237	236	235	234
XX34	242	247	246	245	244	242	241	247	246	244	243	242	241
XX35	249	254	253	252	251	249	248	254	253	251	250	249	248
XX36	256	261	260	259	258	256	255	261	260	258	257	256	255
XX37	263	268	267	266	265	263	262	268	267	265	264	263	262
XX38	270	275	274	273	272	270	269	275	274	272	264	266	262
XX39	277	282	281	280	279	277	276	282	281	279	278	277	276
XX40	284	289	288	287	286	284	283	289	288	286	285	284	283
XX41	291	296	295	294	293	291	290	296	295	293	292	291	290
XX42	298	303	302	301	300	298	297	303	302	300	299	298	297
XX43	305	310	309	308	307	305	304	310	309	307	306	305	304
XX44	312	317	316	315	314	312	311	317	316	314	313	312	311
XX45	319	324	323	322	321	319	318	324	323	321	320	319	318
XX46	326	331	330	329	328	326	325	331	330	328	327	326	325
XX47	333	338	337	336	335	333	332	338	337	335	327	326	325
XX48	340	345	344	343	342	340	339	345	344	342	341	340	339
XX49	347	352	351	350	349	347	346	352	351	349	348	347	346
XX50	354	359	358	357	356	354	353	359	358	356	355	354	353
XX51	361	366	365	364	363	361	360	366	365	363	362	361	360
XX52	368	373	372	371	370	368	367	373	372	370	369	368	367

Refunds processed during these cycles were accelerated.

Source: IDRS Manuals, 1980-1992 (IRS)

F. Demographics/Economics

Singles = [F3]/[C2] *100
PCBirths = [F5]/[F1] *1,000
Under30 = [F6]/[C2] *100
Over64 = [F7]/[C2] *100
AvgPI = [A5]/[C2] *100 [in constant 1992 dollars, per GDP deflator]
AvgPIgrowth = ({[A5]/[C2]} / {[A5]/[C2]}$_{-1}$ – 1)*100
ExclIncomePct = [F8]/[A5] *100
UnemplRate = [F9]
TFSEmplPct = [F10]/[F11] *100
College = [F12]/[C2] *100
Males = [F13]/[F14] *100
PopDensity = [F1]/[F15] *100

[F1] Population

Source: Population Division, Bureau of the Census

Derivation: Population estimates are available on diskette by state by year by age.

[F2] Population Under Age 18

Source: Population Division, Bureau of the Census

Derivation: Population estimates are available on diskette by state by year by age.

[F3] Number of Singles Among Potential Tax Returns

Source: Current Population Survey, Bureau of the Census

Derivation: The number of singles who would be required to file if the filing threshold were zero. (See Appendix B.)

[F4] Number of Marrieds Among Potential Tax Returns

Source: Current Population Survey, Bureau of the Census

Derivation: The number of marrieds who would be required to file if the filing threshold were zero. This combines any married records into the appropriate "Married-Joint" return (i.e., spouses on the CPS are not both counted as separate potential returns). (See Appendix B.)

[F5] Number of Births

Source: Health & Human Services

Derivation: The number of live births is available by year by state in paper reports.

[F6] Number of Potential Primary Taxpayers Under 30 Years Old

Source: Current Population Survey, Bureau of the Census

Derivation: The number of taxpayers under age 30 who would be required to file if the filing threshold were zero. (On married returns, this refers to the oldest spouse. See Appendix B.)

[F7] Number of Potential Primary Taxpayers Over 64 Years Old

Source: Current Population Survey, Bureau of the Census

Derivation: The number of taxpayers over age 64 who would be required to file if the filing threshold were zero. (On married returns, this refers to the oldest spouse. See Appendix B.)

[F8] Income on Potential Returns That is Not Taxable

Source: Current Population Survey, Bureau of the Census

Derivation: Total income reported on the CPS [PINCTOT] for all non-required returns (see Appendix B) plus the non-reportable income on required returns (Supplementary Security Income [I52B] plus an estimate of 2/3 of Veterans, Unemployment, and Workman's Compensation income [I53D]).

[F9] Unemployment Rate Among Those 16 or Older

Source: Bureau of Labor Statistics

Derivation: The percentage of those who are 16 years old or older who are in the labor market, but are unemployed. This is available on diskette by year by state.

[F10] Total Employment in the Trade, Finance & Service Sectors

Source: Bureau of Labor Statistics

Derivation: The number of people employed in these sectors is available on diskette by year by state.

[F11] Total Non-Farm Employment

Source: Bureau of Labor Statistics

Derivation: The total number of people employed in all non-farm sectors is available on diskette by year by state.

[F12] Number Among Potential Returns Having Had at Least Some College Education

Source: Current Population Survey, Bureau of the Census

Derivation: The number of taxpayers having at least some college education who would be required to file if the filing threshold were zero. (See Appendix B.)

[F13] Number of Males Among Potential Single and Head of Household Returns

Source: Current Population Survey, Bureau of the Census

Derivation: The number of males who would be required to file (Single or Head of Household) if the filing threshold were zero. (See Appendix B.)

[F14] Number of Potential Single and Head of Household Returns

Source: Current Population Survey, Bureau of the Census

Derivation: The number of taxpayers who would be required to file as Single or Head of Household if the filing threshold were zero. (See Appendix B.)

[F15] Land Area

Source: Statistical Abstract of the United States.

Derivation: The land area of each state is tabulated.

Appendix B

Method for Using the Current Population Survey To Estimate the Number of Returns Required to be Filed

Step 1: Extract the appropriate variables from each year's file. (See Table B-1.)

Step 2: Combine the two records of married couples into one record.

Step 3: Follow the logic of Table B-2 to estimate:

- which records would seem to qualify as dependents;
- which records would seem to qualify for the Head of Household filing status, and therefore which others should be classified as Married-Joint or Single;
- the following amounts:

  a. Gross Business Income [CGross and FGross]
  Gross Income includes gross business income—before expenses—but the CPS has only net business income. So, I estimated gross business income from net business income using ratios derived from the 1988 TCMP of individuals for 9 net income levels for both sole proprietors (Schedule C) and farmers (Schedule F) as follows:

Ratio of Gross to Net Business Income

Net Income			Ratio of Gross to Net	
			Schedule C	Schedule F
	≤	-$5,000	-1.945	-1.524
-$4,999	-	$0	-3.787	-5.020
$1	-	$500	8.440	24.624
$501	-	$1,000	3.603	11.954
$1,001	-	$2,500	2.874	9.750
$2,501	-	$5,000	2.499	8.114
$5,001	-	$25,000	2.226	5.323
$25,001	-	$50,000	2.138	3.946
	>	$50,000	1.876	3.266

Source: 1988 TCMP of individuals

  b. Gross Income subject to tax [GI]

  GI = Wages + CGross + FGross + Interest + Dividends,Rent,Royalties + SocSec,RR + (Veterans, Unemployment, Workmans Comp)/3 + Pensions

  c. Filing Threshold [FThresh]: standard deduction plus personal exemptions based on the estimated filing status and the tax parameters summarized in Table B-3; and

- which records would seem to be required to file a return:
  a. those with GI > FThresh; OR
  b. those with GI < 0; OR
  c. those required to file Schedule SE; OR
  d. those with CNet<0 OR FNet<0.

Table B-1. An Explanation of the Variables and Codes Used on the Current Population Survey

Variable/Code	Explanation
MARSTAT	Marital Status
1	Married
2	Widowed
3	Divorced
4	Separated
5	Never Married
AGE	Years; if married, age of older spouse
SEX	
1	Male
2	Female
HIGHGRAD	Highest Educational Grade Level
0	Children under 15
1-19	Highest grade attended + 1
TENURE	Housing occupancy
0	Not applicable
1	Owned or being bought
2	Rented
3	No cash rent
FREC10	Number of own children (never married) under 18
FREC26	Labor Force Status of Householder and Spouse
0	Not applicable
1	Both spouses in labor force
2	One spouse in labor force
3	Single Male
4	Single Female
I53CDIV	Dividends recipiency
0	Not applicable
1	Yes
2	No
I53CRENT	Rent & royalties recipiency
0	Not applicable
1	Yes
2	No
I51A	Wage & salary income
I51B	Self-employment income
I51C	Farm income
I52A	Income from SS or RR retirement
I52B	Supplementary Security income
I53B	Interest income
I53C	Dividend, rent & royalty income
I53D	Veterans, unemployment & workman's compensation
I53E	Pension income
I53F	Child support income, other
PINCTOT	Total income
MARSUPPW	Weight x 100 (i.e., two implied decimal places)

Table B-2. Logic for Identifying Income Tax Returns by Filing Status from the Current Population Survey, 1980-1992

Criterion 1: **Dependency Test**	3. [illegible] AND b. AGE > 23 and **[GI]** < Exemption Value from Table A-3 (Part A) one of t he following Income tests is met: : a. AGE < 24 and I51A < 3.5 x Single Standard Deduction (Table A-3,Part A) and I51B =0 and I51C = 0; OR				
	If NO: Non-Dependents			**If YES: Dependents**	
Criterion 2: **Filing Status**	Married-Joint	Head of Household	Single	If Year < 1987 Single	If Year > 1986 Single
	MARSTAT=1	MARSTAT=2, 3, 4 or 5 and FREC10 > 0 and TENURE > 0	MARSTAT=2, 3, 4 or 5 and Filing Status ≠ Head of Household	MARSTAT=2, 3, 4 or 5	MARSTAT=2, 3, 4 or 5
Gross Income [GI] (subject to income tax)	I51A+CGross+FGross+I52A +I53B+I53C+(I53D)/3+I53E	51A+CGross+FGross+I52A +I53B+I53C+(I53D)/3+I53E	51A+CGross+FGross+I52A +I53B+I53C+(I53D)/3+I53E	51A+CGross+FGross+I52A +I53B+I53C+(I53D)/3+I53E	51A+CGross+FGross+I52A +I53B+I53C+(I53D)/3+I53E
Exemptions a. Taxpayer(s) b. If AGE>64 and Year < 1987 c. Total number d. $ per exemption **[EV]** e. Total value **[EX]**	 2 2 a + b Table B-3 (Part A) c x d	 1 1 a + b Table B-3 (Part A) c x d	 1 1 a + b Table B-3 (Part A) c x d	 1 1 a + b Table B-3 (Part A) c x d	 0 0 0 0 0
Standard Deduction [SD]	Table B-3 (Parts A + C)	Table B-3 (Parts A + C)	Table B-3 (Parts A + C)	Table B-3 (Part B)	Table B-3 (Parts B + C)
Filing Threshold [FThresh]	SD+2EV	SD+EV	SD+EV	SD+EV	SD
Criterion 3: **Is a required tax return if:**	1. GI > FThresh OR 2. GI < 0 OR 3. SchedSE=1 OR 4. CNet < 0 OR 5. FNet < 0	1. GI > FThresh OR 2. GI < 0 OR 3. SchedSE=1 OR 4. CNet < 0 OR 5. FNet < 0	1. GI > FThresh OR 2. GI < 0 OR 3. SchedSE=1 OR 4. CNet < 0 OR 5. FNet < 0	1. GI > FThresh OR 2. GI < 0 OR 3. SchedSE=1 OR 4. CNet < 0 OR 5. FNet < 0	1. GI > FThresh OR 2. GI < 0 OR 3. SchedSE=1 OR 4. CNet < 0 OR 5. FNet < 0

Table B-3. Exemption and Standard Deduction Values, TY80-TY92

Tax Year	Exemption value	Standard Deduction, by Filing Status: Married-Joint	Married-Separate	Single	Head of Household
Part A: Non-Dependents					
1980	$1,000	$3,400	$1,700	$2,300	$2,300
1981	$1,000	$3,400	$1,700	$2,300	$2,300
1982	$1,000	$3,400	$1,700	$2,300	$2,300
1983	$1,000	$3,400	$1,700	$2,300	$2,300
1984	$1,000	$3,400	$1,700	$2,300	$2,300
1985	$1,040	$3,540	$1,770	$2,390	$2,390
1986	$1,080	$3,670	$1,835	$2,480	$2,480
1987	$1,900	$3,760	$1,880	$2,540	$2,540
1988	$1,950	$5,000	$2,500	$3,000	$4,400
1989	$2,000	$5,200	$2,600	$3,100	$4,550
1990	$2,050	$5,450	$2,725	$3,250	$4,750
1991	$2,150	$5,700	$2,850	$3,400	$5,000
1992	$2,300	$6,000	$3,000	$3,600	$5,250
Part B: Dependents					
1980-1986: If unearned income (l53B+l53C) is greater than the Exemption value from Part A, then Standard Deduction is equal either to earned income or the amount below—whichever is smaller.					
1980				$2,300	
1981				$2,300	
1982				$2,300	
1983				$2,300	
1984				$2,300	
1985				$2,390	
1986				$2,480	
1987-1992: If earned income (l51A+l51B+l51C) is less than Standard Deduction from Part A, then Standard Deduction is equal either to earned income or the amount below—whichever is larger.					
1987				$500	
1988				$500	
1989				$500	
1990				$500	
1991				$550	
1992				$600	
Part C: Elderly					
Additional Standard Deduction if Taxpayer is 65 or Older (whether dependent or not)					
1987		$2,440	$1,220	$1,210	$2,610
1988		$1,200	$600	$750	$750
1989		$1,200	$600	$750	$750
1990		$1,300	$650	$800	$800
1991		$1,300	$650	$850	$850
1992		$1,400	$700	$900	$900

Appendix C

Three Definitions of Income and Offsets

Rule Changes

During the time period of this study (1982-1991), numerous changes have taken place in the rules governing what income must be reported and what offsets may be claimed—often introducing or eliminating entire items, but usually modifying the amounts somewhat. In particular, the Tax Reform Act of 1986 introduced enormous change in the middle of this time period. All of the changes are summarized by year in Table C-1.

The Problem

The problem that these changes introduce is that they change the amount of income and offsets that taxpayers are required to (or able to) report. Unless we can control for the impact of these rule changes, we might interpret the increase or decrease in the amount reported as reflecting a change in compliance. This is particularly serious if an important explanatory variable underwent changes at the same time as the rule changes; the change in reporting caused by the changed rules could inadvertently be attributed to the change in the explanatory variable.

The Solution

In the case of dividend income, the elimination (beginning in 1987) of the $100 or $200 exclusion was controlled for by modifying the data—creating a constant-law dividend series. This was done by continuing to apply the dividend exclusion in the 1987-1991 micro data from SOI before the data were aggregated to the state level.

Several other rule changes were controlled for by including appropriate explanatory variables that captured the effects of the changes. Major changes (including indexing) of personal exemptions and standard deductions were controlled for with the variables *FThresholdPct* and *ChildExemptsPct*. Changing rules governing the deductibility of state and local taxes were controlled for with the variable *StateTaxPct*.

Most of the rule changes, however, did not lend themselves to either of these two solutions. For that reason, I tested the sensitivity of my results to the remaining rule changes by defining each of my reporting compliance dependent variables (*IncomePct*, *OffsetsPct*, and *NetIncomePct*) according to three distinct definitions:

Definition A: excludes all components of income or offsets whose rules changed during the period, which changes could not be controlled for in any other way.

Definition B: excludes only those components of income or offsets whose rules changed in years other than 1986 (i.e., in the Tax Reform Act of 1986).

Definition C: includes all components of income and offsets regardless of rule changes.

Table C-2 indicates which specific components are included in each of the three definitions. Tables C-3 through C-5 summarize the results of applying the same specification to a given dependent variable defined in each of these three ways.

Table C-1. Summary of Major Income and Offset Rule Changes, 1983-1991

Tax Year	Income Items	Offset Items
1983	+ Withholding on pensions & annuities begun	+ Excess Medicare/Hosp. Ins. Benefits Credit introduced – Threshold for medical expense deduction incr. to 5%; no alternative for medical insurance premiums – Casualty & theft losses subject to 10% of AGI threshold
1984	+ Social Security benefits partly taxable + Disability income exclusion eliminated	+ Limit for non-itemized charitable contrib. increased to $75 + 1% threshold for medicine & drug deduction eliminated – New rules for casualty & theft losses – New rules for credit for the elderly/disabled – General Business Credit begun; limits other credits
1985	+ New alimony definitions apply + Need to provide SSN of alimony recipient	+ Limit for non-itemized charitable contrib. increased to 50% + Indexing begun for exemptions & standard deductions + Increase in Earned Income Credit – Foreign Housing deduction eliminated
1986		+ Limit for non-itemized charitable contributions removed – Residential Energy Credit eliminated – Credit for research subject to General Business Credit
1987 (Tax Reform Act of 1986)	+ Dividend exclusion eliminated + Capital gains exclusion eliminated + Unemployment compensation 100% taxable	+ Self-employed helath insurance adjustment added + Exemption value almost doubled + Blind Heads of Households offset more than Singles + Increase in Earned Income Tax Credit + Introduction of credit for prior-yr. minimum tax + Introduction of Mortgage Interest Credit – IRAs reduced for high-AGI returns – Deduction for married couple when both work eliminated – Non-itemizer deduction for charitable contribution ended – SSN required for dependents ≥ 5 years old – Dependents no longer able to claim their own exemption – Standard deduction for dependents greatly reduced – Threshold for medical expense deduction incr. to 7.5% – Sales tax deduction eliminated – Limits on mortgage interest deduction – Only 65% of personal interest deductible – Moving expense was adjstment, now itemized deduction – 2% of AGI threshold for some misc. itemized deductions – Political contributions credit eliminated – Investment credit generally repealed
1988		+ Head of Household standard deduction greatly increased + Standard deductions greatly increased–except dependents + Less stringent limits on home mortgage interest deduction + Earned Income Credit income eligibility greatly expanded – Aged/blind offset reduced – Only 40% of personal interest deductible
1989		+ New adjustment for qualified performing artist expense + New adjustment for jury duty pay given to employer + New adjustment for employer-provided vehicle – Reimbursed employee business expense eliminated – Age when dependent SSN is required decreased to 2 – Only 20% of personal interest deductible – Child & Dependent Care Credit requires SSN of provider
1990	– Series EE savings bond interest deductible for education expenses	+ Half of self employment tax deductible (adjustment) – Adjustment for employer-provided vehicle eliminated – Only 10% of personal interest deductible
1991		+ Disabled access credit introduced + Enhanced oil recovery credit introduced + Earned Income Credit greatly expanded; Singles eligible – Exemptions reduced for high-income returns – Age when dependent SSN is required decreased to 1 – SSN of child required for Earned Income Credit – Itemized deductions limited for high-income returns – Personal interest deduction eliminated completely

+ denotes increases; – denotes decreases; *italics* denote compliance-oriented requirements

Table C-2. Income and Offset Components Included in the Three Alternate Definitions

Component	Definition A	Definition B	Definition C	Component	Definition A	Definition B	Definition C
Income Items				**Offsets: Itemized Deductions**			
Wages, salaries, tips, etc.	✓	✓	✓	State & local income taxes	✓	✓	✓
Interest (Schedule B)	✓	✓	✓	Real estate taxes	✓	✓	✓
Dividends (Schedule B)	1	✓	✓	Sales taxes	2	2	2
Refunds of state & local income tax	✓	✓	✓	Other taxes	✓	✓	✓
Alimony received	✓	✓	✓	Home mortgage interest	✓	✓	✓
Business income (Schedule C)	✓	✓	✓	Deductible points	✓	✓	✓
Pensions, IRA distributions, etc.	✓	✓	✓	Investment interest	✓	✓	✓
Rents, royalties (Schedule E)	✓	✓	✓	Personal interest		✓	✓
Partnerships, estates, trusts (Sched. E)	✓	✓	✓	Charitable contributions (total)	✓	✓	✓
Farm income (Schedule F)	✓	✓	✓	Medical & dental expenses			✓
Other income	✓	✓	✓	Casualty & theft losses			✓
Capital gain/loss (Schedule D)			✓	Moving expenses		✓	✓
Capital gain distribution			✓	Miscellaneous deductions		✓	✓
Other gain/loss (Form 4797)			✓	**Offsets: Credits**			
Unemployment compensation		✓	✓	Child & dependent care credit	✓	✓	✓
Social Security benefits			✓	Elderly/disabled credit	✓	✓	✓
Offsets: Adjustments				Foreign tax credit	✓	✓	✓
Keogh/SEP deduction	✓	✓	✓	Residential energy credit			✓
Penalty on early withdrawals	✓	✓	✓	Political contributions credit			✓
Alimony paid	✓	✓	✓	General business credit			✓
Reimbursed employee business expenses			✓	Investment credit			✓
IRA contributions		✓	✓	Jobs credit			✓
1/2 of self-employment tax			✓	Alcohol fuel credit			✓
Self-employed health insurance		✓	✓	Research credit			✓
Ded'n for married couple when both work		✓	✓	Low-income housing credit			✓
Disability Income exclusion			✓	Disabled access credit			✓
Foreign Housing deduction (F2555)			✓	Enhanced oil recovery credit			✓
Forestation amortization			✓	Earned Income credit			✓
Repaymnt subpay under T.A. of 1974			✓	Gas & special fuels credit			✓
Qualified performng artist expenses			✓	Regulated invest. co. credit			✓
Jury duty pay given to employer			✓	Nonconventional fuel credit			✓
Employer-provided vehicle			✓	Mortgage interest credit			✓
Offsets: Exemps./Std. Ded./etc.							
Personal exemptions	2	2	2				
Aged/Blind exemption/std. deduction	2	2	2				
Standard deduction	2	2	2				
Charitable contributions for nonitemizers			✓				

1 Rule change controlled for by modifying data.
2 Rule changes controlled for with independent variables.

Table C-3A. Alternate Estimates of the Determinants of Voluntary Reporting Compliance: IncomePct Equation, Definition A

	Reduced Form	AuditRate Exogenous	AuditRate Endogenous
Ln(AuditRate+1)		-0.188470	
		(-0.17)	
Ln(pAuditRate+1)			16.158539
			(3.37)
FilingRate	0.348404	0.349188	0.345586
	(7.61)	(7.58)	(7.64)
FThresholdPct	1.308479	1.314732	1.182627
	(4.45)	(4.43)	(4.04)
MargTaxRate@$15K	1.444192	1.446746	1.221297
	(1.37)	(1.37)	(1.17)
MargTaxRate@$57K	-1.918453	-1.926786	-1.978458
	(-0.96)	(-0.96)	(-1.00)
ChildExemptsPct	1.513413	1.512272	1.475395
	(1.88)	(1.88)	(1.86)
Ln(AvgBurden)	3.620572	3.539073	3.383676
	(0.58)	(0.56)	(0.55)
SoleProps	1.864310	1.851415	1.428688
	(2.40)	(2.37)	(1.84)
SolePropTFS	-3.812498	-3.782585	-2.925128
	(-2.65)	(-2.60)	(-2.02)
PaidPrep	-0.159916	-0.160512	-0.166282
	(-4.58)	(-4.57)	(-4.81)
Ln(IRP+1)	-5.270040	-5.204916	-1.121633
	(-1.09)	(-1.07)	(-0.23)
Ln(CID+1)	0.880407	0.876367	0.932191
	(2.88)	(2.85)	(3.08)
TPS_CallsPC	-0.004920	-0.004909	-0.003994
	(-1.64)	(-1.63)	(-1.34)
TPS_RetPrepPC	0.123438	0.124580	0.130914
	(1.92)	(1.93)	(2.07)
Singles	0.198022	0.199427	0.266954
	(1.07)	(1.08)	(1.46)
Under30	-0.111077	-0.111273	-0.098600
	(-1.22)	(-1.22)	(-1.10)
Over64	-0.092157	-0.092702	-0.075873
	(-0.91)	(-0.92)	(-0.76)
PCBirths	0.867257	0.869632	0.991262
	(3.99)	(3.99)	(4.55)
ExclIncomePct	-0.475907	-0.475467	-0.642278
	(-1.18)	(-1.18)	(-1.60)
UnemplRate	-0.342146	-0.339953	-0.370384
	(-2.71)	(-2.68)	(-2.97)
Adj. R-squared	0.751512	0.750927	0.757573
Hausman Test Statistic			12.60
Significance (percent)			0.04

2SLSDV estimates from state-level panel data for 1982-1991; t-statistics in parentheses.

Table C-3B. Alternate Estimates of the Determinants of Voluntary Reporting Compliance: IncomePct Equation, Definition B

	Reduced Form	AuditRate Exogenous	AuditRate Endogenous
Ln(AuditRate+1)		-0.234193	
		(-0.21)	
Ln(pAuditRate+1)			15.823423
			(3.30)
FilingRate	0.343673	0.344646	0.340913
	(7.52)	(7.50)	(7.55)
FThresholdPct	1.290814	1.298584	1.167572
	(4.39)	(4.38)	(3.99)
MargTaxRate@$15K	1.445924	1.449096	1.227651
	(1.37)	(1.37)	(1.18)
MargTaxRate@$57K	-1.930628	-1.940983	-1.989389
	(-0.97)	(-0.97)	(-1.01)
ChildExemptsPct	1.539394	1.537976	1.502164
	(1.91)	(1.91)	(1.89)
Ln(AvgBurden)	4.034964	3.933693	3.802981
	(0.65)	(0.63)	(0.62)
SoleProps	1.838605	1.822581	1.412017
	(2.37)	(2.34)	(1.82)
SolePropTFS	-3.762151	-3.724981	-2.893184
	(-2.61)	(-2.57)	(-2.00)
PaidPrep	-0.157733	-0.158474	-0.163967
	(-4.52)	(-4.52)	(-4.75)
Ln(IRP+1)	-4.979131	-4.898207	-0.916759
	(-1.03)	(-1.01)	(-0.19)
Ln(CID+1)	0.897546	0.892526	0.948256
	(2.94)	(2.91)	(3.13)
TPS_CallsPC	-0.004919	-0.004905	-0.004012
	(-1.64)	(-1.63)	(-1.35)
TPS_RetPrepPC	0.116803	0.118222	0.124124
	(1.82)	(1.83)	(1.96)
Singles	0.203393	0.205140	0.270896
	(1.10)	(1.11)	(1.48)
Under30	-0.112213	-0.112457	-0.099995
	(-1.24)	(-1.24)	(-1.11)
Over64	-0.103635	-0.104313	-0.087689
	(-1.03)	(-1.03)	(-0.88)
PCBirths	0.876360	0.879312	0.997793
	(4.04)	(4.04)	(4.58)
ExclIncomePct	-0.439579	-0.439033	-0.602500
	(-1.09)	(-1.09)	(-1.50)
UnemplRate	-0.311743	-0.309018	-0.339395
	(-2.47)	(-2.44)	(-2.72)
Adj. R-squared	0.754386	0.753817	0.760122
Hausman Test Statistic			12.22
Significance (percent)			0.05

2SLSDV estimates from state-level panel data for 1982-1991; t-statistics in parentheses.

Table C-3C. Alternate Estimates of the Determinants of Voluntary Reporting Compliance: IncomePct Equation, Definition C

	Reduced Form	AuditRate Exogenous	AuditRate Endogenous
Ln(AuditRate+1)		0.747762	
		(0.54)	
Ln(pAuditRate+1)			17.100229
			(2.79)
FilingRate	0.252764	0.249655	0.249782
	(4.34)	(4.26)	(4.32)
FThresholdPct	0.817877	0.793067	0.684691
	(2.18)	(2.10)	(1.83)
MargTaxRate@$15K	0.623057	0.612927	0.387171
	(0.46)	(0.46)	(0.29)
MargTaxRate@$57K	-4.116766	-4.083704	-4.180268
	(-1.62)	(-1.60)	(-1.65)
ChildExemptsPct	0.494783	0.499309	0.454549
	(0.48)	(0.49)	(0.45)
Ln(AvgBurden)	12.445461	12.768814	12.194758
	(1.56)	(1.60)	(1.54)
SoleProps	1.631045	1.682206	1.170036
	(1.65)	(1.70)	(1.18)
SolePropTFS	-3.625201	-3.743883	-2.686116
	(-1.98)	(-2.03)	(-1.45)
PaidPrep	0.243538	0.245903	0.236801
	(5.48)	(5.50)	(5.36)
Ln(IRP+1)	-12.130604	-12.388988	-7.740435
	(-1.96)	(-2.00)	(-1.22)
Ln(CID+1)	0.353773	0.369801	0.408575
	(0.91)	(0.95)	(1.06)
TPS_CallsPC	0.001980	0.001938	0.002961
	(0.52)	(0.51)	(0.78)
TPS_RetPrepPC	0.056846	0.052313	0.064757
	(0.70)	(0.64)	(0.80)
Singles	0.479234	0.473658	0.552184
	(2.04)	(2.01)	(2.36)
Under30	-0.294153	-0.293373	-0.280949
	(-2.55)	(-2.54)	(-2.45)
Over64	-0.143764	-0.141600	-0.126532
	(-1.12)	(-1.10)	(-0.99)
PCBirths	0.364124	0.354701	0.495356
	(1.32)	(1.28)	(1.78)
ExclIncomePct	-0.004561	-0.006305	-0.180628
	(-0.01)	(-0.01)	(-0.35)
UnemplRate	-0.584247	-0.592947	-0.614131
	(-3.64)	(-3.67)	(-3.85)
Adj. R-squared	0.684621	0.684076	0.689705
Hausman Test Statistic			7.46
Significance (percent)			0.66

2SLSDV estimates from state-level panel data for 1982-1991; t-statistics in parentheses.

Table C-4A. Alternate Estimates of the Determinants of Voluntary Reporting Compliance: OffsetsPct Equation, Definition A

	Reduced Form	AuditRate Exogenous	AuditRate Endogenous
Ln(AuditRate+1)		-0.014032	
		(-0.04)	
Ln(pAuditRate+1)			3.313904
			(2.00)
FilingRate	0.137765	0.137827	0.137683
	(8.70)	(8.64)	(8.72)
FThresholdPct	0.888633	0.889081	0.857427
	(8.70)	(8.64)	(8.33)
MargTaxRate@$15K	-0.587938	-0.587949	-0.663976
	(-1.60)	(-1.60)	(-1.80)
MargTaxRate@$57K	0.555397	0.554778	0.530545
	(0.80)	(0.79)	(0.76)
ChildExemptsPct	0.480910	0.480669	0.457696
	(1.71)	(1.70)	(1.63)
StateTaxPct	0.114491	0.114837	0.145114
	(1.60)	(1.59)	(1.99)
Ln(AvgBurden)	-3.713426	-3.720379	-3.550471
	(-1.85)	(-1.84)	(-1.77)
SoleProps	0.334352	0.333725	0.274123
	(1.19)	(1.19)	(0.98)
SolePropTFS	-0.649006	-0.647430	-0.527449
	(-1.25)	(-1.24)	(-1.01)
PaidPrep	-0.013940	-0.013982	-0.014858
	(-1.15)	(-1.14)	(-1.23)
Ln(CID+1)	0.303220	0.302891	0.314909
	(2.85)	(2.83)	(2.96)
TPS_CallsPC	-0.001075	-0.001073	-0.000742
	(-1.04)	(-1.04)	(-0.71)
TPS_RetPrepPC	-0.008631	-0.008554	-0.007756
	(-0.39)	(-0.38)	(-0.35)
Singles	0.061474	0.061546	0.072872
	(0.96)	(0.96)	(1.14)
Under30	-0.012759	-0.012760	-0.008269
	(-0.40)	(-0.40)	(-0.26)
Over64	-0.044852	-0.044891	-0.042744
	(-1.29)	(-1.28)	(-1.23)
PCBirths	0.243101	0.243143	0.253864
	(3.43)	(3.42)	(3.58)
UnemplRate	0.088225	0.088377	0.078118
	(2.08)	(2.07)	(1.84)
Adj. R-squared	0.918641	0.918445	0.919223
Hausman Test Statistic			4.48
Significance (percent)			3.48

2SLSDV estimates from state-level panel data for 1982-1991; t-statistics in parentheses.

Table C-4B. Alternate Estimates of the Determinants of Voluntary Reporting Compliance: OffsetsPct Equation, Definition B

	Reduced Form	AuditRate Exogenous	AuditRate Endogenous
Ln(AuditRate+1)		0.400095	
		(0.91)	
Ln(pAuditRate+1)			0.838515
			(0.44)
FilingRate	0.156508	0.154739	0.156487
	(8.60)	(8.45)	(8.59)
FThresholdPct	0.971414	0.958624	0.963518
	(8.28)	(8.11)	(8.11)
MargTaxRate@$15K	-0.875888	-0.875585	-0.895127
	(-2.08)	(-2.07)	(-2.11)
MargTaxRate@$57K	1.041951	1.059591	1.035663
	(1.30)	(1.32)	(1.29)
ChildExemptsPct	0.418671	0.425551	0.412797
	(1.29)	(1.32)	(1.27)
StateTaxPct	0.171034	0.161168	0.178782
	(2.08)	(1.94)	(2.12)
Ln(AvgBurden)	-3.888890	-3.690652	-3.847658
	(-1.68)	(-1.59)	(-1.66)
SoleProps	0.306918	0.324786	0.291678
	(0.95)	(1.01)	(0.90)
SolePropTFS	-0.532798	-0.577720	-0.502041
	(-0.89)	(-0.96)	(-0.83)
PaidPrep	-0.028998	-0.027802	-0.029230
	(-2.08)	(-1.98)	(-2.09)
Ln(CID+1)	0.321109	0.330501	0.324067
	(2.62)	(2.69)	(2.64)
TPS_CallsPC	-0.000976	-0.001026	-0.000892
	(-0.82)	(-0.87)	(-0.74)
TPS_RetPrepPC	-0.009437	-0.011617	-0.009215
	(-0.37)	(-0.45)	(-0.36)
Singles	0.090286	0.088233	0.093170
	(1.23)	(1.20)	(1.26)
Under30	-0.041797	-0.041785	-0.040661
	(-1.15)	(-1.15)	(-1.11)
Over64	-0.089209	-0.088098	-0.088676
	(-2.23)	(-2.20)	(-2.21)
PCBirths	0.342355	0.341163	0.345078
	(4.20)	(4.19)	(4.22)
UnemplRate	0.066211	0.061880	0.063654
	(1.36)	(1.26)	(1.30)
Adj. R-squared	0.812830	0.812749	0.812464
Hausman Test Statistic			0.68
Significance (percent)			41.10

2SLSDV estimates from state-level panel data for 1982-1991; t-statistics in parentheses.

Table C-4C. Alternate Estimates of the Determinants of Voluntary Reporting Compliance: OffsetsPct Equation, Definition C

	Reduced Form	AuditRate Exogenous	AuditRate Endogenous
Ln(AuditRate+1)		0.468282	
		(0.93)	
Ln(pAuditRate+1)			-0.525384
			(-0.24)
FilingRate	0.181028	0.178958	0.181041
	(8.70)	(8.55)	(8.70)
FThresholdPct	1.065443	1.050473	1.070390
	(7.95)	(7.78)	(7.88)
MargTaxRate@$15K	-0.572093	-0.571739	-0.560038
	(-1.19)	(-1.18)	(-1.15)
MargTaxRate@$57K	1.332833	1.353478	1.336773
	(1.46)	(1.48)	(1.46)
ChildExemptsPct	0.478022	0.486074	0.481702
	(1.29)	(1.31)	(1.30)
StateTaxPct	0.192764	0.181216	0.187909
	(2.05)	(1.91)	(1.95)
Ln(AvgBurden)	-4.523191	-4.291167	-4.549026
	(-1.71)	(-1.62)	(-1.72)
SoleProps	0.175650	0.196563	0.185199
	(0.48)	(0.53)	(0.50)
SolePropTFS	-0.265604	-0.318182	-0.284876
	(-0.39)	(-0.46)	(-0.41)
PaidPrep	-0.029982	-0.028582	-0.029836
	(-1.88)	(-1.78)	(-1.87)
Ln(CID+1)	0.365774	0.376766	0.363921
	(2.61)	(2.68)	(2.59)
TPS_CallsPC	-0.001308	-0.001367	-0.001361
	(-0.97)	(-1.01)	(-0.99)
TPS_RetPrepPC	-0.007447	-0.009999	-0.007586
	(-0.26)	(-0.34)	(-0.26)
Singles	0.049462	0.047058	0.047654
	(0.59)	(0.56)	(0.56)
Under30	-0.052673	-0.052659	-0.053385
	(-1.27)	(-1.27)	(-1.28)
Over64	-0.066966	-0.065665	-0.067300
	(-1.46)	(-1.43)	(-1.47)
PCBirths	0.504793	0.503399	0.503087
	(5.42)	(5.40)	(5.38)
UnemplRate	0.034251	0.029182	0.035853
	(0.62)	(0.52)	(0.64)
Adj. R-squared	0.771141	0.771063	0.770619
Hausman Test Statistic			0.01
Significance Level (percent)			91.74

2SLSDV estimates from state-level panel data for 1982-1991; t-statistics in parentheses.

Table C-5A. Alternate Estimates of the Determinants of Voluntary Reporting Compliance: NetIncomePct Equation, Definition A

	Reduced Form	AuditRate Exogenous	AuditRate Endogenous
Ln(AuditRate+1)		-0.164486	
		(-0.18)	
Ln(pAuditRate+1)			13.892113
			(3.46)
FilingRate	0.210336	0.211018	0.207853
	(5.57)	(5.55)	(5.57)
FThresholdPct	0.451641	0.456955	0.339758
	(1.86)	(1.86)	(1.40)
MargTaxRate@$15K	2.158997	2.159429	1.921272
	(2.47)	(2.47)	(2.22)
MargTaxRate@$57K	-2.411405	-2.417894	-2.442911
	(-1.46)	(-1.46)	(-1.50)
ChildExemptsPct	1.068153	1.065778	1.000080
	(1.60)	(1.59)	(1.52)
StateTaxPct	-0.198065	-0.194302	-0.101522
	(-1.14)	(-1.11)	(-0.59)
Ln(AvgBurden)	5.327885	5.247584	4.888900
	(1.03)	(1.01)	(0.96)
SoleProps	1.448109	1.440579	1.169128
	(2.19)	(2.17)	(1.77)
SolePropTFS	-2.975202	-2.956409	-2.399908
	(-2.41)	(-2.38)	(-1.95)
PaidPrep	-0.148681	-0.149157	-0.153009
	(-5.14)	(-5.13)	(-5.36)
Ln(IRP+1)	-2.437200	-2.398065	0.675205
	(-0.60)	(-0.58)	(0.16)
Ln(CID+1)	0.556926	0.553086	0.593380
	(2.20)	(2.17)	(2.37)
TPS_CallsPC	-0.004338	-0.004322	-0.003378
	(-1.74)	(-1.73)	(-1.36)
TPS_RetPrepPC	0.131082	0.132038	0.136453
	(2.47)	(2.48)	(2.61)
Singles	0.137200	0.138211	0.190927
	(0.90)	(0.90)	(1.26)
Under30	-0.104943	-0.105004	-0.091372
	(-1.40)	(-1.40)	(-1.23)
Over64	-0.035875	-0.036364	-0.022223
	(-0.43)	(-0.44)	(-0.27)
PCBirths	0.650126	0.651351	0.734990
	(3.54)	(3.54)	(4.02)
ExclIncomePct	-0.789413	-0.788466	-0.917979
	(-2.36)	(-2.36)	(-2.77)
UnemplRate	-0.399323	-0.397605	-0.428625
	(-3.82)	(-3.78)	(-4.14)
Adj. R-squared	0.795961	0.795481	0.801260
Hausman Test Statistic			14.79
Significance Level (percent)			0.01

2SLSDV estimates from state-level panel data for 1982-1991; t-statistics in parentheses.

Table C-5B. Alternate Estimates of the Determinants of Voluntary Reporting Compliance: NetIncomePct Equation, Definition B

	Reduced Form	AuditRate Exogenous	AuditRate Endogenous
Ln(AuditRate+1)		-0.600534	
		(-0.69)	
Ln(pAuditRate+1)			15.751865
			(4.11)
FilingRate	0.189413	0.191902	0.186598
	(5.22)	(5.26)	(5.24)
FThresholdPct	0.346964	0.366365	0.220104
	(1.49)	(1.56)	(0.95)
MargTaxRate@$15K	2.412418	2.413994	2.142868
	(2.87)	(2.87)	(2.59)
MargTaxRate@$57K	-2.953985	-2.977680	-2.989710
	(-1.86)	(-1.88)	(-1.92)
ChildExemptsPct	1.129008	1.120337	1.051822
	(1.76)	(1.75)	(1.67)
StateTaxPct	-0.232960	-0.219220	-0.123492
	(-1.40)	(-1.31)	(-0.75)
Ln(AvgBurden)	5.907853	5.614677	5.410100
	(1.19)	(1.12)	(1.11)
SoleProps	1.465337	1.437846	1.149008
	(2.30)	(2.25)	(1.83)
SolePropTFS	-3.070375	-3.001760	-2.418065
	(-2.59)	(-2.52)	(-2.06)
PaidPrep	-0.132227	-0.133963	-0.137134
	(-4.76)	(-4.80)	(-5.03)
Ln(IRP+1)	-4.427897	-4.285017	-0.898832
	(-1.13)	(-1.09)	(-0.23)
Ln(CID+1)	0.555437	0.541417	0.596771
	(2.29)	(2.22)	(2.50)
TPS_CallsPC	-0.004144	-0.004086	-0.003054
	(-1.73)	(-1.70)	(-1.29)
TPS_RetPrepPC	0.121947	0.125437	0.128037
	(2.40)	(2.45)	(2.56)
Singles	0.104311	0.108001	0.165231
	(0.71)	(0.74)	(1.14)
Under30	-0.073649	-0.073871	-0.058261
	(-1.02)	(-1.02)	(-0.82)
Over64	-0.001542	-0.003329	0.013938
	(-0.02)	(-0.04)	(0.18)
PCBirths	0.517605	0.522077	0.613829
	(2.93)	(2.96)	(3.51)
ExclIncomePct	-0.795554	-0.792093	-0.941330
	(-2.48)	(-2.47)	(-2.97)
UnemplRate	-0.344540	-0.338269	-0.377765
	(-3.43)	(-3.35)	(-3.82)
Adj. R-squared	0.816035	0.815798	0.822874
Hausman Test Statistic			22.68
Significance Level (percent)			0.00

2SLSDV estimates from state-level panel data for 1982-1991; t-statistics in parentheses.

Table C-5C. Alternate Estimates of the Determinants of Voluntary Reporting Compliance: NetIncomePct Equation, Definition C

	Reduced Form	AuditRate Exogenous	AuditRate Endogenous
Ln(AuditRate+1)		0.505934	
		(0.38)	
Ln(pAuditRate+1)			17.449324
			(2.97)
FilingRate	0.075358	0.073262	0.072239
	(1.37)	(1.32)	(1.32)
FThresholdPct	-0.206899	-0.223244	-0.347430
	(-0.58)	(-0.62)	(-0.98)
MargTaxRate@$15K	1.412672	1.411344	1.114075
	(1.11)	(1.10)	(0.88)
MargTaxRate@$57K	-5.496076	-5.476114	-5.535650
	(-2.28)	(-2.27)	(-2.32)
ChildExemptsPct	0.112393	0.119698	0.026889
	(0.12)	(0.12)	(0.03)
StateTaxPct	-0.516877	-0.528452	-0.395614
	(-2.04)	(-2.07)	(-1.56)
Ln(AvgBurden)	15.243426	15.490419	14.692035
	(2.02)	(2.04)	(1.96)
SoleProps	1.126800	1.149961	0.776383
	(1.16)	(1.19)	(0.80)
SolePropTFS	-2.680483	-2.738289	-1.957879
	(-1.49)	(-1.51)	(-1.09)
PaidPrep	0.266116	0.267579	0.260680
	(6.31)	(6.31)	(6.23)
Ln(IRP+1)	-11.014677	-11.135050	-7.105312
	(-1.85)	(-1.86)	(-1.17)
Ln(CID+1)	-0.015120	-0.003308	0.030668
	(-0.04)	(-0.01)	(0.08)
TPS_CallsPC	0.002709	0.002661	0.003916
	(0.74)	(0.73)	(1.08)
TPS_RetPrepPC	0.061035	0.058095	0.067781
	(0.79)	(0.75)	(0.88)
Singles	0.431267	0.428158	0.498752
	(1.94)	(1.92)	(2.25)
Under30	-0.251773	-0.251586	-0.234727
	(-2.30)	(-2.29)	(-2.16)
Over64	-0.060353	-0.058847	-0.043205
	(-0.50)	(-0.48)	(-0.36)
PCBirths	-0.115332	-0.119100	-0.008738
	(-0.43)	(-0.44)	(-0.03)
ExclIncomePct	-0.472732	-0.475647	-0.634217
	(-0.97)	(-0.97)	(-1.31)
UnemplRate	-0.565252	-0.570536	-0.602056
	(-3.70)	(-3.72)	(-3.97)
Adj. R-squared	0.749892	0.749371	0.754544
Hausman Test Statistic			9.37
Significance Level (percent)			0.23

2SLSDV estimates from state-level panel data for 1982-1991; t-statistics in parentheses.

Appendix D

Personal Income Compared With Reportable Income

Table D-1. Summary of Definitional Differences Between Personal Income (National Accounts) and Income Reportable on Individual Tax Returns

	Reportable on Tax Returns	Not Reportable on Tax Returns
Included in Personal Income	Amts paid to individuals required to file a return: • Wages, salaries, and tips • Fees • Proprietor income (with inventory valuation and capital consumption adjustments) • Rental income of individuals (with capital consumption adjustment) • Dividends above exclusion paid to individuals • Most interest income paid to individuals • Taxable military retirement and government pensions • Some Social Security income and unemployment compensation	• Income of individuals not required to file • Income of non-profit institutions serving individuals, private noninsured welfare funds, and private trust funds. • Other labor income except fees (e.g., pension and profit sharing, group insurance, workers' compensation, and supplemental unemployment benefits) • Change in farm inventories for non-accrual returns • Excess of interest accrued over interest paid on bonds • Depletion and depreciation expenditures claimed by nonfarm proprietors, bad debt adjustments, and income of tax-exempt cooperatives • Interest on IRAs and Keogh plans • Other nonfarm proprietors' income adjustments • Dividend exclusions • Interest from Series EE savings bonds used for education (1990 and following) • Most transfer payments to individuals • Imputed income (e.g., payments in kind to employees, and net rental value of owner-occupied dwellings) • Investment income retained by life insurance carriers and non-insured pension funds • Investment income received by nonprofit institutions or retained by fiduciaries • Other income (e.g., interest on state and local bonds, and tax-exempt military pay & allowances)
Not Included in Personal Income	• Personal contributions for social insurance • Excess of tax depreciation over NIPA depreciation • Difference of book value and current value of business inventory used up in production • Capital gains ("net gain from sale of assets," incl. livestock, timber, and certain real estate) • Payments of taxable private pensions • Net income of qualified S Croporations • Miscellaneous taxable income (e.g., noncorporate special assessments, noncash awards, and net gambling receipts)	• Illegal-source income

Sources: Park and Reeb (1989); and BEA (1986).

Appendix E

Descriptive Statistics

Table E-1. National Statistics Over the 1982-1991 Period for All Included Variables (490 Observations)

Variable	Mean	Std. Dev.	Minimum	Maximum	Units
Compliance (Dependent Variables)					
FilingRate	91.15	3.98	78.76	104.17	Percent
IncomePct(A)	68.31	4.62	50.56	84.36	Percent
IncomePct(B)	68.61	4.64	50.59	84.43	Percent
IncomePct(C)	70.21	5.22	50.72	94.95	Percent
OffsetsPct(A)	22.96	2.82	16.07	33.73	Percent
OffsetsPct(B)	25.78	2.13	20.22	36.23	Percent
OffsetsPct(C)	27.24	2.21	22.17	39.65	Percent
Tax Policy					
FThresholdPct	16.82	2.42	11.30	23.92	Percent
Amnesty5	0.31	0.46	0	1	1=yes; 0=no
MargTaxRate@$15K	21.67	0.91	20.25	23.87	Percent
MargTaxRate@$50K	36.91	5.81	29.44	47.66	Percent
ChildExemptsPct	2.29	0.70	0.98	5.49	Percent
StateTaxPct	5.64	1.80	1.57	10.63	Percent
Opportunity/Burden					
AvgBurden	11.47	0.94	9.10	13.96	Hours/return
SoleProps	9.27	2.29	4.50	16.11	Percent
PaidPrep	49.18	7.28	26.99	71.90	Percent
Enforcement					
AuditStartRate	1.10	0.51	0.19	3.51	Percent
DET_pct	53.88	7.93	31.62	73.88	Percent
AvgDET	6.10	1.82	2.08	13.03	Hours/audit
IRPDocRate	6.61	1.40	3.16	10.51	# documents
TDITotRate	4.20	1.60	1.33	9.29	Percent
RefOffsetRate	0.89	0.99	0.00	11.16	Percent
CIDConvRate	9.18	6.03	0.00	51.52	Per million
IRS Responsiveness					
TPS_CallsPC	166.41	75.20	42.96	615.93	Per thousand
TPS_RetPrepPC	4.63	4.41	0.00	23.02	Per thousand
Demographics					
Singles	51.65	3.45	42.19	60.19	Percent
Under30	31.57	3.40	19.29	42.61	Percent
Over64	18.14	2.65	10.02	27.55	Percent
PCBirths	15.86	1.83	11.94	26.66	Per thousand
TFSEmplPct	52.19	5.26	38.71	69.33	Percent
Economics					
AvgPI	34.62	4.67	25.03	51.17	1992 $K
AvgPIGrowth	1.07	2.88	-9.46	10.10	Percent
ExclIncomePct	3.46	0.89	1.47	6.59	Percent
UnemplRate	6.88	2.30	2.15	18.03	Percent

Table E-2. Average Compliance Measures and Ranks By State

State	1982-1991 Average FilingRate	Rank *	1982-1991 Average IncomePct(A)	Rank *	1982-1991 Average OffsetsPct(A)	Rank *	1982-1991 Average NetIncomePct(A)	Rank *
AL	91.43	23	70.39	15	24.10	13	46.29	20
AR	91.52	21	66.34	41	24.82	8	41.52	44
AZ	86.57	49	67.73	32	24.01	16	43.73	37
CA	92.45	15	66.89	38	24.01	15	42.88	40
CO	92.26	16	68.98	23	23.53	23	45.45	30
CT	94.44	4	72.06	9	19.95	49	52.11	1
DE	91.61	20	73.17	1	23.29	26	49.88	4
FL	88.92	38	68.73	25	22.03	39	46.70	16
GA	92.55	14	71.05	14	24.06	14	46.99	15
HI	99.52	1	69.36	21	23.78	18	45.57	29
IA	88.78	40	64.52	46	22.63	33	41.89	43
ID	87.13	47	59.98	48	24.89	7	35.09	49
IL	91.21	27	69.30	22	20.80	47	48.50	10
IN	88.13	43	72.58	5	22.94	28	49.64	5
KS	89.30	33	68.53	28	22.63	31	45.90	25
KY	88.13	44	67.39	35	23.53	22	43.86	34
LA	88.52	41	67.12	36	22.68	30	44.44	33
MA	93.10	11	68.68	27	21.57	43	47.10	14
MD	95.49	3	72.97	2	23.42	25	49.55	6
ME	91.76	17	65.64	43	23.19	27	42.45	41
MI	89.55	32	70.02	17	22.12	37	47.90	11
MN	91.40	25	70.12	16	24.47	10	45.64	27
MO	90.85	29	67.66	33	22.06	38	45.61	28
MS	88.99	37	65.36	44	25.38	4	39.99	46
MT	90.31	30	62.12	47	24.00	17	38.12	47
NC	93.46	8	71.15	13	24.97	6	46.17	22
ND	93.85	6	69.86	18	24.16	12	45.70	26
NE	93.85	5	64.93	45	23.78	19	41.15	45
NH	93.37	9	68.73	26	21.39	44	47.34	12
NJ	97.78	2	71.16	12	20.91	46	50.24	3
NM	92.65	13	68.97	24	25.21	5	43.76	36
NV	91.43	22	72.67	4	22.15	36	50.53	2
NY	91.29	26	68.41	30	21.92	41	46.50	19
OH	91.62	19	69.66	19	22.46	34	47.20	13
OK	87.75	46	67.61	34	23.78	20	43.83	35
OR	87.95	45	67.12	37	24.76	9	42.36	42
PA	89.17	35	67.96	31	21.35	45	46.60	18
RI	89.07	36	66.70	39	22.01	40	44.69	32
SC	92.75	12	72.49	6	26.26	2	46.23	21
SD	90.91	28	58.42	49	22.20	35	36.22	48
TN	93.31	10	72.47	7	23.64	21	48.83	8
TX	91.70	18	68.48	29	21.83	42	46.64	17
UT	90.30	31	72.82	3	29.14	1	43.68	38
VA	93.58	7	72.09	8	23.47	24	48.62	9
VT	88.26	42	71.53	11	25.56	3	45.97	24
WA	91.41	24	71.78	10	22.63	32	49.15	7
WI	89.23	34	69.47	20	24.29	11	45.18	31
WV	86.99	48	65.83	42	22.94	29	42.89	39
WY	88.80	39	66.63	40	20.59	48	46.04	23

* Rank is ordered with 1 = highest state and 49 = lowest state for each variable.

Appendix F

Estimation of the Burden Associated With Tax Forms

Table F-1. Criteria for Inferring From CPS Data Which Forms and Schedules Were Required

The Current Population Survey (CPS) was used to estimate the number of required returns, by state and year (see Appendix B). In addition, when a return was required, the CPS was used to estimate which tax form and schedules were required. The total number of each type of form and schedule (by state and year) were used to calculate the corresponding average burden faced by taxpayers (*AvgBurden*). The following criteria were used to estimate which forms and schedules were required:

Form 1040EZ = 1 only if all of the following are true (= 0 otherwise):

1 Filing status = Single [see Table B-2]
2 Exemptions:
   - FREC10 = 0
   - AGE < 65
3 Maximum Taxable Income (Gross Income – FThresh) < $50,000 [see Table B-2]
4 Income—all of the following must be true:
   - Self-Employment income (I51B) = 0
   - Farm income (I51C) = 0
   - SS or RR Retirement income (I52A) = 0
   - Supplemental Security income (I52B) = 0
   - Interest income (I53B) ≤ $400
   - Dividend, Rent & Royalty income (I53C) = 0
   - Veterans, Unemployment & Workman's Compensation income (I53D) = 0
   - Pension income (I53E) = 0
   - Child Support income (I53F) = 0
5 TENURE ≠ 1

Form 1040A = 1 only if all of the following are true (= 0 otherwise):

1 Not a Form 1040EZ
2 Maximum Taxable Income (Gross Income – FThresh) < $50,000 [see Table B-2]
3 Income—all of the following must be true:
   - Self-Employment income (I51B) = 0
   - Farm income (I51C) = 0
   - SS or RR Retirement income (I52A) = 0
   - Supplemental Security income (I52B) = 0
   - Rent & Royalty Income Recipiency (I53CRENT) ≠ 1
   - Pension income (I53E) = 0
   - Child Support income (I53F) = 0
4 TENURE ≠ 1

Form 1040 = 1 only if return is neither a Form 1040EZ nor a Form 1040A (= 0 otherwise)

Continued on next page...

Table F-1. Criteria for Inferring From CPS Data Which Forms and Schedules Were Required (Continued)

Schedule 1 = 1 only if both of the following are true (= 0 otherwise):
1 Return is Form 1040A
2 Interest income (I53B) > $400 OR
[Dividend, Rent & Royalty income (I53C) > $400 and Div. Recipiency (I53CDIV) = 1]

Schedule A = 1 only if TENURE = 1 (= 0 otherwise)

Schedule B = 1 only if both of the following are true (= 0 otherwise):
1 Return is Form 1040
2 Interest income (I53B) > $400 OR
[Dividend, Rent & Royalty income (I53C) > $400 and Div. Recipiency (I53CDIV) = 1]

Schedule C = 1 only if Self-Employment Income (I51B) ≠ 0 (= 0 otherwise)

Schedule D = 1 only if Dividend Income Recipiency (I53CDIV) = 1 (= 0 otherwise)

Schedule E = 1 only if Rent, Royalty Income Recipiency (I53CRENT) = 1 (= 0 otherwise)

Schedule F = 1 only if Farm Income (I51C) ≠ 0 (= 0 otherwise)

Schedule R = 1 only if both of the following are true (= 0 otherwise):
1 AGE > 64
2 Income—one or more of the following is true:
- SS or RR Retirement income (I52A) > 0
- Supplemental Security income (I52B) > 0
- Veterans, Unemployment & Workman's Compensation income (I53D) > 0
- Pension income (I53E) > 0

Schedule SE = 1 only if both of the following are true (= 0 otherwise):
1 Self-Employment income (I51B) ≥ $400
2 Wage & Salary income (I51A) ≤ the appropriate threshold:

Tax Year	Threshold	Tax Year	Threshold
1980	$25,900	1986	$42,000
1981	$29,700	1987	$43,800
1982	$32,400	1988	$45,000
1983	$35,700	1989	$48,000
1984	$37,800	1990	$51,300
1985	$39,600	1991	$125,000

Table F-2. Estimated* Hours Required to Complete and File Various Forms and Schedules, 1982-1991

	Year				
Form/Schedule	1982[a]	1983[b]	1984[b]	1985[b]	1986[c]
Form 1040	8.42	7.97	8.16	8.78	8.08
Form 1040A	7.55	8.47	8.43	10.26	7.75
Form 1040EZ	3.14	3.00	3.00	3.41	2.03
Schedule A	6.31	8.06	7.20	7.41	7.62
Schedule B	1.71	1.74	1.64	1.64	1.28
Schedule C	9.77	10.39	10.65	10.12	10.53
Schedule D	2.41	2.91	4.11	3.34	3.75
Schedule E	5.05	3.95	4.21	3.26	5.67
Schedule F	6.52	4.11	4.13	6.65	6.40
Schedule R	2.92	2.92	1.34	1.34	1.54
Schedule SE	1.54	1.54	1.99	1.99	1.99

	Year				
Form/Schedule	1987[c]	1988[c]	1989[d]	1990[d]	1991[d]
Form 1040	8.11	9.26	9.40	9.55	9.72
Form 1040A	6.96	6.96	7.08	7.73	6.67
Form 1040EZ	1.51	1.51	1.33	1.98	2.02
Schedule A	5.58	4.59	4.57	4.22	4.55
Schedule B	1.28	1.28	1.28	1.33	1.32
Schedule C	10.87	10.81	9.63	9.65	9.67
Schedule D	3.42	3.26	3.75	3.30	3.55
Schedule E	6.18	6.18	5.83	5.82	5.82
Schedule F	6.47	6.89	6.28	6.23	6.26
Schedule R	1.54	1.54	1.53	1.52	1.53
Schedule SE	1.75	1.75	1.75	1.80	1.77

* Based on a study conducted by Arthur D. Little, Inc. (ADL) for the IRS. These estimates include the time required for: recordkeeping; learning about the law or the form; preparing the form; and copying, assembling, and sending the form.

a Estimated specifically for this study using the ADL model

b Estimates contained in the ADL report

c IRS estimates based on the ADL model

d IRS estimates published in the Form 1040 tax package

Appendix G

Estimated Revenue Impact of the Decline in Audit Coverage

Table G-1. Calculation of Tax Forgone Due to AuditRates Falling Below the 1982 Rate, 1982-1991

Year	Actual AuditRate [A]	Actual Contribution to NetIncomePct* [B=13.89 Ln(A+1)]	Forgone Contribution to NetIncomePct [C=13.38-B]	Personal Income ($ Billions) [D]	Forgone Net Income ($ Billions) [E=C/100*D]	Average Marginal Tax Rate (%) [F]	Forgone Tax ($ Billions) [G=E*F]
82	1.62	13.38	0.00	2,677	0.00	21.90	0.00
83	1.52	12.84	0.54	2,850	15.29	20.11	3.08
84	1.32	11.71	1.67	3,136	52.31	19.65	10.28
85	1.25	11.25	2.13	3,359	71.45	19.67	14.05
86	1.08	10.15	3.23	3,571	115.49	19.70	22.75
87	1.07	10.09	3.29	3,781	124.47	17.93	22.31
88	0.94	9.24	4.14	4,054	167.96	17.75	29.81
89	0.69	7.25	6.13	4,358	267.09	17.86	47.71
90	0.65	6.99	6.39	4,638	296.34	17.81	52.78
91	0.65	6.93	6.45	4,802	309.89	17.69	54.81
Total			33.97	37,225	1,420.29		257.58

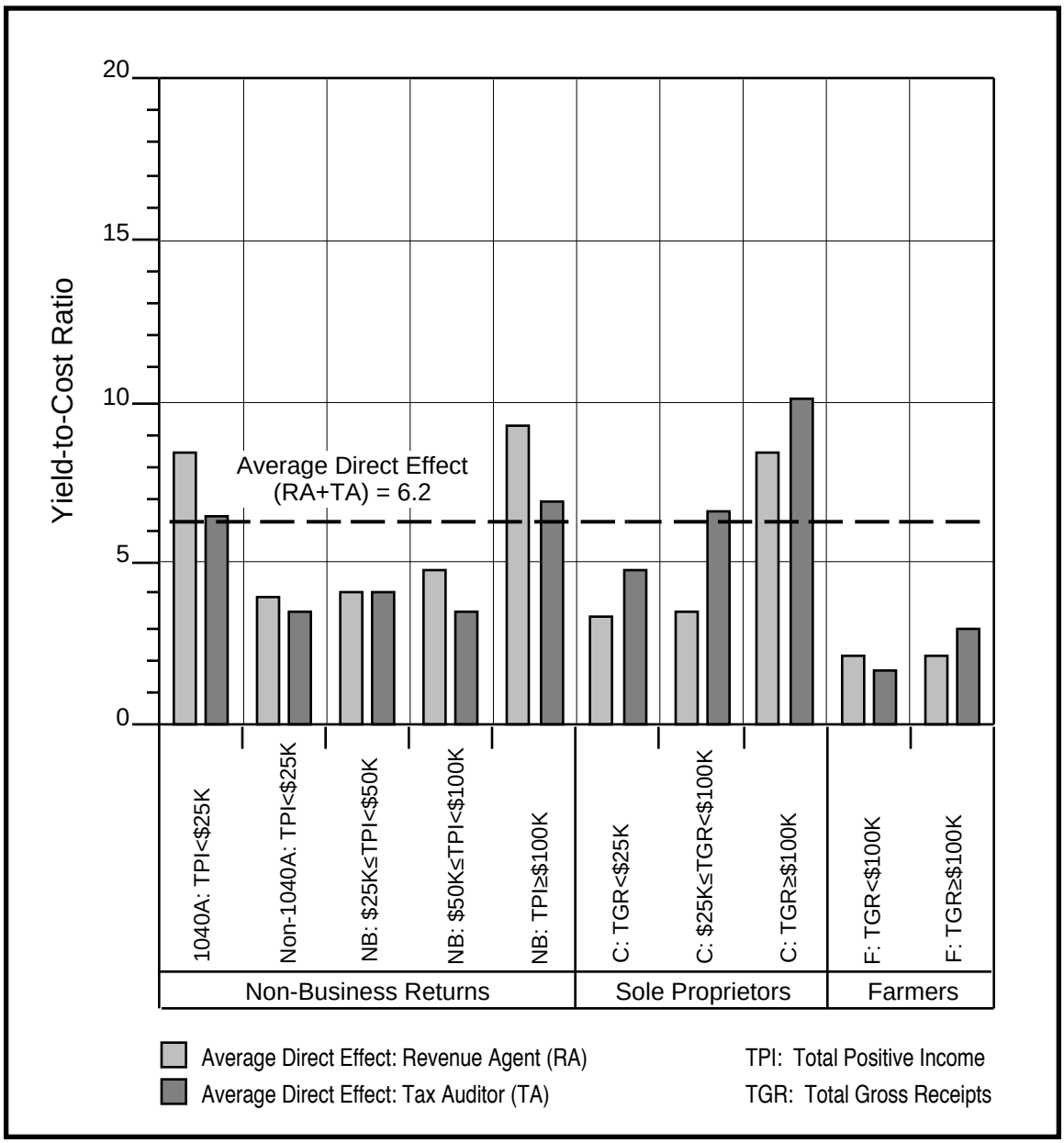
* These calculations are based on the results for Definition A, in which the coefficient on Ln(pAuditRate+1) is 13.892113. (See Table 3 in the text.)

Appendix H

Interactions With the AuditRate Variable

We know from years of auditing experience that the direct revenue effect of audits varies widely according to the class of return that is audited. In fact, audit resources are allocated to these classes based on the unique marginal yield-to-cost potential of each class. Figure H-1 illustrates how average direct yield-to-cost ratios varied across audit classes in 1991.

Figure H-1. Average Direct Audit Yield-to-Cost Ratio by Audit Class, 1991



Given that the direct effect of audits is a function of the type of return audited, it is reasonable to ask whether the indirect effect of audits also varies by class, or whether it is constant across all classes. The specification employed in this study estimates only the average effect of the overall audit rate, and suggests that this average indirect effect is 69.4 compared with the average direct effect of 6.2 (which is based on the adjustments recommended by the auditors—not what is eventually collected). One rather simple way to test whether the impact of audits on voluntary compliance depends on the type of returns being audited is to include as explanatory variables the interaction of *AuditRate* with other variables that help to distinguish the audit classes. Figure H-1 includes a brief description of the ten individual income tax audit classes presently used by IRS. The classes are defined primarily on the basis of two key aspects of a return: whether it is a business return (whose principal source of income is from sole proprietorships [reported on Schedule C], or from farms [reported on Schedule F]), and the amount of income (or gross receipts) reported on the return. It was a straightforward matter to create interaction terms representing these two factors:

1. **Prevalence of Non-Business Taxpayers:** the percent of potential returns (per the Current Population Survey, see Appendix B) not having business income.

$$NB = 100 - (SoleProps + Farmers)$$

Interaction Term #1: LnpASR_NB = [Ln(pAuditRate+1)]*NB

2. **Average Income:** Personal Income (from the Bureau of Economic Analysis), in constant 1992 dollars, per potential return.

$$AvgPI = (Personal\ Income) / (Number\ of\ Potential\ Returns)$$

Interaction Term #2: LnpASR_PI = [Ln(pAuditRate+1)]*AvgPI

Table H-1 summarizes the results of including these two interaction terms along with the original audit rate variable, and compares these with the same specification, but without the interaction terms. (The coefficients on the other variables are not shown, but did not change appreciably from those reported in Table 3 of the text.)

Table H-1. Regression Results Comparing Specifications With and Without Interaction Terms

AuditRate Variables	Mean [A]	IncomePct(A) Coefficient [B]	IncomePct(A) Product [C=B*A]	OffsetsPct(A) Coefficient [D]	OffsetsPct(A) Product [E=D*A]	NetIncomePct(A) Coefficient [F]	NetIncomePct(A) Product [G=F*A]
With Interactions							
LnpASRp1	0.7218	13.0650 (0.81)	9.4309	9.4858 (1.62)	6.8473	6.1198 (0.44)	4.4176
LnpASR_PI	24.7978	-0.0656 (-0.34)	-1.6271	0.0100 (0.16)	0.2483	-0.0886 (-0.56)	-2.1975
LnpASR_NB	63.6708	0.0579 (0.34)	3.6845	-0.715 (-1.17)	-4.5550	0.1169 (0.82)	7.4429
Sum			11.4883		2.5406		9.6629
Without Interactions							
LnpASRp1	0.7218	16.1585 (3.37)	11.6640	3.3139 (2.00)	2.3921	13.8921 (3.46)	10.0280

t-statistics in parentheses

Observations and Conclusions

1. Although none of the audit coefficients is significant in the interaction specification, the results are similar to those in the non-interaction specification.

2. Figure H-2 illustrates how these two estimates of the indirect effect compare with each other and with the direct effect on a state-by-state basis. The average indirect effect seems to be somewhat more sensitive to *AuditRate* than the average direct effect, but both generally decline with increasing *AuditRate,* as one would expect.

3. These results are also inconclusive because it is not clear which is more important—the mix of audits across classes, or the mix of returns in the population. Further study is warranted. For example, it is possible to gather audit rate data for each audit class, and to evaluate the sensitivity of the indirect effect to audit class.

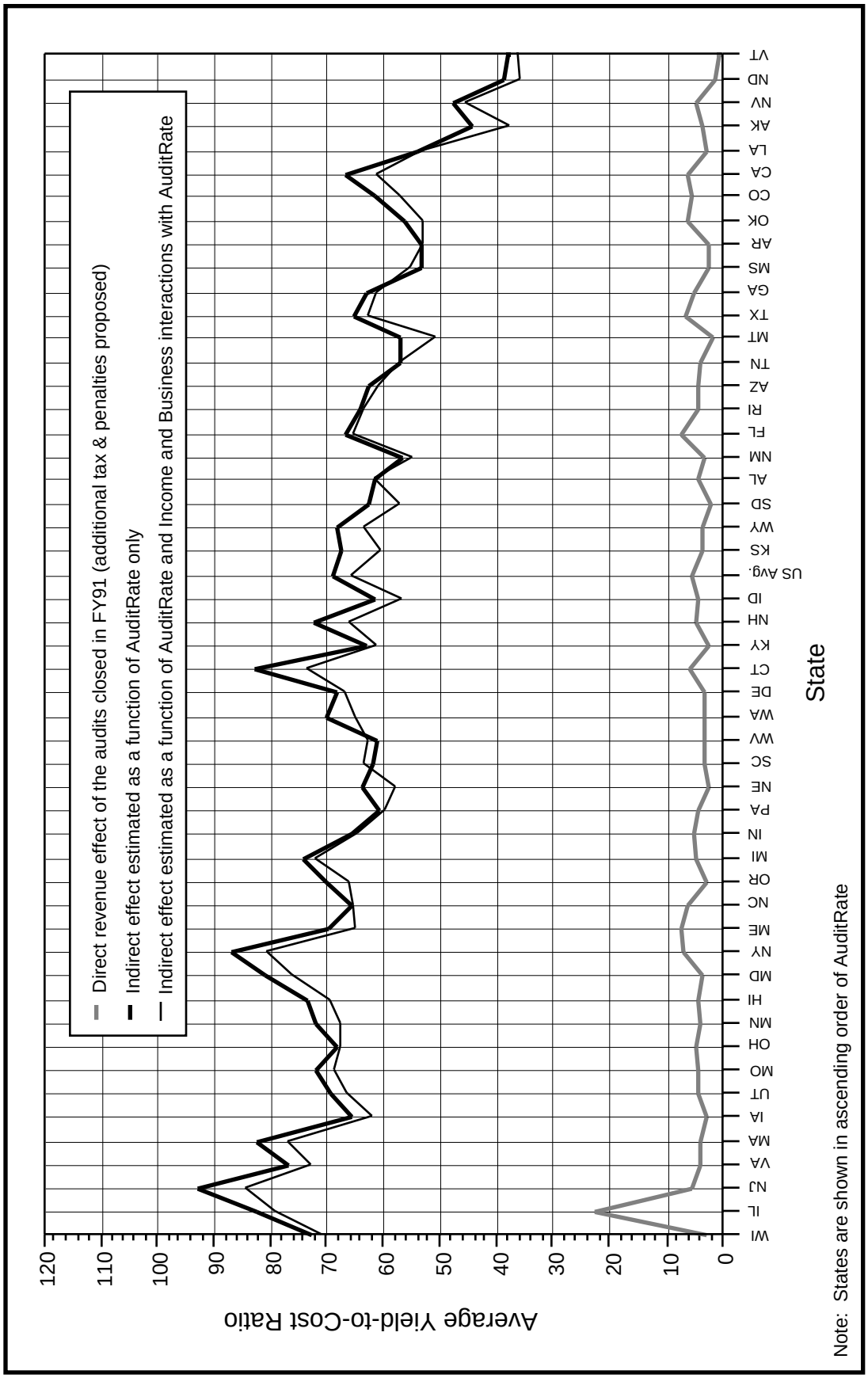


Figure H-2. Two Estimates of the Indirect Effect of Audits Compared With the Direct Effect, By State, 1991

Appendix I
Derivation of Total and Marginal Indirect Revenue Functions For Five IRS Activities

AuditRate

$$\text{NetIncomePct} = \gamma_{11}\,\text{Ln}(\text{AuditRate} + 1)$$

$$\frac{\text{NetIncome}}{\text{Pers. Income}} \times 100 = \gamma_{11}\,\text{Ln}(\text{Audits/ReturnsFiled} \times 100 + 1)$$

$$\text{NetIncome} = \gamma_{11}\,\text{Ln}(\text{Audits/ReturnsFiled} \times 100 + 1) \times \text{Pers. Income}/100$$

$$\frac{\partial(\text{NetIncome})}{\partial(\text{Audit})} = \gamma_{11}\,\text{Pers. Income}/[(\text{AuditRate} + 1) \times \text{ReturnsFiled}]$$

CID_ConvRate

$$\text{NetIncomePct} = \gamma_{13}\,\text{Ln}(\text{CID\_ConvRate} + 1)$$

$$\frac{\text{NetIncome}}{\text{Pers. Income}} \times 100 = \gamma_{13}\,\text{Ln}(\text{Convictions/Population} \times 10^6 + 1)$$

$$\text{NetIncome} = \gamma_{13}\,\text{Ln}(\text{Convictions/Population} \times 10^6 + 1) \times \text{Pers. Income}/100$$

$$\frac{\partial(\text{NetIncome})}{\partial(\text{Conviction})} = \gamma_{13}\,\text{Pers. Income} \times 10^4/[(\text{CID\_ConvRate} + 1) \times \text{Population}]$$

IRP_DocRate

$$\text{NetIncomePct} = \gamma_1\,\text{FilingRate}$$

$$\frac{\text{NetIncome}}{\text{Pers. Income}} \times 100 = \gamma_1\,[\varphi_6\,\text{IRP\_DocRate}] = \gamma_1\,[\varphi_6\,\text{IRP\_Docs/Pot'lReturns}]$$

$$\text{NetIncome} = \gamma_1\,[\varphi_6\,\text{IRP\_Docs/Pot'lReturns}] \times \text{Pers. Income}/100$$

$$\frac{\partial(\text{NetIncome})}{\partial(\text{IRP\_Doc})} = \gamma_1\,\varphi_6 \times \text{Pers. Income}/[100 \times \text{Pot'lReturns}]$$

TDI_TotRate

$$\text{NetIncomePct} = \gamma_1 \text{ FilingRate}$$

$$\text{NetIncomePct} = \gamma_1 [\varphi_7 \text{ Ln}(\text{TDI\_TotRate} + 1)]$$

$$\frac{\text{NetIncome}}{\text{Pers. Income}} \times 100 = \gamma_1 [\varphi_7 \text{ Ln}(\text{Notices/Pot'lReturns} \times 100 + 1)]$$

$$\text{NetIncome} = \gamma_1 [\varphi_7 \text{ Ln}(\text{Notices/Pot'lReturns} \times 100 + 1)] \times \text{Pers. Income}/100$$

$$\frac{\partial(\text{NetIncome})}{\partial(\text{Notice})} = \gamma_1 \varphi_7 \times \text{Pers. Income}/[(\text{TDI\_TotRate} + 1) \times \text{Pot'lReturns}]$$

TPS_RetPrepPC

$$\text{NetIncomePct} = \gamma_1 \text{ FilingRate} + \gamma_{15} \text{ TPS\_RetPrepPC}$$

$$\text{NetIncomePct} = \gamma_1 [\varphi_9 \text{ TPS\_RetPrepPC}] + \gamma_{15} \text{ TPS\_RetPrepPC}$$

$$\frac{\text{NetIncome}}{\text{Pers. Income}} \times 100 = (\gamma_1 \varphi_9 + \gamma_{15}) (\text{ReturnsPrep/Population} \times 10^3)$$

$$\text{NetIncome} = (\gamma_1 \varphi_9 + \gamma_{15}) (\text{ReturnsPrep/Population} \times 10^3) \times \text{Pers. Income}/100$$

$$\frac{\partial(\text{NetIncome})}{\partial(\text{ReturnsPrep})} = (\gamma_1 \varphi_9 + \gamma_{15}) \times 10 \times \text{Pers. Income/Population}$$

Note: The notation in these equations follows the notation introduced in section 2.3 of the text. Since the purpose of these calculations is to derive the marginal effect of changing one key variable at a time (holding all others constant), all other terms have been dropped. The induced revenue can be estimated by multiplying the additional NetIncome by the appropriate average marginal tax rate.