

# *The Impact of Information on Law Enforcement*

Susan B. Long

*School of Management, Syracuse University*

This paper examines Internal Revenue Service allocation of staffing to tax enforcement. It is these decisions on which taxpayer classes should receive IRS attention which shape the odds of audit individuals face.

What factors influence these enforcement decisions? Recent times have seen a fundamental change in the information available to Internal Revenue Service decision-makers. Up until the late sixties, the tax agency had little means of determining the extent of tax violations, or how violation rates varied across regional areas, taxpayer groups or over time. While IRS had information from its regular audit program, this was not necessarily a reliable guide to general violation rates because of the selective nature of the audit process.

In the fall of 1967 this situation changed when detailed information about tax violations, scientifically gathered by the

\* This is an edited version of a paper presented at the Annual Meeting of the Law and Society Association, San Diego, June 7, 1985. Support for this research provided by a grant from the National Science Foundation is gratefully acknowledged. Any points of view or opinions expressed are those of the author and do not necessarily reflect those of Syracuse University or the National Science Foundation.

agency, became available. This information was the result of a multi-million dollar program – a combined effort on the part of research and enforcement staff of the agency – to scientifically measure the extent of taxpayer noncompliance. Using carefully monitored, detailed audits of a stratified random sample of tax returns, the program was designed to estimate the extent of compliance with tax reporting requirements detectable from a thorough tax audit of the taxpayer's return.<sup>1</sup> Since the returns were scientifically selected, results could then be projected to estimate what IRS auditors would find if every return were subject to a detailed tax audit.<sup>2</sup> In succeeding years, these compliance estimates have been regularly updated with additional TCMP surveys.

The purpose of this study is to determine what impact this information on taxpayer compliance has had upon staffing allocations across tax payer audit classes. The major thesis is that access to more accurate and reliable information about reporting errors by individuals filing federal income tax returns enabled the Internal Revenue Service to make more rational and efficient utilization of its audit resources in administering tax laws.<sup>3</sup> Specifically, it is hypothesized that once TCMP data on compliance patterns became available, IRS increased audit coverage of regions and return classes with low compliance and reduced audit resources allocated to high compliance areas.

A variety of attempts have been made to model the determinants of law enforcement actions.<sup>4</sup> A central assumption in many of these models is that increases in enforcement resources are applied in the most rational manner – to those areas most in need of increased enforcement attention. Some have taken this to the extreme, and have assumed that law enforcement agencies "set optimal penalties and produce optimal values of the probability of conviction and . . . penalties so as to maximize net per capita losses from crime" (Ehrlich, 1977a:771). Models by their nature must oversimplify, and few would argue that this view of the "rationale decision-maker" ignores much of the reality of day to day enforcement activity. Studies of discretionary justice often reveal the variability of actual administrative practices, and the multiplicity of competing practical objectives policy makers must weigh (Wilson, 1968; Nonet, 1969, Posner, 1972; Rosett, 1976; Kagan, 1978, Gardiner and Lyman, 1984).

Our focus here is on modeling decision making in tax

enforcement. While there is a growing body of research on taxpayer compliance, this work has almost exclusively focused on the determinants of taxpayer behaviour.<sup>5</sup> Attention given to IRS enforcement policies has been in the context of its potential impact upon taxpayer compliance (see for example, Witte and Woodbury, 1985; Schwartz and Orleans, 1967). This study is the first to attempt to model IRS enforcement reactions in response to taxpayer compliance.

## The Research Design

To address this research question a quasi-experimental research design is used. Three design elements contribute methodological rigor: (1) the experimental intervention, (2) the availability of detailed before and after measures, and (3) the use of the experimental groups as their own controls.

The first central design element is the existence of the "experimental intervention" itself, for it is the introduction of information on taxpayer compliance whose impact we wish to test.<sup>6</sup> The first survey, whose results became available in 1967 (fiscal 1968) involved a stratified random sample of approximately 50,000 returns filed in 1964. Each of the returns in this sample was given a detailed audit by an experienced IRS auditor or revenue agent, and detailed line-by-line information on reporting errors was recorded from each return. The compiled results for each region, giving both the rates and amounts of underreporting errors for various taxpayer groups, were distributed to IRS national and regional officials in the fall of 1967.

For purposes of this analysis, I have divided the post-intervention period starting with 1967 (fiscal 1968) into three periods. What I have called the "initial" period covers the two year period after the information first became available. The results of a second TCMP survey were introduced in fiscal 1970. The second intervention period considered extends the initial period four years to also cover the period between the availability of this new survey information and the next survey. Since it is expected that there could be a considerable lag in time between the availability of compliance estimates and their

potential effects on staffing policies, a third period extending through 1980 will also be examined.<sup>7</sup>

A second key element in the research design is the availability of detailed before and after measures. Time series data covering 1959 through 1980 have been compiled for this study so that any changes in enforcement trends after the introduction of compliance information can be carefully monitored. The series brings together internal agency data from different IRS divisions tracking over time the characteristics of IRS audit personnel (by both position and grade), the changing nature of the return workload, the allocation of IRS audit personnel among specific audit activities, and the conduct and results of federal income tax audits. The variety and scope of the information obtained offer an unusually detailed look at IRS audit practices.

The third element in the research design is the use of experimental groups as their own controls. Since events in addition to the experimental intervention can alter the behavior being monitored, the use of control groups to disentangle the effects of uncontrolled events from that of the experimental intervention itself provides important safeguards to the internal validity of an interrupted time series design.

Two different types of comparison groups are employed (see Table 1). Our first division is based upon the characteristics of the return. The IRS divides federal income tax returns filed by individuals into more homogenous groups based upon the source and level of income reported. These groupings, denoted by IRS as "audit classes," form one of the basic divisions used for IRS planning, resource allocation, and performance analysis. Explicit decisions are made concerning the allocation of audit personnel to each return class, and often estimates are made of the number of audits to be conducted in each category during the upcoming fiscal year. Detailed management statistics are produced on a quarterly, and sometimes more frequent, basis to monitor the implementation of this audit plan.

A second experimental division is based upon the structure of IRS field offices. IRS personnel carrying out tax audits are assigned to seven regional offices (also described in Table 1). The regional commissioners who head each office have a range of delegated authority to carry out the audit program, and in conjunction with the national office, develop their own audit plans for the upcoming year based upon decisions concerning



**Table 1** Study's experimental groups based upon Internal Revenue Service audit classes and regional divisions for income tax returns filed by individuals

Experimental Groups Based Upon:	
IRS audit classes by reported adjusted gross income (AGI)	IRS regional offices**
Wage-earner returns*	North-Atlantic: Connecticut, Maine, Massachusetts, New Hampshire, New York, Rhode Island, Vermont
class 1: less than \$10,000 AGI-using standard deduction	Mid-Atlantic: Delaware, District of Columbia, Maryland, New Jersey, Pennsylvania, Virginia
class 2: less than \$10,000 AGI-itemizing deductions	Southeast: Alabama, Florida, Georgia, Mississippi, North Carolina, South Carolina, Tennessee
class 3: between \$10,000-\$50,000 AGI	Central: Indiana, Kentucky, Michigan, Ohio, West Virginia
class 4: \$50,000 AGI and more	Midwest: Illinois, Iowa, Minnesota, Missouri, Nebraska, North Dakota, South Dakota, Wisconsin
business returns*	Southwest: Arkansas, Colorado, Kansas, Louisiana, New Mexico, Oklahoma, Texas, Wyoming
class 5: less than \$10,000 AGI	Western: Alaska, Arizona, California, Idaho, Hawaii, Montana, Nevada, Oregon, Utah, Washington
class 6: between \$10,000-\$30,000 AGI	
class 7: \$30,000 AGI and more	

\* wage-earner/business return division based upon whether taxpayer reported schedule C (business or professional) or Schedule F (Farm) income

\*\* From 1959-67 IRS had nine regional offices. To preserve comparability with later periods, data from this early period have been re-aggregated from the district office level into the same regional structure.

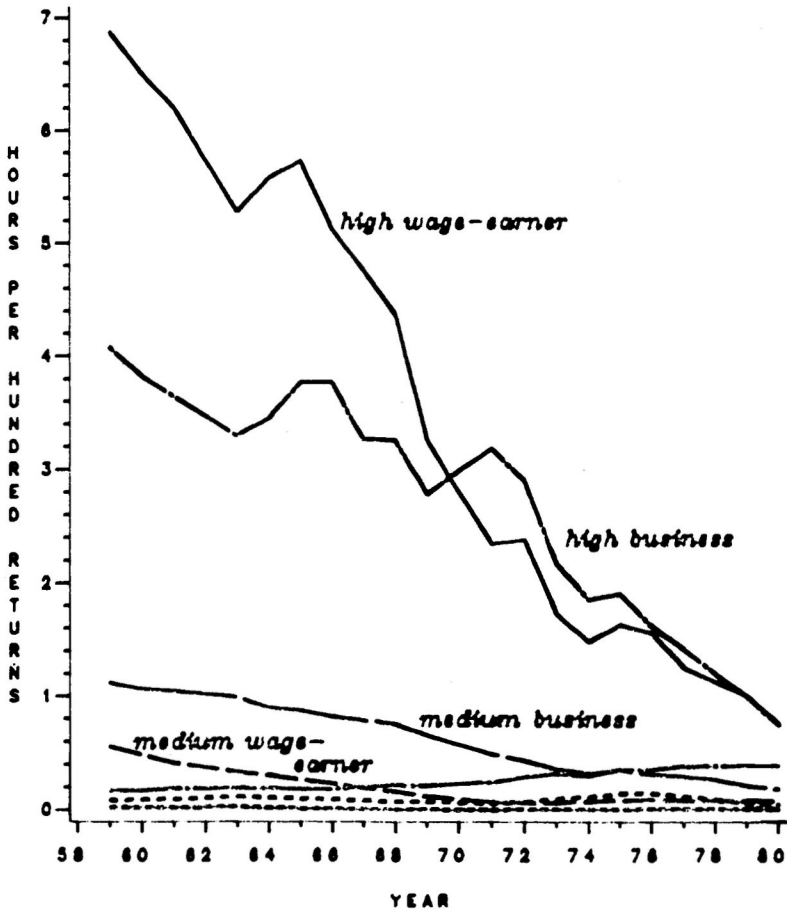
enforcement priorities among the audit classes. Again detailed management statistics are regularly prepared to monitor the audit program in each region.

Our experimental design employs a total of forty-nine experimental groups (the seven audit classes in each of the seven regional offices) in eight sets of experimental comparisons. We first examine the impact of compliance information on the distribution of audit personnel among seven audit classes in the United States as a whole. Next we will examine regional variation within each audit class in the allocation of audit personnel (seven comparisons, one for each audit class).

Two aspects of this multiple comparison approach facilitate treating these experimental groups as their own controls. First, our basic experimental hypothesis predicts differential responses will occur in response to the introduction of compliance data. Audit resources should be transferred to audit classes and regional areas having low compliance, while audit classes and regional areas with high compliance should experience lowered relative staffing allocations. Second, at any point in time, the audit personnel available are fairly fixed. If more personnel time is allocated to group A or region B, less time is available to audit other groups or regions. In essence, the nature of the experimental manipulation provides naturally varying experimental cues which should translate into quite different and opposite behavioral reactions on the part of IRS officials. Since the predicted outcomes of the intervention differ by experimental group, the threat to the study's internal validity from uncontrolled events altering each group in precisely the direction predicted by our model becomes increasingly remote.

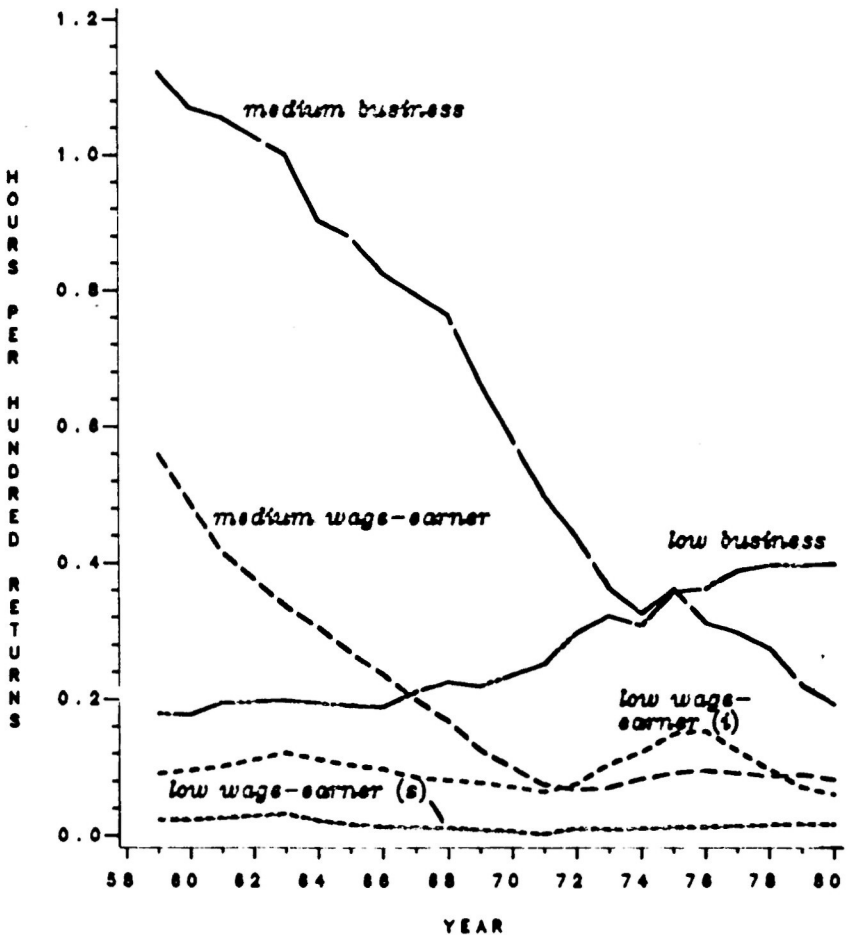
### **The Allocation of Audit Resources to Return Classes**

Three approaches will be used to analyze the results of our experimental tests; (a) trend analysis, (b) the decomposition of audit personnel flows, and (c) interrupted time series causal modeling. We will analyze the results of our first set of experimental contrasts among audit classes at the national level before turning to an examination of regional patterns of federal income tax audit coverage.

Figure 1 *The allocation of IRS audit hours*

#### *National trends in IRS audit coverage*

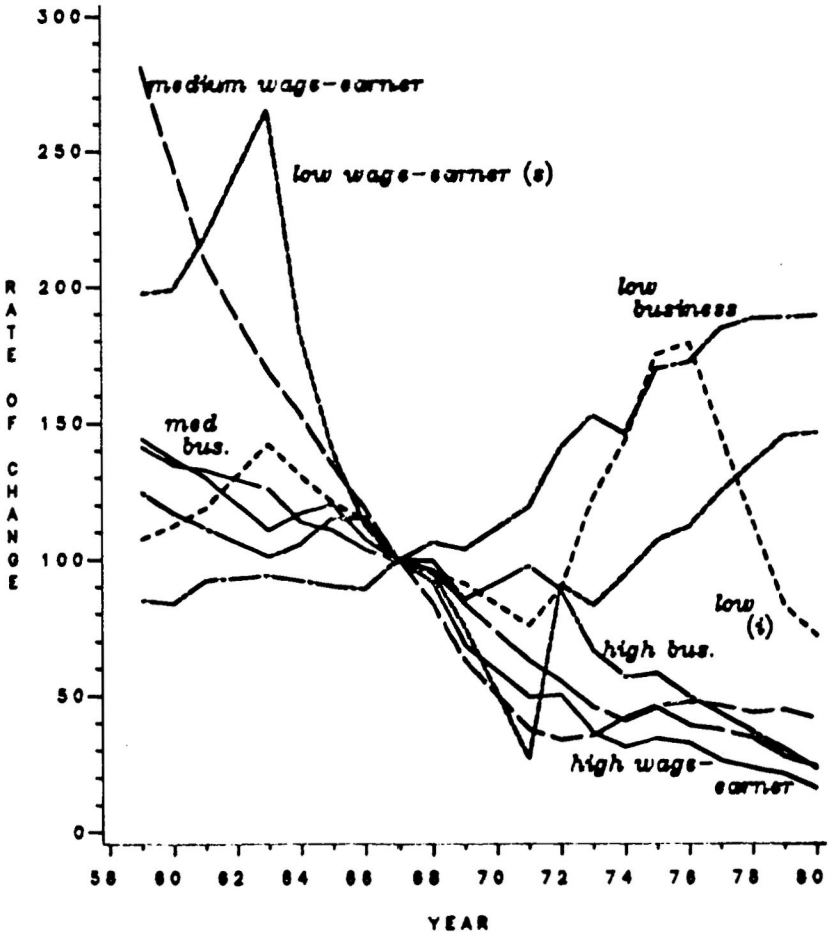
Figure 1 contains a series of time series plots of the allocation of IRS audit hours to the examination of federal income tax returns filed by individuals. Per capita figures based upon the ratio of audit hours to returns filed in each return class are depicted in Figures 1 and 2. While high income business and nonbusiness returns have consistently received the highest concentration of audit attention throughout this period, these groups have also experienced the sharpest absolute declines in audit coverage, falling from as high as nearly seven hours per every hundred

**Figure 2** *The allocation of IRS audit hours (expanded scale)*

returns filed in 1959 to less than a single hour per hundred returns by 1980. This translates into a drop in the proportion of returns examined from 51 percent in 1959 for high income nonbusiness returns and 24 percent for high income business returns to less than 8 percent and 5 percent, respectively, in 1980. Audit coverage has also declined for middle income business and non business returns from 1 in 10 in 1959 to only 1 in 50 in 1980.

In stark contrast, attention devoted to low income returns has increased during the past several decades, particularly for low

Figure 3 Change in IRS audit hours per return (from 1967 base year)



income business returns. Indeed, beginning in the late sixties and early seventies, low income business and itemized non-business returns had greater chances of audit than comparable middle income returns. This meant that during most of the seventies, one's chances of detection for a violation in reporting tax liability was higher for those filing a regular 1040 return making less than \$10,000 than for those making more than \$10,000 (except at the highest income levels).

The rate of growth or decline in audit hours per return is depicted in Figure 3, using 1967 as the base year. (This year was

chosen as the base for comparison since it divides the time series into pre- and post-intervention periods). Table 2 summarizes these trends.

While audit coverage changed dramatically for many audit classes during the years under consideration, the nature of these trends were remarkably stable from the pre- to the post-intervention periods. Those return classes whose audit coverage grew fastest during the pre-intervention period, continued to outstrip others' growth in audit coverage after the intervention. Return classes whose audit coverage declined during 1959-1967, continued to decline after 1967.

If we examine the pre-intervention and the three post-intervention periods, we find only three changes in the rank order of return classes according to their rate of change in audit coverage (see Table 2). Low income standard deduction returns after falling sharply in 1970 and 1971 move sharply upward in the rankings until they are second only to low income business returns.<sup>8</sup> The second shift in rankings occurs for middle income nonbusiness returns which move from last to fourth place. Finally, high and middle income business returns switch places in the late seventies, as high income return coverage continues to decline.

#### *The effect of changing patterns in return filings*

How are we to account for the remarkable stability in these patterns, or the three departures from this mold? An examination of patterns of return filings during this same period provide a partial explanation. Allocations of audit time are slow to react to changes in return filings. For those audit classes with declining numbers of returns, this "inertia effect" produces a growth in audit coverage. Similarly, those return classes experiencing the fastest growth rates in return filings experience the sharpest declines in audit coverage.

The only two audit classes experiencing an absolute decline in the number of returns filed between 1959 and 1980 are low income itemized wage-earner returns and low income business returns. As a result, these return classes experienced the sharpest gains in audit coverage during this period. At the other extreme, the two return classes which experienced the largest growth in return filings between 1959 and 1973 (middle and high income wage earner returns) and between 1974 and 1980 (high income

**Table 2** *Ranking of return classes by their rate of change in audit hours per return*

Rank Order in Change Rate	Pre-intervention		Post-intervention	
	1959-67	1967-69	1967-73	1967-80
1	low business*	low business*	low business*	low business*
2	low nonbus (i)*	low nonbus (i)*	low nonbus (i)*	low nonbus (s)
3	high business	high business	low nonbus (s)	low nonbus (i)
4	med. business	med. business	high business	med nonbus
5	low nonbus (s)	low nonbus (s)	med. business	med. business**
6	high nonbus**	high nonbus**	high nonbus**	high business**
7	med. nonbus**	med. nonbus**	med. nonbus**	high nonbus**

\* only classes with declining return filings

\*\* classes with the highest rate of increase in return filings

indicate change in relative ranking

wage-earner and business returns) are the classes experiencing the sharpest rates of decline in audit coverage during these respective periods. It is the acceleration in the growth rate of return filings of high income business returns, surpassing the growth rates of both business and nonbusiness middle income returns, that appears to explain why the ranking for this class fell below that of the other two in the last period.

The extent to which changes in staffing hours lagged behind changes in return filings throughout these years is summarized in Table 3. Rates of change in the number of audit hours are consistently less than rates of change (whether positive or negative) in return filings for four out of seven audit classes. Generally the larger the rate of change in return filings the larger the disparity between the growth in audit hours and returns.

The remaining three audit classes (low income returns) depart from this pattern in six out of twelve comparisons. In these three return groups, rates of changes in audit hours exceed the rates of change in return filings for four of these six departures. In the remaining two, hours increase while return filings are decreasing. Thus, while much of the stability of audit coverage patterns shown in Table 2 is explained by the stability in filing trends plus a lag in response to changes in return filings, we must search elsewhere to explain some of the patterns among low income returns.

#### *Decomposing audit personnel flows*

IRS's ability to respond to changes in return filings is constrained by both its staffing levels, and the technical mix of its audit personnel. Table 4 describes the character and distribution of audit personnel during the pre-intervention period of 1959-67, while Table 5 provides the same information for the post-intervention periods. Several points are worthy of note. First is the composition of IRS audit personnel. In 1959 roughly two-thirds of available time was provided by revenue agents - a position occupied by the most technically trained and qualified auditors. The remaining one-third were office auditors who examined the less complex returns (almost exclusively low income wage-earner returns in 1959). Almost all of these were concentrated in the lowest grades in 1959 - GS levels 4 through 7.

The second factor of note is how IRS adapted its audit program to the growth and increasing complexity of the new



**Table 3** *Percentage changes in federal return filings and Internal Revenue Service audit hours during pre- and post-intervention periods for United States per audit class*

change in workload and resource allocations	wage-earner income			business, professional, and farm income			returns
	standard	low	medium	low	medium	high	
pre-intervention period change:							
1959-1967							
audit hours	- 40%	- 9%	34%	1%	42%	63%	14.2%
return filing	18	- 2	276	- 14	100	103	14.3
audit hrs/return	- 51	- 7	- 74	17	- 29	- 20	- 0.1
post-intervention period change:							
1967-1969							
audit hours	- 6%	- 21%	- 9%	- 7%	1%	13%	- 7.2%
return filing	25	- 13	46	- 11	21	32	5.9
audit hrs/return	- 25	- 9	- 37	4	- 16	- 15	- 12.4
1967-1973							
audit hours	32%	- 42%	- 9%	+ 8%	- 24%	2%	- 17.8%
return filing	58	- 53	156	- 30	66	52	9.5
audit hrs/return	- 17	+ 24	- 64	+ 53	- 54	- 33	- 24.9
1967-1980							
audit hours	101%	- 69%	92%	- 10%	- 32%	32%	- 3.0%
return filing	38	- 57	361	- 53	178	467	31.6
audit hrs/return	46	- 28	- 58	+ 90	- 76	- 77	- 26.3
base year 1967 levels							
audit hours (000's)	221	2719	1711	1612	1533	1062	9709
return filing (000's)	1793	3181	859	765	193	32	6836
audit hrs/return	1.2	8.5	19.9	21.1	79.5	327.3	14.2

• for audit class detailed definitions see Table 2.

**Table 4** *Changes in the composition of Internal Revenue Service audit personnel during pre-intervention period, 1959-67*

(adjusted gross income)	total	1959 audit hours (000's)		change in audit hours 1959-67 (000's)	
		revenue agent hours	office auditor hours	revenue agent hours	office auditor hours
		grds. 9-11/grds. 4-7	grds. 9-11/grds. 4-7		
wage-earner					
low-standard	371	31	9	-150	74
low-itemized	2973	884	169	-254	1111
middle	1281	1092	33	430	335
high	370	365	3	231	3
business					
low	1591	1459	19	21	194
middle	1082	1058	6	451	43
high	652	648	1	410	5
all return	8505*	5718	242	1204	1769
			2545	827	-1391

\* includes 182(000) hours spent on returns reporting negative adjusted gross incomes in 1959

**Table 5** *Changes in the composition of Internal Revenue Service personnel, during post-intervention periods, 1967-80*

audit class of return (adjusted gross income)	total	1967 audit hours (000's)		
		revenue agent hours	office auditor/ tax technician grds. 9-11/grds. 4-7	
wage earner				
low-standard	221	30	83	108
low-itemized	2719	639	1280	799
middle	600	593	6	2
business				
low	1612	1347	213	51
middle	1533	1472	49	12
middle	1533	1472	49	12
high	1062	1055	6	1
all returns	9709 *	6544	2011	1155
		change audit hours 1967-73 (000's)		
	total change	revenue agent change	office auditor/ tax technician change grds. 9-11 grds. 4-7	
wage earner				
low-standard	70	- 11	50	31
low-itemized	- 1131	- 129	- 701	- 301
middle	- 155	- 407	88	163
high	- 35	- 42	5	3
business				
low	123	- 352	433	41
middle	- 366	- 392	28	
high	17	9	8	1
all returns	- 1727	- 1570	- 93	- 64
		change audit hours 1973-80 (000's)		
	total change	revenue agent change	office auditor/ tax technician change grds. 9-11 grds. 4-7	
wage earner				
low-standard	- 748	- 234	- 330	- 188
low-itemized	- 748	- 234	- 330	- 188
middle	1729	48	619	1049
high	395	230	89	75
business				
low	- 286	- 159	- 106	- 21
middle	- 132	- 360	197	30
high	320	195	107	18
all returns	1432 **	- 168	581	1000

\* includes 251(000) hours spent on returns reporting negative adjusted gross incomes in 1967

\*\* includes 19 (000) hours from tax technicians with GS level above grade 11.

returns that were filed. While increases in auditor hours matched the overall growth rate in return filings during the pre-intervention period (Table 3), the increasing numbers of middle and high income returns required more hours per average to examine than the 1959 workload. To preserve its 1959 audit coverage, IRS would have had to had an increase in staffing hours almost five times as large (an increase of 5.6 million hours instead of the 1.2 million it experienced).

IRS coped by upgrading its office auditor staff. While almost all office auditors were concentrated in the lowest grades (4-7) in 1959, eight years later two-thirds had moved up to GS grades 9 or 11. These were used in two ways. First, they were used to replace revenue agents for some of the more technical examinations of low income itemized and business returns, freeing up revenue agent time which could be used on the more complicated higher income returns. Secondly, they were used in increasing numbers for the examination of middle income wage-earner returns – then the fastest growing return sector. The increases in available revenue agent time were concentrated on the most complicated return classes – upper income wage-earner and middle and upper income business returns.

The post-intervention years present a stark contrast to the pre-1967 period. First, available auditor time did not increase, and in fact dropped sharply (18%) between 1967 and 1973.<sup>9</sup> While increases between 1973 and 1980 largely made up this loss, the number of revenue agent hours continued to decline. Replacements were drawn from the ranks of office auditors, largely those with the least skills (grades 4-7).<sup>10</sup> Thus, not only did audit hours fail to increase during the post-intervention period, but they came increasingly from the less technically skilled office auditor (renamed tax technician) ranks, particularly those drawn from the lowest auditor grades. This occurred despite a growing return population, increasingly concentrated in the more complicated higher income levels.

There was, however, an additional source of audit personnel after 1967 from the redeployment of auditors formerly assigned to examine low income itemized returns. This, the largest return sector, experienced a sharp drop in return filings (down over 50 percent). While the decline in auditor time lagged behind these return declines, this decrease in workload freed over 1.1 million in additional auditor hours by 1973 (and 1.9 million by 1980) that

could be deployed elsewhere. But most of these were office auditors. While they could relieve some of the pressure on revenue agents time by taking over examination tasks in the middle income wage-earner and lower income business returns, the workload in the latter category has also fallen with a thirty percent drop in return filings.

With the increase in office audit staff, combined with the redeployment of auditors from the declining workload of lower income business and itemized wage-earner returns, there was an enormous increase in available lower grade office auditors (tax technicians) after 1973. Since these could only be assigned to the least complicated returns, this appears to be the major reason why audit coverage for low income standard deduction and middle income wage-earner returns increased then. Without upgrading available auditors' skills, these were the only two return classes with growing workloads where these less technically trained personnel could be reassigned.

In contrast, there was no similar solution for higher income wage-earner returns and middle and upper income business returns which typically required highly technically trained audit personnel. The combination of declining revenue agent hours and increasing return filings spelled sharply declining audit coverage for them throughout the post-intervention years.

#### *The advent of compliance measures*

The advent of compliance estimates in 1967 provided IRS with the ability to estimate the amount of tax liability not reported on returns individuals filed, the number of returns underreporting, and a host of indices relating these statistics to activity in its regular audit program.

In general, TCMP estimates confirmed that both the frequency and amount of tax under-reporting increased with income. Noncompliance also increased with the complexity of the return. Those filing a standard deduction return had error rates only one seventh as high as taxpayers with the same income itemizing their deductions, and underreported taxes only one third as high. Returns reporting business, professional or farm income had error rates at least twenty percent greater than wage-earner returns, and the amounts of underreporting were sharply higher on business returns.

IRS's pattern of audit coverage was not wholly dissimilar to the patterns of noncompliance IRS estimates revealed. For example, in fiscal 1967 the number of IRS audit hours per hundred returns also increased as taxpayer income increased, and was more highly concentrated on the more complex returns (with the exception of the high income business returns which received less attention than similar nonbusiness returns). Thus, it was not the case that the information revealed by the compliance study suggested a need for wholesale realignment of audit coverage among its audit classes.

Nonetheless, the new data did suggest some areas where a change in audit coverage seemed warranted. Business as compared with nonbusiness returns as a class received proportionately less attention from IRS given their frequency and amount of underreporting errors. For example, the percent of unreported taxes recovered through audit was consistently less for business than nonbusiness returns at all income levels, and the odds of detection were generally lower despite the poorer compliance record of business returns. The largest disparity occurred for high income business returns, which showed the lowest concentration of audit hours relative to the unreported taxes in that class.

It is clear from the audit coverage trends we have previously examined that these compliance figures produced no clearcut immediate or long term realignment in audit patterns. On the contrary, to the extent that increases in audit coverage occurred these generally favored low income returns (see earlier Table 3), while more complicated high income wage-earner and middle and upper income business returns experienced sharply lower levels of audit coverage.

This pattern, however, is complicated by several additional factors. First, the implications of these compliance figures for audit coverage of return classes depend in part upon what criteria are used. While rates and amounts of underreporting error rise with income, the percent of the "true" tax liability underreported falls with rising income. Judged on this criterion alone, the higher income the more compliant the average taxpayer is.<sup>11</sup> Although it continues to be true that more complex returns (whether comparing itemized to standard deduction returns, or business to wage-earner returns) continue to have higher noncompliance levels on whatever standard used.

Further, while it seems preferable that less compliant tax-

payers should not have smaller odds of detection than more compliant taxpayers, how much higher (if any) their audit chances should be is not clear. Should everyone have the same odds of detection? If odds should increase for less compliant groups, how much higher should their audit coverage become? Further, compliance figures changed over time. How much emphasis should be devoted to return classes with the greatest deterioration in compliance?

Finally, the agency is faced with many trade-offs including the immediate return to the U.S. Treasury from its audit activity. Despite the relatively low compliance problems found on low income standard deduction returns, these are quite profitable to audit. The average return per audit hour from examining these returns in 1967 was \$328 – greater than the per hour return for all but the highest income returns. Since the latter require higher paid audit personnel and also generate more losses and processing burdens from taxpayer appeals<sup>12</sup> the net direct return from audit might be highest from auditing those taxpayers who are most compliant. The average revenue generated per hour of auditor time is also lower for the more complicated business returns, despite its generally higher noncompliance levels. Questions of fairness aside, practical political considerations could favor allocating IRS personnel to those groups who are most "profitable" to audit.

#### *An interrupted time series causal model*

If the impact of compliance information on audit hour allocations was dramatic, its impact would be evident from an inspection of the plots in Figure 1. Clearly this is not the case. As we have observed, changes in auditor availability and return workloads over this period have also had a major impact upon audit coverage. Thus, to isolate the potential effects of compliance information from other causal factors, a formal interrupted times series causal model was developed.<sup>13</sup> If we are successful in modelling staffing allocation decisions, we should be able to estimate the relative importance of different factors on audit coverage, and test the specific impact that the introduction of compliance data had. Our observations consist of annual U.S. data between 1959 and 1980 covering each of the seven audit classes.<sup>14</sup>

The basic unit in resource allocation, auditor hours, is used as our dependent variable. Four classes of causal variables are included: (a) staffing availability, (b) changes in workload, (c) results from the regular audit program, and (d) compliance measures.

Four measures of staffing availability were employed. First, the audit hours available in any one period, should be a function of the audit hours available from the previous period. Second, increases (or decreases) in audit staffing in the nation as a whole set limits on staffing available to any specific audit class. Since auditors vary in their training and the type of returns they are assigned to examine, changes in staffing availability was separately included for (a) revenue agents, (b) office auditors/tax technicians in grades 9-11, and (c) office auditors/tax technicians in grades 4-7.<sup>15</sup>

Two workload variables were used. The first measured the change in staffing hours needed to maintain the same audit coverage given changes in return filings for that audit class. The second measured workload changes summed for the other return classes, since competing workload demands could also effect staffing allocations.

Results from the regular audit program may have been used to guide staffing allocations, particularly prior to the availability of compliance information. Historically IRS has focused upon two basic indices. First is the total additional tax and penalties generated from audits (also examined on a per return and per audit hour basis). These figures should be positively related to increases in staffing levels. Second, is the number (and rate) of audits resulting in no tax change which IRS works to minimize.

For our fourth class of indicators, taxpayer compliance information provides estimates of the total (corrected) tax liability, the tax dollars underreported on tax returns, and the number of tax returns on which tax underreporting occurs. These figures can also be adjusted by figures from the regular audit program to give estimates of the number of returns or tax dollars underreported which are not detected through the regular audit program (the return and tax "gaps"). Thus, five alternative compliance indices representing return errors, taxes underreported, the return gap, the tax gap, and the proportion of tax liability not reported were examined.

This basic model was parameterized on a regular and on a per



capita basis. In the regular model, audit hours was the dependent variable. In the per capita model, audit hours per return filing was the dependent variable, and all the explanatory variables were also converted to "per capita" basis.

The model, even without compliance information, explained a very high proportion of total variation in staffing allocations – 98 percent for the regular model, and 99 percent for the per capita model. Of the compliance measures, the tax gap variable proves to have the highest explanatory power. But at best it increased the total explanatory power of the model only marginally, accounting for an additional 1 to 5 percent of the variation in audit allocations remaining after the other three classes of variables are included in the model. (The addition of other compliance measures did not contribute any additional explanatory power.) If results from the regular audit program are included for both the pre and post-intervention periods, compliance information accounts at most for only an additional one percent of the variation in audit coverage, and the slope fails to reach statistical significance (while in the per capita model (not shown) it turned negative in value). Thus, in general, the introduction of compliance information appears to have had at best only a very modest impact upon audit coverage at the national level. Even varying the onset of when compliance data's impact was expected, did not materially change this general conclusion.

Estimates for our other causal parameters mirrored the picture already seen in the earlier stages of this analysis. First, long-term stability in audit allocations – rising or falling in reaction to changes in staffing availability. Second, changes in workload were a less significant factor (and changes in competing workloads proved to be insignificant). Further, changes in staffing allocations significantly fell behind changes in workload requirements. Each hour of increased (or decreased) return workload resulted in only 0.2 hours change in actual hours of staffing allocated.<sup>16</sup> Nor did subsequent years bring any "catch up" in allocations since neither workload changes lagged for additional periods, or a two or three year moving average of workload changes proved to add any additional explanatory power. The productivity of audits from the regular examination program, and the number or rate of "no change" audits also appear to be related to staffing allocations in the predicted

manner (positive for revenue, negative for no change audits).<sup>17</sup>

Our results were not materially affected by any departures from the underlying assumptions of our regression model. A full residual analysis was conducted to check for linearity, homoscedasticity, normality and independence of disturbance terms, and the presence of unduly influential outliers (Belsey et. al, 1980; Daniel and Wood, 1980; Neter et al., 1983; Bollen and Jackman, 1985). For our regular model, marginal departures from normality (based upon the Kolomogorov D statistic (Stephens, 1974) and normal probability plots) and slight autocorrelation (though not rising to the level of statistical significance in a Durbin-Watson test (1951)) disappeared when three outliers were removed, although their removal did not materially affect regression estimates. For our per capita model, some curvilinearity was introduced for several variables, but was sufficiently small so that after data transformation the same pattern of results remained.

## **Regional Allocation of Audit Resources**

The regional allocation of staffing levels mirror trends at the national level. Plots for audit hours per return by region show a similar pattern to the national trends found in Figure 1. However, even after adjusting for differing return workloads,<sup>18</sup> staffing levels differ by a factor of two or three from one region to the next. As shown in Table 6, some regions have only about half the average number of hours per return as in the United States as a whole, while other regions have half again as many as U.S. levels. The extent of this regional variation remains fairly constant throughout the pre- and post-intervention periods.

While the range of variability remains little changed, there is considerable movement over time in which regions have high or low audit coverage. Correlations between annual audit coverage ratios, also shown in Table 6, are generally lower the longer the time span. There is no correlation at all between audit coverage ratios of 1959 and 1980. Even for periods four or more years apart, correlation levels drop below 0.6, with only one third or less of the variation in common.

**Tabel 6** Regional variation in audit hours allocated per returns filed as a percent of US ratio

Year	N*	Range		Correlation (r)						
		low	high	1959	1967	1969	1973	1980		
1959	49	54%	156%	1.00						
1967	49	61	151	.41	1.00					
1969	49	69	159	.49	.49	.74	1.00			
1973	49	68	152	.30	.47	.56	1.00			
1980	49	58	157	-.01	.31	.3	.37	1.00		

\* seven audit classes for each of seven regions

Differences in regional growth rates of return filings do not appear to account for the pattern of regional differences in audit coverage found. While return filings grew rapidly for the Mid-Atlantic region in the early years, and accelerated for the Southeast, Southwest and Western regions in later years, regional growth rates show no systematic associations with higher or lower relative staffing levels. When we examine correlations within each of the seven audit classes, only two are statistically significant and these are opposite in sign. Thus we must look elsewhere for an explanation for these regional differences in audit coverage.

*Regional variation in compliance levels*

Compliance estimates also show pronounced regional variation. Although variability differs somewhat by measure, regional noncompliance levels ranged from a low of 40 percent of U.S. audit class averages to over 3 or 4 times U.S. figures.

The advent of compliance information, however, did not result in any dramatic realignment of regional audit coverage, even though unlike the earlier national comparisons, these compliance indices showed unambiguously which regions had higher or lower tax compliance. This was true despite the fact that regional differences in audit coverage did not correspond with differences in compliance patterns. Inspection of time series plots by audit class for each region reveals that regions with high noncompliance or return/tax gap figures appear almost as likely to decrease as to increase audit time.

Simple bivariate correlations between staffing ratios and tax-payer compliance relative to U.S. figures summarize these relations. Most of the variation in regional staffing levels cannot be explained by regional differences in compliance levels. Initial TCMP figures indicate that regional variation in compliance levels had some positive correlation (about 10–16 percent variation in common) with the pre-intervention audit coverage ratios of 1967. While there was a modest increase in the degree of association between audit coverage and compliance between 1967 and 1980, most of this was not brought about by changes in audit staffing. Indeed changes in audit staffing more often decreased rather than increased the level of correlation (21 decreases compared to 15 increases in 36 year-to-year com-

parisons) particularly for both the average tax underreported, and the percent of tax liability not reported.<sup>19</sup>

### *Modeling regional staffing allocations*

Just as at the national level, changes in audit coverage do not appear to be dramatically affected by the introduction of regional compliance information. Thus, we turn to an interrupted time series causal modeling approach to better isolate the potential effects of compliance information from other causal factors.

The same model applied earlier to U.S. staffing allocations is also estimated using these regional figures, both with audit hours (regular model) and hours per return (per capita model) as the dependent variable. Results show that again, the model without compliance measures explained a very high proportion of total variation in staffing allocations whether indexed by audit hours, or hours per return. Multiple R<sup>2</sup> varied from a low of 85 percent to a high of 97–99 percent.

Of the five compliance measures examined, the tax gap proved to have the greatest explanatory power across the seven audit classes. Even so, its effects ranged from nonexistent to moderately weak. For no audit category was the slope for tax gap statistically significant for all four of the alternative formulations considered. The strongest effects appeared for the low income business return class. Here the partial R<sup>2</sup> attributable to compliance information varied from 0.021 to 0.126, depending upon the model's formulation. For all audit classes, however, the median partial R<sup>2</sup> was only 0.018. Further, the model's total multiple R<sup>2</sup> values were little affected by the addition or subtraction of the tax gap variable ranged from 0.0000 up to 0.0092, with a median change of only 0.0008 – a change only in the fourth decimal place. Translated this means that compliance information contributed only an additional one tenth of one percent in explaining the variation in audit coverage.<sup>20</sup>

## **Conclusion**

The hypothesis that IRS officials would respond "rationally" to the introduction of data on taxpayer compliance by increasing

auditor time devoted to those regions and audit classes with low compliance levels, and reducing time devoted to high compliance areas was given only minimal support from this study. In general, the introduction of TCMP compliance data did not bring about any dramatic restructuring in audit coverage – even when it disclosed regions or return classes with much lower compliance levels which were receiving less audit attention than more compliant groups.

Inspection of time series plots and trend statistics at both the national and regional levels revealed little, if any, changes in audit coverage attributable to the introduction of compliance data. Although taxpayers reporting business and professional income had much lower compliance rates, these taxpayers as a group continued to receive less relative attention than more compliant taxpayers – particularly middle and upper income business taxpayers who had the highest amounts of under-reporting. Some regions continued to have two or three times the number of audit staff available per return filed though it bore little relationship to regional differences in compliance levels. Even after controlling for other causal factors, at most differences in compliance levels among taxpayer classes and regional areas explained only an additional one to five percent of the variation in audit coverage.

While constraints imposed both by the level of audit staffing at IRS, and the technical mix of its audit personnel, reduced IRS's ability to handle its growing workload, these factors did not explain why IRS failed to more fully utilize compliance information in allocating its limited resources to those regions and audit classes most in need of audit.

## Notes

1. Surveys have also been conducted to estimate noncompliance with filing requirements and delinquency in the payment of tax liabilities (see IRS, 1977).
2. As a measure of tax violations, these figures reflect the strengths and weaknesses of income tax audits for assessing tax violations. Reporting errors arising from any reason – from inadvertent errors to tax fraud – are covered, although simple math errors are excluded. For a discussion of under and over-reporting biases of this method, see IRS, 1979, 1983; Long, 1980.

3. The IRS has also used TCMP data to develop predictive formula for screening returns for audit selection. An examination of the use of these predictive formula in promoting IRS efficiency is outside the scope of this paper.
4. See for example, Carr-Hill and Stern, 1979; Greenwood and Wadycki, 1973; McPheters and Strong, 1974; Pogue, 1972; Votey and Phillips, 1972; Wilson, 1979.
5. For theoretical models of tax compliance see Allingham and Sandmo, 1972; Sandmo, 1981; Srinivasan, 1973; Witte and Woodbury, 1983b; Yitzhaki, 1974. For recent reviews see Boidman, 1983; Witte and Woodbury, 1983a; also Cahalan and Ekstrand, 1980; Clotfelter, 1983.
6. Copies of the actual internal agency documents containing these compliance estimates, as well as internal office memorandum indicating the timing and distribution of this information among IRS national and regional officials have been obtained for use in this study.
7. Years subsequent to 1980 are not included in this study because a fundamental change in the design of IRS's management information system make data on audit classes after this point not comparable with the pre-1980 period.
8. While low income itemized wage-earner returns show a cyclical trend, when change during each post-intervention period is considered as a whole, this return class ranks consistently second until it is passed by low income standard deduction returns in the last period.
9. A hiring freeze, and the later assignment of economic stabilization activities to the Service during the Nixon presidency are some of the factors which contributed to the decline in available staffing time.
10. The reasons for IRS decision to downgrade the technical mix of its employees is outside the scope of this paper. But since 1948, the ranks of office auditors/tax technicians have grown at a much faster rate than that of revenue agents - despite the increasing complexity of return workloads.
11. Because audit classes are based upon reported - not corrected - adjusted gross income levels, these ratios may be misleading since high income taxpayers with large tax writeoffs appear as "low income" taxpayers. This was one of the main considerations which lead IRS to change its audit classification system in 1980 to one based upon total positive income. There is also some evidence that the thoroughness of audits declines as the income and complexity of a return rises (Long, 1980), further biasing comparisons.
12. The larger the dollar claim arising from audit, the more likely the taxpayer is to appeal. Not only do such appeals increase IRS administrative costs, but IRS typically dropped two-thirds of the original dollar claims during this period (with better "deals" going to those with the most at stake). See Long, 1980.
13. The annual nature of return filing behavior and the unavailability of quarterly data by audit class for much of the pre-intervention period

resulted in a time series of insufficient length to use ARIMA modeling methods. The breadth and detailed nature of the information available, plus our interest in the combination of causal factors giving rise to the audit patterns observed, also suggested that an interrupted time series causal modeling approach would be a more appropriate methodological strategy here.

14. Audit breakdowns for 1970 were never produced by IRS because of transition problems with a new management information system. Only six months of data was available for 1962. Thus, both of these years had to be excluded from the analysis.
15. Because return workloads vary, the total change in hours was prorated among audit class based upon the proportion of total audit hours that audit class received in the previous time period.
16. Because of this, the relationship in the per capita model is negative since increases in workload decreased the audit hours available on a per return basis.
17. These estimates were not seriously effected by multicollinearity. With only staffing availability and workload variables in the regular model, the highest variance inflation factor (VIF) was only 1.2, on a scale where 1.0 indicates no correlation among the explanatory variables while 100.0 would occur when the intercorrelation among the explanatory variables rose to 99 percent. Once, the audit and compliance variables were added, multicollinearity in the regular model reached nearly 5 for lagged hours, although it wasn't a problem for our compliance measure which had a VIF of only 1.2. Multicollinearity ran somewhat higher in the per capita model, reaching 11.5 for lagged hours per return although even this level was not a problem given the very high explanatory power of this variable.
18. Regional differences in sources or levels of return income have also been controlled for since these ranges were calculated within audit classes.
19. Changes in staffing levels were associated with increased correlations concerning the estimated percent of return underreporting. But even here one out of three year-to-year changes decreased this correlation, and by the end of 1980 less than 20 percent of their variation was shared.
20. Multicollinearity for our tax gap variable, did not explain the lack of significance of these findings. Residual analysis on results from these models showed some modest departures, but neither data transformations or the omission of potential influential outliers had any substantial impact upon the basic pattern of findings reported in Table 16. While some corrections slightly increased the significance of the tax gap variable, in other instances the slope became non-significant. Nor did varying the onset when compliance data's impact was expected materially change our general conclusions.



## References

- Allingham, Michael G. and Agnar Sandmo (1972) "Income Tax Evasion: A Theoretical Analysis," 1 *Journal of Public Economics* 323.
- Belsley, David A., Edwin Kuh and Roy E. Welsch (1980) *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York: John Wiley & Sons.
- Boidman, Nathan (1983) "A Summary of What Can Be Learned from the Experience of Other Countries with Income Tax Compliance Problems" in P. Sawicki (ed.), *Income Tax Compliance: A Report of the ABA Section of Taxation Invitation Conference on Income Tax Compliance*. Washington, D.C.: American Bar Association.
- Bollen, Kenneth A. and Robert W. Jackman (1985) "Regression Diagnostics: An Expository Treatment of Outliers and Influential Cases," 13 *Sociological Method & Research* 510.
- Calahan, Margaret and Laurie Ekstrand (1980) "Who are the tax cheaters and why do they cheat," Paper presented at the 1980 annual meeting of the Law and Society Association.
- Carr-Hill, R. A. and N. H. Stern (1979) *Crime, the Police and Criminal Statistics*. London: Academic Press.
- Clotfelter, Charles T. (1983) "Tax Evasion and Tax Rates: An Analysis of Individual Returns," 65 *Review of Economics and Statistics* 363.
- Daniel, Cuthbert and Fred S. Wood (1980) *Fitting Equations to Data* (2nd ed.). New York: John Wiley & Sons.
- Durbin, J. and G. S. Watson (1951) "Testing for Serial Correlation in Least Squares Regression," 2 *Biometrika* 159.
- Ehrlich, Isaac (1977) "Capital punishment and deterrence: Some Further Thoughts and Additional Evidence," 85 *Journal of Political Economy* 741.
- Gardiner, John A. and Theodore R. Lyman (1984) *The Fraud Control Game: State Responses to Fraud and Abuse in AFDC and Medicaid Programs*. Bloomington: Indiana University Press.
- Greenwood, Michael J. and Walter J. Wadycki (1973) "Crime Rates and Public Expenditures for Police Protection: Their Interaction," 31 *Review of Social Economy* 138.
- Kagan, Robert A. (1978) *Regulatory Justice: Implementing a Wage-Price Freeze*. New York: Russel Sage Foundation.
- Long, Susan B. (1980) *The Internal Revenue Service: Measuring Tax Offenses and Enforcement Response*. Washington, D.C.: U.S. Government Printing Office.
- McPheters, Lee R. and William B. Strong (1974) "Law Enforcement Expenditures and Urban Crime," 27 *National Tax Journal* 633.
- Neter, John, William Wasserman and Michael H. Kutner (1983) *Applied Linear Regression Models*. Homewood, Ill.: Irwin.
- Nonet, Phillippe (1969) *Administrative Justice: Advocacy and Change in A Government Agency*. New York: Russel Sage Foundation.
- Pogue, Thomas F (1972) "Effect of Police Expenditures on Crime Rates: Some Evidence," 3 *Public Finance Quarterly* 14.

- Posner, Richard A. (1972) "The Behavior of Administrative Agencies," 1 *Journal of Legal Studies* 305.
- Rosett, Arthur and Donald R. Cressey (1976) *Justice By Consent: Plea Bargains in the American Courthouse*. Philadelphia: Lippincott.
- Sandmo, Agnar (1981) "Income Tax Evasion, Labor Supply, and the Equity-Efficiency Tradeoff," 16 *Journal of Public Economics* 265.
- Schwartz, Richard D. and Sonya Orleans (1967) "On Legal Sanctions," 34 *University of Chicago Law Review* 274.
- Srinivasan, T. N. (1973) "Tax Evasion: A Model," 2 *Journal of Public Economics* 339.
- Stephens, M. A. (1974) "EDF Statistics for Goodness of Fit and Some Comparisons," 69 *Journal of the American Statistical Association* 730.
- U.S. Internal Revenue Service (1977) *TCMP Handbook*, Document 6457.
- U.S. Internal Revenue Service (1979) *Estimate of Income Underreported on Individual Income Tax Returns*. Publication 1104.
- U.S. Internal Revenue Service (1983) *Income Tax Compliance Research: Estimates for 1973-1981*. Washington, D.C.: U.S. Treasury Department.
- Votey, Harold L. and Llad Phillips (1973) "Police Effectiveness and the Production Function for Law Enforcement," 1 *Journal of Legal Studies* 423.
- Wilson, James Q. (1968) *Varieties of Police Behavior: The Management of Law and Order in Eight Communities*. Cambridge, Mass.: Harvard University Press.
- Wilson, James Q. and Barbara Boland (1979) *The Effect of the Police on Crime*. Washington, D.C.: U.S. Government Printing Office.
- Witte, Ann D. and Diane F. Woodbury (1983a) "What We Know About the Factors Affecting Compliance with the Tax Laws" in P. Sawicki (ed.), *Income Tax Compliance: A Report of the ABA Section of Taxation Invitational Conference on Income Tax Compliance*. Washington, D.C.: American Bar Association.
- Witte, Ann D. and Diane F. Woodbury (1983b) "The Effect of Tax Laws and Administration on Tax Compliance," Working paper 83-1, Department of Economics, University of North Carolina, Chapel Hill.
- Witte, Ann D. and Diane F. Woodbury (1985) "The Effect of Tax Laws and Tax Administration on Tax Compliance: The Case of the U.S. Individual Income Tax," 38 *National Tax Journal* 1.
- Yitzhaki, Shlomo (1974) "A Note on Income Tax Evasion: A Theoretical Analysis," 3 *Journal of Public Economics* 201.