THE RELATIVE EFFECT OF PENALTY MAGNITUDES ON COMPLIANCE: AN EXPERIMENTAL EXAMINATION OF DETERRENCE

By

RICK LYNN CROSSER

Bachelor of Science
Colorado State University
Fort Collins, Colorado
1970

Bachelor of Science
Colorado State University
Fort Collins, Colorado
1978

Master of Science
Colorado State University
Fort Collins, Colorado
1982

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 1987

Thesis
1987D
C951r
cop. 2

OKLAHOMA STATE UNIVERSITY LIBRARY

THE RELATIVE EFFECT OF PENALTY MAGNITUDES ON COMPLIANCE: AN EXPERIMENTAL EXAMINATION OF DETERRENCE

Thesis Approved:

Thesis Advisor

Dean of the Graduate College

1307078

ACKNOWLEDGEMENTS

I would like to express my appreciation to the members of my committee, Dr. Milton F. Usry, Dr. Janet I. Kimbrell, Dr. William D. Warde, and especially Dr. Lawrence H. Hammer. Their insights and expertise made this study possible. I also appreciate the assistance from other faculty at Oklahoma State University.

I am also indebted to my fellow doctoral students. The support, friendship, and contributions to this study of so many is appreciated. A special acknowledgement is due to Doug Laufer whose contributions to the study and whose perspectives on the doctoral program were immensely valuable.

Most of all I wish to thank my family. The support and understanding of my siblings, Carla, Nancy, Mark, and Brian, and my sister-in-law, Cindy, and the joy surrounding my niece, Hanna Christine, were essential to undertaking and completing the doctoral program. Also, the loyalty of Otis, who sacrificed numerous afternoons and evenings of bird and squirrel watching to lie at my feet while I composed and typed this study, is appreciated. To Mom, Merle Crosser, I dedicate this dissertation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

THE RESEARCH PROBLEM

Introduction

The purpose of this chapter is to present an overview of the tax compliance problem, to introduce the theory of deterrence, to provide a description of the purpose of the research, and to describe the organization of the remainder of the dissertation.

Overview of the Problem of Tax Noncompliance

Noncompliance with United States tax laws includes failure to file a tax return, underreporting of income, overstatement of deductions or exemptions, and failure to pay taxes due. Although noncompliance in some cases is inadvertant, this study focuses on overt noncompliance because overt noncompliance with U.S. tax laws is a severe and growing problem. Estimates of the loss in federal revenue due to such noncompliance were placed at $95.2 billion for 1981 by the U.S. Internal Revenue Service [1983]. Voluntary compliance rates have been estimated to have dropped from 84 percent in 1973 to less than 81 percent in 1981. These estimates may be conservative, based on the

estimates of federal revenue loss by Gutmann [1977, 1980], Feige [1979], and Tanzi [1982].

Gutmann estimated the underground cash economy to be $176 billion for 1976. This estimate constituted about 9.5 percent of the total 1976 reported GNP. In 1980, Gutmann revised the earlier figures upward to $223 billion for 1976. The estimated tax loss associated with this activity was over $50 billion.

Growth trends have been applied to the discrete evasion estimates of Gutmann, the IRS [1983], and others [Feige, 1979; Tanzi, 1982]. The resulting approximations of current and potential future tax revenue lost through evasion are staggering. Point estimates for growth rate vary from study to study, but the consensus is that the current rate of 10 to 15 percent is on the rise. Consensus places the estimate of the 1983 tax gap in excess of $100 billion. Predictions for 1985 are $120 billion [Friedrich, 1983].

If inroads can be made into eliminating the causes of the gap in revenue collections, that is, the amount of tax revenue lost through evasion, significant reductions in the much-discussed federal deficit could result in addition to realizing improvement in the overall U.S. tax environment.

Musgrave and Musgrave [1984, p. 225], elaborating on the requirements for a good tax structure, stressed the importance of the criteria of equity and efficiency. When these criteria are not being effectively carried out, the

entire economy suffers. Equity is diminished by the ability of taxpayers to evade paying taxes. Horizontal equity is violated by a taxpayer's ability to evade paying the taxes paid by a compliant taxpayer with similar income. Vertical equity is breached by the high income tax evader paying no more tax than a low income taxpayer. When the costs of compliance to taxpayers and of collection for the government are reduced in relation to the total tax revenue, efficiency is increased. More capital would remain available for private investment and the government would have more funds available for allocation.

To accomplish these objectives, factors contributing to the tax compliance gap must be identified. The focus of research in the area of overt noncompliance should begin with identification of causal variables and then proceed toward an integration of those variables into a descriptive model of compliance/noncompliance.

The IRS is trying to encourage compliance and provide an equitable base for the tax system. Part of the IRS's solution to the problem of noncompliance lies in the area of detection. The frequency of audits in recent years has declined. In 1976, 2.59% of all taxpayers were audited. By 1982 the rate had dropped to 1.55%, and in 1983 only 1.36% of all taxpayers' returns were audited [IRS, 1983; Murray, 1984]. The trend is expected to continue, especially if the IRS budget cutbacks continue under the Reagan administration [Murray, 1984]. Increasing audit frequency would require a

larger staff and a substantial budgetary increase. Due to budget constraints, the IRS must use other methods to improve taxpayer compliance.

Roscoe L. Egger, Jr. [1983], a former Commissioner of the IRS, affirmed the Service's committment to closing the gap. Currently, the IRS is increasing use of computer technology as a detection tool [Adams, 1985]. Modern technology aids in matching reported earnings with other third party documentation. Detection of discrepancies in taxpayer reporting will permit collection of a greater portion of owed but unpaid taxes. Despite these technological efforts by the IRS, the tax gap appears to be widening.

The current approach appears somewhat myopic. Efforts to improve detection of tax evasion could be supplemented by other methods of increasing the level of compliance. Development of such methods requires a clearer understanding of the factors influencing noncompliance behavior. Studies are needed on the noncompliance effect(s) of tax rates, sanctions, attitudes, etc.

According to Jerome Kurtz [1977], former Commissioner of the IRS, U.S. citizenry appear to perceive current penalties for underreporting income and overstating deductions on tax returns as very low. Kurtz noted that IRS penalties are not high enough to deter noncompliance, "If a person is an economic being and figures out the odds, then there is a very high incentive to cheat." In other words, a

person who can determine the probability of being caught coupled with a low penalty will calculate the expected amount of gain to be obtained through evasion to be greater than the expected amount of the loss from detection and penalties. Although criminal sentences of up to 5 years in prison and $100,000 fines are possible, criminal cases made up less than 1/10 of one percent of total IRS investigations and audits in 1981, which themselves made up less than two percent of total returns filed [Mansfield, 1983, p. 218]. Penalties for the majority of tax noncompliance are surprisingly small. Table I contains a listing of these penalties and their maximum levels. For example, the penalty for substantial underpayment [Internal Revenue Code Section 6661] amounts to only 10% of understated income, plus interest. This penalty became effective with the Tax Equity and Fiscal Responsibility Act of 1982 [U.S. Congress, 1982]. During its first year, the IRS assessed only 149 underpayment penalties under this provision. Other offenders of this law paid only the outstanding taxes owed plus compound interest computed for late payment. The general results of this type of sanction prompted one IRS official to say, "We're still the best place in town to borrow from for many people." [Murray, 1984].

Harry Mansfield, Senior Tax Attorney at Ropes and Gray in Boston, discussed the problem of noncompliance [1983, p. 230]. He stated, "...no one really knows much about the relative deterrent effect upon taxpayers of criminal

TABLE I

MAJOR TAXPAYER PENALTIES

Type	Code Section	Description	Penalty
Criminal:			
Evasion	7201	Willful attempt to evade or defeat a tax	Felony--fine of $100,000; prison for 5 years
Failure to pay or report	7203	Willful failure to pay tax, file return, or keep records	Misdemenor--fine of $25,000; prison for 1 year
False statement	7206(1)	False return or statement	Felony--fine of $100,000; prison for 3 years
Civil:			
Fraud	6653(b)	Deficiency due to fraud	50% of deficiency, plus 50% of related interest
Negligence	6653(a)	Negligence or intentional disregard of rules	5% of deficiency, plus 50% of related interest
Delinquency	6651(a)(1)	Failure to file timely return	5% per month, up to 25%
Delinquency	6651(a)(2)	Failure to pay timely tax	1/2 of 1% per month, up to 25%
Valuation over-statement	6659	Valuation overstatement of property	Variable--10% to 30% of allocated tax deficiency
Substantial Under-statement	6661	Nonexcused underpayment exceeding 10% of correct tax	10% of allocated tax deficiency
Promoting abusive tax shelters	6700	False statement about tax benefits, or valuation overstatement	Greater of $2,000 or 20% of activity gross income

Source: Internal Revenue Code of 1954.

sanctions, civil penalties, and interest charges...we need that element of fear." Deterrence theory fits this idea.

Deterrence Theory

The deterrence proposition is that the perception of certain, swift, and severe sanctions will keep people from engaging in sanctionable behavior [Teevan, 1976]. Deterrence theory is a perceptual behavior theory based on a set of cues. These cues explain why individual persons conform to or deviate from norms, i.e. comply with or do not comply with a set of rules. A number of these cues or variables have been identified by theorists or researchers. The most important of these include: (1) sanction fear, (2) moral beliefs, (3) degree of integration of an individual into the social system, (4) extent of satisfaction with the social and political systems, (5) reinforcement value for the individual of various behaviors, (6) perceptions of the relative advantage or deprivation of one's life circumstances, and (7) various psychological conditions [Tittle and Logan, 1973; Tittle, 1976].

Through the middle 1970s, the most popular focus of researchers had been on motivational components, while interest in constraint as a possible factor toward achieving compliance had been minimal [Goode, 1972; Arvey and Ivancevich, 1980]. In fact, many researchers have been highly skeptical that sanction threats, or even sanctions, induce complying behavior [Salem and Bowers, 1970; Bowers

and Salem, 1971; Waldo and Chiricos, 1972].

Much of the criticism of the sanction threat cue was due to the results of dichotomizing sanction threat into a cue that indicates the individual's probability of noncomplying behavior being detected and a cue which explains the nature and severity of the sanction (penalty). Evidence was presented to suggest that probability of detection has a deterrent effect on noncompliance, while the perceived severity of penalties does not [Waldo and Chiricos, 1972; Bailey and Lott, 1976; Silberman, 1976; Teevan, 1976; Meier and Johnson, 1977].

Apparently, there is a strong interaction between probability of detection and severity of penalty. If the probability of detection is perceived high, the penalty magnitude can be relatively low, and noncompliance will be deterred. Thus, it has been presumed that if the probability of detection is low, the perceived penalty must be high to achieve the same level of deterrence, ceteris paribus [Grasmick and Bryjak, 1980; Scott and Grasmick, 1981]. Many of the critics were focusing on an element of the second interactive proposition, i.e. the issue that a low (or zero) probability of detection will not deter a potential malefactor from noncompliance. But all of the cues and interactions influence an individual's perceptions of the consequences of noncomplying behavior [Saltzman et al., 1982; Minor and Harry, 1982; Paternoster et al., 1983]. These consequences enter into an individual's decision on

whether to comply or not. If the balance between the results of various attitudinal behavior cues and the interactive sanction fear cues outweighs the gain to be derived from noncompliance, deterrence will be achieved. The necessary cues must be in place for noncomplying behavior to be deterred.

Purpose of the Study

The primary purpose of this study was to determine whether relatively higher magnitudes of penalties have a greater deterrent on noncompliance, given a relatively low, constant perceived probability of detection. A laboratory experiment using college students as subjects self-grading and self-assessing exam scores following a set of rules was conducted to attempt to isolate the impact of the factor--magnitude of penalty--in the model of noncompliance.

This study differs from those that have preceded it in several ways. In one way, the use of a laboratory experiment method is relatively new to deterrence studies. For example, Grasmick and Bryjak [1980] and Scott and Grasmick [1981] utilized subject-self-reporting, survey techniques. Also, the experimental task in this study was real, not hypothetical. Thus experimental validity was enhanced. The subjects were truly involved in the task due to an inherent reward structure, the course grade.

Another difference from prior works is that this study utilized a sanction environment analagous to the sanction

environment of income tax compliance. A low (non-zero) probability of detection existed. Much previous research in deterrence has focused on subjects confronted with a high certainty of apprehension [e.g. Grasmick and Bryjak, 1980]. Also, Scott and Grasmick [1981] utilized zero probability of apprehension situations in addition to high certainty levels of apprehension. Although Jackson and Jones [1985] tested penalty magnitudes at low detection levels, their results were inconclusive. They suggested further investigation into the importance of penalties as sanction threats in an income tax compliance environment.

To test the effects of an honor system on cheating, Canning [1956] and Ackerman [1971] used somewhat similar self-grading procedures as used in this study. Canning utilized a pre-test, post-test experimental design. Experimental variables were not intentionally manipulated by the researcher. His concern was with the change in the incidence of cheating after the implementation of an honor system at a university. Ackerman [1971] did not use a direct experimental method. He used a questionnaire asking students if they had cheated while self-grading examinations. In the present study, variables, such as stated penalty magnitudes, were intentionally manipulated within the experiment to study the change in compliance.

Both Canning and Ackerman allowed the classroom instructors to conduct their respective experiments or collect the data. Some bias may have been present in their

data due to the presence of a correspondent condition [Friedland, Thibaut, and Walker, 1973]. This condition results when student-subjects perform or comply for their teachers. In the present study, a person separate from the instructors directed the experimental portion of the study in an attempt to eliminate some of the correspondent condition bias.

Two distinct phases of this study were conducted during two consecutive academic semesters. In Phase One all experimental subjects received the same low, constant probability of detection. Penalty magnitudes were varied to test their effect on noncompliance. Phase Two of the study was conducted during a semester subsequent to the conducting of Phase One. In Phase Two of this study, using the same general methodology, the interactive effect of relatively higher magnitudes of penalties and relatively higher probabilities of detection was observed on noncompliance. Results obtained from the use of the varied levels of the two sanction variables should enhance and provide further definition to the general theory of deterrence.

Organization of Remaining Chapters

Chapter II provides a review of the relevant tax compliance/noncompliance and behavioral experimentation literature. Chapter III details the methodology employed in the study. Chapter IV offers an analysis of the results of Phase One of the study, and Chapter V provides an analysis

of the results of Phase Two of the study. Chapter VI contains a summary and conclusions of the study.

CHAPTER II

REVIEW OF THE LITERATURE

Introduction

In recent years, a number of researchers have addressed the issue of the factors that influence noncompliance. Many varied research methods have been utilized. Some authors have explored the relationship between such factors as probability of detection, magnitude of penalties, economic factors, fairness, social acceptability, perceived government inefficiencies, demographics, and legal complexities. This literature review will present findings of compliance research categorized by the general research method. Results from analytical models will be described first. That section will be followed by a discussion of empirical studies of factors affecting noncompliance. The third and final section will describe results of behavioral studies.

Analytical Models of Noncompliance

The first attempt at characterizing the evasion (noncompliance) problem in a theoretical/analytical context was done by Allingham and Sandmo [1972]. The individual taxpayer is assumed to make decisions on compliance by

maximizing his/her expected utility for personal disposable income plus gains. Two key relationships were developed: the dollar amount of tax evasion is inversely related (1) to the probability of detection, and (2) to the magnitude of penalties on owed but unpaid taxes. The analysis was extended to include social status as an argument of the von Neumann-Morgenstern utility function. If social standing has a positive utility to a taxpayer and is adversely affected by detection of evasion, tax evasion would tend to decrease due to the possible loss of social standing. This implies that social standing has a substitution effect for income. Extensions and refinements of the Allingham and Sandmo model have been produced by Kolm [1973], Srinivasan [1973], Yitzhaki [1974], and Witte and Woodbury [1985].

Models such as these allow researchers to examine the effect(s) of tax rates, probability of detection, and penalties, on amount of unreported income. In these models the utility function is assumed to be strictly concave, i.e., $U' > 0$ and $U'' < 0$. The individual is assumed to be risk averse. The measure of the degree of risk aversion is the Arrow-Pratt measure ($R = - [U''/U']$). R is larger for more risk averse individuals and is often assumed to be a decreasing function of income. Therefore, an increase in the penalty rate or the probability of detection will increase declared income. The effect of an increase in the tax rate on declared income is ambiguous in this model.

Kolm's note on the optimum evasion model [1973] uses Allingham and Sandmo's model from the government sector viewpoint. Allingham and Sandmo's model implies that the penalty rate and the probability of detection are substitutes for each other, and they assume a risk averse utility function. But Kolm notes that utility is separable in private and public goods. His results indicate that the usual public good condition (i.e., the marginal rate of substitution between public and private goods equals the marginal rate of transformation) is modified to allow for the effects of tax rate on amount of reported income. In such a scenario the optimal tax rate depends on the tradeoff between public and private goods.

Srinivasan's model [1973] indicates that the optimal level of enforcement will depend on how the probability of detection varies with the amount spent on detection and how government revenue (tax revenue plus penalties) changes with the probability of detection. Srinivasan found that evasion will decrease as the probability of detection increases. This finding is due to his model specification that the level of enforcement should be set higher if the probability of detection rises more rapidly with enforcement expenditure and if government revenue increases rapidly as the probability of detection rises.

Yitzhaki's analysis [1974] shows that if a penalty is imposed on the evaded tax (as is the case in the United States), there are no contradictory effects. That is, in

Allingham and Sandmo's previously-mentioned model the assumption existed that the taxpayer should pay a fixed penalty rate on the undeclared income. This assumption leads to the conclusion that when the tax rate increases there will be two opposing effects, an income and a substitution effect. The substitution effect encourages underreporting because a higher tax rate makes it more profitable to underreport income. In contrast, the income effect discourages underreporting because a higher tax rate reduces net income, causing individuals to become more risk averse. In Yitzhaki's model there is no substitution effect. Only an income effect is present. Therefore, as the tax rate increases, underreported income decreases.

The previous models fail to reflect actual tax structures. For example the models assume tax rates that are not progressive but are applied at a fixed rate to all income. Witte and Woodbury [1985] take a step forward in their model development. Their model reflects the tax laws and tax administration policies of the United States (during the 1960s and 1970s). A progressive tax structure is assumed and the taxpayer is faced with three possible IRS actions--audit, civil penalty, and criminal sanction. Also, it is assumed that the penalty for noncompliance depends on the amount of tax underpayment and that the absolute value of the penalties increase with the extent of IRS action. The authors assumed, as have the previously-mentioned analyses, that the taxpayer is risk averse.

Witte and Woodbury empirically tested their model using data relating to 1969 tax returns that the IRS released for outside research during 1981. The data set contained such variables as audit rates, prosecution rates, seriousness of sentences imposed for conviction of criminal tax fraud, socio-economic, and demographic factors. The data were aggregated to the three-digit zip code. The research results indicated that for some taxpayer classes the probability of civil fraud penalty is negatively related to voluntary compliance. Additionally, the results indicated that the measure of probability of criminal sanctions is not significantly related to compliance. These findings are contrary to the model specification (i.e., increased penalties increase compliance) and to the general theory of deterrence. These results indicate that more research is necessary to validate deterrence theory in the context of United States tax law.

Empirical Studies on Factors Affecting Noncompliance

In an early empirical work Schwartz and Orleans [1967], with IRS cooperation were able to relate degree of tax compliance to sanction threat and conscience appeal. Subjects were randomly assigned, after selection from census data, to treatment and control groups. Prior to submission of 1962 tax returns, one group was subjected to an interview containing questions that suggested the possibility of

sanction for dishonesty in reporting income. Another group was asked questions designed to remind them of their moral obligations, while a third group was interviewed but asked no conscience appeal or sanction threat questions. A fourth group was not interviewed. The IRS then supplied adjusted gross income, tax deduction, and tax payment figures for the experimental and control groups as a whole for 1961 and 1962.

The results showed that both sanction threat and conscience appeal could induce greater conformity (both treatment groups had significantly higher reported income than the control groups), but conscience appeal was found to be more effective. The degree of effectiveness of each of the inducements was found to vary by social characteristics of the respondents, particularly socio-economic status. It was further discovered that sanction threat apparently generated, among a minority of subjects (35%), attempts to make up through greater deduction what they had lost in an honest reporting of income. Thus, despite the fact that the independent variables were not really threats or conscience appeal, and there was no measurement of the perceived reality or perceived probability of imposition of the sanction, the study strongly suggests that reminding individuals of the possibility of negative sanctions does help secure conformity. But it seems that bringing to mind possible sanctions may not be as effective in achieving compliance with norms as bringing to mind other things.

Clotfelter [1983] measured the effect of marginal tax rates and other factors on compliance. He was employed by the Office of Tax Administration of the Treasury Department. His data was obtained from the Taxpayer Compliance Management Program (TCMP) program for 1969 from approximately 47,000 tax returns. The author acknowledged that variables related to IRS activity such as audit rates, penalty rates, and the level of sanctions must influence taxpayer reporting behavior, but he left them out of his analysis because of the simultaneity problem (described earlier). Clotfelter's results indicate that tax compliance increases with age and the proportion of income derived from wages and salaries. Compliance was lower for those audited in more recent years, and evidence indicated that marginal tax rates had a positive relationship to evasion.

Behavioral Studies on Factors Affecting Noncompliance

Since most researchers in the field of tax compliance do not have access to an adequate data base from which to apply questions, behavioral methodologies have become widely used in compliance factors studies. The general approaches to the behavioral studies have taken three directions--surveys, simulations, and laboratory experiments.

To date, surveys have been the most widely used. For example Mason and Calvin [1978, 1984] surveyed about 800

Oregon households to examine demographic factors, the motivations relating to noncompliance, the level of noncompliance, the loss of revenue to the Oregon Department of Revenue because of noncompliance, and public confidence in the tax system. They found that lower and higher income groups had the greatest amount of unreported income and that the middle income group had the least. Underreporting income was more likely among younger persons, those with more education, employed persons and students, the self-employed, and newer residents. The strongest motivation for noncompliance was the low perceived probability of being caught. Further analysis summarized in their 1984 paper was concluded with the indication that both taxpayers who are satisfied and those who are dissatisfied with the tax system remain honest because they are afraid of being caught.

Spicer and Lundstedt [1976] used survey data to construct two indexes--a tax evasion index and a tax resistence index--to test their premise that the decision to evade is dependent not only upon perceived penalties, but on a set of attitudes and norms. They found that perceptions of tax inequity, the number of evaders known to the respondents personally, and previous experience with tax audits were all associated with a higher level of admitted tax evasion. They also found that the propensity to evade declined with age, income level, and perceived probability of detection. In addition they found that the propensity to

evade increased as the proportion of income in wages, salaries, or pensions increased.

Song and Yarbough [1978] constructed an index of tax ethics on the basis of a 1975 survey of taxpayers in eastern North Carolina. It was concluded that the most important factor governing tax compliance was fear of detection. In general, they also concluded that married persons and homeowners had a higher level of tax ethics than single persons or renters. A higher level of income and a higher level of education were also related to good tax ethics. (This finding seemingly contradicts Mason and Calvin's [1978] result: Underreporting income was more likely among those with more education.) Tax ethics were found to be worse among those people who believed that tax evasion by others was common and among people who felt alienated, powerless, and distrustful.

Lewis [1979] analyzed a 200-male taxpayer survey in Bath, U.K., during the summer of 1977. The measurement instrument consisted of 16 Likert attitude statements measured on 5-point scales. Factor analysis was applied. Lewis concluded that self-interest was the primary motivating force behind noncompliance. Also, the results indicated that people with higher incomes have less favorable attitudes toward income taxes.

Dean, Keenan, and Kenney [1980] examined attudes toward taxpaying by questioning a nonrandom sample of adults in Scotland. The respondents felt that people evade taxes

primarily because taxes are too high or unfair and for economic reasons. Most felt that evasion was neither good nor bad, that opportunities for reducing one's taxes by a small amount through evasion were widespread, and that people would try to reduce their taxes by at least a small amount if they were unlikely to be caught.

Grasmick and Bryjak [1980] suggested that a rational individual would not disregard their perceptions of the severity of legal penalty when confronted with a potential compliance decision. They relied on expected utility theory to consider the association of interactive factors of noncompliance, severity and probability of sanctions, in the context of individuals perceiving detection probability to be relatively high. The authors surveyed 400 Oklahoma City, Oklahoma, residents collecting information about the respondents' past involvement in eight illegal activities: theft of an item worth less than $20, theft of an item worth $20 or more, illegal gambling, purposely hurting someone, cheating on tax returns, littering, illegal fireworks use, and driving while under the influence of alcohol. A perceived certainty of arrest was measured in addition to a perceived severity of punishment. The data appeared to suggest that perceived severity of punishment, when properly defined, operationalized, and interacted with certainty of arrest, is as significant a deterrent as perceived certainty of detection and arrest.

Using the simulation approach Friedland, Maital, and

Rutenberg [1978] examined the impact of different probabilities of audit, amount of fines, and tax rates on the decision to evade taxes. Fifteen undergraduate psychology students at Tel-Aviv University served as subjects in the experiment. The experiment consisted of the subjects reporting an income and paying tax accordingly on a monthly basis for ten simulated months. The objective was for each subject to maximize his/her net income (gross income less fines). Fines were administered to randomly audited returns for evaded taxes. At the end of the experiment a small money prize was distributed in proportion to each subject's net income. Audit frequencies were maintained at the inverse of fine magnitude. Therefore the expected value of gains from evasion was zero. It was concluded that large fines tended to be more effective deterrents than frequent audits. Also, the authors concluded that even when audit frequencies are reduced proportionately, large fines appear to be more effective deterrents than small fines. This result was satisfying, but due to limitations of the study, it is not all-conclusive. Some of the limitations are that the sample size was very small, the probability of audit was an uncertainty and could have been a deterrent in its own right, the subjects had nothing real to lose in the experimental context, and that the authors feel that evasion is simply more acceptable in Mediterranean countries than in Anglo-Saxon countries. They recommended replication and

furthur study of the variables in question with regard to compliance.

Spicer and Becker [1980] using the same simulation methodology as Friedland, et al. [1978], examined the relationship between perceived inequities in the tax system. They used fifty-seven University of Colorado at Colorado Springs students as subjects. The sample was stratified using Spicer's evasion index [Spicer, 1974]. The authors concluded that those who perceived their tax rates to be higher than average evaded the highest proportion of their taxes. Those who perceived their tax rates to be lower than average evaded the lowest proportion of taxes. Gender was also a significant determinant of tax evasion, with males evading a larger proportion of taxes than females, all other factors being equal. This finding concurs with that reached by Friedland, et al. [1978] but is contrary to the findings of Canning [1956] and Ackerman [1971].

Canning [1956] and Ackerman [1971] tested the effects of an honor system on classroom cheating. They used a self-grading procedure for students taking examinations. Canning conducted his experiment annually at Brigham Young University from 1948 (one year before an honor system was established) through 1953. He concluded that rates of cheating declined 63 percent over this period. Additionally, he determined that the average magnitude of cheating decreased by 33 percent. Before the honor system, male students cheated out of proportion to their number in

the total group. After five years of the system, the male proportion was reduced and females cheated disproportionately. Ackerman [1971] explored the effects of self-grading as compared to a conventional procedure on test-score outcomes. He concluded the effect of the honor procedure did not significantly have an effect on the test performances of the experimental classes. The difference in cheating between males and females was not significant.

Tittle and Rowe [1973] focused on the comparison of the relative effects of a moral appeal and a sanction threat in deterring classroom cheating. The experimental task was self-grading a series of quizzes by students in sociology classes. The experiment demonstrated that the cheating could be deterred by the combination of a threat of detection and punishment. The moral appeal had no significant effect. The results are strongly supported by deterrence theory.

Laufer [1985] used a generic task of self-assessment and reporting to determine if complexity has a significant impact on compliance. Two forms of complexity were examined: computational complexity and rule complexity. Undergraduate students were used as subjects and graded their own examinations in an actual classroom setting. They were provided with instructions (rules) and forms for reporting their exam scores. The instructions and forms were constructed to capture two levels of computational complexity (simple and complex) and two levels of rule

complexity (objective and subjective). The grade in the course was based upon the self-reported scores of the subjects. The author determined an exam score for each student (unbeknown to the students) and compared it to the reported score. It was concluded that complexity does not appear to be a significant factor in affecting the overall level of compliance. The presence of a strong motivation to comply may have been the cause for a lack of significant overt noncompliance. It was further concluded that there appears to be a significant relationship between the degree of computational complexity and the variability between reported and actual scores. Frequency and degree of errors increase as complexity increases. The conclusion drawn from the study is that a change in tax law which decreased complexity would likely promote a decrease in variability of errors but not affect the aggregate level of overt noncompliance.

In a recent laboratory experiment Jackson and Jones [1985] measured the relative importance of the risk of detection and magnitude of monetary penalty in the tax evasion decision. Relying on prospect theory the authors posited that people do not differentiate accurately between small probabilities and focus instead on the magnitude of the gamble. A laboratory experiment was conducted using student-subjects at the University of Colorado at Boulder, the University of Oklahoma, and the University of Texas at Austin. The experiment was broken into two parts: a

noncontextual part where subjects were given questionnaires involving loss alternatives with small stated probabilities and a contextual part where subjects were given choices with regard to alternative detection/penalty structures in a tax evasion opportunity questionnaire. Both elements of the sanction structure were varied in the choice alternatives: choice one included a low probability of detection with the highest penalty; choice two contained a higher probability of detection with a lower penalty; choice three was neither sanction element. Results indicate that people prefer greater risk when the potential magnitude of loss is less. Also, the results suggest the taxpayer will be more sensitive to magnitude of a penalty than a probability of detection, and higher risk of detection was seen to be a weaker deterrent than increased penalties.

The authors suggested several extensions to this work; continued foundational research in tax evasion, utilization of other research designs in the evasion/compliance environment, and further research on the importance of the magnitude of penalties.

Summary

This chapter presented an overview of prior research in compliance. Analytical and empirical research have not conclusively validated the relative magnitude of the primary sanction threat variables of deterrence--penalty and probability of detection. Most research on the probability

of detection variable supports the theory that high probabilities of detection deter noncompliance. When a penalty variable has been studied in compliance or deterrence research, conclusions have not definitively supported the theory of deterrence--higher penalties result in greater compliance. In tax compliance research specifically, the importance of the magnitude of penalties as a sanction threat is still questioned. Therefore, this study has manipulated both primary sanction threat variables to observe their effects on compliance.

The current study has utilized experimental generic task methodology to directly measure compliance/ noncompliance while varying an independent variable, penalties. The frequency of audit variable was fixed at a low discrete level in Phase One of this study to allow the controlled observation of the effect of varied magnitudes of penalties on compliance to a set of rules. In Phase Two of this study both the audit variable and the penalty variable were varied to observe their interactive effect on compliance.

CHAPTER III

DESIGN OF THE STUDY

Introduction

This study relies upon deterrence theory. A laboratory experiment was conducted in two phases. The first phase was to determine whether different magnitudes of penalties will affect the aggregate amount of compliance to a set of rules for self-grading examinations in an actual classroom setting with the frequency of audit held constant at a low level. Phase One was conducted during one complete semester of an academic year. The second phase of the experiment was to determine whether different magnitudes of penalties will affect the aggregate amount of compliance to a similar set of rules in a similar setting with the frequency of audit manipulated. Phase Two of the experiment was conducted during the academic semester following Phase One.

This chapter will explain the experimental testing by describing the hypotheses that were tested, the experimental design, and the specific structure of the experiment.

Statement of Hypothesis

Deterrence theory predicts that with a higher penalty a greater proportion of a group should comply. A low

probability of detection (audit) will allow the potential cheater to focus on the magnitude of penalty that he or she must then compare against the possible gain from the cheating. If the perceived penalty is severe enough, the potential cheater may be deterred. Therefore, the null hypothesis tested in the first phase of the experiment is:

H_{o1}: There will be no difference in the mean of noncompliance groups receiving relatively large penalties and groups receiving relatively small penalties given a constant low level of audit frequency for all groups.

The alternative hypothesis is:

H_{a1}: The mean of noncompliance will be less for groups with relatively large penalties than for groups with relatively small penalties given a constant low audit frequency for all groups.

The theory of deterrence predicts that as the probability of detection increases the magnitude of penalties becomes less important as a factor in the decision to comply or not to comply. Therefore, the null hypothesis that was tested in the second phase of this study is:

H_{o2}: There will be no difference in the mean of noncompliance groups receiving relatively large penalties and groups receiving relatively small penalties given a relatively larger level of audit frequency for all groups.

The alternative hypothesis is:

H_{alt}: The mean of noncompliance will be less for groups with relatively large penalties than for groups with relatively small penalties given a relatively larger level of audit frequency for all groups.

Design of the Laboratory Study

The ideal experiment to study the tax compliance effects of different penalty structures would require an enormous sample. Ideally, groups of taxpayers in demographically matched populations would be subjected to different penalty structures under the Internal Revenue Code, followed by the measurement and comparison of the levels of noncompliance among the groups over a period of several years. Such an immense study would be unmanageable, inequitable, and would violate current statutes on confidentiality. Furthermore, actual levels of compliance for items not subject to third party reporting would still be unobservable in this scenario.

Laboratory experimentation offers several advantages over field study. The researcher can mitigate potentially confounding effects of many extraneous variables. The independent variable can be manipulated with relative ease. By careful design of the research, a cause-effect relationship can be established between the independent and dependent variables [Swieringa and Weick, 1982].

This study examined compliance to a set of rules in a generic task laboratory setting. Generic tasks are used in laboratory studies of real world tasks which cannot be practically duplicated. A generic task is composed of common everyday behaviors. The task chosen in a laboratory experiment must clearly capture the essential interrelationships between the key real world variables. Birnberg and Nath [1968, pp. 44-45] stress the importance of having both explicit rewards and implicit rewards within an experiment to strengthen internal validity. The internal validity of a generic task is dramatically affected by the reward structure. It is important that the reward structure inherent in the generic task have valences similar to that of the real world task. Actual levels of compliance were observed in this experiment.

The experimental task was the self-grading, self-assessing of scores on course examinations by students enrolled in introductory accounting courses. The task is an extension of the experimental approaches used by Canning [1956], Tittle and Rowe [1973], and Laufer [1985]. Although the task is similar to that employed by other researchers, it was used in this study for the purpose of measuring the relative effects of different magnitudes of penalties and different probabilities of detection on classroom cheating. The task allows measurement of compliance to a set of grading and reporting rules. The inherent reward structure has similar valences to the real world task of interest. In

this experiment, the student has graded his/her own examination. The student has the incentive to complete the task since his/her reported scores earns a portion of the final course grade. In the real world task, the taxpayer can potentially increase his/her disposable income by noncompliance. The student can potentially increase his/her actual exam score by violation of grading rules. A student who cheats runs the risk of a loss in grade if audited.

The study used undergraduate students as surrogates for the U.S. taxpaying public. The context of the experiment is a decision making task. The subjects were required to grade their examinations by making decisions in following a set of rules. The subject decided whether to cheat or not to cheat.

There has been a considerable amount of research on the use of students as subject surrogates. (See, for example, [Copeland, Francia, Strawser, 1973], [Abdel-khalik, 1974], [Ashton and Kramer, 1980], and [Krogstad, Ettenson, and Shanteau, 1984]). There has not been a concensus as to the overall research effect of students as surrogates. However, explicit consideration must be given to the appropriateness of students for the task.

Ashton and Kramer [1980] used student subjects as surrogates in a study of auditors' internal control judgements. They concluded that students were acceptable surrogates for auditors making similar framed decisions. The authors generalized their conclusion to situations where

students process information and make decisions. The experimental task in their study was an information processing and decision making task.

The students in this study were informed that a monitor would prepare and administer the midterm exams. The regular classroom instructor did not appear to be directly involved in the examination and grading process. Since the total classroom environment is a surrogate for the taxpaying environment, a reasonable analogy to the noncorrespondent relationship of taxpayers to the IRS was created. The classroom instructor was in a correspondent condition with the students and could have been a biasing factor if allowed to administer the exams and the rules for compliance. Friedland, Thibaut, and Walker [1973] based on an experiment using 96 undergraduate students playing a management game, concluded that a significantly higher level of compliance to a legislator's or enforcer's rules resulted when a strong correspondent condition was present. Laufer's [1985] results appeared to be biased by the correspondent condition. The fact that the researcher served as both the instructor and the exam monitor may have contributed to the higher than expected degree of compliance. Official communication from the monitor to students was handled via the U.S. mail. This communication procedure helped mitigate the correspondent condition.

The proposed experiment began on the first day of class of a semester. The instructor distributed the course

syllabus and an information sheet which was completed by the students. The information sheet contained data such as address, schedule of classes, and grade point average. The instructor informed the students midterm exams will be self-graded. The students were told a monitor would prepare and administer the exams. The monitor was introduced, distributed an "Examination Procedures" document (Appendix A), and explained the self-grading procedures and the general rules for compliance to the students.

During the experiment, the students' semester letter grade was determined from four midterm exams and one comprehensive final. The midterm exams were graded by the students, and these self-reported scores were used toward grade determination. The midterm scores were weighted 66.67% of the semester grade. The final exam, worth 33.33% of the course grade, was graded by the instructor. Semester letter grades, A, B, C, D, and F, were awarded on the basis of the traditional 90%, 80%, 70%, 60%, and below 60%, respectively, of total points possible (600) on five exams. To control for the possible effects of other reward variables on the decision to cheat or not, students were informed no other elements, such as quizzes or homework assignments, were used as determinants of the final course grade.

The first exam was be used as a training run to acquaint the students with the actual procedures for self-grading. It was hoped that this practice would

eliminate some of the bias caused by inexperience and uncertainty with what is expected. Additionally, it was hoped that it would strengthen the analogy of students in a compliance setting with experienced taxpayers in a compliance setting. Krogstead, Ettenson, and Shanteau [1984] have shown that experience in task performance affects reliability of decisions. Students were informed via grading instructions that the first self-graded midterm exam would also be 100 percent investigator-graded. No risk of penalty for grading errors was incurred by the students. If there was a substantial difference between any student's reported and actual score, he/she met with the monitor to discuss the difficulty. The investigator-graded score was recorded for exam one. The experimental penalty treatments (and the experimental audit treatments in the second phase of the experiment) were applied to the three subsequent midterm examinations.

Examinations were administered in a traditional classroom setting. The exams consisted of objective style questions and problems. Upon completion of the exam, the students submitted their exam papers and answer sheets to the monitor. During the time period from the end of the examination period until beginning of the next class meeting, all answer sheets were photocopied. The photocopied answer sheets were used by the researcher to determine a correct score. The students were not aware of the copying procedure. The exam papers were returned to the

students in the class period following the administration of the examination. Attached to the student's exam was a copy of each of the following: Grade Report Form Instructions, Grade Report Form, and Answer Key with explanations. (See Appendix B for examples).

The Grade Report Form Instructions were the primary experimental instrument. Within that instrument, statements were made on the penalties for failure to follow the rules for self-grading. The penalty for noncompliance, negligence, or fraud was the independent variable of Phase One of the experiment and was varied across randomly selected subject groups. The penalties varied from zero, to two times the difference between reported and actual scores, to ten times the difference. The largest magnitude was intended to appear relatively severe to a potential cheater, but there has not been any evidence presented in the scientific literature establishing definite penalties that deter noncompliance.

The percentage of exams audited remained constant at 4% throughout Phase One of the experimental stage of the study. A varied percentage may be a deterrent itself. Therefore, the second phase of the study was conducted to observe the interactive effect of varied percentages of audits and varied penalty magnitudes. The percentage of exams audited was varied between 12% and 24% during Phase Two. The Grade Report Form Instructions stated that a discrete number of exams and Grade Report Forms would be audited. This number

will correspond to a certain percentage of the class size, and was adjusted downward as student withdrawals required. The computation of this number was demonstrated on the Grade Report Form Instructions. Tversky and Kahnemen [1971] have suggested that students have difficulty in assessing probabilities of event occurrences. Percentage figures alone could possibly bias the treatment variables. The combination of percentage and discrete number information gave the subjects a fairly concrete perception of their audit potential.

A manipulation check [Sudman and Bradburn, 1982, p.247] was incorporated into the Grade Report Form to help insure that the student subjects were aware of their penalty treatments and the percentage of exams to be audited. The subjects were asked to write their penalty rate and the percentage of exams to be audited, as stated in the Grade Report Form Instructions, on the Grade Report Form.

The students were given approximately one week to perform the grading task. The due date was stated in the Grade Report Form Instructions. Students submitted their Grade Report Forms to the monitor at the beginning of class in their usual room at their regular class time. Students were not required to prepare other documentation. If students were not timely in submitting their forms, penalties were assessed as per the instruction instrument. Students were notified by mail if such a condition applied. The exam scores as reported on the Grade Report Forms were

compared to the actual score computed by the researcher.

To reinforce the probability of detection, audits were actually conducted. Exams were randomly chosen for audit. (The amount of cheating as indicated by the difference between reported and actual scores had no bearing on the audit selection). The term 'random' was not used to describe this procedure to the students. Although a random selection technique was used, the wording was important to convey the desired meaning. Slovic, Kunreuther, and White [1974, p. 203] have indicated that individuals have a difficult time in rationalizing the idea of randomness. Students were told that exams would be selected for audit by drawing names from a hat.

When a student was randomly selected for an audit, she/he was notified by mail of the time and place of a meeting with the monitor. The student returned a confirmation card enclosed with the audit letter. The actual audit was conducted much like an IRS office audit. Actual answer sheets and complete exam papers were copied for use by the monitor during the audit. The audited student was asked to provide support for the way in which the exam was graded. The penalties applied for defalcations were stated in the students' Grade Report Form Instructions. The penalties were deducted from the experimenter-graded (true) score. Since the concern of the experiment is to observe deterrence of overt acts of noncompliance, inadvertant errors were given favorable status when

reasonable explanations were provided by the student. This objective may have been difficult to achieve, but was given consideration.

After completion of the audit phase of each examination, an announcement was made to the students by each class's instructor as to the general results of the audit. An example statement might be, "Eighteen students had their grading checked by the monitor and fourteen were found to have no discrepancies; four received penalties." This announcement reinforced the concept that audits potentially existed, and it was especially important for those students who were not directly affected by audits. The announcements in the experimental settings were analogous to newspaper, radio, and television reports of IRS activity.

Experimental Ethics

Keys and Hendricks [1984] define 'ethics' as the area of intentional human behavior that affects the well-being of others and can be categorized as right or wrong. In the academic research environment ethics may refer to a set of moral principles which govern the conduct of research. Keys and Hendricks discuss the ethical researcher's obligations to protect the subject's right to privacy and his right to receive results. They also discuss the roles of deception and debriefing in laboratory research.

The right to privacy includes obtaining the willing

consent of the subjects and maintaining the confidentiality of the data. In the current study the student-subjects were given the opportunity to opt out of the self-grading scheme or transfer sections with no stigma attached. Students were not coerced into remaining in the test group. Since each student reported his/her own exam results, the task should have been taken seriously.

The second requirement in insuring privacy includes maintaining confidentiality of collected data. In the current research, differences between a student's reported exam score and his/her actual exam score were available only in aggregate form. The instructors were apprised only of subjects' reported scores. This information was used for the purpose of intrasemester withdrawals and for issuing final semester letter grades.

By necessity, a degree of deception was incorporated into the experiment. Full disclosure of the procedure by which the researcher has a copy of each exam would extinguish the very behavior of interest. The deception factor was not used in a way that would jeopardize a student's semester grade.

After the study was completed each semester, a debriefing session was conducted. A questionnaire was used to determine whether the experimental task was performed as intended. Students were given the opportunity to ask questions about the study. The deception of copying all examination answer sheets was maintained after Phase One,

since that deception was necessary in Phase Two. After Phase Two was completed, the copying procedure was disclosed.

Students' rights to receive results was satisfied by the internal experimental design. The student knows his/her reported exam scores. She/he knows if any overt act of noncompliance was committed.

In compliance with University and College of Business policy, all necessary forms were timely filed by the researcher. These were duly accepted, giving approval for the use of human subjects in this laboratory experiment. Copies of the approval are on file in the Office of Business and Economic Research of Oklahoma State University. The aforementioned ethics issues were addressed within the approved documentation.

Results of a Pilot Study

A pilot study of the experiment was conducted during the summer semester of 1985 on one principles of accounting class section of 23 students. The primary purpose was to test the validity of the instruments and experimental treatment methods. On the basis of preliminary findings some minor modifications were made during the test period and/or for the subsequent proposed experiment.

During the pilot study the class instructor reported that the students were quite aware of the separation between instructor and monitor. Feedback in the form of expressions

and verbal responses indicated that the correspondent condition is strong with the instructor, but weak with the monitor. The monitor apparently was perceived as an adversary. The instructor was perceived as more of a friend by the students. The analogy of the monitor to the IRS, as far as separation is concerned, appears to have been in place.

The pilot study uncovered the fact that the students had difficulty grading exam problems where partial credit was allocated with some degree of subjectivity through sequential steps. Although detailed grading explanations were in the answer keys, students who were audited and informal conversations with students related the tendency to be more conservative and overcautious when scoring with subjective rules. The majority of both frequency and magnitude of cheating occurred on the objective exam questions. Apparently, the inexperience of making decisions when confronted with subjective rules dominated the decision of whether to overtly cheat. Objective question rules were understood by the subjects. The only decision left for the subjects was the cheating decision. To eliminate the potential effects on compliance that some subjective grading may cause, only objective style questions and grading rules were used (and were used on the two final midterm exams of the pilot study).

Students in the pilot study group appeared to be aware of the low, constant probability of audit. In conversations

with both the instructor and the monitor, students reported they were aware of the magnitude of penalty on various exams and the constant probability of audit. These observations are crucial, since they suggest the experimental treatments are being effectively applied. Conversations also revealed the students were not aware of the copying procedure of the exam answer sheets. Therefore, the students' decisions to comply or not to comply with the grading instructions were not biased by this potential deterrent.

After administration and grading of the fourth midterm exam, debriefings were conducted and an exit questionnaire was administered. Several personal interviews were also conducted to verify the structure of the test instruments and to uncover any other possible problems. No other problems were found.

Statistical Design

The indication of compliance to rules is the difference between the correct score, determined by the researcher, and the student's self-reported scores. Therefore this difference factor will be the measure used to determine noncompliance in the experiment.

A counterbalanced experimental design [Campbell and Stanley, 1963, pp. 50-52] was used. The specific arrangement was a Latin square. The general plan for applying treatments was as follows in Figure 1, although the arrangement of the treatments was varied for each class section.

Class Section One

First Exam Score	Exam #2	Exam #3	Exam #4
Top Group	X	Y	Z
Mid Group	Y	Z	X
Low Group	Z	X	Y

Figure 1. Latin Square Design

Each individual class section was subjected to one probability of random audit throughout the entire semester in which the experiment was conducted. In Phase One of the experiment ten class sections were utilized. All ten class sections were audited at the 4% level. In Phase Two, eight class sections were utilized; four class sections were audited at a level of 12% and four class sections were audited at a level of 24%.

The rows of the square represent the partitioning of a class by relative rank of scores earned on the first midterm exam of the semester. That is, the one-third of the students scoring highest on Exam I became one experimental unit (Top Group). The middle third of the class became a second experimental unit (Mid Group), and the bottom third became the third experimental unit (Low Group). Due to a high class drop rate, the initial actual treatment

applications were to partitioned groups comprising approximately 25 percent, 25 percent, and 50 percent of the students to Groups Top, Mid, and Low, respectively.

Partitioning the subjects by Exam I score aids in controlling for the effects of variables such as, opportunity to cheat and perceived difficulty of exam on the dependent variable, noncompliance. A difference in opportunity to cheat exists due to relatively different scores on examinations. For example, a score of 45 out of 100 points allows a student-subject greater opportunity than a score of 90 out of 100. Also, the incentive to cheat would apparently differ among groups. For example, a student earning a very high score honestly would have no incentive to cheat. A student earning a failing score might have a powerful incentive to cheat. The pilot study indicated that there are significant differences between score-partitioned group means of noncompliance.

The columns of the square represent the three different midterm exams upon which the treatments were applied. The partitioning by exam aided in controlling for a maturation effect and significant differences in perceived difficulty between exams. The letters X, Y, and Z represent the treatments (different magnitudes of penalty). X denotes a penalty of zero. Y indicates a penalty of two times the difference between actual and reported exam scores. Z indicates a penalty of ten times the difference between actual and reported exam scores.

ANOVA was used to test the hypothesis of no difference in treatment means in the Latin square design. To test all possible comparisons of the treatment means the Duncan multiple range test was applied [Steele and Torrie, 1980, p. 186]. The Duncan procedure controls the Type I comparisonwise error rate [SAS, 1982, p. 174]. By using ten replications of the 3 x 3 Latin square design adequate degrees of freedom are available to pass significance judgments on the statistical results of the Phase One experiment. Eight replications of the 3 x 3 Latin square design were used in Phase Two of the experiment. Four of the replicates were audited at the 12% level, and four of the replicates were audited at the 24% level. The replications are different class sections of the first Principles of Accounting course. The number of replications was constrained by the number of class sections offered by administration in a semester. The Phase Two design was a split, split-plot with sub-unit treatments in a Latin square [Cochran and Cox, 1957, pp. 306-311].

Summary

This chapter presented the design and procedures used in conducting the experiments. Also discussed were the results of a pilot study and statistical experimental designs utilized. External validity generally can not be achieved with certainty when laboratory experiment methodology is used. But the external validity of the

current experiment, in addition to the internal validity, was apparently enhanced through experimental realism. The specific generic task methodology used allowed for a high degree of experimental realism to be achieved. The task was real, not hypothetical. The task contained an inherent reward structure, the course grade. Also, the subject involvement in the experiment was similar to the decision making processes taxpayers go through when determining their tax liability.

A pilot study indicated that the experimental methodology employed was believable to the subjects and they took the task seriously. Also, it was apparent that subjective type exam questions and answers were a biasing factor in measuring noncompliance. Therefore, only objective type questions were used in the examinations during the experimental periods.

For the purpose of determining the impact of different magnitudes of penalties and different probabilities of detection on noncompliance, a Latin square experimental design was used. Internal validity of the experiment should be enhanced through use of this design. Also, this design allowed for statistical analysis of the effect of different magnitudes of penalties and different probabilities of detection on noncompliance using ANOVA procedures.

CHAPTER IV

ANALYSIS OF DATA FROM PHASE ONE

Introduction

In this chapter, the results of Phase One of the experiment are provided. The results of the experiment are analyzed and discussed. An interpretation of the results is presented.

Results of Analysis of Data From the Experiment

The objective of Phase One of this study was to determine whether relatively higher magnitudes of penalties have a greater deterrent on noncompliance, given a relatively low (4 percent), constant perceived probability of detection. Toward this end, an experiment was designed to allow for manipulation of independent variables so the experimental effect of the treatments could be measured. The dependent variable in the model was a measure of noncompliance, the difference between reported exam scores and actual exam scores. The independent variables within the model were (1) examinations, (2) partitioned groups of student subjects, and (3) the treatment, penalty magnitude.

Penalty magnitudes of zero, twice the difference between the actual and reported exam scores, and ten times that difference were applied. By using ten different class sections concurrently, the experiment was replicated ten times. A total of 483 students were enrolled in the classes. The results of the ANOVA for the Latin square arrangement are presented in Table II.

TABLE II

ANALYSIS OF VARIANCE STATISTICS--
FULL MODEL OF NONCOMPLIANCE
PHASE ONE

Source of Variation	Degrees of Freedom	F Ratio	Observed Significance Level
Penalty	2, 20	56.06	.0001
Groups	2, 20	12.95	.0002
Examinations	2, 20	5.10	.0162
Class Sections	9, 20	2.75	.0286
Section x Exams	18, 20	0.87	.6100
Section x Groups	18, 20	1.89	.0849
Section x Penalty	18, 20	1.73	.1175

Use of ANOVA with three levels of the penalty factor

resulted in a two-tailed test of the hypothesis, rather than the one-tailed test indicated in the hypothesis statement. Based upon the data collected and procedures utilized, the null hypothesis of no treatment effect is rejected because the F ratio of 56.06 has an observed significance level less than 1 percent with 2 and 20 degrees of freedom. The significant difference between the magnitude of penalty treatments indicates the level of noncompliance was affected by the magnitude of penalties--relatively lower noncompliance to rules resulted in conjunction with relatively greater statutory penalties.

The treatment effect (penalty) is interpreted to be independent. That is, the differences in the responses, the means of noncompliance, to the three treatments would be the same whether other factors are present or not. This conclusion is derived from Table II where the results of the test of the interaction between class sections and penalties is presented. The F ratio of 1.73 with 18 and 20 degrees of freedom has an observed significance level of 11.75 percent. The presence of an interaction is not greatly supported. Furthermore, the graph in Figure 2 (Appendix D) depicts the independence of the penalty magnitude treatments from the class section factor with regard to the mean of noncompliance. Independence is not conclusively supported by the graph. Some low level of interaction may be present. Therefore, Duncan's multiple range test was conducted to compare the penalty means separately for each class section.

The ranking of penalty means does not differ for any of the class sections. In short, the nature of the differences between penalty treatments does not significantly depend on which class section is investigated. The main effects of the treatment, penalties, are presumed to be independent.

ANOVA indicated the penalty means are different from each other but not which means differ from which other means. Table III displays the comparison of the penalty magnitude means of noncompliance.

TABLE III

DUNCAN'S MULTIPLE RANGE TEST--
PENALTY MAGNITUDES
PHASE ONE

Penalty Magnitude	Observations	Mean of Noncompliance	Test Grouping
Zero	30	9.14	A
2 x noncompliance	30	4.72	B
10 x noncompliance	30	1.96	C
Overall	90	5.27	

Based upon the Duncan procedure for multiple comparisons, the null hypothesis of equal penalty treatment means is rejected at the testing level of alpha equal to

five percent. The different letters (i.e., A, B, and C) show that the three penalty means are significantly different from each other at the five percent testing level for significant differences. These results validate a portion of the general theory of deterrence--relatively greater penalty levels deter noncompliance.

The experiment was designed to account for possible effects on noncompliance due to different subject groups, different examinations, different class sections (replications), and different penalty magnitudes. From Table II it can be seen that all of variables produce an effect on the dependent variable, noncompliance, at the significance level of five percent.

The three subject groups represent the partitioning of a class by relative rank of scores earned on the first midterm exam of the semester. That is, the one-third of the students scoring highest on Exam I became one experimental unit (Top Group). The middle third of the class became a second experimental unit (Mid Group), and the bottom third became the third experimental unit (Low Group). The results of this partitioning of the subjects by Exam I score indicates that factors such as, opportunity to cheat, perceived difficulty of exam, and the amount of incentive to cheat would apparently differ among groups on their effect on the dependent variable, noncompliance. This result was anticipated and is not construed to be abnormal. A difference in opportunity to cheat exists due to relatively

different scores on examinations. For example, a score of 45 out of 100 points allows a student-subject greater opportunity than a score of 90 out of 100. Also, the incentive to cheat apparently differs among groups. For example, a student earning a very high score honestly would have no incentive to cheat. A student earning a failing score might have a powerful incentive to cheat. The mean amounts of noncompliance relative to the partitioned groups and the results of the Duncan test are shown in Table IV.

TABLE IV

DUNCAN'S MULTIPLE RANGE TEST--
PARTITIONED GROUPS
PHASE ONE

Partitioned Subject Groups	Observations	Mean of Noncompliance	Test Grouping
Top Group	30	3.42	A
Mid Group	30	5.56	B
Low Group	30	6.84	B

The table indicates that the Mid and Low Groups are not significantly different from each other with regard to noncompliance. Both are significantly different from the Top Group. Apparently, the Top Group reflected their lower

incentive and opportunity to cheat through the significantly lower mean noncompliance i.e., less cheating. The Mid and Top Groups had more opportunity and more incentive to cheat.

The Examination variable represents the three different midterm exams upon which the treatments, penalties, were applied. Table II indicates examinations did produce an effect upon noncompliance. The F ratio of this effect, 5.10, with 2 and 20 degrees of freedom says that the effect is not significant at the one percent level; however, at the level of five percent, the effect due to different examinations is significant. Because of repeated testing, maturation, experience, and cumulative carryover, this variable could be expected to be significant. Table V presents the means of noncompliance relative to this independent variable, Examinations, and the results of the Duncan test.

TABLE V

DUNCAN'S MULTIPLE RANGE TEST--
EXAMINATIONS
PHASE ONE

Examinations	Observations	Mean of Noncompliance	Test Grouping
Exam 2	30	4.10	A
Exam 3	30	5.48	A B
Exam 4	30	6.24	B

Table V shows the increasing trend in the means of noncompliance relative to the examinations. Exams 2 and 4 are shown to be significantly different from each other but not from Exam 3. The maturation or experience factor seems to have a significant effect upon noncompliance at the level of alpha equal to five percent in this experiment.

TABLE VI

DUNCAN'S MULTIPLE RANGE TEST--
CLASS SECTIONS
PHASE ONE

Class Section	Teacher	Observations	Mean of Noncompliance	Test Grouping
7	2	9	7.715	A
4	1	9	6.638	A B
5	2	9	6.571	A B
1	3	9	6.068	A B
3	1	9	5.093	A B
10	3	9	4.879	A B
6	2	9	4.473	A B
9	3	9	4.213	A B
8	3	9	4.198	A B
2	1	9	2.874	B

The variable, Class Sections, represented the different class sections or replications on which the experiment was conducted. From Table II it can be seen this variable, with an observed significance level less than five percent, has a significant effect on the dependent variable, noncompliance. This result was unexpected. Since the same examinations were given in all ten classes and experimental procedures were administered in a like manner for all class sections, those factors should not have produced a significant effect. Table VI presents the mean amounts of noncompliance for each section and the results of the Duncan test for differences among the ten class sections. It shows that class sections 7 and 2 are significantly different from each other but not from the other eight sections at the level of alpha equal to five percent. Possibly factors such as peer pressure within a specific section, time of day, different class instructors, or other factors aided in producing the effect upon noncompliance.

Three different instructors covered the 10 class sections. Whether an instructor was liked (or disliked), perceived to be good (or bad), or perceived fair (or unfair) could have biased compliance to a set of rules and sanctions. Possibly, students who liked an instructor complied, in part, for the instructor. Conversely, it is possible that students may have had higher noncompliance, if the instructor was disliked. For some students, good or fair may have been the perception(s) that had biasing

effects upon compliance. Whether, or not, a specific instructor was liked may not have been relevant. An instructor may have been perceived fair or good, but was disliked. Compliance may have been enhanced by the perceptions of fairness or goodness. Conversely, students perceiving their instructor(s) to be unfair, or bad, may have felt a compulsion to cheat. These students may have felt disadvantaged. Cheating, or not following the rules, was a means to gain, or regain, perceived proper advantage in their class. Table VI shows that the two outlier class sections, 2 and 7, were taught by different instructors. There does not appear to be a discernible pattern of noncompliance by instructor.

Summary

Based on analyses of the results it appears that differing levels of the treatment, penalties, will effect noncompliance. However, differing levels of other independent variables also had significant effects on the independent variable. Subject groups partitioned by relative exam score, different examinations, and replications represented by different class sections indicated effects upon noncompliance.

These results are not surprising. Individuals reach decisions processing many, many factors. Even if all factors applicable to one decision setting could be identified, it is very difficult to isolate one factor from

all of the other relevant factors.

The following chapter will discuss the analysis of data of an experiment where an additional variable, audit frequency is addressed. Also, the possible interaction of treatment variables will be analyzed.

CHAPTER V

ANALYSIS OF DATA FROM PHASE TWO

Introduction

In this chapter, the results of the second phase of the experiment are provided. The results of this experiment are analyzed and discussed. An interpretation of the results is presented.

Results of Analysis of Data From the Experiment

The objective of Phase Two of this study was to determine whether relatively higher magnitudes of penalties have a greater deterrent on noncompliance, given a relatively greater perceived probability of detection. Toward this end, the experiment conducted as Phase One was modified to include an additional independent variable, audit levels. Other independent variables within the model were (1) partitioned groups of student subjects, (2) examinations, (3) class sections, and (4) penalty magnitude. The penalty magnitudes applied were zero, twice the difference between the actual and reported exam scores, and

ten times that difference. Eight class sections contained the experimental units on which the experiment was conducted. A total of 451 students were enrolled in these sections. Four of the sections had an audit level of 12 percent applied, and four sections received an audit level of 24 percent. The results of the ANOVA are presented in Table VII.

TABLE VII

ANALYSIS OF VARIANCE STATISTICS--
MODEL OF NONCOMPLIANCE
PHASE TWO

Source of Variation	Degrees of Freedom	F Ratio	Observed Significance Level
Audit	1, 6	16.52	0.0066
Penalty	2, 52	34.08	0.0001
Audit x Penalty	2, 52	0.36	0.6989
Examinations	2, 52	1.65	0.2027
Audit x Examinations	2, 52	0.82	0.4449
Groups	2, 52	29.60	0.0001
Audit x Groups	2, 52	5.17	0.0090

ANOVA is used to test the independence of the treatment effect, penalty magnitude. The test examines whether a tradeoff exists for subjects between penalty magnitude and the level of audit. That is, the test examines whether the relative effect of the penalty treatment on noncompliance would be the same regardless of the level of audit. The Penalty x Audit interaction is the measure of independence. Since the observed significance level of the interaction is 0.6989, apparently the interaction is not present. A comparison of Penalty means is not dependent on the level of Audit in this experiment. This result is surprising. Grasmick and Bryjak [1980] hypothesized that rational individuals who perceive the certainty of apprehension as high will be influenced by the seriousness of punishment if apprehended. Whereas, regardless of the perceived consequence of being caught, individuals do not regard it as a potential cost, if they believe they will not be caught. The results of these authors' analyses supported their prediction of the interactive effect between perceived severity of penalty and perceived certainty of apprehension in their study. The means of noncompliance to penalty level by audit level for the present study are shown in Table VIII. The pattern of the response variable means is as anticipated from the general theory of deterrence. The mean of noncompliance decreased as both the penalty magnitude and audit frequency increased, but statistically the interactive

relationship is not significant.

TABLE VIII

MEANS OF NONCOMPLIANCE--
PENALTY BY AUDIT
PHASE TWO

Penalty Magnitude	Audit Frequency 12%	24%
Zero	9.934	6.259
2 x noncompliance	5.475	2.888
10 x noncompliance	3.654	0.991

Responses to the post-experimental questionnaire indicate the subjects in this experiment appeared to be aware of the audit and penalty levels and the interaction of these variables. That is, the subjects reflected that the importance of the penalty level with regard to noncompliance depended on the level of audit. There apparantly was a trade-off taking place. Fifty-nine percent of the subjects indicated that larger-sized statutory penalties caused them to refrain from cheating. Also, 80 percent of the subjects stated that their probability of being audited increased their compliance to the rules and instructions. Fifty-seven percent of the subjects indicated that if their probability

of audit was close to zero they would not be very concerned about the size of the possible penalty. Whereas, 71 percent of the subjects stated that at a probability of audit near 100 percent they would be very concerned with the size of the possible penalty. Although the subjects were aware of the penalty and audit relationship, based on these responses, it appears that many other factors were present in the individuals' decision processes in this experiment. In a compliance or noncompliance setting the significance of the interaction of factors such as audit and penalty levels may be lessened due to many of these other factors. Some of these factors may be maturation, experience, peer pressure, or celerity in detection or penalty imposition.

Since the interaction was not present, it is possible to make an overall comparison of Penalty treatments averaged over all levels of Audit frequency. Use of ANOVA given three levels of the Penalty factor resulted in a two-tailed test of the hypothesis. Based upon the data collected and procedures utilized, the null hypothesis of no penalty treatment effect is rejected because the F ratio of 34.08 has an observed significance level of 0.0001 with 2 and 52 degrees of freedom. The significant difference between the magnitude of penalty treatments indicates the level of noncompliance was affected by the magnitude of penalties. Table IX displays the average magnitude of noncompliance by the three penalty magnitudes and the results of Duncan's Multiple Range Test.

Based upon the Duncan procedure for multiple comparisons, the null hypothesis of equal penalty treatment means is rejected at the testing level of alpha equal to five percent. The average amount of noncompliance for each penalty magnitude was significantly different from each of the others in the Phase Two test. This result, as with Phase One's result, validates a portion of the general theory of deterrence--relatively greater penalty levels deter noncompliance. Individuals are apparently deterred from cheating as the statutory penalty level is increased in this experimental setting.

TABLE IX

DUNCAN'S MULTIPLE RANGE TEST--
PENALTY MAGNITUDES
PHASE TWO

Penalty Magnitude	Observations	Mean of Noncompliance	Test Grouping
Zero	24	8.097	A
2 x noncompliance	24	4.182	B
10 x noncompliance	24	2.323	C
Overall	72	4.865	

The second treatment variable of interest in Phase Two

of the experiment is audit frequency. This variable had two levels, 12 percent and 24 percent. (Phase One utilized a constant audit level of four percent). As shown in Table VII, this variable, audit, is significant at a level of 0.0066. The means for the two levels of audit are shown in Table X.

TABLE X

DUNCAN'S MULTIPLE RANGE TEST--
AUDIT FREQUENCY
PHASE TWO

Audit Frequency	Observations	Mean of Noncompliance	Test Grouping
12 percent	36	6.355	A
24 percent	36	3.379	B

The means of the two audit frequency levels are significantly different at the level of alpha equal to five percent. As the audit frequency increased, the mean of noncompliance decreased. Individuals were apparently deterred from cheating as their probabilities of detection increased in this experiment.

An F ratio of 29.60 with an observed significance level Of 0.0001 shows that the variable Groups has a significant

effect on the response variable, noncompliance. The results of the partitioned Groups of the subjects by Exam I score indicates that variables such as, opportunity to cheat, perceived difficulty of exam, and the amount of incentive to cheat would apparently differ among groups on their effect on the dependent variable, noncompliance. This result was anticipated and is not construed to be abnormal. The mean amounts of noncompliance relative to the partitioned groups and the results of the Duncan test are shown in Table XI.

TABLE XI

DUNCAN'S MULTIPLE RANGE TEST--
PARTITIONED GROUPS
PHASE TWO

Partitioned Groups of Subjects	Observations	Mean of Noncompliance	Test Grouping
Top Group	24	1.998	A
Mid Group	24	5.129	B
Low Group	24	7.474	C

Table XI indicates that the three groups are significantly different at the level of alpha equal to five percent. Apparently, the Top Group reflected their lower incentive and opportunity to cheat through the significantly

lower mean noncompliance i.e., less cheating. The Mid and Low Groups had more opportunity and more incentive to cheat. This result is similar that of Phase One, with one exception. In the present test all three means are significantly different. In Phase One the Mid and Low Groups were significantly different from the Top Group but not from each other. There is not a definitive explanation for this difference.

TABLE XII

DUNCAN'S MULTIPLE RANGE TEST--
EXAMINATIONS
PHASE TWO

Examinations	Observations	Mean of Noncompliance	Test Grouping
Exam 2	24	5.429	A
Exam 3	24	5.013	B
Exam 4	24	4.159	C

The Examination variable represents the three different midterm exams upon which the treatments, penalties, were applied. Table VII indicates examinations did not produce an effect upon noncompliance. The observed significance level of this effect is greater than 5 percent. Because all

of the examinations were constructed to attain an equivalent difficulty and could simply be construed as occasions, it is reasonable to expect this factor to be insignificant. At the same time, due to repeated measures and the effects of maturation, it would be reasonable to anticipate a significant difference as was discovered in Phase One (Table V). The conflicting results between the two phases of the experiment do not have an explanation beyond experimental error. Table XII presents the means of noncompliance relative to this independent variable, examinations, and the results of the Duncan test.

Summary

Based on analyses of the results it appears that differing levels of the treatment, penalties, will effect noncompliance. Also, differing levels of the other independent variable of concern, audit frequency, had significant effects on the independent variable. These results were anticipated. They verify the general theory of deterrence--lower noncompliance follows higher magnitudes of penalties and higher probabilites of audit.

The interaction of penalties and audit frequencies was not determined to be significant. This result is surprising. It is apparently contrary to the theory of deterrence. Prior research has shown the existence of an interactive effect in certain situations. The interactive affect of the two factors in the present study may have

become lost by the subjects. That is, since individuals reach decisions processing many factors, the interactive effect of penlties and audit frequencies may have become confounded with other factors. Even if all factors applicable to one decision setting could be identified, it is very difficult to isolate the relative importance of one factor from all of the other relevant factors.

The following chapter will discuss the conclusions that have been reached from these experiments. Also, the limitations of this research and suggestions for future research will be presented.

CHAPTER VI

SUMMARY AND CONCLUSIONS

Introduction

The objective of this chapter is to provide a brief overview of the research, provide conclusions drawn from the study, discuss the limitations of the study and offer recommendations for future reseach.

Overview

The purpose of this research was to determine if increased penalty magnitudes have a significant impact on noncompliance with a self-assessing and reporting system. Increasing levels of penalties were examined under two scenarios regarding audit frequency. The first phase of the study utilized a low, constant audit level. This situation is of interest because the increase in taxpayer noncompliance has been, in part, attributed to low levels of IRS auditing. Therefore, by increasing the level of penalties through reliance on the general theory of deterrence, greater compliance (or lower noncompliance) should result. The second phase of the study examined the noncompliance and penalty magnitude relationship under an

increasing audit scenario.

Practical and legal constraints do not allow for a direct study of tax noncompliance existing in the society of the United States. Therefore, this research project made use of a generic task laboratory experiment methodology. The experiment was constructed with undergraduate students as surrogates for taxpayers. The student subjects graded their own examinations. To allow for this task, the students were provided with instructions and forms to use in determining and reporting their examination scores. This was the task surrogate for determining and reporting a tax liability. The instructions were detailed to explain the varying treatment levels. The forms contained a manipulation check to insure the subjects were aware of their specific treatments. The examination scores reported by the students were used for course grade determination purposes. Unknown to the students was the fact that the researcher also determined an exam score (i.e., the correct score) for each student. The differences between the reported and correct scores (i.e., noncompliance) were statistically analyzed. Also, a post experiment questionnaire completed by the student subjects was analyzed to determine the subjects' perception of various experimental factors.

Conclusions

The results of the study support four conclusions which merit further discussion. The conclusions are as follows:

(1) As statutory penalties are increased, noncompliance to rules decreases given a low, constant probability of detection.

(2) As statutory penalties are increased and the probability of audit increases, noncompliance to rules decreases.

(3) Penalty magnitude levels do not appear to be significantly dependent upon the probability of audit, with regard to decreasing noncompliance. That is, an interaction between the factors apparently does not exist.

(4) Subject groups with more incentive and opportunity to cheat than other subjects, recorded higher average noncompliance.

In reality, the IRS audits approximately one percent of the individual tax returns [IRS, 1983]. Surveys have indicated that individual taxpayers perceive their probabilites of being audited are between two and twelve percent [Westat, 1980]. Therefore, conclusions from this reseach offer realistic validity with regard to the audit variable. Also, the conclusion that an increase in statutory penalties is accompanied by a decrease in noncompliance is a validation of the general theory of deterrence. Congress relied on this theory in a portion of

the Tax Reform Act of 1986 [U.S. Congress, 1986]. Several penalties were increased. Some of these included the penalty for failure to pay tax which increased from one-half to one percent per month, the negligence penalty which has been expanded to cover all taxes, and the fraud penalty which has been increased from 50 percent to 75 percent of the underpayment.

The conclusions also support the part of the theory of deterrence that indicates that increasing the audit probability will be accompanied by a decrease in noncompliance. The IRS is attempting to follow this conclusion through enhanced detection capabilities. Their primary means to achieve higher detection probabilities is computer technology. Due to budget constraints this enhancement has been slow in coming. In comparing the two primary factors of this research, penalty magnitude and audit probability, the penalty factor has the quicker practical application. It is not significantly dependent on budget monies.

The third conclusion suggests that penalty magnitudes can be increased for purposes of decreasing noncompliance to U.S. tax law without concern about increasing audit frequency. One factor appears not to be dependent on the other. Regardless of the probability of audit Congress should be able to achieve a decreased noncompliance rate by increasing the magnitude of penalties for noncompliance.

The fourth conclusion is supported by the results of a

study by Madeo, Schepanski, and Uecker [1987]. Higher income taxpayers have a greater opportunity to cheat on their taxes than low income taxpayers. These researchers used IRS TCMP data to discover that high income taxpayers report a lower portion of their actual income than low income taxpayers. The present study examined groups of subjects partitioned by relative score on Exam I. Subjects in the lowest scoring group had the greatest opportunity to cheat (the difference between the actual score and the maximum potential score). This subject group reflected the greatest average amount of cheating. These findings imply that improvements in compliance might be brought about by a stronger scrutiny of the tax returns of high income taxpayers. Possibly, the IRS could shift budgetary dollars away from examinations of low income taxpayers to enhance examinations of high income taxpayers. This shift would increase the probability of detection of noncompliance.

Limitations

Behavioral experimental studies raise questions with regard to internal and external validity by their very nature. The methodology and design of Phase One of the present study were such that the internal validity should be insured. In Phase Two the design was weakened due to the addition of the independent variable, audit frequency. An incomplete, but adequate, design was used. There were not sufficient class sections available to have a full factorial

design due to administrative constraints.

The greatest limitation of any laboratory experimentation in human behavior is external validity. Generalization of findings to other settings must be done with caution. Although care has been taken to design the task to be analogous to the real world taxpayer decision, it is possible that students use a different frame for their decisions on exam scores than taxpayers use for income reporting. Also, studies have indicated that underreporting income was more likely among younger persons, those with more education, and students [Mason and Calvin, 1978, 1984]. The experimental results may be biased toward those groups.

Another limitation of the study is that statutorily equal penalties may not be perceived to be personally equally severe. Confounding factors such as social stigma, wealth status, gender, and moral beliefs, may affect one's propensity to cheat. The present study did not positively eliminate the effects from any of these potential sources.

Additionally, in each phase of the study three different teachers covered the tested class sections. It is possible a teacher bias existed in the collected data on compliance. Whether an instructor was liked (or disliked), perceived to be good (or bad), or perceived fair (or unfair) could have biased compliance to the set of rules and sanctions. Possibly, students who liked an instructor complied, in part, for the instructor. Conversely, it is possible that students may have had higher noncompliance, if

the instructor was disliked. For some students, good or fair may have been the perception(s) that had biasing effects upon compliance. Whether, or not, a specific instructor was liked may not have been relevant. An instructor may have been perceived fair or good, but was disliked. Compliance may have been enhanced by the perceptions of fairness or goodness. Conversely, students perceiving their instructor(s) to be unfair, or bad, may have had a compulsion to cheat. These students may have felt disadvantaged. Cheating, or not following the rules, was a means to gain, or regain, perceived proper advantage in their classes. Also, during the semester(s) an instructor could have made accidental or causal remarks related to the self-grading process. It is unlikely that this type of occurrence existed, but there is a possibility. Cheating may have been enhanced, or may have been curtailed due to the comments. The teacher factor was not directly examined or controlled in the study. The degree of significance of this effect upon compliance is unknown.

Also, the possibilty exists that a portion of the students used in this study did not take the task seriously. In the past, there has been a significantly high drop rate in introductory accounting courses. The high drop rate may be due in part to students who do not adequately apply themselves to the course subject. If a significant proportion of the students failed to take the task seriously, the data collected may be biased. No research

has established the degree to which taxpayers approach tax filing seriously. The extent to which this problem might inhibit external validity is indeterminable.

A final limitation concerns an often discounted element of the general theory of deterrence, celerity of detection and penalty imposition [Teevan, 1976]. In the present study detection occurred quickly. Penalties, where applicable, were swiftly administered. In contrast, the IRS currently operates at a fairly slow pace. Some detection procedures can take up to eighteen months to implement [IRS, 1983]. Due to the many various appeals procedures available to taxpayers, penalty payments could occur at a time far removed from the noncompliance event. Celerity could, in part, account for the significant effects penalties and audit frequencies had on noncompliance in the experiment. This factor could constrain external validity.

Future Research

The results of this study provide additional insights into the factors that affect overt noncompliance. Research indicates that increased magnitudes of penalty can lead to optimal levels of compliance. Therefore, administrative agencies such as the IRS may pressure Congress to modify current laws. Improved efficiency and increased equity in the federal tax system might result.

The next step in compliance research is to identify the minimum penalty magnitudes that will result in compliance,

or noncompliance, settling at acceptable or optimal levels. This is important due to the fact that it is possible that the minimum penalty is so great that society would rebel against it. Therefore, other factors should be investigated further. For example, future research should address the optimal level of auditing necessary to bring about optimal compliance.

Additionally, research should be conducted to explore the interaction between the penalty and audit factors. Although the present study did not uncover a significant relationship between these factors, the general theory of deterrence seems to indicate there is such an interactive relationship. Finally, studies should be conducted to investigate how the celerity factor fits into the general theory of deterrence.

SELECTED BIBLIOGRAPHY

Abdel-Khalik, A. R., "On the Efficiency of Subject Surrogation in Accounting Research," Accounting Review (October 1974), pp. 743-750.

Ackerman, P. D., "The Effects of Honor-Grading on Students' Test Scores," American Educational Research Journal (March 1971), pp. 321-333.

Adams, M., "IRS Warns Cheaters: Here We Come," USA Today (May 14, 1985), pp. 1A, 6B.

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

Arvey, R. D. and J. M. Ivancevich, "Punishment in Organizations: A Review, Propositions, and Research Suggestions," Academy of Management Review (Vol. 5, No. 1, 1980), pp. 123-132.

Ashton, R. H. and S. S. Kramer, "Student Surrogates in Accounting Research," Journal of Accounting Research (Spring 1980), pp. 1-15.

Bailey, W. C. and R. P. Lott, "Crime, Punishment and Personality: An Examination of the Deterrence Question," Journal of Criminal Law and Criminology (Vol. 67, 1976), pp. 99-109.

Birnberg, J. G. and R. Nath, "Laboratory Experimentation in Accounting Research." The Accounting Review (January 1968), pp. 38-45.

Bowers, W. J. and R. G. Salem, "Severity of Formal Sanctions as a Repressive Response to Deviant Behavior," Law and Society Review (Vol. 6, 1971), pp. 427-441.

Campbell, D. T. and J. C. Stanley, Experimental and Quasi-Experimental Designs for Research (Houghton Mifflin Company, 1963).

Canning, R. R., "Does an Honor System Reduce Classroom Cheating? An Experimental Answer," Journal of Experimental Education (June 1956), pp. 291-296.

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

Cochran, W. G. and G. M. Cox, Experimental Designs (John Wiley & Sons, 1957).

Copeland, R. M., A. J. Francia and R. H. Strawser, "Students as Subjects in Behavioral Business Research," The Accounting Review (April 1973), pp. 365-372.

Dean, P., T. Keenan and F. Kenney, "Taxpayers' Attitudes to Income Tax Evasion: An Empirical Study," British Tax Review (1980), pp. 28-44.

Egger, R. L.,Jr., "Report of the Commissioner of the IRS for 1983," (U.S. Department of Treasury, 1983).

Feige, E., "How Big is the Irregular Economy?" Challenge (November/December 1979), pp. 5-13.

Friedland, N., J. Thibaut and L. Walker, "Some Determinants of the Violation of Rules," Journal of Applied Social Psychology (Vol.3, No.2, 1973), pp. 103-118.

_____, S. Maital and A. Rutenberg, "A Simulation Study of Income Tax Evasion," Journal of Public Economics (1978), pp. 107-116.

Friedrich, O., "Cheating by the Millions," Time (March 28, 1983), pp. 26-33.

Goode, W. J., "Force in Human Society," American Sociological Review (October 1972), pp. 507-519.

Grasmick, H. G. and G. J. Bryjak, "The Deterrent Effect of Perceived Severity of Punishment," Social Forces (December 1980), pp. 471-491.

Gutmann, P., "The Subterranean Economy" Financial Analysts Journal (November/December 1977), pp. 26, 27, 34.

_____, "Latest Notes From the Subterranean Economy," Business and Society Review (Summer 1980), pp. 22-28.

Internal Revenue Code, Section 6661.

Internal Revenue Service, Income Tax Compliance Research: Estimates for 1973-1981 (U.S. Department of Treasury, 1983).

Jackson, B. R. and S. M. Jones, "Salience of Tax Evasion Penalties Versus Detection Risk," Journal of the American Taxation Association (Spring 1985), pp. 7-17.

Keys, D. E. and J. A. Hendricks, "The Ethics of Accounting Research," Journal of Accounting Education (Fall 1984), pp. 77-88.

Kolm, S., "A Note on Optimum Tax Evasion," Journal of Public Economics (1973), pp. 265-270.

Krogstad, J. L., R. T. Ettenson and J. Shanteau, "Context and Experience in Auditors' Materiality Judgments," Auditing: A Journal of Practice & Theory (Fall 1984), pp. 54-73.

Kurtz, J., "Kurtz on Disclosure of Questionable Positions," Tax Advisor, (August 1977), pp. 496-498.

Laufer, D. M., "An Experimental Study of the Impact of Rule Complexity on Rule Compliance," unpublished Ph.D. dissertation (Oklahoma State University, 1985).

Lewis, A., "An Empirical Assessment of Tax Mentality," Public Finance (1979), pp. 245-257.

Madeo, S. A., A. S. Schepanski and W. C. Uecker, "Modeling Judgments of Taxpayer Compliance," Accounting Review (April 1987), pp. 323-342.

Mansfield, H. K., "The Role of Sanctions in Taxpayer Compliance," in P. Sawicki (ed.), Income Tax Compliance: A Report of the ABA Section of Taxation Invitational Conference on Income Tax Compliance (American Bar Association, 1983), pp. 217-233.

Mason, R. and L. D. Calvin, "A Study of Admitted Income Tax Evasion," Law and Society Review (Fall 1978), pp. 73-89.

_____ and _____, "Public Confidence and Admitted Tax Evasion," National Tax Journal (December 1984), pp. 489-496.

Meier, R. F. and W. T. Johnson, "Deterrence as Social Control: The Legal and Extralegal Production of Conformity," American Sociological Review (April 1977), pp. 292-304.

Minor, W. W. and J. Harry, "Deterrent and Experiential Effects in Perceptual Deterrence Research: A Replication and Extension," Journal of Research in Crime and Delinquency (Vol.19, No.2, 1982), pp. 190-203.

Murray, A., "Cheating Uncle Sam," Wall Street Journal (April 10, 1984), pp. 1, 20.

Musgrave, R. A. and P. B. Musgrave, Public Finance in Theory and Practice (McGraw-Hill, Inc., 1984).

Paternoster, R., L. E. Saltzman, G. P. Waldo and T. G. Chiricos, "Perceived Risk and Social Control: Do Sanctions Really Deter?" Law and Society Review (Vol.17, No.3, 1983), pp. 457-479.

Salem, R. G. and W. J. Bowers, "Severity of Formal Sanctions as a Deterrent to Deviant Behavior," Law and Society Review (1970), pp. 21-40.

Saltzman, L. E., R. Paternoster, G. P. Waldo and T. G. Chiricos, "Deterrent and Experiential Effects: The Problem of Causal Order in Perceptual Deterrence Research," Journal of Research in Crime and Delinquency (Vol.19, No.2, 1982), pp. 172-189.

SAS Institute Inc., SAS User's Guide: Statistics, 1982 Edition, (Carey, NC: SAS Institute, 1982).

Schwartz, R. and S. Orleans, "On Legal Sanctions," The University of Chicago Law Review (Winter 1967), pp. 274-300.

Scott, W. J. and H. G. Grasmick, "Deterrence and Income Tax Cheating: Testing Interaction Hypotheses in Utilitarian Theories," The Journal of Applied Behavioral Science (Vol.17, No.3, 1981), pp. 395-408.

Silberman, M., "Toward a Theory of Criminal Deterrence," American Sociological Review (June 1976), pp. 442-461.

Slovic, P., H. Kunreuther and G. F. White, "Decision Processes, Rationality and Adjustment to Natural Hazards," in G. F. White (ed.), Natural Hazards, Local, National, and Global (Oxford University Press, 1974), pp. 187-205.

Song, Y. and T. E. Yarbrough, "Tax Ethics and Taxpayer Attitudes: A Survey," Public Administration Review (September/October 1978), pp. 442-452.

Spicer, M. W., "A Behavioral Model of Income Tax Evasion," unpublished Ph.D. dissertation (Ohio State University, 1974).

_____ and L. A. Becker, "Fiscal Inequity and Tax Evasion: An Experimental Approach," National Tax Journal (June 1980), pp. 171-175.

_____ and S. B. Lundstedt, "Understanding Tax Evasion," Public Finance (Vol.31, No.2, 1976), pp. 295-305.

Srinivasan, T. N., "Tax Evasion: A Model," Journal of Public Economics (1973), pp. 339-346.

Steel, R. G. D. and J. H. Torrie, Principles and Procedures of Statistics (McGraw-Hill, 1980).

Sudman, S. and N. M. Bradburn, Asking Questions (Jossey-Bass, Inc., 1982).

Swieringa, R. J. and K. E. Weick, "An Assessment of Laboratory Experiments in Accounting," Journal of Accounting Research (Vol. 20, Supplement 1982), pp. 56-101.

Tanzi, V., "Underground Economy in the United States: Estimates and Implications," Banca Nazionale Del Lavoro, Quarterly Review (Vol. 33, 1982), pp. 427,428.

Teevan, J. J., "Subjective Perception of Deterrence (Continued)," Social Forces (March 1976), pp. 155-164.

Tittle, C. R., "Sanction Fear and the Maintenance of Social Order," Social Forces (December 1976), pp. 579-596.

_____ and C. H. Logan, "Sanctions and Deviance: Evidence and Remaining Questions," Law and Society Review (Spring 1973), pp. 371-392.

_____ and A. R. Rowe, "Moral Appeal, Sanction Threat and Deviance: An Experimental Test," Social Problems (Vol.20, 1973), pp. 488-498.

Tversky, A., and D. Kahneman, "Belief in the Law of Small Numbers," Psychological Bulletin (August 1971), pp. 105-110.

U.S. Congress, Tax Equity and Fiscal Resonsibility Act of 1982 [P.L. 97-248], 97th Cong., 2nd Sess. (September 3, 1982).

U.S. Congress, Tax Reform Act of 1986 [P.L. 99-514], 99th Cong., 2nd Sess. (October 22, 1986).

Waldo, G. P. and T. G. Chiricos, "Perceived Penal Sanction and Self-Reported Criminality: A Neglected Approach to Deterrence Research," Social Problems (Spring 1972), pp. 522-540.

Westat, "Individual Income Tax Compliance Factors Study: Qualitative Research Results" (Report to the IRS on Contract No. TIR-78-50, 1980).

Witte, A. D. and D. 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, (March 1985), pp. 1-13.

Yitzhaki, S., "A Note on Income Tax Evasion: A Theoretical Analysis," Journal of Public Economics (Vol.3, 1974), pp. 201-202.

APPENDIXES

APPENDIX A

EXAMINATION PROCEDURES

EXAMINATION PROCEDURES

ACCTG 2103

1. A self-grading procedure will be used in this course to grade exams.

2. Your course grade will be based solely upon examination scores.

3. Exam days will follow traditional procedures:

   (a) Closed-book exams will be taken during regular class time.

   (b) Calculators may be used. Notes or other study aids are not permitted during the exam.

   (c) Upon completing the exam or expiration of time, the exam will be submitted to your monitor.

   (d) Any student suspected of cheating during the exam will be given a zero for exam score. If such student wishes to withdraw from the course, he/she will be given an "F".

4. The very next class meeting, the exams will be returned. You will receive the following items:

   (a) Answer Key
   (b) Grading Instructions
   (c) Grade Report Form

5. Using the answer key and following the grading instructions, you will complete the Grade Report Form. Be sure to follow directions carefully. Answer all questions. Fill in all blanks. Failure to follow directions may result in a penalty and a lower exam score.

6. Only the completed Grade Report Form will be returned to your monitor. You will keep your exam, the answer key and other materials. The completed Grade Report Form will be due approximately one week after the exam day. The specific day, time, and place will be explained in the grading instructions.

Loss of Points

1. Late Report: Failure to submit the Grade Report Form when due will result in a loss of 5% of the reported exam score daily for the first three days. One day is charged beginning with the due time. For example, if the due time is 10:30 AM, and the grade report form is submitted at 10:34 AM one day late is charged. Be sure your grade reports are submitted on time. Turn in your report early, if you want.

   The third day after the due date, an additional loss of points of 25% of the reported score will be charged, above the 5% daily charge.

2. Report not turned in: If a student fails to submit a grade report form, a zero will be recorded for that exam.

3. If a student does not take the examination at the regularly scheduled time, a zero will be recorded for that exam.

   See your instructor's syllabus for details and exceptions.

4. KEEP YOUR EXAMS for review and to support your grading computations, if they are possibly checked.

   DO NOT LOSE OR MISPLACE YOUR EXAMS !!!!!!!!!

5. There may be different exam versions. Different answer keys, grading instructions, and grade report forms may be used. Do not use any materials of another student. Confusion, unnecessary questioning, and incorrect reporting may result. This could lead to a loss of points.

   During the semester all students will be treated equally as to exams, answer keys, grading instructions, and Grade Report Forms.

Compliance Check

1. During the semester a few students' Grade Report Forms will be randomly selected to check for grading errors and compliance with grading instructions. Names will be drawn from a hat. The monitor will make xerox copies of only those selected students' tests to confirm grading.

   Approximately 12% to 24% of the Grade Report Forms will be checked on each exam. The precise number will be stated in your Grading Instructions for each self-graded test.

2. If, by chance, your report form is randomly selected to be checked, you will be notified by mail to meet with the monitor. It is imperative that the monitor has your correct local address. If you move, be sure to notify the monitor.

   Time and place of the meeting will be stated in the notification letter. When you get the notification letter, return the enclosed confirmation card to the monitor's office immediately.

3. At the meeting, bring your original exam. Checking your grading will be mandatory at the meeting. Do not lose your original exams! If you lose the original exam, it may result in a zero score for the portion checked.

4. If negligent or fraudulent errors are discovered during the meeting, penalty points may be deducted from your actual test score as determined by the monitor.

   Specific penalties will be described in the Grading Instructions for each self-graded test.

   University policies on cheating will not apply to self-graded tests due to other penalties.

Improved Learning

1. The grading procedure should result in an enhancement of learning. Previous semesters using similar methods have suggested that most students enjoyed grading their own exams. Students may learn more from their mistakes when grading their own exams.

2. If you have strong objections to the self-grading procedure, you may request that the monitor grade all your exams. To be excluded, you must inform the monitor immediately (today!).

   Your Examination Monitor: Rick Crosser, 107 Business Bldg., 624-5115.

****** DO NOT LOSE THIS DOCUMENT. ******
IT WILL BE USEFUL LATER IN THE SEMESTER.

APPENDIX B

GRADE REPORT FORM AND GRADE REPORT FORM INSTRUCTIONS

GRADE REPORT FORM
Exam III

Name_____ Student Number_____

Local Address_____ Telephone Number_____

Class Time_____ Instructor_____

Reminder: FILL IN ALL BLANKS!

Carefully read both pages of the GRADE REPORT FORM INSTRUCTIONS. From the INSTRUCTIONS, answer the next two questions.

What is the size of penalty #1, page 2 _____?

The percentage of exams to be checked is _____?

EXAM SCORING:

Part I. (18 points)
Questions 1- 9, points earned _____

Part II. (15 points)
Questions 10-14, points earned _____

Part III. (9 points)
Questions 15-17, points earned _____

Part IV. (42 points)
Questions 18-31, points earned _____

Part V. (16 points)
Questions 32-35, points earned _____

EXAM SCORE: Add all of the lines above _____.

I have reviewed this Grade Report Form and to the best of my knowledge and belief the information reported herein is true, correct, and complete. I understand that in the event of a possible check of this exam, if a difference is found between the score as reported here and the true score, penalties may be applied.

_____ _____
(Your Signature) (Date)

GRADE REPORT FORM INSTRUCTIONS

ACCTG 2103 - EXAM III

USE YOUR OWN GRADE REPORT FORM AND INSTRUCTIONS. USE OF ANOTHER STUDENT'S FORMS WILL CAUSE MISTAKES AND YOU WILL BE PENALIZED. (SEE PENALTY #1, PAGE 2.) REFER TO YOUR HANDOUT, EXAMINATION PROCEDURES.

1. Fill out the information section on the Grade Report Form. Failure to complete the entire form will result in penalties. (See Penalty #3, page 2 of this instruction sheet.)

2. Use the answer key to grade your exam.

3. Read your teacher's explanations carefully, even if you answered the question correctly. This will help you learn as you grade your test.

4. Record your exam score on the Grade Report Form. Be sure to follow the directions carefully !!!!!

5. When you have completed the Grade Report Form, read the statement at the bottom. Sign your name and write the date in the spaces provided. Be sure to sign your Grade Report Form. Failure to sign the form will result in penalties. (See Penalty #3, page 2.)

TURN IN YOUR COMPLETED GRADE REPORT AT THE BEGINNING OF CLASS TIME, IN YOUR REGULAR CLASSROOM, ON WEDNESDAY, NOVEMBER 6. LATE FORMS WILL BE PENALIZED. (See page 2 of EXAMINATION PROCEDURES)

BE SURE TO FINISH READING THE INSTRUCTIONS ON THE NEXT PAGE

PENALTIES

#1. 12 percent of EXAM III PAPERS WILL BE CHECKED FOR FOLLOWING GRADING INSTRUCTIONS. (In your class, 12% x 39 = 4.68 = 5 exams). The exams will be chosen randomly, (drawn from a hat) prior to receipt of your Grade Report Form. If your exam is chosen, you will be contacted for an office visit.

TWO TIMES the difference between the reported score and true score will be deducted as a penalty from the true score, if you are discovered to have been careless or fraudulent in following instructions. University policy on cheating will not apply for self-graded exams.

#2. IF MATH ERRORS ARE FOUND ON THE GRADE REPORT FORM, the math mistake will be corrected, and the corrected score will be recorded. There will be no penalty for math errors. Students will be notified by mail at their local address of any math errors on their grade report form and of the corrected score.

#3. IF THE GRADE REPORT FORM IS NOT PROPERLY AND COMPLETELY PREPARED (including signature and date), one (1) point will be deducted for each omission and error.

#4. FOR REPORTS SUBMITTED AFTER THE DUE DATE AND TIME ABOVE, see numbers 1, 2, and 3 on page 2 of the handout, EXAMINATION PROCEDURES.

If you have questions or serious problems in grading your exam, contact RICK CROSSER, room 107, Business Bldg. Phone: 624-5115. Office Hours: Monday, Wednesday, Friday 12:10-2:30.

APPENDIX C

POSTEXPERIMENT QUESTIONNAIRE AND RESULTS

Grading Questionnaire
Accounting 2103

Please answer the following questions about the self-grading procedure for exams. Darken the space on the answer sheet for the one answer which best describes your feelings. Your answers will be anonymous. Do not put your name on the answer sheet.

Please use the following scale to answer the questions: (Except for Question 26)

[1] Definitely Yes
[2] Yes
[3] Not applicable to me
[4] No
[5] Definitely No

1. I enjoyed grading my own exam.

2. I learned more from grading my own exam than I would have otherwise.

3. The explanations on the answer key were beneficial in aiding me learn from my mistakes.

4. I studied less for this course because I graded my own exams.

5. It was very time consuming to grade my own exams.

6. My effort level in complying with the grading instructions was consistent from one exam to the next.

7. It was very hard to grade my own exams.

8. The first exam was much harder to grade than the others.

9. If I was unsure how to grade a particular question, I gave myself the benefit of the doubt.

10. I bent the grading rules where I thought my total exam score was too low.

11. I would have had the same exam scores if the instructor had graded my exams.

12. I would have had lower exam scores if the instructor had graded my exams.

13. I would have had higher exam scores if the instructor had graded my exams.

14. At the beginning of the semester, I expected at least one of my exams would be checked for grading accuracy.

15. If my chance of being caught was almost zero, I would not be concerned about the size of the possible penalty for cheating.

16. The chance of having my exam checked made me more careful in following the grading instructions.

17. If my chance of being caught is almost 100%, I would not be concerned about the size of the possible penalty for cheating.

18. I was very concerned about making a mistake in grading my exam.

19. The larger-sized penalties caused me to refrain from cheating.

20. If the penalty was zero, I felt safe to cheat.

21. I was careful about grading my exams because of the University policy for academic dishonesty.

22. I think the monitor was close to the day-to-day activities of my class.

23. I felt a lot of extra stress from trying to follow the grading instructions.

24. I felt that the test questions were unfair.

25. I feel that letting students grade their own tests is unfair.

26. The percentage of exams actually checked for the class as a whole was:

[1]	[2]	[3]	[4]	[5]
0%	1%	4%	10%	100%

TABLE XIII

POSTEXPERIMENT QUESTIONNAIRE--
SUMMARY RESULTS

Question Number	Mean
1	2.851
2	3.018
3	2.155
4	4.186
5	3.924
6	3.752
7	4.038
8	3.617
9	3.012
10	4.112
11	2.057
12	4.081
13	3.899
14	2.127
15	2.791
16	2.140
17	3.635
18	2.450
19	2.553
20	3.119
21	3.362
22	3.229
23	3.513
24	2.892
25	3.253
26	3.044

APPENDIX D

PLOT OF NONCOMPLIANCE BY CLASS SECTIONS

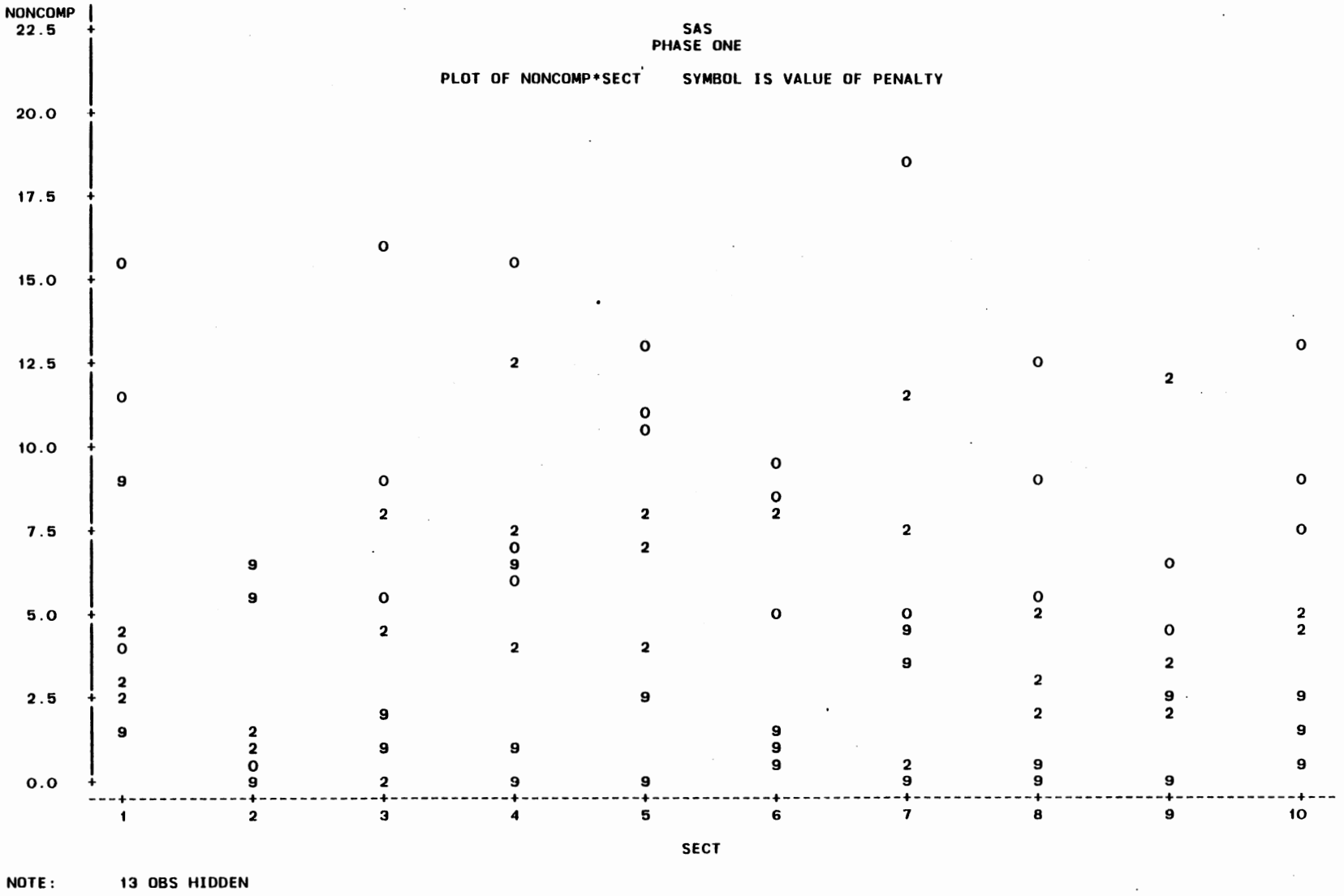

Figure 2. Plot of Noncompliance by Class Sections

VITA

Rick Lynn Crosser

Candidate for the Degree of

Doctor of Philosophy

Thesis: THE RELATIVE EFFECT OF PENALTY MAGNITUDES ON COMPLIANCE: AN EXPERIMENTAL EXAMINATION OF DETERRENCE

Major Field: Business Administration

Biographical:

Personal Data: Born in Cherokee, Iowa, March 10, 1949, to Elmer LeRoy and Merle Elizabeth Leidy Crosser.

Education: Graduated from North Platte High School, North Platte, Nebraska, in June, 1966; received Bachelor of Science in Mechanical Engineering from Colorado State University in August, 1970; received Bachelor of Science in Business Administration from Colorado State University in December, 1978; received Master of Science in Taxation from Colorado State University in May, 1982; completed requirements for the Doctor of Philosophy degree at Oklahoma State University in December, 1987.

Professional Experience: Merchandising Manager, J.C. Penney Co., 1970 to 1976; General Manager and Controlling Officer, JK Sporting Goods Corporation, 1976 to 1978; Staff Accountant, Armstrong and France, CPAs, 1979 to 1980; Tax Specialist, Kruchten and Company, CPAs, 1980 to 1982; Sole Practitioner, CPA, 1982 to 1983; Graduate Teaching Associate, Oklahoma State University, 1983 to 1987.

Professional Memberships: American Accounting Association; American Taxation Association; American Institute of Certified Public Accountants; Colorado Society of Certified Public Accountants.