DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA 91125

ARE WE A NATION OF TAX CHEATERS? NEW ECONOMETRIC EVIDENCE ON TAX COMPLIANCE*

Jeffrey A. Dubin
California Institute of Technology

Michael J. Graetz
Yale Law School

Louis L. Wilde
California Institute of Technology

*Forthcoming in *American Economic Review, Papers and Proceedings*, May 1987.

SOCIAL SCIENCE WORKING PAPER 626

January 1987

ABSTRACT

The theoretical basis for the economic approach to tax compliance has, at least until recently, been inadequate, and the limited empirical work based on it is seriously flawed. In this paper we briefly review both, as well as new theoretical and, especially, empirical work on the tax compliance problem. With respect to the latter we present preliminary results based on a state-level, time-series, cross-section data set drawn in part from the annual reports of the Commissioner of Internal Revenue.

ARE WE A NATION OF TAX CHEATERS? NEW ECONOMETRIC EVIDENCE ON TAX COMPLIANCE

Jeffrey A. Dubin, Michael J. Graetz, and Louis L. Wilde*

In 1982, then Commissioner of Internal Revenue Roscoe Egger reported to Congress that legal sector noncompliance with the Federal Income Tax statutes generated an "income tax gap" of $81 billion in 1981, up from $29 billion in 1973. He further projected a gap of $120 billion for 1985 (U.S. Congress, 1982). Perceptions of accelerating noncompliance inspired a crisis mentality within the Internal Revenue Service, Congress, and the tax bar.

The IRS responded in part by funding a major independent study of tax noncompliance via the National Academy of Sciences, and the American Bar Foundation initiated an investigation of its own in 1984. Congress enacted compliance legislation in 1981, 1982, and 1984, and completely overhauled the federal income tax laws in 1986. These enactments added a wide variety of new penalties for noncompliance and strengthened others, dramatically expanded requirements for third party reporting of information to the IRS, added to the IRS's arsenal of procedural weapons and adopted everyone's favorite vehicle to combat noncompliance—lower tax rates.

All this clamor and action has taken place in the absence of any solid factual foundation (Graetz and Wilde, 1985). We are not at all certain of the actual decline in tax compliance during the past decade, and even if noncompliance has increased significantly, its causes, and thus appropriate remedies, simply are not known. For example, recent unpublished Internal Revenue Service estimates have significantly reduced Commissioner Egger's projections for 1985—to $92 billion; in fact, the real income tax gap for individual returns is now thought to have fallen from $39.1 billion in 1981 to $36.8 billion in 1986, measured in 1972 dollars. These figures do not support the widespread claims that the American public is becoming a nation of tax cheaters or that the integrity of the tax system is seriously at risk, but the complete story is much more complex. Not only must there be additional efforts to determine what circumstances imply increased noncompliance, but the effects of recent tax law and penalty changes as well as changes in IRS budgets and audit capacity must also be taken into account. Ultimately this is an empirical story, but valid empirical work must be based on the proper theoretical foundation.

The theoretical basis for the economic approach to tax compliance has, at least until recently, been inadequate, and the limited empirical work based on it is seriously flawed. In this paper we briefly review both, as well as new theoretical and, especially, empirical work on the tax compliance problem.

---

*Assistant Professor of Economics, California Institute of Technology; Professor of Law, Yale; and Professor of Economics, California Institute of Technology

I. THE ECONOMIC THEORY OF TAX COMPLIANCE

The contemporary revival of the economic analysis of crime began in 1968 with Becker's classic article "Crime and Punishment: An Economic Approach." While Becker mentioned tax evasion as a potential application of his general model, Allingham and Sandmo (1972) published the first formal analysis. In their model, the taxpayer's actual income is exogenously given and known only to the taxpayer. A constant proportional tax is applied to reported income, with such reported amounts chosen by the taxpayer to maximize expected utility of net wealth. With some exogenous and constant probability, the taxpayer is "audited." If the taxpayer is discovered to be underreporting income, a penalty proportional to the amount of undeclared income must be paid in addition to the proportional tax rate.

The bulk of the remainder of the theoretical economics literature on tax compliance consists largely of extensions and refinements of Allingham and Sandmo's model. While many ambiguous results are produced by these analyses, one prediction is universal: an exogenous increase in the probability of detection and conviction or in the penalty rate will increase compliance.

More recent theoretical innovations have attempted to move out of the decision-theoretic framework characteristic of the early tax compliance literature. Of particular interest here are the principal-agent models of Border and Sobel (1985) and Reinganum and Wilde (1986) and the game-theoretic model of Graetz, Reinganum, and Wilde (1986). In both of these approaches the IRS is allowed to act strategically, conditioning its audit rules on the information it receives from taxpayers.

Whether the IRS should be included as a strategic actor in theoretical models of tax compliance is of more than technical interest. In assessing empirically the deterrent effects of audits, it is crucial whether the IRS audit selection process turns on taxpayer compliance behavior. If it does, then any empirical model meant to explain taxpayer compliance behavior that treats audit rates as exogenous may be seriously misspecified. In fact any deterrent effect of audits may be outweighed by the (presumed) countervailing incentive of the IRS to audit most heavily those returns for which expected compliance is the lowest (in light of information received) and thereby produce an observed negative relationship between audits and compliance.

II. EXISTING EMPIRICAL WORK

To date, there has been a surprisingly small amount of empirical work on the determinants of tax compliance. Not counting survey work or work attempting to measure aggregate noncompliance, we have found only four relevant studies; Clotfelter (1983), Witte and Woodbury (1984, 1985), Slemrod (1986), and Dubin and Wilde (1986).

Clotfelter analyzed a data set collected originally as part of the 1969 IRS Taxpayer Compliance Measurement Program (TCMP). The TCMP involves detailed "line-by-line audits" of a stratified random sample of taxpayers, which result in income tax assessments regarded by the IRS as "correct." IRS uses TCMP audits in developing a scoring mechanism (the "Discriminant Index Function," or "DIF") to establish and refine the audit selection decisions it applies to the larger population of taxpayers. TCMP is a far better technique for learning about the effectiveness of IRS audits than about aggregate noncompliance (Graetz and Wilde, 1985), but it is nevertheless one of the best sources of data currently available for estimating noncompliance.

Using raw TCMP data, Clotfelter investigated the relationship between marginal tax rates and tax evasion for three classes of taxpayers (nonbusiness, nonfarm business, and farm). For each group, he regressed the log of underreported income on a measure of the effective marginal tax rate, after-tax income, wages as a proportion of adjusted gross income, interest and dividends as a proportion of adjusted gross income, and several socio-demographic variables. The average audit rate for each taxpayer class was not included as an independent variable since, as Clotfelter put it, "the probability [of audit] for any tax return in a given class is a function of its reported items"; in other words, there is a potential simultaneity problem that makes it inappropriate to use audit rates as exogenous explanatory variables in an equation meant to explain compliance with the tax laws.

Clotfelter found that both the level of after-tax income and marginal tax rates have significant negative effects on compliance. While these results are interesting, they should be used with caution. Clotfelter tried to avoid the simultaneity issue by leaving audit rates out of his model, but his model is still misspecified if audit rates affect compliance. In any event, since he left audit rates out of his analysis, Clotfelter's work implies nothing about their deterrent effects.

Witte and Woodbury (1985) explicitly attempt to analyze the effects of audit rates and sanction levels on compliance using a data set provided to them by the IRS. This data includes a percentage compliance variable related to 1969 returns filed in 1970 (estimated by the IRS from DIF scores, not actual IRS audits), IRS agency variables such as audit rates and sanction levels, and a host of demographic and socio-economic variables, all aggregated to the three digit zip code level. Separate equations were estimated for each of seven audit classes, defined by income level (low, medium, or high) and by type of return (1040 only, Schedule C or F present, Schedule C and F not present), using seemingly unrelated regression. In particular, the estimated 1969 percentage compliance variable was regressed on a constant term and 36 explanatory variables, including audit rates for 1967, 1968, and 1969 within the audit class, and for all other audit classes.

A detailed discussion of Witte and Woodbury's work can be found in Dubin and Wilde (1986). Two major problems with it are (1) the numerical properties of their data set are unsatisfactory and (2) many of the agency variables are likely to be endogenous so that their model is misspecified.

These problems perhaps explain some of the peculiar results obtained by Witte and Woodbury. In their published paper, for example, they report selected results for three of the seven audit classes. For these audit classes reported mean elasticities of percentage compliance with respect to "audit rates" range from .002 to .02, approximately. However, by referring to their 1984 working paper, one finds first that these elasticities are obtained by summing the coefficients, when significant, on all six of the audit variables (1967, 1968, and 1969 audit rates within each audit class and for all other audit classes). Second, only one of the 1969 within-class audit rate variables is significant and it has a negative sign, six of the seven 1968 within-class audit rate variables are significant and half have a negative sign, and six of the seven 1967 within-class audit rate variables are significant but all have a positive sign. It is difficult indeed to conclude from these results that increases in audit rates increase compliance.

Slemrod (1985) takes a different approach in his analysis of tax avoidance. He notes that tax liability is a step function of taxable income for most taxpayers, the step size being $50 in 1977. He shows that noncompliers, theoretically, have an incentive to report income levels near the upper end

of the relevant step range. Using 1977 TCMP data, Slemrod regresses the taxpayer's position within the relevant $50 bracket (a number from 1 to 50) on several factors. A tendency to be located higher in the relevant $50 bracket is shown to be positively associated with higher marginal tax rates, being less than 65 years old, being married, and the presence of certain "fungible items." But Slemrod's approach is of limited value at best; it cannot get at the degree of tax evasion even if it is present, as he hypothesizes.

The most recent empirical study of tax noncompliance using microeconomic data is by Dubin and Wilde (1986). These authors analyze a subset of the 1969 data set used by Witte and Woodbury, augmented with data taken from the 1969 *Annual Report of the Commissioner of Internal Revenue*. The dependent variable is the same as that used by Witte and Woodbury, an IRS estimate (based on DIF scores) of the percentage compliance rates for individuals in each of the seven audit classes described above.

The explanatory variables are the 1969 within-class audit rate; three variables that have been thought to reflect opportunities to evade: the unemployment rate, the percentage employed in manufacturing, and the percent self-employed; and three variables that the literature (principally surveys) suggests are important: the percentage of the population over 65 years old, the percentage of persons over 25 with at least four years of high school education, and the percentage of nonwhite population. Following recent theoretical work (and empirical work elsewhere in the economics of crime literature), Dubin and Wilde treat the audit rate as being potentially endogenous. This hypothesis is tested using an instrumental variables procedure, the "instruments" being the number of criminal fraud investigations initiated in 1970 per 1968 return filed in 1969, the percentage of taxpayers receiving a first or second notice in 1969 indicating that taxes were due, and the IRS budget per tax return filed. Dubin and Wilde regard the last of these as a good instrument but discuss possible shortcomings of the first two.

In four of the seven audit classes (low-income nonbusiness with a standard deduction, low-income business, and both high-income classes) the audit rate was found to be endogenous. In all audit classes a deterrent effect of audits on noncompliance was found but in three of the four cases in which audits were found to be endogenous, the deterrent effect was dominated by the countervailing incentives for the IRS to audit most heavily those returns with the greatest expected noncompliance, so that in equilibrium audit rates were negatively related to compliance for these audit classes.

III. NEW EMPIRICAL WORK

The *Annual Report of the Commissioner of Internal Revenue* contains a wealth of data that has not to our knowledge been exploited by researchers. The typical report gives state-level information for each type of tax (individual and corporate income, estate and gift, etc.) regarding total collections, number and amount of refunds, number of returns filed, number of returns examined, additional tax and penalties recommended after examination, and costs incurred by the IRS. We have assembled data from these *Annual Reports* for 1977 through 1985, and we expect eventually to add additional years. We have also obtained for these years data on socio-economic and demographic variables similar to those used by Dubin and Wilde (1986).

We have just begun to explore this rich data set, but are able to report here three preliminary results. Following Dubin and Wilde (1986), we use the percentage return per audit for individual

returns (additional tax and penalties from audits ÷ total collections, per million audits, in 1972 dollars) (PBANG) as the dependent variable. Our independent variables are lagged values of the audit rate (examinations per hundred returns filed) (IAR); percent of the adult population with a high school education (PERED); percent of the population over 45 (PEROLD); per capita income (PICAP) and its square (PICAP2); the unemployment rate (UR); percent of the work force employed in manufacturing (PMAN); and a time trend (TIME). We allow for endogeneity of the audit rate using the budget per return (BPR) and the percent of individual returns filed (PIRF) as instruments. The time-series results are broadly consistent with Dubin and Wilde's cross-section results: (1) the audit rate is endogenous as indicated by the significant coefficient of the predicted audit rate (PIAR) in the compliance equation (Table 1, column 2); (2) there is a deterrent effect associated with increases in the audit rate, but in equilibrium it is dominated by the IRS's incentive to audit according to expected yield; and (3) compliance increases with per capita income but at a decreasing rate, peaking below the maximum per capita income. In addition, there is a significant negative time trend in the audit rate and in compliance (see Table 1, columns 2 and 4).

This last result appears quite significant; after allowing for a variety of economic and demographic factors and changes over time in the state-level IRS budget per return filed, we still find a significant negative time trend both in the audit rate and in compliance. In an effort to assess the impact of these negative trends on the overall performance of the tax system, we also analyzed the time structure of individual collections per return filed.

Since one finds a deterrent effect of audits in both the 1969 cross-section data set and in the 1977-85 time-series, cross-section data set, the audit rate should be positively related to collections. To test this hypothesis we use individual collections per return (ICR) as the dependent variable in a model which is otherwise exactly the same as the one described above. The audit rate again turns out to be endogenous (based on the coefficient of PIAR in column 3 of Table 1), and is, as predicted, positively related to individual collections per return. Surprisingly, there is no significant time trend (see Table 1, column 3).

The lack of a residual time trend in individual collections per return is surprising and provocative given the negative time trend in audits and noncompliance. A variety of explanations are possible: for example, (1) audits may have become more "efficient" over time and thus have had an increasing deterrent effect, offsetting the decrease in compliance; (2) penalty revisions since 1981 may be improving compliance; (3) shifts in real tax rates over time may have increased collections per return even in the face of declining compliance; or (4) increased use of third-party reports and the "information matching" program may have increased collections per return independent of actual audits.

Further investigation of the time-series data set should help sort out these issues. But already we have learned a great deal. For example, the recent IRS estimates now show the real compliance gap for individuals to have increased from $22 billion in 1978 to $36.8 billion in 1985, a difference of $14.85 billion. If, however, the audit rate had not fallen during this period, our individual-collections-per-return equation indicates that real individual collections would have risen by $15.17 billion in 1985, actually lowering the estimated tax gap in comparison to 1978.

REFERENCES

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

Becker, Gary S., "Crime and Punishment: An Economic Approach," *Journal of Political Economy* 76 (1968):169-217.

Border, Kim, and Joel Sobel, "A Theory of Auditing and Plunder," forthcoming in *Review of Economic Studies*.

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

Dubin, Jeffrey A., and Louis L. Wilde, "An Empirical Analysis of Federal Income Tax Auditing and Compliance," SSWP No. 615, Caltech, October 1986.

Graetz, Michael J., Jennifer R. Reinganum, and Louis L. Wilde, "The Tax Compliance Game: Toward an Interactive Theory of Law Enforcement," *Journal of Law, Economics, and Organization* 2 (1986):1-32.

Graetz, Michael J., and Louis L. Wilde, "The Economics of Tax Compliance: Fact and Fantasy," *National Tax Journal* 38 (1985):355-363.

Reinganum, Jennifer F., and Louis L. Wilde, "Income Tax Compliance in a Principal-Agent Framework," *Journal of Public Economics* 26 (1985):1-18.

Slemrod, Joel, "An Empirical Test for Tax Evasion," *The Review of Economics and Statistics* (1985):232-238.

U.S. Congress, *Tax Compliance Act of 1982 and Related Legislation*, Hearings before the Ways and Means Committee, 97th Cong., 2d Sess. (GPO, 1982).

Witte, Ann D., and Diane F. Woodbury, "A Test of an Economic Model of Tax Compliance," working paper, Wellesley College (September 1984).

Witte, Ann D., and Diane F. Woodbury, "The Effect of Tax Laws and Tax Administration on Tax Compliance: The Case of the U.S. Individual Income Tax," *National Tax Journal* 38 (March 1985):1-14.

TABLE 1

COMPLIANCE, COLLECTIONS, AND AUDIT EQUATIONS (1978-1985)

Independent Variable[a]	Dependent Variable PBANG	ICR[c]	IAR(-1)
ONE	-12.94471	-5.14238	10.76544
(1.00)	(-4.68820)	(-3.80874)	(3.59506)
IAR(-1)	-0.00774	0.00261	—
(1.65)	(-1.44852)	(1.00019)	
PIAR[b]	0.88060	0.18126	—
(1.741)	(5.40323)	(2.27448)	
PERED(-1)	3.60711	-0.16745	-0.81338
(0.68)	(4.22410)	(-0.40102)	(-0.97201)
PEROLD(-1)	-0.94030	1.67478	-2.29299
(0.42)	(-0.65915)	(2.40092)	(-1.75911)
UI(-1)	-6.14816	2.81068	0.90842
(0.073)	(-2.64503)	(2.47286)	(0.41711)
PICAP(-1)[c]	-1.16237	1.54763	0.00920
(5.33)	(-5.36695)	(14.61351)	(0.45836)
PICAP2(-1)	0.00598	-0.00724	-0.00090
(29.73)	(4.43124)	(10.96262)	(-0.73252)
PMAN(-1)	-1.14494	3.04922	-1.46025
(0.19)	(-1.40866)	(7.67214)	(-2.15393)
TIME	0.18706	-0.00053	-0.00836
(81.5)	(6.30054)	(-0.36678)	(-2.96073)
BPR(-1)[c]	—	—	289.9653
(0.00042)			(6.74527)
PIRF(-1)	—	—	-3.26681
(0.587)			(-2.92698)
Number of Observations	400	400	400
R-squared	0.29827	0.63217	0.24709
Mean of Dependent Var	0.684	2.234	1.741

[a] Mean values in parenthesis below variable names. t-statistics (below coefficients) while qualitatively similar to instrumental variable estimates, are not identical.

[b] PIAR is the predicted value of IAR(-1) from the audit equation.

[c] Measured in thousands of 1972 dollars.