

Spring 3-13-2017

The Effects of an Employment Tax Enforcement Regime on US Small Business and Proprietor Payment Compliance

Rafael Dacal

Follow this and additional works at: http://scholarworks.gsu.edu/bus_admin_diss

Recommended Citation

Dacal, Rafael, "The Effects of an Employment Tax Enforcement Regime on US Small Business and Proprietor Payment Compliance." Dissertation, Georgia State University, 2017.
http://scholarworks.gsu.edu/bus_admin_diss/78

This Dissertation is brought to you for free and open access by the Programs in Business Administration at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Business Administration Dissertations by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.

PERMISSION TO BORROW

In presenting this dissertation as a partial fulfillment of the requirements for an advanced degree from Georgia State University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote from, copy from, or publish this dissertation may be granted by the author or, in her absence, the professor under whose direction it was written or, in his absence, by the Dean of the Robinson College of Business. Such quoting, copying, or publishing must be solely for scholarly purposes and must not involve potential financial gain. It is understood that any copying from or publication of this dissertation that involves potential gain will not be allowed without written permission of the author.

Rafael Dacal

NOTICE TO BORROWERS

All dissertations deposited in the Georgia State University Library must be used only in accordance with the stipulations prescribed by the author in the preceding statement.

The author of this dissertation is:

Rafael Dacal
600 Boxwood Terrace
Johns Creek, GA 30005

The director of this dissertation is:

Dr. Detmar W. Straub, Jr.
J. Mack Robinson College of Business
Georgia State University
Atlanta, GA 30302-4015

The Effects of an Employment Tax Enforcement Regime on US Small Business and
Proprietor Payment Compliance

by

Rafael Dacal

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree

Of

Executive Doctorate in Business

In the Robinson College of Business

Of

Georgia State University

GEORGIA STATE UNIVERSITY

ROBINSON COLLEGE OF BUSINESS

2017

Copyright by
Rafael Dacal
2017

ACCEPTANCE

This dissertation was prepared under the direction of the *RAFAEL DACAL* Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Business Administration in the J. Mack Robinson College of Business of Georgia State University.

Richard Phillips, Dean

DISSERTATION COMMITTEE

Dr. Detmar W. Straub, Jr.

Dr. Pam Scholder Ellen

Dr. Alexander H. Turk

ACKNOWLEDGEMENTS

I would like to thank my committee for all their help. I was extremely lucky of being able to have worked with the three of you. A special thanks to my advisor Professor Dr. Detmar Straub for all your tutelage, mentoring and patients. I thank you for all the advices that made my research possible and for helping me to grow as a researcher. I would also like to thank my committee members Dr. Pam Scholder Ellen and Dr. Alex Turk for serving as part of my committee. Dr. Pam Scholder Ellen, thank you for all your insight and suggestions, which enhanced my understanding of experimental design. Dr. Alex Turk, thank you for helping me structure the taxation components of this paper. I would also like to thank you for letting my defense be an enjoyable moment. You were all very instrumental in my research and I owe a debt of gratitude to you.

I would also like to thank my family for their help and understating. To Nicole, Joshua, and María thank you for putting up with the time, the process, and me. To my mother, father, and María for always encouraging me and telling me “¡termina!” I could not have done it with you out your patients, love and encouragement.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iv
LIST OF TABLES	ix
LIST OF FIGURES	xi
I INTRODUCTION	1
II PRACTICAL MOTIVATION	6
II.1 Employment Tax Motivation	6
II.2 Compliance Motivation	8
III RESEARCH QUESTION	11
IV LITERATURE REVIEW	12
IV.1 Tax	12
IV.2 Employment or Trust Tax	19
IV.2.1 <i>Employment Tax Defined</i>	19
IV.2.2 <i>Employment Tax and the Tax Incidence</i>	20
IV.3 Compliance	22
IV.3.1 <i>Filing and Reporting Compliance</i>	25
IV.3.1.1 <i>Sources of Compliance (or Noncompliance)</i>	30
IV.3.1.2 <i>Characteristics of Noncompliant Taxpayers</i>	36
IV.3.1.3 <i>Audit Regimes and Compliance</i>	38
IV.3.2 <i>Payment Compliance</i>	41
IV.3.3 <i>Income Shock</i>	44
IV.4 Literature Review Summary	44
V THEORY	46
V.1 Theoretical Motivation	46
V.2 General Deterrence Theory	47
V.2.1 <i>Background</i>	47
V.2.2 <i>Principles of General Deterrence Theory</i>	53

V.2.2.1	<i>Celerity</i>	53
V.2.2.2	<i>Severity</i>	53
V.2.2.3	<i>Certainty</i>	54
V.3	Research Model	55
VI	LITERATURE REVIEW CONCLUSION	57
VII	HYPOTHESES	59
VIII	METHODS	60
VIII.1	Experimental Design	61
VIII.1.1	<i>Validity</i>	61
VIII.1.2	<i>Experiment</i>	62
VIII.1.3	<i>Social Desirability Bias</i>	66
VIII.1.4	<i>Probability of Detection or Enforcement</i>	68
VIII.1.5	<i>Tax Liability Thresholds Selection – the Tax Rate</i>	70
VIII.1.6	<i>Population of Interest and Sampling Design</i>	71
VIII.1.7	<i>Sample size</i>	72
VIII.1.8	<i>Unit of Observation and Unit of Analysis</i>	72
VIII.1.9	<i>Implementing the Penalties and Interest</i>	72
VIII.1.10	<i>Safe Harbor and Self-Assignment</i>	73
VIII.2	Data Analysis	75
VIII.3	Test Phase	75
VIII.3.1	<i>Phase 1</i>	76
VIII.3.2	<i>Phase 2</i>	76
VIII.3.3	<i>Phase 3</i>	78
VIII.3.4	<i>Phase 4</i>	79
VIII.4	Measuring Risk Appetite	79
VIII.5	Additional Test Details	81
VIII.5.1	<i>Recruitment</i>	81
VIII.5.2	<i>Funding and Compensation</i>	81
VIII.6	Deviations from Original Design	82

IX FINDINGS.....	84
IX.1 General Demographics of Subjects.....	84
IX.2 Risk Appetite Analysis and Enforcement Perception.....	85
IX.2.1 Reliability.....	86
IX.2.2 Validity.....	86
IX.2.3 Probability of Enforcement and Penalty Rate	88
IX.2.3.1 Changes to Probability of Enforcement.....	88
IX.2.3.2 Changes to Penalty Rate.....	88
IX.3 Hypothesis testing	89
IX.3.1 Hypothesis 1	89
IX.3.1.1 Round 4 Test Results.....	90
IX.3.1.2 Round 5 Test Results.....	93
IX.3.1.3 Round 6 Test Results.....	95
IX.3.1.4 Hypothesis 1 ANOVA Test Results.....	97
IX.3.2 Hypothesis 2a	98
IX.3.2.1 Round 4 Test Results.....	99
IX.3.2.2 Round 5 Test Results.....	101
IX.3.2.3 Round 6 Test Results.....	103
IX.3.2.4 Hypothesis 2a ANOVA Test Results.....	105
IX.3.3 Hypothesis 2b	106
IX.3.3.1 Round 4 Test Results.....	107
IX.3.3.2 Round 5 Test Results.....	109
IX.3.3.3 Round 6 Test Results.....	111
IX.3.3.4 Hypothesis 2b ANOVA Test Results.....	113
IX.3.4 Hypothesis 3	113
IX.3.4.1 Round 4 Testing	114
IX.3.4.2 Round 5	116
IX.3.4.3 Round 6	119

IX.3.4.4	<i>Hypothesis 3 ANOVA test</i>	121
IX.3.5	<i>Hypothesis 4</i>	121
IX.3.5.1	<i>Hypothesis 4 ANOVA test</i>	124
IX.3.6	<i>Additional Payment Compliance Analysis</i>	125
IX.3.6.1	<i>Payment Compliance ANOVA</i>	125
IX.3.6.2	<i>Logistic Regression Control Compliance</i>	126
IX.3.6.2.1	Employment Tax Payment Compliance	127
IX.3.6.2.2	Subscription Rate.....	129
X	DISCUSSION	133
X.1	Contribution to Theory	133
X.2	Contribution to Practice	134
X.2.1	<i>Employment Tax Motivation</i>	134
X.2.2	<i>Compliance Motivation</i>	135
XI	LIMITATIONS	137
	APPENDICES	140
XI.1	Appendix A: Test Questionnaire	140
XI.2	Appendix B: IRS Tax Form 944	142
XI.3	Appendix C: Recruitment Script and Hand-out	145
XI.4	Appendix D: Additional literature review	147
XI.4.1	<i>Corporate responsibility</i>	147
XI.4.2	<i>Employment Tax and Tax Incidence</i>	147
XI.4.3	<i>Characteristics of non-filers</i>	151
	REFERENCES	156
	VITA	177

LIST OF TABLES

Table 1: Employment Taxes Civil Penalties Assessed, Fiscal Year 2014.....	8
Table 2: Enforcement Probability for the Random Assignment Section	69
Table 3: Enforcement Probability for the Self-Assignment Section	69
Table 4: Enforcement Probability for the Experiment	69
Table 5: Risk Construct.....	80
Table 6: Enforcement, Timing and Penalty Questions.....	81
Table 7: Involvement with Employment Tax Filing and Payment Process	84
Table 8: Geographical Breakdown of Participating Subjects	85
Table 9: Cronbach’s Alpha Results.....	86
Table 10: Factor Analysis for Risk.....	87
Table 11: Results by Hypothesis.....	89
Table 12: Safe Harbor Subscription Rate by Tax Liability Thresholds.....	90
Table 13: Hypothesis 1 Round 4 Statistical Tests	92
Table 14: Hypothesis 1 Round 4 Effect Size Tests	93
Table 15: Hypothesis 1 Round 5 Statistical Tests	94
Table 16: Hypothesis 1 Round 5 Effect Size Tests	95
Table 17: Hypothesis 1 Round 6 Statistical Tests	96
Table 18: Hypothesis 1 Round 6 Effect Size Tests	97
Table 19: Hypothesis 1 ANOVA tests for	98
Table 20: Enforcement Treatments Frequency and Percent.....	99
Table 21: Hypothesis 2a Round 4 Statistical Tests	100
Table 22: Hypothesis 2a Round 4 Effect Size Tests	101
Table 23: Hypothesis 2a Round 5 Statistical Tests	102
Table 24: Hypothesis 2a Round 5 Effect Size Tests	103
Table 25: Hypothesis 2a Round 6 Statistical Tests	104
Table 26: Hypothesis 2a Round 6 Effect Size Tests	105
Table 27: Hypothesis 2a ANOVA tests	105
Table 28: Hypothesis 2b Round 4 Statistical Tests	108
Table 29: Hypothesis 2b Round 4 Effect Size Tests.....	108
Table 30: Hypothesis 2b Round 5 Statistical Tests	110

Table 31: Hypothesis 2b Round 5 Effect Size Tests.....	110
Table 32: Hypothesis 2b Round 6 Statistical Tests.....	112
Table 33: Hypothesis 2b Round 6 Effect Size Tests.....	112
Table 34: ANOVA tests for Hypothesis 2b.....	113
Table 35: Compliance and Noncompliance Rate by Rounds.....	114
Table 36: Hypothesis 3 Round 4 Statistical Tests.....	116
Table 37: Hypothesis 3 Round 4 Effect Size Tests.....	116
Table 38: Hypothesis 3 Round 5 Statistical Tests.....	118
Table 39: Hypothesis 3 Round 5 Effect Size Tests.....	118
Table 40: Hypothesis 3 Round 6 Statistical Tests.....	120
Table 41: Hypothesis 3 Round 5 Effect Size Tests.....	121
Table 42: Hypothesis 3 ANOVA Tests.....	121
Table 43: Hypothesis 4 Statistical Tests.....	124
Table 44: Hypothesis 4 Effect Size Test.....	124
Table 45: Round 6 ANOVA Test.....	125
Table 46: Payment Compliance frequency by Round.....	125
Table 47: ANOVA Test between Compliance Rates under the Safe Harbor Provision.....	126
Table 48: Compliance Rates under a Safe Harbor Provision Effect Size Test.....	126
Table 49: Payment Compliance logistic Regression Results Given Risk control....	127
Table 50: Goodness of fit analysis.....	127
Table 51: Payment Compliance logistic Regression Results.....	128
Table 52: Goodness of fit analysis.....	129
Table 53: Payment Compliance logistic Regression Results Given Risk control....	130
Table 54: Goodness of fit analysis.....	130
Table 55: Subscription to Safe Harbor Regression Results.....	131
Table 56: Goodness of fit analysis.....	131

LIST OF FIGURES

Figure 1: United States Federal Surplus or Deficits from September 1977 to September 2013	10
Figure 2: Literature Review Space.....	12
Figure 3: Enforcement Compliance Cycle.....	50
Figure 4: Deterrence Model	52
Figure 5: Research Model	56
Figure 6: Test Phase Process.....	66
Figure 7: Phase 1 Random Assignment Design.....	76
Figure 8: Phase 2 Random Assignment Design.....	77
Figure 9: Phase 3 Randomize Design with a Cutoff	78
Figure 10: Round 4 Distributions of the Safe Harbor Provision Subscription Rate Given Employment Tax Liability Threshold	91
Figure 11: Round 5 Distributions of the Safe Harbor Provision Subscription Rate given Employment Tax Liability Threshold	93
Figure 12: Round 6 Distributions of the Safe Harbor Provision Subscription Rate given Employment Tax Liability Threshold.	95
Figure 13: Round 4 Distributions between Enforcements Treatments and the Safe Harbor Provision Subscription Rate.....	99
Figure 14: Round 5 Distributions between Enforcement Treatments and the Safe Harbor Provision Subscription Rate.....	102
Figure 15: Round 6 Distributions between Enforcement Treatments and the Safe Harbor Provision Subscription Rate.....	103
Figure 16: Round 4 Distributions between Enforcement Treatments and Payment Compliance Rate.	107
Figure 17: Round 5 Distributions between Enforcement Treatments and Payment Compliance Rate.	109
Figure 18: Round 6 Distributions between Enforcement Treatments and Payment Compliance Rate.	111
Figure 19: Round 4 Distributions between Payment Compliance and the Safe Harbor Provision Subscription Rate.....	115

Figure 20: Round 5 Distributions between Payment Compliance and the Safe Harbor Provision Subscription Rate.....	117
Figure 21: Round 6 Distributions between Compliance Payment and the Safe Harbor Provision Subscription Rate.....	119
Figure 22: Round 6 Distributions between Payment Compliance and the Safe Harbor Provision Subscription Rate.....	123

ABSTRACT

The Effects of an Employment Tax Enforcement Regime on US Small Business and
Proprietor Payment Compliance

by

Rafael Dacal

April 2017

Chair: Dr. Detmar W. Straub, Jr

Major Academic Unit: J. Mack Robinson College of Business

This study attempted to identify ways to improve voluntary compliance and minimize taxpayer burden, but also tries to understand the behavior of taxpayers' compliance given the compliance regimen. Most explicitly, it attempted to identify ways to improve payment compliance using regimens already utilized in other parts of the tax code. The research question was whether different tax regimes, such as safe harbor, can change the behavior of employment tax payment for small business or self-employed taxpayers. The idea was to determine if a safe harbor provision can reduce the proclivity of authorized individuals to implement a payroll tax dilemma strategy and whether or regimen can reduce payment noncompliance in time of economic distress. To answer the research question, an online experiment was employed. The experimental design was an impact study. The population of interest in this study was all authorized individuals from small and self-employed firms. The sample size totaled 205, and it was based on the a-priori sample size calculation. Analysis of variance (ANOVA) was chosen as the data analysis technique, but other nonparametric test and logistic regression models were used to further analyze the data. This study showed that for subjects who did not subscribed to

safe harbor provision but experienced an increased probability of apprehension increased their payment compliance. Also, the availability of a safe harbor provision lead to a large numbers to a safe harbor provision subscription in order to avoid enforcement. This study was able to show that individuals were willing to improve their payment compliance rate when enforcement was increased. The General Deterrence Theory explains that increased deterrence will lead to higher compliance. The study showed a 10 percent improvement in payment compliance when safe harbor was implemented. The results from this study also suggest that provisions such as a safe harbor can be a method of reducing filing costs and audit costs and ultimately taxpayer burden. On the other hand, the results of this study were inconclusive in determining if such provisions can improve payment compliance. Nevertheless, the outcome of this study can improve timing and accuracy of employment taxes payments and it may improve the accuracy of employment tax payment.

INDEX WORDS: Employment tax, compliance, payment compliance, employment tax enforcement regime, Small Business, experimental design, ANOVA, nonparametric test, General Deterrence Theory, audit, tax enforcement, risk attitude scale, online experiment

I INTRODUCTION

Taxes are as old as humanity and so is tax noncompliance. Therefore, even though the concept of noncompliance is ancient, the in-depth study of tax compliance only began around the late 1960s. One definition of tax compliance was articulated in *Gregory v. Helvering* 293 U.S. 465 (1935) decision, which states that as long as information is complete, timely, and accurate under the tax code, no taxpayer is bound under any sense of “patriotic duty” (69 F.2d 811) to pay more than their fair share. Conversely, the act of not filing taxes with information that is complete, timely, and accurate is noncompliance, whether the act is deliberate or accidental.

On a positive note, the vast majority of taxpayers comply properly with their complete tax obligation, but noncompliance remains the major reason for the existence of the US Internal Revenue Service (IRS). In general, the IRS enforces compliance via the National Research Program (NRP). This program randomly selects a stratified sample of approximately 50,000 to intensively audit. Using this audit information, IRS develops internal algorithms for case scoring and workload case selection and for calculating the tax gap. However, the number of audits over the last several years has been decreasing mainly due to budgetary constraints and employee attrition. In an optimal environment, random selection would be used in all audit cases to maximize compliance (Bloomquist, 2009). However due to the aforementioned constraints, the Service is trying to find ways to improve voluntary payment, filing, and reporting compliance, including variants on employment tax.

Taxpayer compliance can be sub categorized into three types: payment, filing, and reporting compliance (Brown & Mazur, 2003). Payment compliance is the accurate

payment of tax liabilities. Filing compliance is the complete, timely, and accurate filing of tax forms. Reporting compliance is the correct reporting of tax amounts owed to the US government. Of the three sub categories, this study focused on the first type, payment compliance of employment tax. While there have been numerous experimental studies on compliance, none has focused on employment tax. This is important because one of the major focuses for the Service has been to improve voluntary compliance of employment tax, which accounts for about 72 billion dollars of the tax gap, which is any misreporting and eventual underpayment of tax liability. Currently, the tax gap associated with employment tax is equivalent to the budgets of the Energy Department, the Department of Homeland Security, and the Commerce Department combined.

In addition to the financial importance of understanding this tax base, this study focused on payment compliance because most of the employment tax noncompliance arises from payment noncompliance. This phenomenon occurs because employment tax has a very high level of information matching, so it is relatively easy to detect filing and reporting noncompliance. This is a primary motivation for the current study in that results here could be used to improve voluntary compliance efforts and reduce taxpayer burden. In addition, efforts that advance our understanding of compliance in this circumstance could reduce the 72 billion dollar US tax gap.

Employment tax and income tax withheld from employees' pay is reported using Form 941 "Employer's QUARTERLY Federal Tax Return" or one of the forms in the 94x series. To assist employers in completing and complying with the necessary forms, the IRS publishes Publication 15 (Circular E). This publication provides the guideline employers should follow when filing and depositing employment tax.

One of the requirements is to file employment tax at a regular interval, depending on different tax liability thresholds. In most cases, firms are required to deposit their tax liability on a bi-weekly or monthly basis to avoid incurring penalties and falling behind on payments. However, there are times when companies facing economic distress will use employment tax withholding to overcome liquidity shortcomings (Grady, 2013). This business decision is defined as the “payroll tax dilemma” (pg.1) and the legal liability is associated with an authorized individual. An authorized individual is someone who is personally liable for the payment of employment tax withholdings deposits.

The payroll tax dilemma is when a firm cannot meet its operational liabilities and illegally uses employment tax withholdings to pay against the firm’s liabilities (Godfrey, 2004; Mauldin & Wilder, 1997). The problem that arises from the payroll tax dilemma strategy is that many of these businesses continue to illegally borrow from withheld taxes for lengthening periods of time and Grady (2013) found this strategy will inevitably lead to firm failure.

The present study first assumes that authorized individuals will succumb to the payroll tax dilemma during business crises. We define business crises as an income shock, which reduces firms’ revenues. The argument is that individuals decide to use employment tax withholding as a short-term loan without clearly understanding the negative implications. Some of the implications could be increased interest and penalties and personal financial responsibilities on amounts due. To examine downstream effects, we introduced a safe harbor¹ provision to determine if an intervention minimized the incidence of the payroll tax dilemma.

¹ A safe harbor provision specifies conditions that protect individuals or firms from being deemed in violation of the Internal Revenue Code and subject to an IRS audit

Lastly, this study sets out to enhance the body of knowledge on employment tax payment compliance by studying the employment tax payment behavior exhibited by sole-proprietors and small business owners. Although the literature on tax compliance is extensive, the literature has not taken a great interest in the issue of employment tax compliance and even less on the issue of employment tax payment compliance. It is not clear from the literature as to why this topic has not been addressed given its large economic implications, but unavailability of data and lack of record matching have been some of the possible reasons. With that said, there have been some changes to the tax code and advances in the technology, both of which may allow for additional studies of this particular segment of the tax base.

Nevertheless, the dearth of studies is puzzling given there is a consensus in the literature that small business and sole proprietors are more likely to be noncompliant than other filing groups (Rasholnikov, 2006). The current study allowed subjects to misreport, but it also instituted an audit regimen and imposed a penalty. By testing the changes in behavior of authorized individuals given the requisite reporting compliance regimen, we attempted to identify ways to improve voluntary compliance and minimize taxpayer burden, but also to better understand the behavior of employment tax payment compliance.

The dissertation was divided in six sections. Section 2 provides the practical motivation that drove this research study. Section 3 lists the central research question that this study set out to answer. Section 4 reviews the current body of knowledge, starting from the broad scope of taxes to a narrower scope of employment taxes and tax compliance. Section 5 provides the theoretical motivation and theoretical model used to

test the research question. Section 6 provides a synopsis to the literature review. Section 7 list hypotheses in this study. Section 8 details the methods used to achieve the research objective. The section will include experimental design and statistical tests. Section 9 provides the finding obtained by the experiment. Section 10 provides discussion and conclusion identified from the findings section. Section 11 list the limitations associated with the design and potential future research that could improve the body of knowledge relevant to employment tax payment compliance. All relevant forms and additional information are included in Appendices.

II PRACTICAL MOTIVATION

The practical motivation for this study was to improve our understanding of reporting and voluntary compliance of employment tax payment. It is possible that from this research study the Internal Revenue Service will modify certain employment tax reporting and payment processes. These changes could lead to a lower taxpayer burden, such as a reduction in filing costs and audit costs, thus enhancing and adhering to the Taxpayer Bill of Rights. On the other hand, this project should also provide the IRS with improvements to its tax compliance and tax collection process and policies. Therefore, the proposed outcome could improve timing and accuracy of employment tax payment, and could improve the accuracy of employment tax Federal Tax Deposits.

II.1 Employment Tax Motivation

The current tax code allows business owners to deduct legitimate expenses from the owners' tax income. This process is normally done using Form 1040 Schedule C – Profit or loss from small business (or sole proprietorship) or Form 1120 - U.S. Corporation Income Tax Return.² However, these legitimate expenses do not preclude business or nonprofit entities from having employment tax liabilities. The general misstatement of expenses by business owners may lead to underreporting and underpayment of all tax liabilities. These actions costs the US government and, ultimately the citizenry, billions of dollars and it is termed the tax gap³ (IRS, 2012b), and it unfairly impacts those who voluntarily pay their fair share of taxes. Of the estimated tax gap of \$450 billion dollars, \$122 billion are from Business income, \$67 billion are from corporate income taxes and \$72 billion are from employment taxes (IRS, 2012b).

² Examples of expenses that could be deducted are advertising, car expenses, and depletion.

³ The tax gap is the difference between taxes owed and taxes paid.

The tax gap is a good indicator of tax non-compliance (Andreoni, Erard, & Feinstein, 1998) and it accounts for approximately 15 to 17 percent of noncompliance (Mazar & Ariely, 2006; Winchester, 2009), but a larger percent of the tax gap is a direct result of non-filing and underreporting (Blumenthal, Christian, & Slemrod, 1998). Given the cost associated with enforcement, the IRS believes that improving voluntary compliance is necessary to reduce the tax gap (IRS, 2009).

In addition, the IRS is responsible for ensuring that taxpayers understand and meet their tax obligations. Among those obligations is the filing and payment of employment tax. Employment tax is the levies incurred by both an employer and employee when the employee is paid for services rendered. These levies, or tax liabilities, include social security, Medicare, Federal Unemployment (FUTA) and individual income tax. For all the aforementioned taxes, the two parties are not always responsible to pay or to match the amount owed. For example, the Internal Revenue Code (IRC or the code) requires employers to match the social security and Medicare taxes the employee pays after earning income. On the other hand, the code only requires employers to pay for FUTA, but it is not a requirement for employees. This filing process requires sole-proprietors, partnerships, and corporations to file and deposit employment tax every time an employee works and incurs taxes beyond a certain tax liability threshold.

Lastly, the IRS obtains employment tax filing and reporting information from employees (through W-2 or W-3 filing) and employers (through Form Series 94X filing). This set of information is used to assess the accuracy of employment tax reporting. To do so, the IRS matches the information on the W-2 forms against the

information reported on the 94X forms. Then, the IRS can identify cases where discrepancies exist (i.e., W-2 amount \neq Form 94X amount). For those cases with discrepancies, the accuracy of the employment tax filing by the firms are suspect and could lead to some form of an audit (field or correspondence).

Table 1: Employment Taxes Civil Penalties Assessed, Fiscal Year 2014

Employment taxes:	Penalties assessed (in \$1000)
Civil penalties, total	6,946,702
Accuracy	2,593
Bad check	480,174
Delinquency	1,429,205
Estimated tax	6,155
Failure to pay	3,495,287
Federal tax deposits	1,532,364
Fraud	455
Other	469

Source SOI Tax Stats - IRS Data Book⁴

As the data in Table 1 suggests, reporting for employment tax by businesses is relatively accurate and one of the contributing factors is the matching of information to determine inaccuracies and the corresponding high-level audits. The Table 1 also shows that accuracy only accounts for a negligible percent of cases where penalties were assessed.⁵ On the other hand, failure-to-pay accounts for 50.3 percent of penalties assessed. This is why this study focused on payment compliance of employment tax and not filing compliance.

II.2 Compliance Motivation

According to Bloomquist (2009) the best way to accurately determine taxpayer compliance would be to use random audits. Although small corporation audit coverage has increased over the last decade, there has been a substantial drop over the past two

⁴ This table is an extract from “Table 17. Civil Penalties Assessed and Abated, by Type of Tax and Type of Penalty, Fiscal Year 2014” obtained on 06/20/2015 at <http://www.irs.gov/uac/SOI-Tax-Stats-Civil-Penalties-Assessed-and-Abated-by-Type-of-Tax-and-Type-of-Penalty-IRS-Tax-Stats-Table-17>

⁵ For penalties to be assessed an audit must be conducted and changes to the filing information were made by revenue agent.

years (IRS, 2014a). Given the current IRS noncompliance identification methods, shrinking enforcement personnel budgets, and the cost of random audits, the IRS is trying to find ways to improve employers' voluntary payment, filing, and reporting compliance for employment tax.

On the other hand, TIGTA (2011, 2014a, 2014b) found that there are issues with the NRP employment tax database making it harder for the Service to identify compliance issues with filing and payment. This shortcoming makes it more difficult for the Service to identify compliance issues. Nevertheless, the current system is believed to be the most reliable method currently available (Andreoni et al., 1998). Furthermore, the IRS is currently using record matching criteria to identify potential field audit⁶. While a small percent of employer go through a filed audit, close to 100 percent of the employment tax filings go through an automated (computer base) audit. These audits could be characterized as computerize verification or automated audits, and taxpayers are seldom contacted or may only receive an IRS letter or notice. For this research, the focus was not on the automated process, but on the most labor-intensive audits that require a revenue agent to conduct a physical audit. The reasoning for this narrower focus was that physical audits are costly and IRS budgets are shrinking.

As the These initiatives included revising the collection due process for employment taxes and studying employment tax reporting and filing compliance (IRS, 2009) which included topics such as the Federal Tax Deposit (FTD) Alerts and using NRP to improve the employment tax compliance characteristics. This study built on the

⁶ IRS.gov defines a filed audit as an interview that may take place “at the taxpayer's home, place of business, or accountant's office”

simplification of employment tax payment, which Bloomquist (2003b) states it could lead to increased compliance.

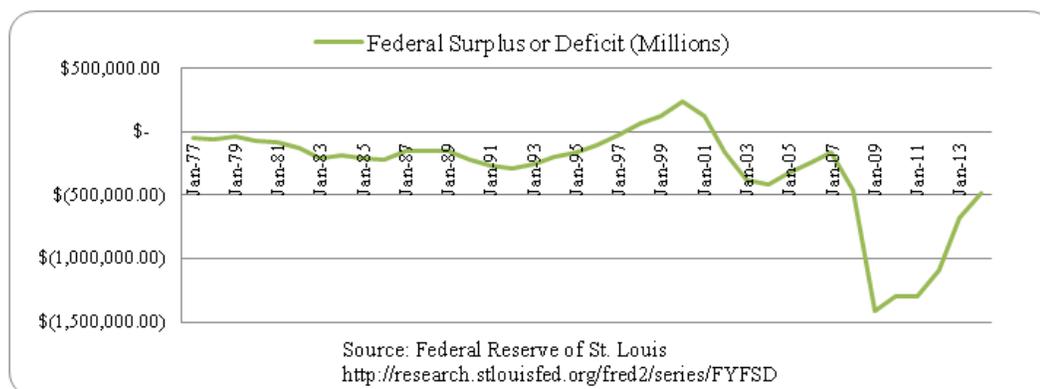


Figure 1 shows, the US government has been incurring deficits since the 1970s. Over most of that time, deficits did not account for a sizable percentage of GDP and during the late 90s surpluses were incurred. As a result of these long running deficits, the IRS began calculating and identifying ways to reduce the tax gap (Brown & Mazur, 2003; Holland, 1958; IRS, 1996). As a result of the ‘great recession’ and the raise of the US deficits to 9.7 percent of GDP (St. Louis Federal Reserve Bank, 2014), the IRS instituted initiatives to curtail the \$72 billion Tax gap from employment tax (IRS, 2009).

These initiatives included revising the collection due process for employment taxes and studying employment tax reporting and filing compliance (IRS, 2009) which included topics such as the Federal Tax Deposit (FTD) Alerts and using NRP to improve the employment tax compliance characteristics. This study built on the simplification of employment tax payment, which Bloomquist (2003b) states it could lead to increased compliance.

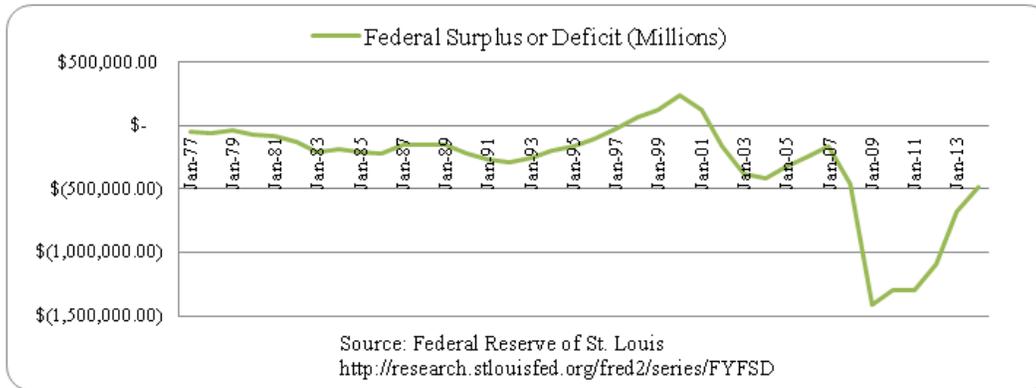


Figure 1: United States Federal Surplus or Deficits from September 1977 to September 2013

III RESEARCH QUESTION

The research question is whether different tax regimes, such as safe harbor, can change the behavior of employment tax payment for small business or self-employed taxpayers. The research is important in order to better appreciate the incentives that government agencies worldwide might be able to use to encourage voluntary tax compliance.

IV LITERATURE REVIEW

This section covers the key literature of employment (or trust) tax, reporting compliance and payment compliance. The literature streams associated with this study are extensive. The following highlight will provide information on the current state of knowledge and how the current research could increase the body of knowledge. The figure below depicts the study literature space and how employment tax, payment compliance, and reporting compliance overlap. Next, a general review on the tax literature will be provided. That will followed by an extensive literature review on employment tax, reporting compliance, and payment compliance.

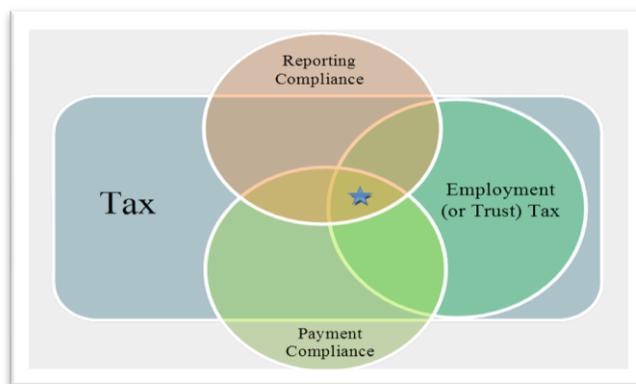


Figure 2: Literature Review Space

IV.1 Tax

Taxation is as old as human society, and corvée was one of its earliest forms. The concept of corvée required individuals in antiquities to provide labor in lieu of taxes. One of the first references of tax payment comes from the Bible – specifically Exodus⁷ – in the form of a poll tax. As one can imagine, the tax literature is extensive, with academic journal articles going as far back as 1815 (Jay, 1815). Needless to say, the

⁷ The amount was half a shekel for every grown man as stated in Exodus 30:11-16 from <https://www.bible.com/bible/1/exo.30>.

depth of this literature could be attributed to the direct impact that taxation has on aspects of everyday life. Some of these aspects are corporate strategy, personal income tax, and employment tax. Since this study is positioned within the tax literature, a general review of the tax literature provides some current issues discussed in the literature and how they may relate to the present scope of study.

Given the broad scope of the tax literature, our attention will be focused on issues in corporate and income tax because of their relationship to the literature in this study. In the domain of corporate tax, there has been plenty of work on tax planning, tax avoidance, and corporate responsibility⁸. The tax planning literature is related to law literature and tax avoidance and evasion are more deeply discussed in the economics literature (Hanlon & Heitzman, 2010). Corporate responsibility typically can be found in the ethics literature.

The corporate tax planning literature covers the impact that global tax jurisdiction has on multinational tax strategy. Buettner and Wamser (2013) define tax planning as the employment of “profit-shifting techniques that only require the adjustment of the internal structure of the multinational firm” (pg. 63). Essentially, companies are employing these strategies to reduce tax liability (Donohoe, McGill, & Outslay, 2014; Markie & Shackelford, 2012) with mixed or minimal results (Buettner & Wamser, 2013; Donohoe, McGill, & Outslay, 2013).⁹ Nevertheless, these tax strategies provide some competitive advantage and firms defend their use (Kubick, Lynch, Mayberry, & Omer, 2015; Manzon, Sharp, & Travlos, 1994).

⁸ Corporate responsibility is discussed in detail in the appendix.

⁹ It is important to note that tax avoidance is legal whereas tax evasion is not.

Tax avoidance, is the legal minimization of one's tax liability. Tax evasion is the illegal practice of not filing or paying taxes. Both of these tax strategies, which takes on different forms, has been around since the dawn of taxes. Some common avoidance techniques are tax shelters¹⁰ and complex hybrid securities (Gallemore, Maydew, & Thornock, 2014). Salihu, Sheikh Obid, and Annuar (2013) have defined tax avoidance as the action of tax planning or managing of tax liabilities. Tax avoidance can be measured in two ways: reported tax versus financial statements and the proportion of reported tax to business income (Hanlon & Heitzman, 2010; Salihu et al., 2013). Donohoe (2015) also stated, "Derivatives allow planners to negate, extend, or expand the formal arrangements and results on which tax liability is based" (pg.37), creating further tax avoidance.

As with corporate tax, the personal income literature has investigated critical issues related to taxation. Within this domain, topics include tax reform and tax credits. Of all the issues found in the income tax discourse, tax reform is normally at the top. The current Internal Revenue Code has more than 70,000 pages and over 9,000 sections, not including regulations and informative publications. Therefore, it is safe to say that the US tax code is complex and its citizenship is interested in reforming it. Barney, Tschopp, and Wells (2012) explain that "frequent change, excessive detail, ambiguity, burdensome recordkeeping, numerous calculations, and confusing forms" (pg. 9) makes the tax system more complex. Marcuss et al. (2013) estimated that the cost of tax complexity on taxpayers exceed \$150 billion in the US. The complexity is so sizeable that it remains the largest cost burden for small business, even after decades of simplification and deregulation initiatives (Chittenden, Kauser, & Panikkos, 2003). Ironically, even

¹⁰ Shelters may or may not be legal, but authors did imply legal action in their paper.

simplification measures develop into even more complex tax schemes once deductions and exceptions are considered.

Because of this, studies have established that complexity and vagueness of the law are good reasons for taxpayer noncompliance (Beck, Davis, & Woon-Oh Jung, 1992; Erard & Ho, 2001; Feinstein, 1991; Martinez-Vazquez & Rider, 2005). Therefore, ways to simplify the US tax code short of a zero-based approach should be welcome.

One approach critics have offered to simplify the code is the Value Added Tax (VAT). VAT is a consumption tax collected by retailers. The U.S. House of Representatives Bill 25¹¹ would impose a national sales tax of 23 percent.¹² However, economic research has shown that a simple sales tax could be less effective than the current income tax system. A Heritage Foundation report also explains how the introduction of a VAT tax would “expand the cost of government,” “inadvertently increase income tax rates,” and “slow economic growth and destroy jobs” (no page) (Mitchell, 2005). Jones, Thomas, and Lang (2012) argue that implementation and enforcement of VAT policy would still be complex and costly.

So why is there such an interest in simplifying the code and subsequently the filing process? Economist over centuries have asserted the need for tax policy to be easy to understand and easy to pay in order to minimize economic distortion, also known as deadweight loss (Pressman, 2014). Data analyzed from Peter, Buttrick, and Duncan (2010) has shown that tax complexity in high income nations has not changed over the

¹¹ The bill was introduced to the House on January 6th, 2015 by Georgia Representative Rob Woodall.

¹² The bill does have exceptions on “intangible property, for property or services purchased for business, export, or investment purposes, and for state government functions” (information obtained from <https://www.congress.gov/bill/114th-congress/house-bill/25>)

last two decades¹³ whereas tax rates have gone down during that same timeframe. They showed that progressivity of taxes has declined over the selected period, a trend which is indicative of a flatter tax policy (Peter et al., 2010).

Another way to improve the efficiency of the US Tax Code would be the reduction of deductions combined with lower tax rates. This would increase the tax base while reducing the marginal tax rate, a change which should lead to more compliance (Gordon, 1989). Another regime for achieving economic efficiency is a global cap on all deductions (Warren, 2014). A global cap on deductions would not eliminate the expenses, but it would reduce the maximum deduction amounts. By doing this, every taxpayer would have the same availability of deduction regardless of expenses (Nagin & Pogarsky, 2001). However, such a global cap would limit those taxpayers with high levels of individual deductions.

Another literature stream in the personal income tax literature deals with credits. One of the biggest antipoverty policies for working Americans has been the Earn Income Tax Credit (EITC) (Athreya, Reilly, & Simpson, 2010; Blumenthal, Erard, & Chih-Chin, 2005; de la Vega, 2013). Much of the EITC literature focuses on how it distorts¹⁴ the economy in terms of income, labor market, and health care. The EITC creates a distortion because individuals with the same yearly income have to pay distinctly different effective tax rates (Ordover, 2014). Thus, they would have different after tax income. Because of the distortion to income and continued compliance malfeasances (Blumenthal et al., 2005), there are a few noncompliance audit selection studies which

¹³ They explained that the tax complexity has not changed because adjustments to tax laws are made continuously in order to avoid bracket creep. They define bracket creep as the lack of inflationary adjustments to the tax code which push taxpayers into higher income brackets due to an increase in nominal income.

¹⁴ The literature showed that the distortions are both positive and negative.

use EITC as an indicator for improving IRS personal income tax audit selection (Erard & Ho, 2002; Martinez-Vazquez & Rider, 2005).

Blumenthal et al. (2005) describe some of the characteristics of EITC non-compliance. In their study, they determined that approximately one third of all EITC filings are from an ineligible EITC filer. Many of these unqualified filers were only able to pass some of the program's requirements and improperly claimed EITC in spite of this. The authors proposed that cheating filers perceived that auditors could portray their actions as an honest mistake.

In addition to affecting the marginal tax rate of qualifying taxpayers, EITC affects health care (Arno, Sohler, Viola, & Schechter, 2009; Averett & Wang, 2013; Kenkel & Schmeiser, 2014; Strully, Rehkopf, & Ziming, 2010). Averett and Wang (2013) argue that the increase in income, through a reduction in the effective tax rate, leads to an artificial improvement in the health behaviors of EITC recipients. In addition, research has shown that the EITC impact the marginal propensity to consume, a tendency which acts as an economic stimulus (Edwards, 2004). Lastly, the EITC improves labor participation and unemployment rates due to increased income, which improves willingness to work and marginal return to labor (Neumark & Wascher, 2011; Scholz, 1994). Overall, the credits, such as EITC, have an impact on the wellbeing of qualifying taxpayers, but could lead to noncompliance through miss use or interpretation.

Lastly, there is an extensive literature studying the impact of elasticity of taxable income with respect to a change in tax rates¹⁵ (Creedy, 2010; Gordon, 1989; Mosberger, 2011; Saez, Matsaganis, & Tsakloglou, 2012a; Saez, Slemrod, & Giertz, 2012b). Creedy

¹⁵ Many of the studies use average or marginal tax rate or net-of-tax rate. The net-of-tax is income after adjusting for taxes.

(2010) shows that the level of responsiveness to a change in tax rates (i.e., elasticity) is a good measure of income tax burden and an indicator of underreporting of taxable income. Dwenger and Steiner (2012) emphasized that the elasticity is a good measurement of income tax administration efficiency, but it is not a robust measurement of corporate tax efficiency. Their study reported that the tax profits did not decline with a decrease in the corporate tax rate and an inelastic taxable income. Unless the assumption that the market is efficient is relaxed, this counters usual economic theory.

In conclusion, tax literature is broad and deep. It is also subtly subdivided by specific taxation literatures, such as the income and corporate tax literatures. The present literature review provides some understanding of current topics in the general tax literature associated with this study. The literature stream provided in this section shows how legal corporate tax planning is linked to tax avoidance. The personal income literature showed how credits and deduction affects the behavior of individual taxpayers and how elasticity of income tax can lead to underreporting.

These literature streams are also related to employment tax and the compliance literature, which is at the core of this study. Specifically, much of the early literature of tax compliance looked at the impact taxes have on labor market and compliance, which is a major concern in the employment tax literature. In the next section, the scope of the literature review will be further narrowed to employment tax. Although the relationships between income and corporate tax may not be directly linked to the literature of this study, smaller businesses do try to minimize their tax liability in order to obtain competitive advantage; they engage in tax avoidance and corporate responsibility

(Kubick et al., 2015); and they report net operating loss (NOL) deductions like their larger counterparts (Planita, 2015).

IV.2 Employment or Trust Tax

The Internal Revenue Code refers to the U.S. Code: Title 26, which is divided into 11 Subtitles (A to K). There also are rulings and regulations associated with every subtitle further complicating enforcement, compliance, and filing of taxes. In addition, Congress continuously changes the code. Given this context, we next review the literature related to Title 26 U.S. Code Subtitle C (employment taxes).

The employment tax literature is extensive. However, a large portion of the literature focuses on the legal interpretation and consequences of rulings and law changes. For the purpose of the present study, legal reviews will be considered to be out of scope so that we can instead focus on empirical studies. The idea is to understand the current body of knowledge gained from empirical studies in order to understand the characteristics of this market segment and to find methods that would be most appropriate for scientific analysis. Much of the empirical employment tax literature concentrates on the tax incidence¹⁶. The tax incidence is the analysis of who ultimately has to pay any enacted taxes. This literature has focused predominately on how these types of taxes affect the labor market and employee wages, and very little tackles the issue of taxpayer compliance behavior.

IV.2.1 Employment Tax Defined

Employment tax references FUTA, self-employment, social security, and Medicare tax withheld from income. Employment tax is also referred to as trust tax. It

¹⁶ Tax incident is related to elasticity

takes the name trust because employees' contributions are held in-trust by the employer until their contributions are deposited with the IRS (Murthy, 2014). In the literature review, employment tax is also synonymous with payroll tax and these labels will be used interchangeably throughout the study.

An important definition related to employment tax and this study is the assess penalty and the duly authorized individual. Publication 15 states that a recovery penalty of 100 percent may be assessed to the firm (and/or duly authorized individuals) if filings and deposits do not meet the required obligations and if he/she willfully fails to pay (IRS, 2013; Murthy, 2014). Thus the assess penalty is the interested penalty a firm (or officer) must pay when it does not meet its tax obligations. A duly authorized individual is defined as an individual or officer who oversees the collection, accounting, and payment of employment tax (IRS, 2013). This definition is important since this was the group of individuals that we solicited to participate in this study. Furthermore, these individuals are legally responsible for any liability that occurs. Last, there is a dearth of research attempting to determine why authorized individuals approve illegal tax transactions, given the possible negative legal and financial implications.

IV.2.2 Employment Tax and the Tax Incidence

In economics, the term-of-art tax incidence¹⁷ is used as a framing for who ultimately is responsible for paying the tax. The general perception is that the individual consuming the product pays the taxes. However, that is not always the case. In terms of labor markets, the party with the highest price elasticity of labor will pay a smaller

¹⁷ tax incidence is further discussed in the Appendix D

portion of any employment tax. Past research have shown that this is more commonly “borne by labour” (pg. 189) (Hamermesh, 1979; Holmlund, 1981; Vroman, 1974).

Given employers are also subject to employment tax matching, employers may exhibit similar behaviors when facing the tax incident. Empirical evidence suggested that a sub-segment of small business – self-employed taxpayers – decreased filing compliance and decreased their payment of employment when employment taxes increased (Heim, Lurie, & Pearce, 2014).

Beyond the tax incidence and its impact on labor and wages, payroll tax dilemma occurs when a firm cannot meet its operational liabilities and it uses employment tax withholdings to pay for these liabilities (Godfrey, 2004; Mauldin & Wilder, 1997). Companies normally face a liquidity crisis during economic downturns or economic distress, such as the great recession (Grady, 2013). Grady (2013) indicates that many of these businesses continue to illegally borrow from withheld taxes for expanded periods of time, and found that the majority of those businesses will eventually fail. Bloomquist (2003a) also argues that a financial strain is one of the determinants of noncompliance for all taxpayers. One of the arguments of this study is that some of the authorized individual will succumb to the payroll tax dilemma when confronting a liquidity crisis (or strain).

A different argument from the payroll dilemma is that individuals decide to use employment tax withholding as a short-term loan without clearly understanding the negative financial implications. Another explanation is that they are simply taking a high interest loan from the federal government. As these business owners may see it, they prefer paying their supplier and staying in business than paying the IRS and go out of

business. There are cases where taxpayers consciously (or unconsciously) make late payment. Given the IRS is unable to identify these individuals until quarter's end, the withholdings could be used to pay other responsibilities. Then, the employers make withholding deposits at the end of the quarter, putting them back to legal standing with their federal tax withholding deposits for employment tax and extending their accounts payables.

Yet another explanation for this phenomenon could be derived from individual's perception of the outcome. This perception could be rooted in self-positivity bias where the individual expects a positive outcome or image (Mezulis, Abramson, Hyde, & Hankin, 2004) or underestimating risks (Dengfeng & Sengupta, 2013). In other words, these business owners are unable to see how their decision will lead them or their business to failure in the long run.

IV.3 Compliance

Like taxes, the issue of tax compliance is as old as the Egyptians (Erard, 1997). However, the in-depth study of tax compliance began around the late 1960s. Compliance can be defined as conforming or complying with a process. Roth, Scholz, and Witte (1989) define tax compliance as the complete, timely and accurate filing of tax returns. Compliance, or lack thereof, is the *raison d'être* for the IRS. Taxpayer compliance can be sub categorized into three areas: payment, filling, and reporting compliance (Brown & Mazur, 2003). According to Brown and Mazur (2003) these types are exclusive and exhaustive.

So how does the IRS define and enforce compliance? The definition of compliance or tax evasion can be derived from the Second District Court of Appeals and

eventual Supreme Court decision of Gregory v. Helvering 293 U.S. 465 (1935). In their decisions, the courts stated that individuals are not bounded under any patriotic duty to pay the Treasury or the IRS anything other than what is legally due.¹⁸ Thus, under the accounting principle of substance over form derived from the Supreme Court ruling, compliance is met as long as the information is complete and acceptable under the tax code. Otherwise, the act is noncompliant whether the act is deliberate or accidental.

In general, the IRS measures reporting compliance using the National Research Program (NRP).¹⁹ NRP is a stratified random sample of approximately 50,000²⁰ intensively audited federal income tax returns (Erard, 2002). From this data, the IRS develops internal algorithms for case scoring (known as DIF scores) and case selection (Andreoni et al., 1998). The IRS uses the following procedure to select potential high-risk noncompliant cases and to enforce compliance:

“The Unreported Income DIF (UIDIF) score rates the return for the potential of unreported income. IRS personnel screen the highest-scoring returns, selecting some for audit and identifying the items on these returns that are most likely to need review” (no page number) (IRS, 2006).

It is clear that improving DIF scores will lead to improve case selection and eventual voluntary compliance. Currently, the compliance in the US is estimated to be 83.1 percent (IRS, 2012a) and probability of an audit is relatively low (IRS, 2014a). Kalambokidis, Turk, and Blumenthal (2012) showed that approximately 10 percent of

¹⁸ The Supreme Court states: “The legal right of a taxpayer to decrease the amount of what otherwise would be his taxes, or altogether avoid them, by means which the law,” The Second District Court of Appeals states: “Anyone may so arrange his affairs that his taxes shall be as low as possible; he is not bound to choose that pattern which will best pay the Treasury; there is not even a patriotic duty to increase one's taxes” (69 F.2d 811).

¹⁹ The Taxpayer Compliance Measurement Program (TCMP) is the predecessor of the NRP program.

²⁰ Number obtained from [http://www.irs.gov/uac/National-Research-Program-\(NRP\)](http://www.irs.gov/uac/National-Research-Program-(NRP)).

subjects in an experiment would be fully non-compliant. Even though the majority of individuals will do the right thing, a consistent reduction in audit rates can erode the long-term voluntary compliance rates. To avoid the potential erosion of voluntary compliance and maintain the high level of voluntary compliance, some researchers propose increasing the use of record matching, such as the W-2s and the Information Reporting and Document Matching (IRDM) (Andreoni et al., 1998; Martinez-Vazquez & Rider, 2005; Scholz, 2003).

Record matching strategies has proven to be a useful to alleviate the tax evasion. For example, over 97 percent of wage earners are compliant (Clotfelter, 1983; Martinez-Vazquez & Rider, 2005; Slemrod, 1985). This segment of the taxpaying population has a very high level of record matching. In addition, 90 percent of investment income is compliant (Martinez-Vazquez & Rider, 2005; Slemrod, 1985) and this number is likely to increase with the introduction of the Form 1099-K in 2009. Again, this high level of compliance behavior can be attributed to a high level of record match. On the other hand 47 percent of small business and partnerships are compliant (Slemrod, 1985), and this low compliance rate is in part attributed to the low level of record matching. Some studies have shown that out of all business entities studied, partnerships are more complaint because it would require collusion from all the partners to underreport (Martinez-Vazquez & Rider, 2005).

This issue is not exclusive to type of taxpayer or form utilized to report. Noncompliance is also more evident in individual line items reported within certain form. For example, the majority of Form 1040 line items have some method of record matching. The one exception is cash contributions and this area exhibits a higher rate of

noncompliance (Blumenthal, Kalambokidis, & Turk, 2012; Turk, Muzikir, Blumenthal, & Kalambokidis, 2007).

In addition to of record matching, IRS Commissioner Koskinen (2015) expressed concern in a key note address at the Urban-Brookings Tax Policy Center about the potential erosion of voluntary compliance because of the impact of budgetary cuts on detection and enforcement. His concern is possibly due in part to empirical evidence, which has shown that increased probability of detection of noncompliance and enforcement will reduce the intentional and accidental noncompliance rates (Alm, Jackson, & McKee, 1992; Bloomquist, 2009; Fischer, Wartick, & Mark, 1992; Hasseldine, Hite, & Toumi, 2007). Indeed, Joulfaian and Rider (1998) calculated a seven percent compliance improvement when the probability of audit increased ten percent. In addition, studies have shown that there is a lag effect on compliance when the audit rate decreases (White & Woodbury, 1985). In addition Alm and Yunus (2009) have found that the presence of a continual tax audit (termed “persistence”) has a short and long term effect on compliance behavior. The sections below will provide a more focused review on filing and reporting compliance and payment compliance.

IV.3.1 Filing and Reporting Compliance

In general, taxpayers will report their incomes, expenses, credits, and deductions. Expenses, credits, and deductions will reduce taxpayers’ taxable income while compensations will increase them. Filing compliance is the complete, timely, and accurate submission of forms to authorities and the amounts disclosed to authorities on filed tax forms are considered reported amounts. Therefore, reporting and filing compliance is the complete, timely, and accurate disclosure of tax information on all

sources of compensation. Underreporting, or reporting noncompliance, is the difference between reported amount and actual amounts incurred during the year. It is simply the understated amount of tax liability owed at the time of filing (Plumley, 2006). Reporting noncompliance may be a product of overstating expenses, credits, and deductions or the understating of incomes.

One of the cornerstone of tax compliance literature is the simple model of rational crime theorized by Becker (1968). Becker's (1968) crime and punishment optimization model illustrates how the cost of the crime, the cost of the apprehension and conviction, and the number of offences work as determining factors for crime deterrence. This model is based on expected utility theory. Expected utility provides a model that explains attitudes toward risk. In this case, it is a model that explains attitudes towards the risk of an audit.

This model shows that as the social cost increases, deterrence increases. Thus, a higher number of offenses lead to a higher probability for deterrence because of the cost associated or incurred by each offence. On the other hand, if individuals are not punished for their offences, the cost of noncompliance decreases. This motivates taxpayers using "tax rate, detection probability, and penalty structure" (pg. 2) to reassess compliance behavior and potentially become less compliant (Fischer et al., 1992).

Conversely, the higher the cost of apprehension and conviction (or our synonym punishment) lead to a lower probability of deterrence. This negative correlation assumes, quite rightly, that detection is not perfect. This is due to the fact that the social cost of committing a crime can, at times, be less costly than the cost of punishing the perpetrator (i.e., cost of crime < cost of punishment) and must take in to account the fact that the

majority of noncompliant are marginally noncompliant making enforcement less advantageous for authorities. Therefore, authorities, such as the IRS, should at a minimum attempt to optimize enforcement where the cost of the crime is at least equal to the cost of punishment. Preferably, they should optimize enforcement by maximizing the dollars generated per hour of enforcement from each audit class.

The IRS could pursue this optimization via a model such as the one developed by Becker. However if detection and punishment were the only predictors of compliance, noncompliance would be much lower (Alm, 1991). Fischer et al. (1992) provide four categories that more completely measure of compliance and emphasize how detection probability is at the core of all the models. These categories of detection probability are:²¹

- Demographic: “age, gender” (pg. 2)
- Socioeconomic: “education, income level, income source, and occupation” (pg. 2)
- Normative Beliefs and Subjective Norms: “ethics, perceived fairness of the tax system, peer influence” (pg. 2)
- Tax law: “complexity of the tax system, IRS contact, sanctions, detection probability, and tax rates” (pg. 2)

The second cornerstone of the tax compliance literature was advanced by Allingham and Sandmo (1972), which was later refined in Yitzhaki (1974). As with Becker, this model is also based on expected utility theory. Allingham and Sandmo argued that rational individuals, when facing multistage situation, will analyze their filing decision using outcomes from prior filing and expected future cost of apprehension. Their model identified a positive relationship between the magnitude of a tax penalty and tax compliance. In other words, higher penalties will lead to higher compliance.

²¹ All variables mentioned in the categories are provided by the Fischer et al. paper. Other papers state that there are 14 variables, but none mentions which ones they are.

The Allingham & Sandmo model breaks down the decision process into several steps. First, taxpayers must decide how much to report, and that decision is followed by the decision to pay (Erard & Ho, 2001). Their original model was:

$$E(U) = (1 - p)U(W - \theta X) + pU(W - \theta X - \pi(W - X))^{22}$$

Where $E(U)$ is the expected utility; p is the probability of detection (or audit); W is income; X is reported income; θ is the tax rate; and π is the penalty rate. If W is greater than X , there is a tax gap. Simply, the right side of the equation provided a formula of the gains $[(1 - p)U(W - \theta X)]$ and the cost $[pU(W - \theta X - \pi(W - X))]$ associated with the final reporting decision. As mentioned above and using Formula, when the probability of detection (p) or penalty (θ) increases, reporting improves because the cost increases beyond the gains from underreporting.

Using this model, Clotfelter (1983), Gordon (1989), Alm et al. (1992), and Andreoni et al. (1998) argued that an increase in the tax rate will increase noncompliance when the probability of getting apprehended remains constant. For example, the Andreoni et al. line of reasoning showed that the after-tax income benefits increase as a result of underreporting while the expected cost of apprehension (i.e., audit) remain constant. If the tax policy's goal is to minimize tax evasion, any increase in taxes should be accompanied by an increase in tax enforcement or an increase in penalties to restrain any changes to the compliance rate. The result expressed by Andreoni et al. is similar to those observed with other assets or normal goods (Stiglitz, 1969) and what is expected using expected utility maximization under uncertainty (Alm et al., 1992; Clotfelter, 1983; Lee, 2001; Stiglitz, 1969). This notion is consistent throughout the entire literature.

²² There many more refinements of this model in the paper and in the literature, but investigator will use this model for simplicity. However, another simple model could have been used: the multistage model $E(U_t) = (1 - p)U(Y_t) + pU(Z_t)$ where Y_t is the gain and Z_t is the cost associated with being discovered.

Using the Allingham and Sandmo model, Lee (2001) asseverates that rational individuals will purposely become noncompliant and will self-insure in the event of an audit. He argues that rational individuals are willing to engage in tax evasion when the probability of detection is low (Lee, 2001) and he suggests that noncompliant taxpayers will use some of the proceeds from not paying taxes as a self-insurance policy in the event of an audit. This strategy becomes viable provided the proceeds from underreporting exceed the cost of audits over an extended period. On the other hand, if the audit rate increases, the scheme becomes less lucrative and the self-insured tax evaders will increase their compliance behavior.

In addition to tax rate and penalties, Allingham and Sandmo's model also explored the issue of asymmetry of information (taxpayers have more information about their transactions than the enforcing authority). They formulate a relationship between the uncertainty associated with tax declarations and asymmetric information. Their premise was that the presence of asymmetric information increases the level of uncertainty. This effect is more prevalent with the absence of third party matching. Their model showed how higher uncertainty coupled with the fact that underreporting by the taxpayer does not automatically trigger a reporting penalty or an audit will inevitably lead to underreporting.

Their model was able to measure the relationship between higher detection and improved income reporting, but was not able to find the relationship between income and reporting (Allingham & Sandmo, 1972; Alm et al., 1992). In addition, the Alm et al. (1992) study showed that due to the lack of reliable information on individual compliance choices and the presence of asymmetric information in any tax transaction may lead

taxpayers to take on riskier behaviors. This finding is aligned with other studies discussed above and the results could be rooted in an individual's perception of the likelihood of sanction (i.e., probability of detection). The Andreoni et al. (1998) and Lee (2001) studies indicate that detection and punishment are not the only predictors of compliant behavior. Alm (1991) explained that if detection and punishment were the only predictors of noncompliance, noncompliance should have increased in the US for all sources of income without third party matching. Furthermore, he explains that studies have shown taxpayers in many circumstances have failed to maximize their expected utilities. Given these shortcomings in the models (both Becker's and later Allingham and Sandmo's model), researchers have undertaken numerous studies to understand and explain the sources and characteristics of noncompliance. The next two subsections discuss the topics of sources and characteristics of noncompliance, which should provide a clear understanding of the discrepancies and inconsistencies observed in tax compliance behavior.

IV.3.1.1 Sources of Compliance (or Noncompliance)

There are different sources of tax compliance. Some of these differences are very subtle, and in other cases, the literature has not reached a consensus because of contradictory results. For example, game theory and principle agent models have shown that taxpayers' perceptions regarding the probability of detection are relatively accurate. Whereas, Scholz (2003) argued that the taxpayers' have "inaccurate beliefs about" (pg. 186) the probability of detection, and Alm et al. (1992) explained that in most case taxpayers will over estimate audit rate. Furthermore, those individuals who have a risk-seeking personality will modify their probability of detection perception more

significantly than those who are risk-averse (Schulz, 2014). Nevertheless, Andreoni et al. (1998) suggested that taxpayers are unable to determine the sources that triggered the audits or the magnitude of the penalties related to noncompliance. Moreover, individual with different risk aversion levels will modify their probability of detection after experiencing tax filing results or audits. This is the notion of persistence.

In addition, compliance may be impacted by the presence of a refund or an amount owed. Robben et al. (1990) suggested that taxpayers who are receiving a refund are more risk adverse because they are already owed money back and would be subject to a loss if cheating were identified. On the other hand, those who owe money have the lower perceived risk of noncompliance because they have already incurred a loss. If not apprehended, the losses are minimized.

Another source of noncompliance is asymmetry of information. Studies mentioned above have claim that individuals have asymmetry information and that the cost of punishment has an impact on compliance. However, up to this point, they have not explained what some of the sources of underreporting are. In most cases, firms and individuals will underreport their source of income wherever taxpayers perceive the sources are harder for authorities to identify (Johns & Slemrod, 2010). For example, the IRS reports that most of its tax gap underreporting is income understating instead of the overstating deductions or credits (Erard, 2002; Johns & Slemrod, 2010; Winchester, 2009). Studies have shown that income underreporting is subtle, and can materialize in different forms.

For example, Beck, Davis, and Jung (1991) explained that introduction of penalties and changes in audit probability impacts behavior. This is consistent with other

studies, but they determined that the changes are dependent on risk aversion. Alm et al. (1992) study also showed that the frequency of audit also has a positive impact on taxpayer compliance, and that reward could encourage compliance. In Robben's et al. (1990) study, businesses that were expecting a refund (a reward) in excess of \$1,000 were compliant 95 percent of the time. Whereby, businesses with balance due of \$1,000 were compliant 70 percent of the time. The same study showed that individual taxpayer compliance rate was 96 percent and 89 percent respectively. This example is called a "balance due effect" (pg. 684) (Joulfaian & Rider, 1998), and it implies that the tax liability threshold have an impact on taxpayers' decision to underreport.

Another source of underreporting was associated with the complexity of the return or high filing burden. Individual filling with additional schedules (especially Schedule C, D, and F) and forms were more likely to underreport (Blumenthal et al., 2005; Clotfelter, 1983; Feinstein, 1991). One explanation may be a consequence of increased burden. The burden on compliance has been estimated at \$100 billion dollars. Burden is calculated using the IRS Individual Burden Model, which accounts in dollars the hours and moneys spent on completing and filing tax forms. Assuming the taxpayers are willing to self-insure, as Lee explained, and they use a model similar to Allingham and Sandmo's model, it would not be hard to conclude that noncompliance in a highly complex return is a practical solution. Clotfelter also propose that different taxpayer segments will underreport differently. For example a wage earner, whose income is matched, will over state deduction and credits. On the other hand, businesses are more likely to underreport by boosting their operating expenses.

Similar to Robben et al. and Clotfelter, Martinez-Vazquez and Rider (2005) found that increase enforcement is an effective way of increasing overall compliance. Martinez-Vazquez and Rider (2005) study showed that the increased enforcement on a mode²³ will increase income compliance, but it will have adverse effect on the untreated modes. Their study illustrated how 46 percent of returns understated income and about half of these had overstated deduction, which was in accordance with other studies. Martinez-Vazquez and Rider argued that taxpayers will determine which modes authorities are focusing their enforcement (i.e., understating income or overstating deductions) to reassess their probability of detection. They argue this is an application of a substitution effect. Once filing results are reassessed, taxpayers switch from underreporting income to over reporting deduction or vice versa. Therefore, gains in compliance are driven exclusively by graters gains in the targeted mode, which have a larger dollar impact than the loses from the mode not targeted.

In addition, they argue that taxpayers' decision to be noncompliant is driven in part by the presence of third party verification, which is in line with other studies presented in this paper. Blumenthal et al. (1998) also found that income level has an impact on reporting when there is enforcement and recode matching. They found that lower income taxpayers were more likely to increase the income reporting after being informed of an impending audit. On the other hand, higher income filer reduced the tax income reporting after being informed of an impending audit. Blumenthal et al. (1998) and Erard (1993) argue that this may be due to the hiring of tax professionals, which prefer more aggressive reporting activities.

²³ These tax modes are defined as different methods of filing, overstate, correctly state and understate income and deductions. The modes used in the Martinez-Vazquez and Rider are income and deduction reporting compliance.

With respect to business compliance, this too is a source of noncompliance. Although, there has been relatively little research completed exclusively on business noncompliance (Alm, 1991), the presence of self-employment income indicates higher probability of noncompliance. Newer studies do provide some insight on this segment's compliance behavior. For example, sole proprietors have shown to underreport employment taxes liability as well as their income (Joulfaian & Rider, 1998; Kukk & Staehr, 2013; Pissarides & Weber, 1989) and the type of industry also has an impact on the compliance level (Nor, Ahmad, & Saleh, 2010).

More persuasive evidence that self-employed business are a source of noncompliance is that the reporting compliance for this group is substantially lower than those of wage earning tax payers. Their voluntary compliance is between 40 and 80 percent depending on the income source,²⁴ and much lower than wage earners. One of the study showing the voluntary compliance of employment tax around 50 percent lower than personal income (Joulfaian & Rider, 1998). While the manifestation of underreporting income falls in line with the rest of the filing population (Blumenthal et al., 2005; Erard & Ho, 2001), the magnitude of underreporting is greater for small business and sole proprietors (Joulfaian & Rider, 1998; Slemrod, 1985).

As with individual tax filing, the complexity of the code has an impact on small business and self-employment filings compliance. Small businesses use tax preparers because of the complexity of the code. Erard (1993) found that tax professionals, especially CPA and tax lawyers, are “more likely to approve of aggressive reporting activities” (pg. 167), which are more likely to end up audited and noncompliant. Also, Erard (1997) argues that taxpayers' perceived probability of detection is diminished

²⁴ Sources could rent farm and self-employment income.

when employing a tax preparer to complete and file their tax returns. He argues the hiring of tax professionals further increasing their chances of noncompliance of business and self-employment tax. Ironically, he proposed that voluntary compliance would increase if taxpayers would complete their own tax forms. His reasoning is that the results showed self-prepared returns are more compliant in subsequent years after an audit. Whereas, returns completed by tax professionals were less compliant in subsequent years after an audit.

Given that individual filings include some employment tax requirements through the filing of Schedule C or Schedule F, some studies have shown that these forms are an indicator of underreporting of employment tax and general noncompliance (Erard, 1997; Joulfaian & Rider, 1998). A recurring reason for the underreporting has been linked to the lack record matching in the past. Interestingly, sole proprietors and farmers, both of which are considered small business by the IRS, were exempt from record matching (Andreoni et al., 1998). This exemption implied that the probability of detection was lower. *Ceteris paribus*, the changes in the code and improved technological capabilities for these taxpayers²⁵ should lead to improved detection of noncompliance and eventually improve compliance rate for this population.

Another important small business self-employment finding relates to the balance due effect. The balance due effect is the amount that taxpayers has not paid by April 15 on his/her tax liability. This finding is important because how much a taxpayer owes on April 15 is an indication of noncompliance. Meaning, the greater the balance due amount

²⁵ The 2009 changes to Sections 6041, 6041A, 6050W, and 6051 requires third party payment settlement organizations to report payment transactions to the IRS for the purpose of record matching, and employment and income tax enforcement.

the more likely they are to underreporting and become noncompliant (Joulfaian & Rider, 1998).

Finally, noncompliance detection is dependent on systematic and unsystematic components and the results of audits are imperfect (Feinstein, 1991). All the variables discussed above could be considered systematic components. They are generalization of the population. On the other hand, unsystematic components are dependent on both the individual behavior of the taxpayer and the knowledge of the IRS' examiner. Feinstein (1991) correctly explain that the revenue agents have different grades and different capabilities between grades and the nature of the return they examine may be substantially different.²⁶ All these differences in returns and examiner level will lead to different probability of detection. The model developed by Feinstein is useful when developing a partial detection model because it takes into account complexity of the return and the ability of the examiner when assessing probability of detection. It is also useful IRS policy because the author explains that training and better employment screening would lead to better probability of detection. This is an issue the Service continues to have two decades after the publication of Feinstein's paper.

IV.3.1.2 Characteristics of Noncompliant Taxpayers

In the literature, there are two groups on noncompliant taxpayers: those who file and underreport and those who do not file. First, review of the characteristics of underreporters will be provided and then provide a brief overview the non-filers.²⁷

Collins and Plumlee (1991) and Feinstein (1991) studies provide evidence that underreporting increased as income level, effort or marginal cost increases. In addition,

²⁶ For example, one case may have a simple Form 1040. At the other end of the complexity scale, a case may have a Form 1040 plus several schedules, including schedule K-1 flow-through.

²⁷ A more in-depth discussion of non-filers is available in Appendix D.

studies have found that age, more specifically students and working adults, have a different “tax mentality” (pg. 20), implying different taxpayer behavior and sophistication (Beck et al., 1991; Blumenthal et al., 2005; Fischer et al., 1992). Lastly, many of the studies have found that prior filing is a good indication of future filing behavior (Erard & Ho, 2001).

Beck et al. (1991) also provide evidence that risk neutral individuals will increase their reported income “when penalty rates, audit probabilities” (pg. 536) and uncertainty increase. Their model also showed that risk neutral taxpayers behave the same way²⁸ as risk adverse taxpayers. In addition, they determined that the risk-neutral model is more robust than prior risk preference models. Furthermore, other studies have shown that filing status could be a predictor of noncompliance. More specifically, studies have shown that married individuals are more likely to underreport (Andreoni et al., 1998; Clotfelter, 1983).

With respect to employment tax, very little is mentioned in the literature about the characteristics of noncompliance. However, there is sufficient information about owners of small businesses to assess their compliance behavior. However to do so, self-employee compliance characteristics will be used as proxy to understand employment tax filing behavior because of the similarity in the population segment. Feinstein (1991) found that taxpayers’ filing with Schedule C (Self-employed) and Schedule F (Farming) are more likely to be noncompliant than other taxpayers. With respect to farmers, he found that much of the underreporting occurs with income item (adjusted gross income). Further complicating matters and increasing the probability of noncompliance, Erard

²⁸ Investigator used the word the same because there was no statistical difference between groups. Beck et al. used “marginally significant” when the p-value was .077.

(1997) found that many returns with schedules such as schedule C, D, E, F and K-1 will disproportionately use tax preparers and increasing the probability of noncompliance. Joulfaian and Rider (1998) also found that by including employment tax (i.e. Schedule C and Schedule F) the probability of noncompliance increases with an increase in the tax rate. Also if an individual has evaded taxes in the past without any enforcement response, there will be an “element of persistence” (pg.101) in the taxpayer’s compliance behavior (Alm & Yunus, 2009).

Now that some characteristics of underreporters have been discussed, a brief overview of non-filers is essential because their presence (or lack thereof) increase biases on the outcome. This subgroup of taxpayers also creates the appearance lower fraudulent levels, and understanding their behaviors improves our understanding of compliance (Feinstein, 1991). Erard and Ho’s (2001) explained that certain characteristics, such as hard to find, youth, married and blue collar, are traits of non-filers. Inevitably, complexity of filing and business income are at the core of non-filers compliance.

IV.3.1.3 Audit Regimes and Compliance

An audit is a “review/examination of an organization's or individual's accounts and financial information to ensure information is being reported correctly, according to the tax laws, to verify the amount of tax reported is substantially correct” (no page) (IRS, 2015). According to the IRS (2015), there are three ways and individual or organization may be selected for and audit: random, document matching, and related examinations. Audits could take place in person or simply a review of records. There are three results from an audit: no change, change with agreement, and change with disagreement.

The experimental design literature shows that this is a key variable in any compliance model for both the filer and the authorities (Alm et al., 1992). The probability of detection is analogous to the probability of an audit. Many studies in this literature have shown that individuals will alter their reporting and paying behavior based on probability of audit, and the severity and celerity of enforcement. As a matter of fact, the impact of an audit that results in an amount change is six times larger because audited taxpayers will interact with other taxpayers and it will cause other taxpayers to voluntarily comply (Alm, Jackson, & McKee, 2004). While this is a great signal to taxpayer by authorities, the audit rate in the United States had decrease from six percent to about one percent over the last four decades, but it depends on different criteria. The budgetary results associated with this reduction to the audit rates have been calculated to be approximately seven billion dollars (Alm et al., 2004).

Before providing further information on the relationship between audit and compliance, a definition of audit is in order. Collins and Plumlee (1991) study is one of the first studies to define alternative audit rules, and they have categorized them into three distinct rules:

- Random audit
- Cut-off audit
- Conditional (or strategic) audit

A random audit rule is when all taxpayers have the same probability of being selected (Collins & Plumlee, 1991). Cut-off audit rule establishes in advance a dollar threshold that triggers an audit. So any time a taxpayer reports an amount below the established threshold, an audit will be conducted (Andreoni et al., 1998). Conditional

rules use selection criteria and population segmentation to modify the probability of detection, and it uses the taxpayer's reported information for case selection (Clark, Friesen, & Muller, 2004). Selection criteria could be line items such as filing status, number of dependents or reported amounts, and population segmentations are "pools"²⁹ that can be created using prior audit results or case type (e.g., different forms – 1040-EZ, 1040 with additional schedules, and 1040 with partnership income).

As mentioned in the tax literature, the best way to accurately determine taxpayer compliance would be to use random audits. However, random audits are too costly. Thus, it could not be operationalized. The cut-off rule has shown to be useful, but only when taxpayers are risk neutral. Therefore, the cut-off audit is not viable as an optimization strategy because it is too restrictive to operationalize. Lastly, conditional "are designed to achieve regulatory compliance with fewer inspections than required by random auditing" (pg. 60) (Clark et al., 2004) and the impact of conditional audits decreases underreporting as income increases (Beck et al., 1992; Collins & Plumlee, 1991).

Another positive implication of audits is the existence of a lag effect, which makes taxpayers compliant over several periods (White & Woodbury, 1985). On the other hand, the impacts of audits are not estimated to last over a long period of time (Andreoni et al., 1998), but they do persist (Alm & Yunus, 2009). They are short-term events. So in terms of tax policy, using simple but frequent record matching audits that would signal 'surveillance' to average taxpayers may be a good policy for the IRS to

²⁹ As defined in Clark et. Al (2004) (pg. 70)

incorporate as part of their compliance efforts³⁰ and would minimize persistence of tax evasion.

In essence, the IRS audit strategies are similar to conditional (or strategic) audit rule used to determine the probability of detection and appear to be the best plausible option available for the authorities. On the other hand, Alm et al. (1992) warns that audit levels similar to those employed by the IRS have a low impact on reporting compliance in the long-run. This low audit rate combined with short term behavioral effects of audits suggest that enforcement has to be continual (again addressing the importance of frequency), especially for those individual who are risk loving or have not been audited for some time. Nevertheless, conditional tax audit appear to be the most successful audit regime.

Lastly, Beck et al. (1992) provide some insights on the important of implicit and explicit costs of an audit. Even when no change is the result of an audit, taxpayers have to incur implicit and explicit costs. These costs may include the hiring of a tax professional, increased hours in preparation and increase stress. Beck et al. (1992) claim that this increased cost due to an audit, which cannot easily be tested in an experimental setting, will increase the risk aversion of taxpayers. They expect these costs to be a source for less aggressive compliance behavior than those observed in their experiments.

IV.3.2 Payment Compliance

Filing compliance means the complete, timely, and accurate filing of taxes. If taxes are accurately filed, then payment compliance implies no underpayment and a timely payment. On the other hand, the literature has provided countless evidence to the

³⁰ This suggestion would be similar to the current IRS Automated Underreporter (AUR) and Correspondence Examinations.

existence of such examples are the balance due effect and the payroll tax dilemma. In the section below, a review of the payment literature is presented though it is important to stress that much of the payment compliance literature is linked to the general tax compliance literature and the models covered above. In addition, most of the literature focuses on filing and not exclusively on the payment of taxes. Many times, payment compliance appears to be a secondary element in studies rather than a primary research focus.

Nevertheless, the concept of payment compliance is important because it is essential to the notion of equity and efficiency. As some of the studies mentioned (Andreoni et al., 1998; Erard, 1997; Erard & Ho, 2001; Fischer et al., 1992; Gordon, 1989), the perceived fairness of the tax code will affect compliance in general. Without ensuring that all (or almost) taxpayers pay their legally incurred tax liabilities, a sense of inequity and inefficiency becomes pervasive in the consciousness of the tax base. In 1988, 40 percent of households underpaid their tax liabilities according to a tax gap analysis (Andreoni et al., 1998). This statistic along with the assertion that slightly over 90 percent of households voluntarily report their taxes (itself a low rate, we would argue) shows a disconnect between people's perception of filing compliance and payment compliance.

The payment of taxes on earned income is a requirement and taxpayers must make payments by year's end (IRS, 2014b). However, many individuals make ongoing payments using withholding or estimated tax payments in order to avoid potential penalties and interest. Therefore, taxpayers make calculated assessments of their expected withholdings and estimated tax payments on their income as the tax year

progress in order to avoid the penalties and interest. As part of this calculated assessment, taxpayers also seek to minimize the risk of providing an interest free loan to the government (i.e. federal and state). In doing so, taxpayers may underreport income and may underpay on the actual tax liability accumulated.

In the Feltham and Paquette (2002) study, the authors argue that individuals who do not adequately report and pay their tax withholdings or estimated tax payments were more likely to underreport. This idea is central to the current study because it could be anticipated that employers will underreport their employment tax liability whenever they are faced with a liquidity problem. This underreporting behavior should eventually morph into a behavior of persistent underpayment given the presence of uncertainty of an audit and perceived behavior control.

In conclusion, studies have shown that individuals will comply with payments when enforcement is increased (Alm et al., 1992; Bloomquist, 2009; Hasseldine et al., 2007). We also know from the literature mentioned above (Allingham & Sandmo, 1972; Andreoni et al., 1998; Becker, 1968; Clotfelter, 1983; Fischer et al., 1992; Lee, 2001) that authorities make calculated assessment to determine the benefits associated with the prosecution of crimes. On the other hand, the literature (Allingham & Sandmo, 1972; Andreoni et al., 1998; Becker, 1968; Fischer et al., 1992) also mentions that individuals make similar assessment to determine the benefits associated with reporting and payment compliance. Both of these calculations are formulated, in essence, by using asymmetric information and probability of uncertainty.

IV.3.3 *Income Shock*

Berloffa and Modena (2013) define negative income shock as the loss of revenues over a period. Hryshko (2014) suggest that households are only able to adjust to a negative income shock as long as the reduction of income is between 8 to 25 percent. Carroll (2009) further explains the difference between transitory and permanent income shock, where transitory income shock relates to short-term changes in income and permanent income shock relates to long-term income shock. In this study, the income shock was identified a transitory and it consists of a 16.5 percent reduction in revenue. The percentage selected is the median of the range provided by Hryshko (2014).

IV.4 Literature Review Summary

In conclusion, taxpayer compliance, like many other human behaviors, can take on subtle differences. Research has shown that marginal tax rate, probability of detection, and penalties will have all an impact on compliance. Additionally, social characteristics of the taxpayer will affect compliance. Some of these social characteristics are marital status and age. Furthermore, studies have found some economic characteristics that are associated with noncompliance. Most of these characteristics deal with the filing of self-employment or business tax.

In addition, there is consensus in the literature that compliance is also driven by the level of satisfaction with the government and the perceived fairness of the tax system. If individuals believe the government is properly working and everyone is paying their fair share of taxes, compliance increases. Lastly, the complexity of the tax system and the presence of third party reporting have an impact on the filing compliance. In other words, complexity introduces ambiguity. This ambiguity leads to overstating deductions

and credits and underreporting income. The lack of third party reporting makes it harder for an authority to verify and because of this asymmetry of information, taxpayers comply less often.

With respect to audit methods, the consensus is that conditional audit rules are the most effective way to address noncompliant taxpayer and to increase enforcement. These audit methods are similar to those used by the IRS. Even though the literature recognizes conditional audit rules as the best option, all the experimental studies reviewed in this section have used a random audit design. Moreover, many of these studies have random audit rates that are operationally unrealistic because of high levels of audit rates. None of the studies have addressed or recommended a standard to correctly address this issue and the tax compliance literature has identified this is as a limitation of prior experimental designs. Joulfaian and Rider (1998) have calculated the predicted probability of an audit to be between 4.3 and 5.7 percent. This range accounts for different conditional audit rule of income tax, but does not account for employment tax audit rates. Furthermore, these rates vary because of the IRS usage of discriminant functions analysis for the selection of cases (Beck et al., 1992), and this is a reason for why taxpayers are unsure about the probability of detection.

With respect to research methods, there is consensus in the literature that econometric models have been useful in identifying characteristics of noncompliance. There is also a consensus that experimental design can be used to better understand the subtle differences that complex econometric models are unable to identify due to redundancy (or multicollinearity).

V THEORY

V.1 Theoretical Motivation

The theoretical motivation of this study is to enhance our understanding of taxpayers' antisocial behavior using General Deterrence Theory. Deterrence is the impact that punishment has on those who have not committed a criminal act (Blumstein, Cohen, & Nagin, 1978). In terms of tax enforcement, deterrence is the impact tax audits and their accompanying punishments, such as interest, penalties, and imprisonment, have on those individuals who have failed to report and pay taxes their tax requirements. General Deterrence Theory has been used in different fields such as: criminal justice (Lawes, 1927; Schelling, 1958; Tittle, Botchkovar, & Antonaccio, 2011), international relations (Kalplowitz, 1973; Quackenbush, 2010, 2011b), economics (Homburg, Fürst, Ehrmann, & Scheinker, 2013; Rhee, 2012), and information systems (Chen, Ramamurthy, & Wen, 2012; Johnston, Warkentin, & Siponen, 2015; Lijiao, Wenli, Qingguo, & Smyth, 2014; Straub & Welke, 1998). The theory is based on the general idea of *homo economicus* – people are rational and they respond to incentives (Blumstein et al., 1978).

Using this basic economic principle, it would be safe to assume that rational individuals will partake in certain tax behaviors if the marginal benefit is equal or greater than the marginal cost. Consequently, tax decisions are determined at the margins (Phillips, 2014). The simple model of rational crime (SMORC) states that there are two methods to increase individuals' marginal cost: (1) higher probability of being apprehended and (2) increased magnitude of punishment (Allingham & Sandmo, 1972; Ariely, 2012; Upadhyay, 2013). Furthermore, studies have shown that individuals will improve their compliance rate when enforcement is increased or the enforcement of

punishment is creditable (Alm et al., 1992; Bloomquist, 2009; Hasseldine et al., 2007; Kalplowitz, 1973; Lijiao et al., 2014). Individuals, firms, and authorities will undertake a cost-benefit analysis to determine whether they should misbehave or prosecute, respectively (Allingham & Sandmo, 1972; Rhee, 2012). In addition, asymmetric information will play a part in the final analysis (Alm et al., 1992; Ariely, 2012; Mazar & Ariely, 2006; Vitell, 2003) and noncompliance will increase if transaction visibility decreases (Andreoni et al., 1998; Bloomquist, 2003b; Roth et al., 1989).

General Deterrence Theory provides an explanation as to how different levels of punishment influence antisocial behavior. Its economic foundation is based on optimization of expected utility, which is one of the most common theories used to explain tax compliance behavior. In general, the measurement of deterrence can be achieved by calculating the change in criminal rates at different levels of punishment (Blumstein et al., 1978). The current study will test the change in payment behavior when different levels of certainty (i.e., probability of detection) are imposed.

V.2 General Deterrence Theory

V.2.1 Background

In the older compliance literature, most of the theoretical focus was on economic theory centered on labor and utility theory. The newer papers refer to decision-making and maximization of expected utility function under uncertainty and risk aversion (Alm, 1991; Alm & Yunus, 2009; Lee, 2001; Martinez-Vazquez & Rider, 2005). Most of the studies on tax compliance refer to the Allingham and Sandmo (1972) model which is based on expected utility derived from a 1944 von Neumann-Morganstern (1974), which includes the inclusion of risk. The economic basis for General Deterrence Theory is thus

based on utility theory. Therefore, General Deterrence Theory, within the context of utility theory and the simple model of rational crime, will help to guide the study design, development of the research instrument, and explanation of results.

There are three principles in the General Deterrence Theory. They are celerity, severity, and certainty. The celerity is the frequency of punishment.³¹ Evidence has shown that increased in actual audits reduce noncompliance rates. (Alm et al., 1992; Bloomquist, 2009; Fischer et al., 1992; Hasseldine et al., 2007; Joulfaian & Rider, 1998). Severity is the magnitude of the punishment. Studies have shown that imposing penalties of 900 percent will deter individuals from becoming noncompliant (Morris, 2010). Certainty is the perceived probability of detection. Please keep in mind that, depending on selection criteria, the current IRS probability of detection is between 1 and 10 percent (Johns & Slemrod, 2010; Joulfaian & Rider, 1998). Rasholnikov (2006) suggested that decisions to evade are based on perceived probability of detection. As the literature on audits explained (Alm et al., 2004; Clark et al., 2004; Collins & Plumlee, 1991), higher perceived audit rate yield higher compliance.

Even though General Deterrence Theory is partially based on utility theory, it is also derived from Exchange Theory, which is often used in sociology (Grasmick & Green, 1980). This theory has been predominantly used in the fields of political science and criminology, but has seen wide use in information systems security and economics as well.

In the field of *political science*, General Deterrence Theory applies to the use of military threats, normally in terms of nuclear threat, to deter other global players from engaging in armed confrontation (Quackenbush, 2011a). More broadly, the theory assist

³¹ Celerity is not within the scope of these studies.

political scientist in understanding the resolution and de-escalation of different international and interstate conflicts (Zagare, 2007)

With respect to *information systems*, General Deterrence Theory is used to identify the antecedents that lead to system risks. Straub and Welke (1998) identified four areas that minimize “system risks” (pg. 441). They classified these areas as deterrence, prevention, detection, and recovery. Willison and Warkentin (2013) define these areas as:

Deterrence is the attempt to preclude unwanted behavior through incentives or disincentives.

Prevention is any effort used to preclude the commission of a crime.

Detection is the identification of committed antisocial actions.

Remedy (or recovery) refers to methods or strategies developed to block future commission of crimes.

Of the four areas identified above, deterrence and prevention are the best way to decrease noncompliance (Willison & Warkentin, 2013). Using these areas to minimize risks within an IT setting, an enforcement compliance cycle was constructed to illustrate where deterrence would apply to the tax noncompliance context.

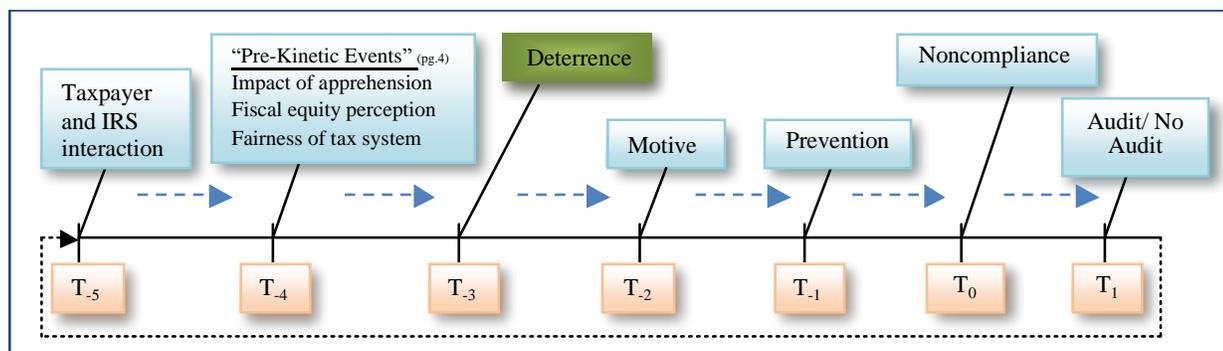


Figure 3 depicts the tax enforcement compliance cycle based on Willison and Warkentin (2013) and the extended security action cycle of Straub and Welke (1998).

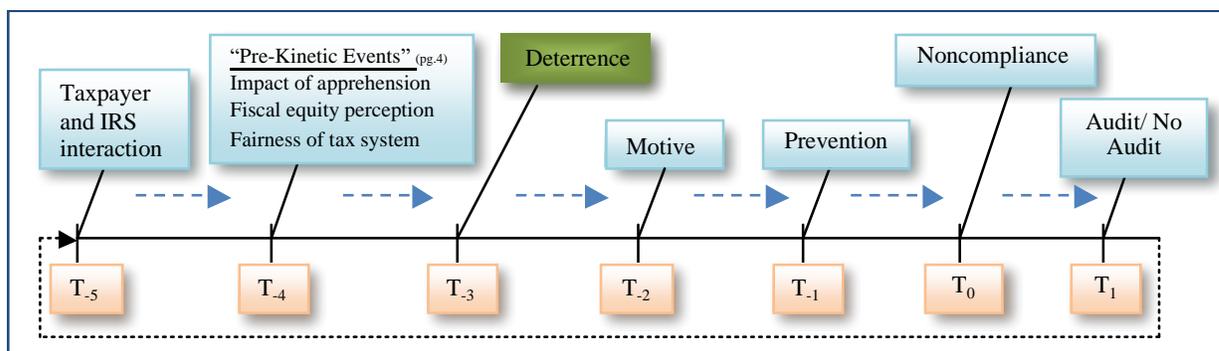


Figure 3: Enforcement Compliance Cycle

As the figure shows, deterrence occurs after an individual has interacted with the authority and has formed an opinion about the system. The idea behind implementing a deterrence strategy is to curtail the motivation to engage in antisocial behaviors. This process repeats itself every time when taxes are filed and paid.

With respect to *criminology*, General Deterrence Theory explains the use of punishment as a method to control antisocial or criminal behavior. This strategy is in line with the strategy proposed by the current IRS commissioner Koskinen, which relies on providing strong disincentives for taxpayers to comply with the code (Koskinen, 2015). For example, the IRS uses different deterrence methods for those who do not comply. Possibly the most common methods currently in use are penalties and interests (P&I). Depending on the violation committed, the severity of those penalties and interest could increase and can reach up 100 percent of the amount owed (IRS, 2013; Morris, 2010). In addition, there are costs associated to undergoing an audit and these increase the net cost of noncompliance and lead to improved compliance regardless of audit outcome. Noncompliant taxpayers can also incur interest on penalties, which further increases the

cost of punishment. These and other deterrence tools, including imprisonment (which is considered to be incapacitation or specific deterrence), encompass the arsenal available to the IRS in deterring individuals from becoming noncompliant.

The notion of deterrence is broadly accepted and the literature explicates individuals' embarrassment of apprehension (Gordon, 1989; Grasmick & Green, 1980) and "moral commitment to the law" (pg. 325) (Grasmick & Green, 1980) as being motives for compliant behavior. Embarrassment and commitment combined with the possibility of monetary and criminal punishment can coalesce to reduce antisocial behavior. In this study, the criminal or antisocial behavior is the deliberate or accidental tax payment noncompliance and much of the compliance literature argues the benefits of deterrence in curbing these antisocial acts.

The model below, derived from Apel (2013) and Grasmick and Green (1980) perceptual deterrence model, provides a visualization of how punishment deters individuals' antisocial behavior. In order to apply this model to employment tax compliance some relevant modifications had to be made, but the model gives a mental picture of how the current study will proceed to research the issues being raised.

The model shows how punishment, through sanctions, subjective norms, and normative beliefs, impact a taxpayer's perceived probability of detection, and how that change in perception further impacts compliance levels. The model is split into two general areas: the aggregate and the individual levels. The first construct is prescribed punishment. This is nothing more than the laws that govern then enforcement of tax noncompliance. The second column constructs are the general public's observation and perception of punishment. In essence, it is the perception the average taxpayer has once

he/she acquires second hand experience of the punishment received by lawbreakers or the interaction they have with other people in their lives.

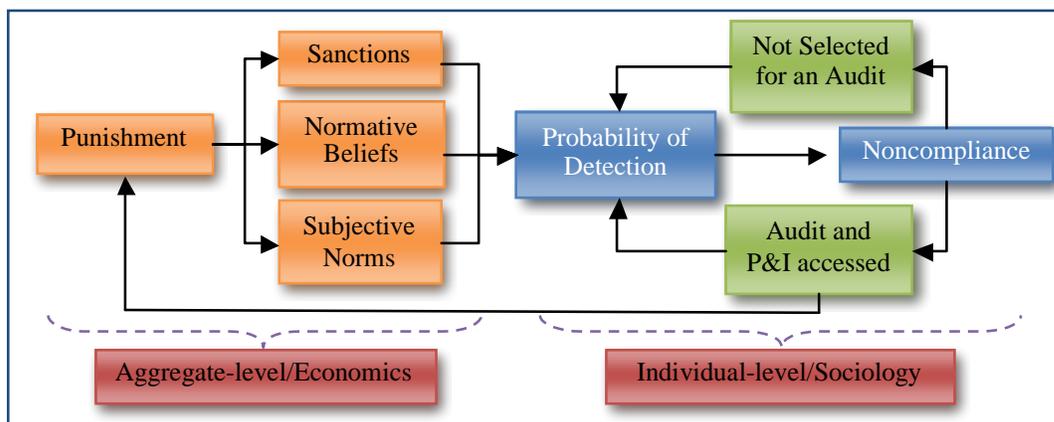


Figure 4: Deterrence Model

The next construct is perceived probability of detection. This construct distinguishes the aggregate-level from the individual-level. This construct is nothing more than the perceived audit rate, which most taxpayers perceive it to be higher than the actual rate (Alm et al., 1992; Scholz, 2003; Schulz, 2014). In this study, the introduction of a safe harbor provision will explicitly alter the probability of detection for one group. Therefore, it is expected that at the individual-level there will be an improved compliance once the provision is chosen. The next column reflects whether the taxpayer was selected for an audit. These two objects are dependent on whether or not the individual is compliant (the last set of constructs in the flowchart).

Now that an explanation of how General Deterrence Theory fits this research, a more in-depth discussion will investigate the three principles of General Deterrence Theory – celerity, severity, and certainty.

V.2.2 Principles of General Deterrence Theory

V.2.2.1 Celerity

Celerity is how quickly an individual receives punishment after committing a crime. To put it in a quotidian context, celerity is the time it takes for a parent to castigate a child after disobeying. The consensus is that the faster the punishment is enacted (especially with small children and pets), the more effective is the deterrence. In terms of the General Deterrence Theory, it is the imminence of punishment. Nagin and Pogarsky (2001) propose that celerity, unlike severity and certainty, does not deter antisocial behaviors. The argument is that sensible adults, unlike children and pets, do understand the connections between delayed consequences of punishment and the associated antisocial behavior. In other words for deterrence to work, immediate punishment is not required.

In this study, the effects of celerity are not being studied. However, its effects could be studied by charging penalties and interest at different time intervals (e.g., 6, 9, 12 months) after an audit identifies noncompliance.

V.2.2.2 Severity

The Oxford Dictionary defines severity as the fact or condition of being strict or harsh (Oxford Dictionaries, 2015). Within General Deterrence Theory, severity of punishment is one of the biggest factor that influence the commission of antisocial behavior (Grasmick & Green, 1980). According to Apel (2013), the severity of punishment is the first principle that must be established in order to create deterrence. The severity of an offence is determined by the law. The assessed punishments are “remedies” for offending. Turk et al. (2007) explain that individuals’ filing behaviors are

altered once a penalty and interests are imposed and Morris (2010) says that the magnitude of the punishment has an impact on filing compliance.

This principle has been measured in prior experiments using three methods. The first is to indicate what the law prescribes. The second method is for participants to select the penalty they are most likely to receive. Lastly, participants would estimate the probability of a penalty given an offence committed.

In the case of employment tax, severity of punishment includes all the different deterrence tools the IRS currently has at its disposal and these are prescribed by the code. However, a duly authorized individual, as defined by the code, cannot avoid personal liability from these taxes. In other words, even when the firm dissolves, the employment tax liabilities do not dissolve, but are simply transferred to duly authorized individuals. This transfer of employment tax liabilities is an additional punishment of the duly authorized individuals for failing to pay employment taxes.

For the purpose of this study, punishments assessed will also be prescribed as in the tax code. Because the code dictates punishment and there is no expected change on the horizon, there will be no adjustment to the severity of punishment as specified in the code.

V.2.2.3 Certainty

Certainty references the degree of certainty or a probability that an eventuality will occur (as in the certainty of getting caught cheating on taxes). In other words, it is the likelihood that an event will take place, given prior information about this event. Within the scope of this research, certainty is the risk of detection (Apel, 2013). Within the tax literature, this is defined as the probability of detection or audit. Within General

Deterrence Theory, severity of punishment is the other biggest factor influencing antisocial behavior (Grasmick & Green, 1980), but, according to Nagin and Pogarsky (2001), it is the most influential factor in deterring antisocial behaviors.

This principle simply expresses the notion that the higher the degree of certainty of being detected, the less likely an individual is to commit an antisocial act. Currently, the probability of an audit is very low in the US. However, the probability of an audit for employment tax is extremely high.

Federal Tax Deposit (FTD) Alerts, which identify employment tax payment noncompliance, are done on a quarterly basis. These alerts are based on a systematic review of employment tax payment over one year. Given the systematic review, employers are more likely to be audited by the IRS. Apel (2013) concluded that when an individual experiences a form of punishment (e.g., an audit), this experience will have an impact on the perceived certainty of punishment.

V.3 Research Model

The research model provided in this study is based on the Allingham and Sandmo (1972). As stated above, their original model was:

$$E(U) = (1 - p)U(W - \theta X) + pU(W - \theta X - \pi(W - X)) \quad (1)$$

Where $E(U)$ is the expected utility; p is the probability of detection (or audit)³²; W is income; X is reported income; θ is the tax rate; and π is the penalty rate. If W is greater than X , there is a tax gap. Simply, the right side of the equation provided a formula of the gains $[(1 - p)U(W - \theta X)]$ and the cost $[pU(W - \theta X - \pi(W - X))]$ associated with the final reporting decision. As mentioned above, and using formula,

³² Since the probability of automated/computerized audit is close 100 percent for employment tax, the enforcement rate will be used to develop a more accurate constructs experienced by taxpayers.

when the probability of detection (p) or penalty (θ) increases, reporting improves because the cost increases beyond the gains from underreporting. This same model could be restated as a multistage model using the following equations:

$$E(U_t) = (1 - p)U(Y_t) + pU(Z_t) \quad (2)$$

Where Y_t are the gains over t-time and Z_t is the cost associated with being discovered during that same period. By introducing a safe harbor provision, the probability of detection is zero (0); therefore, penalties would also be zero. On the other hand X would equal W , and the new equation for those how choose the safe harbor provision will be:

$$E(U) = U(W - \theta W) \quad (1a)$$

$$E(U_t) = U(Y_t) \quad (2a)$$

It is important to remember that for those who do not choose the safe harbor provision, Allingham and Sandmo (1972) model would still apply.

The model below, derived from Allingham and Sandmo (1972) model, provides a visualization of how the certainty of punishment deters individuals' antisocial behavior, holding celerity and severity constant.

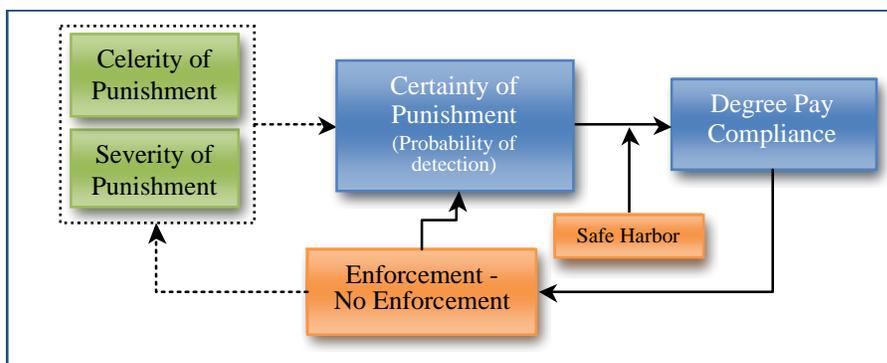


Figure 5: Research Model

VI LITERATURE REVIEW CONCLUSION

Taxpayer compliance can be sub categorized into three groups: payment, filling, and reporting compliance (Brown & Mazur, 2003). Of these sub categories, this study will focus on payment compliance. While there has been a plethora of reporting and filing compliance research, the same cannot be said about payment compliance (relative to the other two sub categories). The main reason for focusing on payment compliance is that the majority of employment tax noncompliance arises from payment noncompliance. This is a direct result of high level of information matching from a third source, which makes it easy to detect filing and reporting noncompliance. As the literature has shown, noncompliance occurs where transaction visibility decreases.

The practical motivation of this study is to better understand this compliance sub category and market segment, which account for \$72 billion of the tax gap. Furthermore, the Service has indicated an interest in improving voluntary compliance of employment tax, and this research study could provide some insights on how to collaborate with this important market segment in an epoch of diminishing budgets.

The theoretical motivation of this study is to enhance our understanding of taxpayers' antisocial behavior using General Deterred Theory. Deterrence is the impact punishment has on those who have not committed an antisocial act (i.e., the externality of punishment). In terms of tax enforcement, deterrence is the impact tax audits and consequential punishments have on those individuals who have failed to meet their tax requirements.

The General Deterrence Theory provides explanation as to how difference levels of punishment influence antisocial behavior. Its economic foundation is based on optimization of expected utility, which is one of the most common theories used to

explain tax compliance behavior. In general, the measurement deterrence can be achieved by calculating the change in criminal rate at different levels punishments (Blumstein et al., 1978). This study will test the change in payment behavior when different levels of certainty (i.e., probability of detection) are imposed

The research question is whether different tax regimes change the payment of employment tax behavior for small business or self-employed taxpayers. The idea is to determine if a safe harbor provision can reduce the proclivity of authorized individuals to implement a payroll tax dilemma strategy and whether or it can reduce payment noncompliance in time of economic distress (i.e., income shock).

The experiment will follow a similar design to Alm et al. (1992). Subjects will be able to earn income. In addition, subjects will have a set number of employees who will work a set number of hours at a set wage price. Using this information, subjects will voluntarily report and pay employment taxes. Once filing and payment of employment tax is completed, subjects will be subject to a random enforcement. If underreporting and/or underpayment are detected during enforcement, the subject will be instructed to pay the owed amount plus a 100 percent penalty on the amount owed (as prescribed by the code). After several rounds, individuals will be asked to participate in a safe harbor program. After several rounds under a safe harbor provision, an income shock (i.e., lower revenues) will be introduced to test the viability of an employment tax safe harbor provision.

VII HYPOTHESES

From the prior literature and theory development sections, there are four hypotheses:

Hypothesis 1: Business owners with higher tax liability thresholds will subscribe to a safe harbor provision at a lower rate than business owners with the lower tax liability thresholds.

Hypothesis 2a: Business owners who experience enforcement will have higher subscription rate to a safe harbor provision than those who do not experience enforcement.

Hypothesis 2b: Business owners who have experienced enforcements³³ will have higher compliance rates than those who do not experience enforcement.

Hypothesis 3: Business owners who chose the safe harbor provision will be more compliant, with respect to payment.

Hypothesis 4: Business owners who chose the safe harbor provision will remain more compliant after an income shock³⁴ is experienced.

³³ In this study, experiencing enforcements means undergoing an audit.

³⁴ Income shock is defined as a 50 percent drop in firm revenue

VIII METHODS

Many of the models created over the past several decades have allowed researchers to advance our learning about taxpayer behavior. Of those methods used to acquire understanding, experimental design appears to be well suited to identify and explain the discrepancies and inconsistencies associated with the economic models. This method allows researchers “control over extraneous influences” (pg. 107) (Alm, 1991).

The experiment followed a similar design to Alm et al. (1992), using scenarios to examine options and make choices in the laboratory setting with the hope of mimicking a real life scenario. To achieve this, subjects were able to earn money. However, the subjects were not able to make it based on how well they performed throughout the experiment because of the delivery method and IRB constraints, as originally prescribed. In addition, subjects were informed of important information about labor (e.g., number of employees, number of worked hours, employee wages). Using this information, subjects were asked to voluntarily report and pay employment taxes. Once filing and payment of employment tax was completed, participants were subject to a random enforcement.³⁵ If underreporting and/or underpayment are detected during enforcement, the subject was informed of audit results (change/no change), and if applicable, was instructed to pay the amount owed as prescribed by the code. After several rounds, subjects were asked if they wish to participate in a safe harbor program. After several rounds under a safe harbor provision, an income shock (i.e., lower business revenues) was introduced to test the viability of an employment tax safe harbor provision.

³⁵ Since the conditional audit strategy for employment tax enforcement is unknown, a simple random enforcement rate will be used.

It is important to know that this study did not attempt to identify different modes of compliance amongst taxpayers. The interest of this study was focused on detection of payment noncompliance and to potentially identify ways to encourage taxpayers to voluntarily comply. In addition, this study was trying to determine if a safe harbor could reduce the taxpayer burden, which accounts for \$100 billion dollars per year in deadweight loss to the economy.

VIII.1 Experimental Design

To answer the research question, an experiment³⁶ was employed to examine business owners' behavior when making employment tax payments. The experimental design was an impact study. An impact study measures the effects of enforcement and/or the safe harbor provision on the subjects partaking in the experiment. This method has been selected because it tends to have strong internal and conclusion validity (Bhattacharjee, 2012; Fischer et al., 1992; Trochim & Donnelly, 2007c). This type of experiment would only have strong external validity if the study took place in its normal environment, as in a field experiment. Some realism is enacted by having the subjects are real world business people and this enhances external validity even in a lab experiment.

VIII.1.1 *Validity*

As mentioned earlier, the method utilized in this study was selected because it tends to have strong internal, external and conclusion validity. **Internal validity** indicates whether observable changes to compliance are in fact affected by the

³⁶ In an ideal setting, this study would be “field” experiment by randomly selecting actual business owners from know IRS data sources. It would not use students in a lab nor would it use a purposive sampling. The business owners would follow the lab experiment, and they would be randomly selected.

introduction of a safe harbor provision. To ensure internal validity, Bhattacharjee (2012) suggest meeting three conditions – covariation of cause and effect, temporal precedence, and no plausible alternative explanation. By meeting these conditions, the experiment is able to manipulate the independent variables through treatments and observe its effects on the dependent variable and control for extraneous variables. To do so, random assignment was used. By utilizing random assignment, it is expected that this limitation is diminished and external validity is improved (Heinsman & Shadish, 1996; Ong-Dean, Huie Hofstetter, & Strick, 2011; Shadish, Clark, & Steiner, 2008). In this process, two groups were created in which subjects with similar employment tax characteristics had the same probability of assignment and neither study administrators nor subjects had the capability of choosing which group they are assigned.

External validity indicates whether observable changes can be generalized to the population. The easiest way to achieve external validity is to draw a random sample from the population. Although the targeted population is known for a given point in time, this information is not easily accessible. Therefore, the finding will be based on the group of subject that participated in this experiment.

VIII.1.2 *Experiment*

For the pretest, eight known associates completed the study. Subjects were recruited from the Georgia State University EDB student and alumni body and known associates from the accounting field (e.g., CPAs) and business owners. These individuals first completed the experiment. Then, they were asked a few open-ended questions about the flow of the experiment, and the content to minimize

misunderstanding by future subjects about employment tax, the safe harbor provision, and Form 94X.

Themes did materialize with the initial interviews about the information provided in the balance sheet ([Appendix A](#)). Minor changes were made to provide some relevant information pre-test subjects thought were necessary to make informed decisions. Before the web experiment went live, a Qualtrics team evaluated the experiment's logic flow to ensure proper random assignments throughout the experiment.

The other step is the experiment. In this step, subjects were randomly assigned into one of two groups (high tax liability threshold and low tax liability threshold). There were 105 subjects randomly assigned to low tax liability threshold and 100 subjects randomly assign to high tax liability threshold for a total of 205 subjects.

For subjects to continue their participation in the experiment, they had to agree to the informed consent, and selected that they reside in the United States. They also had to identify if they ever had any personal involvement with the employment tax process by doing any of following:

- Filed employment taxes with state or federal agencies
- Made employment taxes payments to state or federal agencies
- Supervised the filing or payment of employment taxes to the state or federal agencies
- Assisted a business with the filing or payment of employment taxes to the state or federal agencies

This question satisfied experiment criteria: (1) Has the recruit filed employment taxes or paid employment taxes in the last year and/or (2) Has the recruit supervised the filing of employment tax or supervised the payment of employment taxes in the last year?

Once the subject has decided to participate and fulfilled the informed consent and the criteria requirements, he/she would be able to complete the web-based test. Before

random assignment, the test requires subjects to responding to a series of questions to assess their perception of audit, misreporting, and penalties.

After that, subjects were randomly assigned to one of two groups: high or low tax liability thresholds. Research has shown that individuals with higher tax liability threshold exhibit higher risk seeking behaviors, and individuals with lower tax liability threshold exhibit risk adverse behaviors (Robben et al., 1990) and these results are expected to be exhibited in this study. After random assignment, all subjects had to complete a practice round, which required subjects to respond to a battery of questions. These questions were intended to create a realistic scenario with respect to filing and payment of employment tax. Questions mimicked the filing and payment of a simple IRS Form 940. The intention of a simplified Form 940 was to minimize confusion and boredom.

Since the literature has indicated that tax decisions evolve using information obtained over the course of several years, this process of responding to the battery of questions was repeated six more times. This was done to follow the recommendations provided by Alm (1991), which recommends utilizing several rounds to enhance the sense of realism. With that said, Alm (1991) does suggest that lab experiments are unable to perfectly replicate real life compliance decisions that can be influenced by normative beliefs and subjective norms. Although this study was intended to simulate six employment tax filing and payment cycles, subjects were told that there are an unknown number of rounds in this experiment.

It is expected one of the most important pieces of information an individual will continuously assess is the probability of an audit. As the literature showed, this

probability will shift depending on whether or not the individual was audited after underreporting and inevitably underpaying (Andreoni et al., 1998; Turk et al., 2007). Since the probability of audit (at least automated audit) is approximately 100 percent for employment tax, the enforcement rate will be used to develop the construct experienced by taxpayers. The enforcement rate indicates to taxpayer when the tax authorities physically (or through a letter) interact with taxpayers.

During the experiment, subjects had the opportunity receive income and pay taxes for their firm. In addition, individual had the opportunity to report (or underreport) their employment tax liability and tax payment. If enforcement was undertaken and subjects were found to have underreported employment tax, penalties were imposed. No enforcement took place in the first round, but all other rounds were subject to a random enforcement. The experiment is divided into four phases: random assignment, random assignment with enforcement, self-assignment, and income shock. The first phase consisted of one round and it provided the baseline measure. The second and third phases consist of two rounds each. The last phase consists of one round. [Figure 6](#) illustrates the round process (see [Section 8.3](#) for more details). After the third round was completed, individuals had the option to accept or reject a safe harbor provision. A safe harbor provision specifies conditions that protect individuals or firms from being deemed in violation of the Internal Revenue Code and subject to an IRS audit. Examples of the safe harbor rules are 401(k) rule, W-2 wages rule, telecommunication conversion of capitalized assets under §263(a), and electric transmission and distribution property under §263(a) or §162. The following two rounds will provide the control parameters for the safe harbor provision.

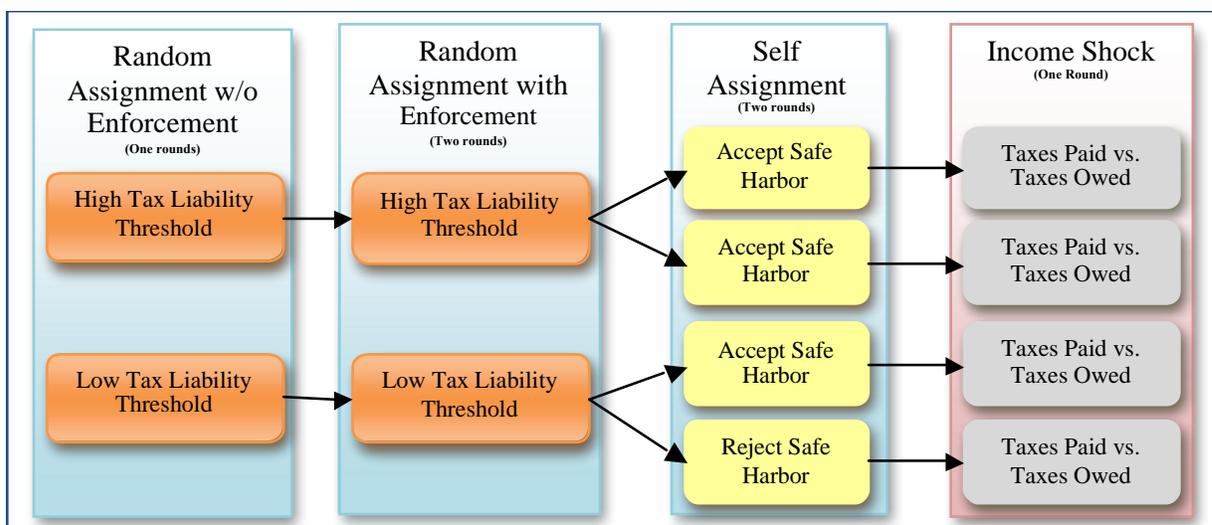


Figure 6: Test Phase Process

Once subjects have completed five rounds of employment tax filing and payment simulations, subjects completed one additional round in which they were exposed to an income shock (intervention). This intervention was to determine if subjects fell into the payroll tax dilemma under a safe harbor provision, and to see whether the provision minimizes the discrepancy between taxes paid and taxes owed for those who subscribe to it.³⁷ Secondly, the intervention allowed the investigator to test the differences between those taxpayers who have large and low tax liability.

VIII.1.3 *Social Desirability Bias*

Social desirability bias is associated with individuals participating in surveys. In essence, individuals try to respond to questions in a manner that satisfies the proctor of the survey or respond to questions in a socially desirable way – “fake good or fake bad” (pg. 394) (Crowne & Marlowe, 1960). Krumpal (2013) suggest that the individual is engaging in the self-preservation of his/her image. Social desirability bias tends to occur

³⁷ The discrepancy will be defined and the “willingness to pay” in the Data Analysis section of the prospectus

in areas of personal or social discomfort, such as, personal weight, voting, taxes, drug use, and sexual preference (Gittelman et al., 2015; Myung-Soo, Nelson, & Kiecker, 1997). This bias tends to magnify the end of the scale responses (e.g. very satisfied) (Myung-Soo et al., 1997). It becomes necessary to guard against social desirability bias because of its impact of survey responses and because this experiment engages on a socially sensitive issue – taxes.

Some of the research studies have shown that social desirability bias takes two forms: individuals characteristics and survey characteristics (Krumpal, 2013). Since the investigator was able to test for individual characteristics, investigator focused on improvising the survey characterizes in order to reduce social desirability bias. One method that has shown to reduce the social desirability bias is through use of indirect (Fisher, 1993) and direct questions (Myung-Soo et al., 1997). The difference between indirect and direct questions is that indirect questions ask the subject to project on other people versus making statements about him/her. The idea of asking both we check for social desirability bias while maintaining measurement validity.

Another method for reducing social desirability bias is to portrait some conjectural distance to the tasks in the study. For example, the instructions in the experiment could tell the subject that they are a contractor for a small business or a typical small business owner. In both cases, the instructions are setting the subject to think in a certain manner. This method was used in this experiment.

Burton-Jones (2009) also explains that social desirability measurement can be distorted from three sources: raters, instrument, and process. The rater error could be minimized by clearly defining the rater (Burton-Jones, 2009). For this experiment, the

rater was a duly authorized individual of a small business. For this experiment, the raters were asked, “to play the role of the typical small business owner with several fulltime employees.”

VIII.1.4 *Probability of Detection or Enforcement*

The IRS audits selection is not random (IRS, 2006) and its probability is not constant (Andreoni et al., 1998). As the literature shows, the IRS audit strategies are similar to conditional or strategic audit rule used to determine the probability of detection and appear to be the best plausible option available for the authorities, in terms of maximizing compliance. Although the actual audit rate for employment tax is unknown to the public, the IRS audit levels of employment tax can be expected to be relatively high³⁸ and does have a high level of third party reporting. Most of the third party matching is completed automatically and only those automated audits that appear to have irregularities are more closely examined and experience enforcement.

The examinations that require closer review are the type examinations taxpayers associate as undergoing an enforcement action. Since taxpayers do not experience any form of automated audit unless selected for further examination, enforcement rate will be used when discussing employment tax audit. Enforcement rate is a better construct for this experiment since taxpayers are more likely to experience the enforcement process than the audits.

Given audit rates are high and the expected enforcement rate also relatively high for the remaining identified taxpayers, a 50 percent random enforcement regime was used in this experiment. The

³⁸ The employment tax deposits are analyzed for payment compliance in all businesses that have a weekly and semiweekly deposit requirement. This analysis is tantamount to an audit and it constituted a large portion of employers.

Table 2 shows the expected probability of enforcement for the random assignment section using a 50 percent random enforcement selection criterion over three rounds.

Table 2: Enforcement Probability for the Random Assignment Section

	1st Rounds	2nd Rounds	3rd Rounds	Probability
2 Enforcements	No Enforcement	Enforcement	Enforcement	25%
One Enforcement	No Enforcement	No Enforcement	Enforcement	25%
One Enforcement	No Enforcement	Enforcement	No Enforcement	25%
No Enforcement	No Enforcement	No Enforcement	No Enforcement	25%

The [for enforcements](#).

Table 3 below provided the probability of enforcement for the self-assigned safe harbor section using a 50 percent random enforcement selection criterion for enforcements.

Table 3: Enforcement Probability for the Self-Assignment Section

	Probability
Two Enforcement	25%
One Enforcement	50%
No Enforcement	25%

The Table 4 below provides the expected probability of enforcement for the entire experiment. However, a sizable number of subjects subscribed to the safe harbor provision, and the probability decreased substantially for two to five enforcements and the probability of no enforcement increased. Further information is provided in the [Findings](#) section.

Table 4: Enforcement Probability for the Experiment

	Probability
Four Enforcements	15.6%
Three Enforcements	31.3%
Two Enforcement	31.3%
One Enforcement	15.6%
No Enforcement	6.2%

VIII.1.5 Tax Liability Thresholds Selection – the Tax Rate

The tax liability threshold selection was derived using the suggestions of prior research on income threshold selection. In some studies, researchers have suggested using thresholds of total gross receipts greater \$25,000 or total gross receipts greater than \$100,000, including Erard (2002). To understand the reasoning, analysis of different individual and business tax data were conducted. The IRS 2003 SOI Data “Table 2: Selected financial data on businesses” data showed approximately 25.4 small businesses existed during that point in time (including approximately 3.3 million small corporations). Of the 25.4 small business 60.7 percent of them made less than \$25,000 in total gross receipts and had an average tax liability of \$160, and 81.2 percent of them made less than \$100,000 in total gross receipts with an average tax liability of \$168.

Although this information is useful in identifying characteristics of smaller business, such as their general earnings potential and business tax owed, it does not provide any indication of employment tax liability, but total gross receipts thresholds. To obtain a proxy of employment tax liability some calculations were made using different IRS data tables. IRS 2013 “Table 1.1 All Returns: Selected Income and Tax Items, by Size and Accumulated,” shows an average adjusted gross income of \$61,700 and an average individual income tax liability of \$13,065. Using the aggregate data and the employment tax rates, an average in trust amount was calculated and it is approximately \$22,500.³⁹ From this amount, employers owe an average of approximately \$4,700 for each employee. Using this information, a definition of high and low tax liability threshold was defined. This amount would be the same as the imposed tax rate. The

³⁹ The estimations were calculated using income data. Income data is positively skewed, so it safe to assume that the median amounts are lower than the average amounts provided. Nevertheless, they provide an approximation of thresholds.

lower employment tax rate required employers to pay \$2,500 per employee and the high require a payment of \$6,750 per employee. Hence, the lower tax liability threshold pays a lower tax rate than individuals assigned to the high tax liability threshold.

VIII.1.6 *Population of Interest and Sampling Design*

The population of interest in this study was all authorized individuals from small and self-employed firms. A duly authorized individual is defined as individuals or officers who oversee the collection, accounting, and payment of employment tax (IRS, 2013). This could be restated as someone who is personally liable for the payment of employment tax withholdings deposits even after the firm becomes insolvent.

Although the population of this study is well defined, the sampling frame for this study was not accessible. Because of this limitation, the sampling design was a purposive sampling. This sampling technique was used to ensure that subjects undertaking this experiment meet the described criteria.

Using a non-probability sampling technique will most likely lead to invalid confidence intervals and sampling errors (Alreck & Settle, 1995), and it does not allow researchers to make inferences about the population. However to overcome this limitation, subjects were randomly assigned to one of two groups using a table of random numbers. By randomly assigning subjects into one of the two groups, a randomized experiment was created and internal validity was ascertained (Trochim & Donnelly, 2007b). Internal validity allows researchers to establish causal relationship between dependent and independent variables.

VIII.1.7 *Sample size*

The sample size of this study is based on the a-priori sample size calculation for analysis of variance (ANOVA). The sample size calculation was computed using G*power. A-priori sample size calculation is use to reduce type II error. Using this calculation, the minimum sample size to obtain a power of 0.83 is 16 observations per cell. The larger the sample size for each cell will lead to smaller standard errors and larger power. The sample size anticipated for this study is 200 subjects. This allows for approximately 50 observations per cell, which result in a statically power of over 0.95. Given that the safe harbor provision is self-assigned, it possible that cells size may differ.

VIII.1.8 *Unit of Observation and Unit of Analysis*

The research design collected data at the individual taxpayer level. Therefore, the unit of observation was small business self-employed taxpayers who filed or paid employment tax in prior years. Since study was identifying behavioral differences between those who subscribed and those who did not subscribe to a safe harbor provision, the unit of analysis was also small business self-employed taxpayers.

VIII.1.9 *Implementing the Penalties and Interest*

The code provides two ways to calculate civil tax penalties for tax evasion: fixed dollar amount or fixed percentage (Rasholnikov, 2006). The idea for fixed dollar amount is to signal certainty and to remove any ambiguity as to the dollar amount of the assessed penalty and interest. In general, this amount increases depending on the severity of the violation. Fix dollar amounts are more likely to be used with violations that are more serious. The fixed percent is a fix percent of underreported amount. As with fixed dollar amount, the percentage imposed is dependent on the severity of the infraction. For

example, employment tax has a recovery penalty of 100 percent while penalty for negligent understatements is 20 percent.⁴⁰

In addition, the fix penalty amount is useful in cases where the amount owed can be significantly different. To illustrate, suppose you have two small business owners. One owes \$10,000 and the other owes \$1 million. It would be draconian to charge \$50,000 to the one owing \$10,000 and a mere slap on the wrist for the one owing \$1 million. Given the potential variation in noncompliance amounts, using a fixed percentage for underreporting amounts maintains a fair and equitable treatment of all individuals before the law. Since a fixed percentage is dictated by Publication 15, this research study will do the same. The mandated percentage is 100 percent.

VIII.1.10 *Safe Harbor and Self-Assignment*

In this study, subjects were instructed to decide whether to participate in a safe harbor program. Under certain conditions, a safe harbor provision will minimize taxpayers' legal responsibilities and should reduce the cost of compliance. In this study, subjects will self-assign their participation in the safe harbor provision by simply checking a box. This decision will take place after several rounds of the experiment. By the time subjects were confronted with this decision, approximately half of all subjects would have been selected and experienced enforcement. Furthermore, for those who had experienced enforcement, many of them were informed that enforcement had taken place and sanctions (i.e., penalties and interest) were assessed.

As the literature has suggested, the audit rate will have an impact on the compliance rate, and for employment tax this rate is relatively high. Rasholnikov (2006) also

⁴⁰ § 6662(a)–(b)(1)

suggested that taxpayers will voluntarily report their filing violations if the sanctions imposed by tax authorities fall below their expectations. He also proposed that taxpayers would voluntarily comply with their tax obligations if their expected penalties were equal or greater to those imposed by the authorities. Furthermore, Alm (1991) suggested that amnesty will increase compliance in the short-run.

Two examples that illustrate these concepts are Voluntary Worker Classification Settlement Program and the 2012 Offshore Voluntary Disclosure Program. In both of these programs, the IRS lowered the penalties to improve voluntary compliance. The Voluntary Worker Classification Settlement Program allowed employers, who are not currently being audited, to reclassify their workers and make minimal payments on prior tax obligations without having to pay penalties or interest. The 2012 Offshore Voluntary Disclosure Program allowed taxpayers who have undisclosed offshore income to report their income by a given due date to reduce their penalties and eliminate their chances of a criminal prosecution. In this two examples illustrate how Rasholnikov (2006) suggestion apply.

From our knowledge of taxpayer behavior, it can be concluded that taxpayers who face high audit or enforcement rate and the possibility of reduced penalties are more likely to subscribe to a safe harbor provision. Kalambokidis et al. (2012) illustrated that individuals were willing to pay a fee for a burden reduction. In addition, it can be concluded that individuals who are not compliant but expectations of penalties are higher than those associated with a safe harbor provision are also expected to subscribe to the provision. Therefore, it is expected that only those who are not compliant and have lower expectation of sanctions will not subscribe to the safe harbor provision.

VIII.2 Data Analysis

Analysis of variance (ANOVA) was chosen as the data analysis technique. ANOVA is beneficial because it allows for the violation of equal cell variances and differences in-group means (Lindman, 1970). In this study design, subjects were randomly assigned into either low or high tax liability threshold groups. These groups are defined by the amount owed to the tax authorities at the end of each round. The benefit of random assignment is that investigators can assume probabilistic equivalence (Trochim & Donnelly, 2007a). Therefore, ANOVA is able to assess the causal relationship between the two independent variables (tax rate, enforcement regime, safe harbor provision, and income shock) and the dependent variable compliance by measuring the mean differences amongst groups. Furthermore, ANOVA is an appropriate tool for analyzing binary data (Lunney, 1970), and it was used in this study. Since the data allowed, regression modeling was used to identify the safe harbor subscription drivers and the employment tax payment compliance drivers.

VIII.3 Test Phase

This study was split into four phases. In all phases, subjects' firm was able to earn income and had a set number of employees who worked a set number of hours at a set wage price. Subjects were also informed that enforcement was going to be conducted throughout the experiment. On the other hand, subjects were not informed as to the probability of audit or enforcement. Using this information, subjects voluntarily reported and paid employment taxes. Compliance was measured at the completion of each round.

Phase 1 consists of randomly assigning subjects to a group. Phase 2 introduces enforcement regime. Phase 3 introduces the safe harbor provision. Phase 4 introduces

income shock. The following sections will provide additional information to the mechanic and data obtained in each phase. A general layout of all rounds is provided in [Figure 6: Test Phase Process](#).

VIII.3.1 *Phase 1*

Phase 1 set out to answer the first hypothesis of this study. To do so, subjects were randomly assigned to a group – denoted by R. The random assignment was determined by high or low tax liability threshold, which are denoted as subscripts H and L respectively. The lower employment tax liability threshold was set at \$2,500 and the high was set at \$6,750 per employee.

The O denotes the measured outcome, and it is the difference between reported amount and actual amounts reported and paid. For this phase/round, it identifies the baseline measure for the experiment. [Figure 7](#) provides a visual assessment of the two rounds within Phase 1. This round is a 2X0, and it will be used as a control group for all analysis.

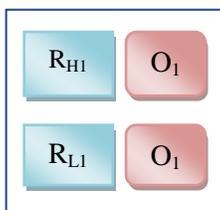


Figure 7: Phase 1 Random Assignment Design.

Once subjects were assigned to a group, subjects remained in that group for the remainder of the experiment.

VIII.3.2 *Phase 2*

Phase 2 consisted of two rounds, and it sets out to answer hypotheses 2a and 2b. As in Phase I, subjects' firm earned income and it had a set number of employees who

will work a set number of hours at a set wage price, which was used to voluntarily report and pay employment taxes.

Unlike Phase 1, subjects were subject to a random enforcement (please see [Section 6.1.4](#) for more details on the probability of enforcement). The subject was subsequently informed of the audit results. If underreporting and/or underpayment were detected during the enforcement process, the subject were instructed to pay the owed amount plus a 100 percent penalty on the amount owed as prescribed by the code.

[Figure 8](#) provides a visual assessment of the two rounds within Phase 2. The letter R denotes random assignment and subscript H and L denote high and low tax liability thresholds. The number under each letter identifies rounds. For example, round 2 high tax liability threshold random assignments is denoted as R_{H2} . The X is enforcement treatment. The measure outcomes – O – was difference between reported and paid amount and actual reported and paid amounts. The opportunity to earn income is based on the scenario of the experiment (see [Appendix A](#)).

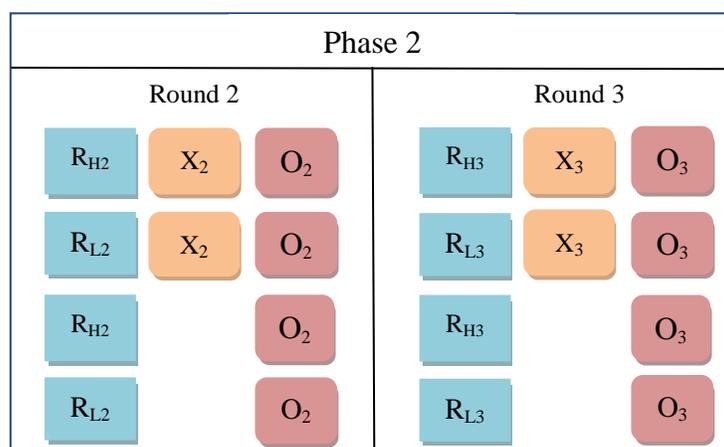


Figure 8: Phase 2 Random Assignment Design.

VIII.3.3 Phase 3

Phase 3 consists of two rounds, and it set out to answer hypothesis 3. To answer the hypothesis, Phase 3 introduced into the decision process the safe harbor provision under the same enforcement regime. As with prior phases, subjects undertook the same tasks to voluntarily reporting and paying of employment taxes.

Subjects received a random enforcement treatment, but only if they did not subscribe to the safe harbor provision at the end of Phase 2 and adhered to requirements. If underreporting and/or underpayment were detected during the enforcement process, the subjects were instructed to pay the owed amount plus a 100 percent penalty on the amount owed as prescribed by the code.

Figure 9 provides a visual assessment of the two rounds within Phase 3. As in Phase 2, the letter R denotes random assignment and subscript H and L denote high and low tax liability thresholds. The number under each letter identifies rounds. The C denotes individuals in who subscribed to the safe harbor provision, which uses a cutoff score. The X implies the enforcement treatment. The O is measured outcomes.

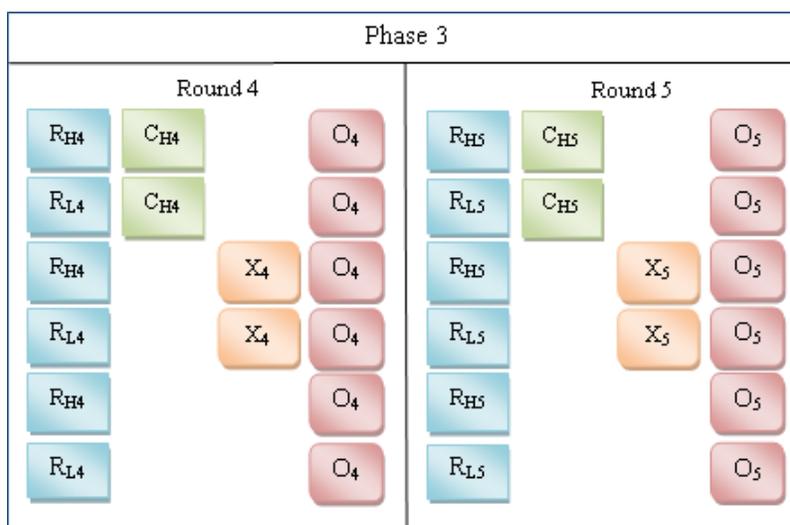


Figure 9: Phase 3 Randomize Design with a Cutoff

VIII.3.4 *Phase 4*

Phase 4 consists of one round, and it sets out to answer hypothesis 4 by introducing an income shock. This manipulation will allow testing difference between all prior treatments. As with prior phases, subjects were able to earn income and had set number of employees who worked a set number of hours at a set wage price, which can be used to voluntarily report and pay employment taxes.

In this phase, subjects were not subject to a random enforcement. It is expected that with and income shock some of the subjects may not adhere to the safe harbor provision, and may undertake noncompliant behavior (Please see [Findings](#) section for details). Once this section was completed, the phase portion of the study ended.

VIII.4 Measuring Risk Appetite

To control for risk appetite in the selection of the safe harbor provision and potential control for compliance behaviors, a risk attitude scale was constructed. This scale is based on Joost and Smidts (2000) and Lampenius and Zickar (2005). The questions were asked to all subjects, and the questions were staggered between speculative risk and risk control. These questions are provided on [Table 5](#).

The data will be analyzed for reliability using scale analysis (Cronbach's alpha), which is the central measurement of internal consistency. To analyze validity of the risk appetite construct, factor analysis was used. Although factor analysis is not primarily used for validity, the results in the component metric and orthogonal rotational matrix could be robust indicator of validity. Given these questions have been used in prior studies, it is expected that the Cronbach's alpha results will yield acceptable internal consistencies ($\alpha > 0.7$ (Nunnally, 1994)). Furthermore, it is expected that the variable

should load into two distinct groups (Speculative Risk and Risk control), validating the construct. (Please see [Finding](#) section for results)

Table 5: Risk Construct

Risk	Questions	Scales
Speculative Risk	I see risk as an opportunity to make money	Agree-Disagree
	I like taking big financial risks	Agree-Disagree
	I get a thrill out of investing	Agree-Disagree
	When I complete my taxes forms I am willing to make aggressive reporting activities	Agree-Disagree
Risk Control	When I invest money a safe return is very important to me	Agree-Disagree
	When I invest I plan on having a specific amount at a future date	Agree-Disagree
	It is important to know with how much money my investment will provide me in the future	Agree-Disagree
	I consult a tax adviser when completing my taxes	Agree-Disagree
	I pro-actively manage tax filing process with the aim of minimizing risk of enforcement	Agree-Disagree

In addition, there were questions on perception of the probability of enforcement and the penalty rate. These questions are provided in

Table 6. These particular questions on the probability of enforcement were requested at the start and the end of the experiment. The table below provides the questions provided at the end of the experiment. The penalty rate was only asked at the end of the experiment to determine change in perception to audit and penalty to employment tax payment.

Table 6: Enforcement, Timing and Penalty Questions

	Questions	Scales
Probability of Enforcement	From your interaction with the IRS, what do you think is the likelihood of employment tax enforcement? <i>(please provide a percent)</i>	Percent
	If you miss report on your employment tax filling, what do you think is the likelihood that IRS will question the amount reported? <i>(please provide a percent)</i>	Percent
	After completing this experiment, what do you think was the enforcement rate for this experiment? <i>(please provide a percent)</i>	Percent
Penalty Rate	If caught and successfully prosecuted, what do you think the employment tax penalties would be? <i>(please provide a percent)</i>	Percent
	After completing this experiment, what do you think was the penalty for failure to pay employment tax? <i>(please provide a percent)</i>	Percent
Timing	If I am caught and successfully prosecuted, how much time do you think it would take the IRS to prosecute?	Open ended (rating scale)

VIII.5 Additional Test Details

VIII.5.1 *Recruitment*

The source for recruitment pre-test was investigator-initiated in-person contact and emails. The recruitment script has been developed to ensure consistent screening and to provide potential subject the ability to opt-out of the research study. In addition, a handout was developed with pertinent information for those individual who accept and qualify to be subjects. Some of the information provided includes study web page and the expected duration of study. An email script has been attached. Online recruitment for the test portion of the experiment was done using Qualtrics' assistance.

VIII.5.2 *Funding and Compensation*

All the compensation was monetary. The funding source for this study was the student PI's own personal income. The student PI did not receive any grants or sponsorship of any kind. The student PI paid Qualtrics for each participant, and Qualtrics compensated subjects five dollars for their participation as approved by the IRB. This payment system was easy, fast, and free for those receiving moneys. This payment

process diverged from the Google Wallet, but I permitted subjects to receive payment without the need to create a Google Wallet account. Furthermore, without the student PI's knowledge, Qualtrics also compensated the participants an additional \$4 from the student PI's payment to Qualtrics. This payment means that all participants received a flat fee of \$9 regardless of treatment/answers for participating in the study.

VIII.6 Deviations from Original Design

This subsection discusses the deviation from the proposed methods identified in [Section 8](#) of this document.

Although ANOVA analysis was conducted as prescribed by the measures section, additional nonparametric statistics were ran for each round of the experiment. The Pearson's χ^2 was used for categorical variables while the Fisher Exact test was used for binary variables. These different nonparametric tests were used because they are more appropriate for the data obtained in this experiment. Nevertheless, the ANOVA analysis and the non-parametric test yielded similar results and both were used in the discussion of our findings.

Another deviation from the methods section was payment to subjects. The prescribed payment for participating in this experiment was five dollars paid using Google wallet. However, Qualtrics directly paid the subjects and Google wallet was not used in the payment. In addition, participants were paid nine dollars instead of five.

A third deviation was the number of subject who participated. The original method section indicated that 15 subjects were to participate in the pre-test. However, only eight subjects participated in the pre-test. For the test, 200 subjects were to complete the experiment. However, the way Qualtrics cutoff the participation, it led to

an additional five subjects to participate. The total number of participants was intended to be 215, but ultimately 213 participated. No pre-test participants were included in the analysis of the results since changes were made.

IX FINDINGS

IX.1 General Demographics of Subjects

The data below shows how subjects who participated in the study are involved in the employment tax process. The results are provided on [Table 7](#). Approximately 40 percent of participants had been involved in the filing and or payment process in a single capacity, with “Employee” accounting for over 50 percent of those. At the other extreme, 6.8 percent of participants were involved in the process in all capacities.

Table 7: Involvement with Employment Tax Filing and Payment Process

	Frequency	Percent
Accountant	8	3.9
Lawyer	1	0.5
Business Owner	21	10.2
Manager	7	3.4
Employee	45	22.0
All Capacities	14	6.8
Two Involvement Capacities	54	26.3
Three Involvement Capacities	33	16.1
Four Involvement Capacities	15	7.3
Other	7	3.4
Total	205	100

Of those who chose “Two Involvement Capacities,” they selected employee 66.7 percent of the time, and managers 59.3 percent of the time. All other options were selected less than 50 percent of the time.

Of those who had chosen “Three Involvement Capacities,” they selected employee 90.9 percent of the time, managers 87.9 percent of the time, business owner 63.6 percent of the time and Accountant 51.5 percent of the time. All other options were selected less than 50 percent of the time.

Of those who had chosen “Four Involvement Capacities,” they selected accountant 100 percent of the time, manager 93.3 percent of the time, and business owner 86.7 percent of the time. All other options were selected less than 50 percent of the time. Table 8 shows the geographical breakdown of participants. Although this information cannot be used to infer to the population of interest, that is, all small business owners and tax managers, it shows that participation came from throughout the United States and it was not concentrated in one particular area in the US. Therefore, we believe that the sampling is reasonable and will provide interesting observations even though it is not, strictly speaking, and a random sample.

Table 8: Geographical Breakdown of Participating Subjects

	Frequency	Percent		Frequency	Percent
Alabama	3	1.5	Mississippi	3	1.5
Arizona	6	2.9	Missouri	2	1
California	15	7.3	Nevada	3	1.5
Colorado	5	2.4	New Jersey	6	2.9
Connecticut	2	1	New York	14	6.8
Delaware	2	1	North Carolina	5	2.4
Florida	25	12.2	Ohio	8	3.9
Georgia	12	5.9	Oklahoma	3	1.5
Hawai'i	1	0.5	Oregon	4	2
Illinois	9	4.4	Pennsylvania	15	7.3
Indiana	4	2	Rhode Island	1	0.5
Iowa	1	0.5	South Carolina	2	1
Kansas	1	0.5	South Dakota	1	0.5
Kentucky	6	2.9	Tennessee	7	3.4
Maryland	3	1.5	Texas	11	5.4
Massachusetts	4	2	Vermont	1	0.5
Michigan	7	3.4	Virginia	3	1.5
Minnesota	2	1	Wisconsin	4	2

IX.2 Risk Appetite Analysis and Enforcement Perception

To control for risk appetite in the selection of the safe harbor provision and potential control for compliance behaviors, a risk attitude scale was constructed. The

data was analyzed for reliability using scale analysis, which is the key statistic for of reliability. To analyze the construct validity of the risk appetite construct, factor analysis was used.

IX.2.1 Reliability

Using questions derived from prior studies, a scale for risk was created and was tested to control for risk appetite in the selection of the safe harbor provision and potential control for compliance behaviors. Table 9 shows the results for the Cronbach's alpha tests. The results in this study yielded acceptable internal consistencies ($\alpha > 0.7$ (Nunnally, 1994)), and Cronbach's alpha and the Tukey analysis show no difference between items. Therefore, the analyses rely on high level of internal consistency in the measures.

Table 9: Cronbach's Alpha Results

Risk type	Cronbach's alpha	Number of items
Speculative control	0.776	4
Risk control	0.765	5

IX.2.2 Validity

Table 10 shows the results for construct validity testing for both risk components (Speculative Risk and Risk control). These results were obtained using principal component methods and Varimax rotation. As the table shows, the variables of interest loaded into two distinct components providing evidence that the components are construct-valid. Variable loading below 0.4 were suppressed to easily identify the loading factors (Hair, 1998, 2014).

One of the variables, "I get a thrill out of investing," loaded into both components. However, it has a 0.466 loading score, which indicates a weaker measure

for risk control. Nevertheless because of its score, this variable will be kept in both components. The total variance explained using this model was 58.8 percent.

Table 10: Factor Analysis for Risk

	Components	
	Risk Control	Speculative Risk
I see risk as an opportunity to make money.		0.827
When I invest money, a safe return is very important to me.	0.770	
I like to take big financial risks.		0.887
When I invest, I plan on having a specific amount at a future date.	0.763	
I get a thrill out of investing.	0.466	0.629
It is important to know how much money my investment will provide me in the future.	0.769	
I consult a tax advice when completing my taxes.	0.539	
When completing my taxes, I am willing to make aggressive reporting activities.		0.661
I proactively manage the tax filing process with the aim of minimizing risk of audit.	0.726	

Additional tests were run to ensure sampling adequacy was suitable for the test. KMO statistics provide analysis on variance and it provides guidance on the usefulness of factor analysis. In essence, it tests whether or not the sample is sufficient given the number of variables used in the factor analysis conducted. Results should be greater than 0.70 for factor analysis to be useful (Kaiser, 1974; Stewart, 1981). The KMO result was 0.736 indicating that the data structure is adequate.

The Bartlett's test of sphericity checks to see if the correlation matrix is an identity matrix. The presence of an identify matrix implies no correlation between the variables used in the factor analysis conducted. Since there has to be some relationship between the variables used to construct a component, an identify matrix would not be suitable for factor analysis. For suitable results, a number below 0.05 is expected (IBM, 2017; Stewart, 1981). In this case, the expected value was below 0.05 (a statistic of 0.000). Consequently, our data is suitable and our components are valid. Predictive

analyses using these results are located in Section [Logistic Regression Control Compliance](#)

IX.2.3 Probability of Enforcement and Penalty Rate

IX.2.3.1 Changes to Probability of Enforcement

The average perceived pre-test audit rate was 40.3 percent, and the post-test results showed a 74.9 percent perceived audit rate. This is a 34.9 percent increase. To determine whether the pre-test audit rate had an impact on the post-test audit rate, a one-way ANOVA was run using a 5 percent alpha protection level. The results showed that there was no statistical difference (p-value = 0.113) between perceived pre and post-test audit rate. The results indicate that the pre-test audit rate (independent variable) had no substantial affected the post-test audit rates (dependent variable).

IX.2.3.2 Changes to Penalty Rate

The average perceived pre-test penalty rate was 262.0 percent, and the post-test results showed a 92.9 percent perceived audit rate. This was a 169.1 percent decrease, in absolute terms. This information implies that payment compliance is improved once the safe harbor provision became available to subjects. To determine if there was a statistical difference in the subjects' perception of the audit rate, a one-way ANOVA was also run using the community-standard 5 percent alpha protection level. The results showed that there was no statistical difference (p-value = 0.805) between perceived pre and post-test rate of audit rate indicating that compliance did not increase. Although the reduction in penalty rate seemed large, the insignificant results from the ANOVA could be in part attributed to the large standard deviation within the data (142.2 and 48.7 percent

respectively). Given the results, additional logistic regression analyses were conducted.

Those results are found in Section IX.3.6.

IX.3 Hypothesis testing

From the prior literature and theory development sections, there are four hypotheses tested. There are three possible outcomes: accepted, rejected, and partial acceptance. Accepted indicates that all the results indicate that the hypothesis is supported. Rejected indicates that all the results indicate that the hypothesis is not supported. Partial acceptance indicates that some of the results do support the hypothesis, but not all them.

Table 11: Results by Hypothesis

Null Hypothesis	Supported?
Hypothesis 1: Business owners within the higher tax liability thresholds will have lower subscription to a safe harbor provision than business owners within the lower tax liability thresholds.	Partial Acceptance
Hypothesis 2a: Business owners who experience enforcement will have higher subscription rate to a safe harbor provision than those who do not experience enforcement.	Accepted
Hypothesis 2b: Business owners who experience enforcements will have higher payment compliance rate than those who do not experience enforcement.	Rejected
Hypothesis 3: Business owners who chose the safe harbor provision will be more compliant, with respect to payment.	Partial Acceptance
Hypothesis 4: Business owners who chose the safe harbor provision will remain more compliant after an income shock is experienced.	Partial Acceptance

The next four subsections provide the results for each hypothesis test. The results to all the hypotheses are presented in

[Table 11](#). The alpha protection levels for all results in the finding sections are 0.05.

IX.3.1 Hypothesis 1

The hypothesis 1 states that, “Business owners within the higher tax liability thresholds will have lower subscription to a safe harbor provision than business owners within the lower tax liability thresholds.” The tests in this subsection will identify if there is any statistical differences in the subscription rate between business owners randomly assigned to high and low tax liability thresholds. The dependent variable for Hypothesis 1 is Subscription rate and the independent variable is tax liability thresholds.

$$SR = \alpha + \beta(LT) \quad (3)$$

where SR is the subscription rate to safe harbor provision and LT (subscript) is the binary tax liability thresholds (low or high),

Formula 3 states that the subscription rate is expected to be greater for subjects with low liability than for those subjects having higher tax liability thresholds.

As the results show on

Table 12, the subscription rate for the voluntary safe harbor participation was large. The subscription rate dropped consistently from one round to the next for the high tax threshold subjects, but it fluctuated for the low tax liability threshold. Nevertheless, both high and low experienced an overall drop of roughly 10 percent (8.4 percent and 10.5 percent respectively).

Table 12: Safe Harbor Subscription Rate by Tax Liability Thresholds

	Low	High
End of Round 3 Subscription Rate	94.3% (99/105)	90.0% (90/100)
End of Round 4 Subscription Rate	96.2% (101/105)	87.0% (87/100)
End of Round 5 Subscription Rate	86.7% (91/105)	81.0% (81/100)

The next four subsections focus on each round of non-parametric test results and the ANOVA test.

IX.3.1.1 Round 4 Test Results

Figure 10: Round 4 Distributions of the Safe Harbor Provision Subscription Rate Given Figure 10 shows the distribution of subscription to the safe harbor provision given their employment tax liability threshold (Low versus High). As the picture shows, a larger number of participants subscribed to the safe harbor provision. However, the figure shows slightly higher subscription rate to the safe harbor provision for those subjects who were randomly assigned to the lower tax liability threshold than those subjects who were randomly assigned to the higher tax liability threshold. The statistical tests for significance are shown below.

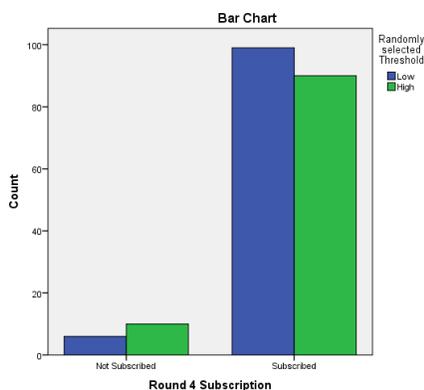


Figure 10: Round 4 Distributions of the Safe Harbor Provision Subscription Rate Given Employment Tax Liability Threshold

To determine the relationship between tax liability threshold and subscription to a safe harbor provision in the fourth round, two similar nonparametric tests were used⁴¹: the Fisher Exact test and the Pearson's Chi-Square (χ^2). The Fisher Exact test is a χ^2 that examines the relationship between the two dimensions. This test results are found on Table 13. The hypothesis is:

⁴¹ Both test were run because there was a low number participants who were audited because of the high level of safe harbor provision subscription. The non-parametric tests were better fit for this scenario than ANOVA.

H_0 : The tax liability threshold and safe harbor provision are independent in the 4th round.

H_a : The tax liability threshold and safe harbor provision are not independent in the 4th round.

The Fisher Exact test p-value is 0.189, so the null hypothesis was not rejected.

These results confirm that there is no association between tax liability threshold and subscription to safe harbor provision. This conclusion is further corroborated by the Pearson's χ^2 (p-value of 0.253), which shows that we can accept the null hypothesis and conclude that compliance and subscription to the safe harbor provision are not related at the community-standard 5 percent alpha protection level. The results confirm higher tax liability thresholds may not lead to higher subscription to safe harbor provision the fourth round.

Table 13: Hypothesis 1 Round 4 Statistical Tests

	Value	df	Asymp. p-value (2-sided)	Exact p-value (2-sided)	Exact p-value (1-sided)
Pearson's χ^2	1.307 ^a	1	0.253		
Continuity Correction ^b	0.780	1	0.377		
Likelihood Ratio	1.317	1	0.251		
Fisher's Exact Test				0.303	0.189
Linear-by-Linear Association	1.301	1	0.254		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 7.80.

b. Computed only for a 2x2 table

Furthermore,

Table 14 shows the results from an effect size analysis using η^2 . The results found an η of 0.080 and an η^2 of 0.006. Using Cohen's interpretation guidelines (Ruscio, 2008), the results show a small effect size between compliance and subscription to the safe harbor provision. This implies that the employment tax liability threshold had a small effect on subscription rate in the fourth round. These results are in line with the Phi and Cramer's V test results, which show a small effect. However, since the results were not statistically significant at the 0.05 alpha scientific community protection level, we can only generalize these effect sizes to the population if the statistical power is above 0.8. In this case, the calculated post-hoc statistical power was 0.19. Therefore, we cannot generalize these results to the population.

Table 14: Hypothesis 1 Round 4 Effect Size Tests

Test	Value	Approx. p-value
η	0.080	-
Phi	-0.080	0.253
Cramer's V	0.080	0.253

IX.3.1.2 Round 5 Test Results

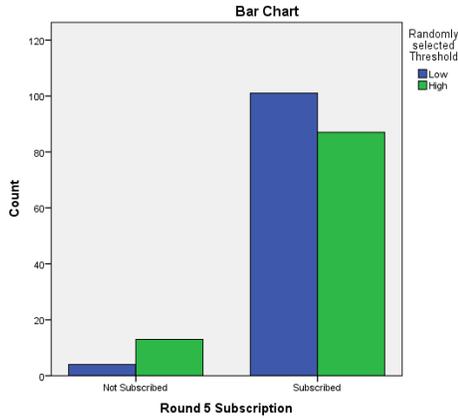


Figure 11 shows the distribution of subscription to the safe harbor provision given their employment tax liability threshold (Low versus High). As the picture shows, the results are similar to those in round 4. There were a larger number of subscriptions to the safe harbor provision and higher subscription rate for subjects assigned lower tax liability threshold than those subjects assigned to the higher tax liability threshold than with the other two rounds observed. The statistical tests for significance are shown below.

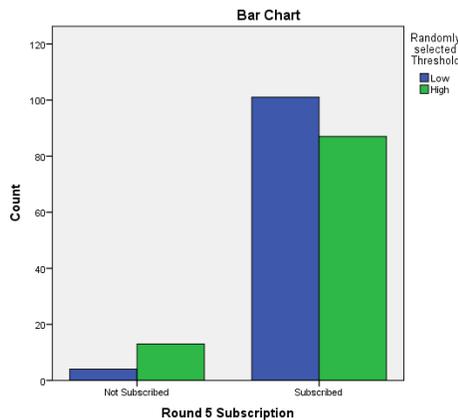


Figure 11: Round 5 Distributions of the Safe Harbor Provision Subscription Rate given Employment Tax Liability Threshold

To determine the relationship between the tax liability threshold and subscription to a safe harbor provision in the fifth round, the same two nonparametric tests were used (Fisher Exact Test and Pearson's χ^2). The hypothesis is provided below and results are provided on [Table 15](#).

H₀: The tax liability threshold and safe harbor provision are independent in the 5th round.

H_a: The tax liability threshold and safe harbor provision are not independent in the 5th round.

The Fisher Exact test p-value was 0.015. From this result, the null hypothesis is rejected, confirming that there is association between the tax liability threshold and subscription to a safe harbor provision in the fifth round. These results confirm that low tax liability thresholds subjects subscribed to the safe harbor in higher statistical numbers by high liability threshold subjects. This conclusion is further collaborated by the Pearson's χ^2 (p-value of 0.017), which shows that we can reject the null hypothesis and conclude that tax liability threshold and subscription to the safe harbor provision are related at the community-standard 5 percent alpha protection level in the fifth round . The results confirm higher tax liability thresholds may lead to higher subscription to safe harbor provision the fifth round.

Table 15: Hypothesis 1 Round 5 Statistical Tests

	Value	df	Asymp. p-value (2-sided)	Exact p-value (2-sided)	Exact p-value (1-sided)
Pearson's χ^2	5.689 ^a	1	0.017		
Continuity Correction ^b	4.544	1	0.033		

Likelihood Ratio	5.938	1	0.015		
Fisher's Exact Test				0.022	0.015
Linear-by-Linear Association	5.661	1	0.017		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 8.29.

b. Computed only for a 2x2 table

Furthermore, [Table 16](#) shows the results from an effect size analysis using η^2 . The results found an η of 0.167 and an η^2 of 0.027. Using Cohen's interpretation guidelines (Ruscio, 2008), the results showed a very small effect size between employment tax liability threshold and subscription to the safe harbor provision. This implies that the employment tax liability threshold had a small effect on the safe harbor provision subscription rate during the fifth round. These results are in line with the Phi and Cramer's V test results, which are statically significant at the community-standard 5 percent alpha protection level.

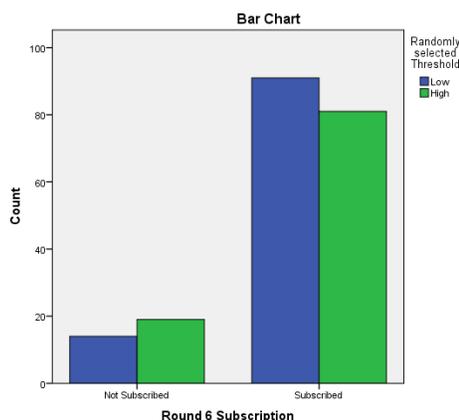
Table 16: Hypothesis 1 Round 5 Effect Size Tests

Test	Value	Approx. p-value
η	0.167	-
Phi	-0.167	0.017
Cramer's V	0.167	0.017

IX.3.1.3 Round 6 Test Results

[Figure 12](#) shows the distribution of subscription to the safe harbor provision given their employment tax liability threshold (Low versus High). As the picture shows, the results are similar to rounds 4 and 5. However, there is a lower subscription rate for this round than the other rounds (in absolute terms), and a larger jump in the low tax liability threshold. The statistical tests for significance are shown below.

Figure 12: Round 6 Distributions of the Safe Harbor Provision Subscription Rate given Employment Tax Liability Threshold.



To determine the relationship between tax liability threshold and subscription to safe harbor provision in the sixth round, the same two nonparametric tests were used (Fisher Exact Test and Pearson's χ^2). The hypothesis is provided below and results are provided on [Table 17](#).

H_0 : The tax liability threshold and subscription to safe harbor provision are independent in the 6th round.

H_a : The tax liability threshold and subscription to safe harbor provision are not independent in the 6th round.

The Fisher Exact test p-value is 0.181, so the null hypothesis is not rejected.

These results confirm there is no association between tax liability threshold and subscription to safe harbor provision. This conclusion is further collaborated by the Pearson's χ^2 (p-value of 0.270), which shows that we can accept the null hypothesis and conclude that compliance and subscription to the safe harbor provision are not related at the community-standard 5 percent alpha protection level. The results confirm higher tax liability thresholds may not lead to higher subscription to safe harbor provision the sixth round.

Table 17: Hypothesis 1 Round 6 Statistical Tests

	Value	df	Asymp. p-value (2-sided)	Exact p-value (2-sided)	Exact p-value (1-sided)
Pearson's χ^2	1.218 ^a	1	0.270		
Continuity Correction ^b	0.834	1	0.361		
Likelihood Ratio	1.220	1	0.269		
Fisher's Exact Test				0.342	0.181
Linear-by-Linear Association	1.212	1	0.271		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 16.10.

b. Computed only for a 2x2 table

Furthermore, [Table 18](#) shows the results from an effect size analysis using η^2 . The results found an η of 0.077 and an η^2 of 0.006. Using Cohen's interpretation guidelines (Ruscio, 2008), the results show a very small effect size between compliance and subscription to the safe harbor provision. These results are also in line with the Phi and Cramer's V test results. The results would imply payment compliance had a small effect on the safe harbor subscription rate the sixth round. However, since the results were not statistically significant at the 0.05 alpha scientific community protection level, we can only generalize these effect sizes to the population if the statistical power is above 0.8. In this case, the calculated post-hoc statistical power was 0.20. Therefore, we cannot generalize these results to the population.

Table 18: Hypothesis 1 Round 6 Effect Size Tests

Test	Value	Approx. p-value
η	0.077	-
Phi	-0.077	0.270
Cramer's V	0.077	0.270

IX.3.1.4 Hypothesis 1 ANOVA Test Results

An ANOVA test was also run as indicated in the methods section. The ANOVA test evaluates the mean difference between employment tax liability threshold and subscription to the safe harbor provision. Table 19 shows the results of the test. The p-value results for rounds 4 and 6 were greater than 0.05, but they were less than 0.05 for round 5. Thus, the results confirm no dependency between the safe harbor subscription rate and the liability threshold subjects (high or low) for the fourth and sixth round. However, there was a mean difference in the safe harbor subscription rate between high and low tax liability threshold subjects in the fifth round, which confirm dependency between the safe harbor subscription rate and the liability threshold subjects (high or low) in the fifth round.

Table 19: Hypothesis 1 ANOVA tests for

	Sum of Squares	df	Mean Square	F	p-value
Round 4	0.094	1	0.094	1.303	0.255
Round 5	0.433	1	0.043	5.794	0.017
Round 6	0.005	1	0.005	0.021	0.884

IX.3.2 Hypothesis 2a

Hypothesis 2a sets out to determine whether Business owners who experience enforcement will have higher subscription rate to a safe harbor provision than those who do not experience enforcement. This section is also segmented into four areas of results based on each round of tests and ANOVA analysis.

The tests in this subsection will identify if there is any statistical differences in the subscription rate and the number of enforcement treatments. The dependent variable for

Hypothesis 2a is once more the subscription rate, but the independent variable is the number of enforcement treatments undergone during the experiment.

$$SR = \alpha + \beta(ET) \quad (4)$$

where SR is the safe harbor subscription rate and ET is the number of enforcement treatments.

Formula 4 states that the subscription rate is expected to improve as subjects encounter higher levels of enforcement.

Table 20 shows the distribution of enforcement treatments during the experiment and a comparison to the expected enforcement rate calculated prior to the experiment. The percent shown in this table can be used as the probability of enforcement as described in Section 8.1.5 (page 45). The table shows a higher number of not selected and “1 enforcement encounter” than previously expected. This shows that participants subscribed to the safe harbor provision at a higher rate than anticipated in the design.

Table 20: Enforcement Treatments Frequency and Percent

Number of Enforcement Treatments	Frequency	Probability	Expected Probability
No treatment	49	23.9%	15.6%
1 treatment	96	46.8%	31.3%
2 treatments	56	27.3%	31.3%
3 treatments	4	2.0%	15.6%
4 treatments	0	0%	6.2%

IX.3.2.1 Round 4 Test Results

Figure 13 shows the distribution of subjects’ decision to subscribe to the safe harbor provision given the number of enforcement treatments (1 thru 4 treatments). As the picture shows, a larger number of participants subscribed to the safe harbor provision. The figure also shows similar distributions for both groups (subscribed and not subscribed), and a high number of subscriptions to the safe harbor provision for those

who experienced the higher enforcement treatments. The statistical tests for significance are shown below.

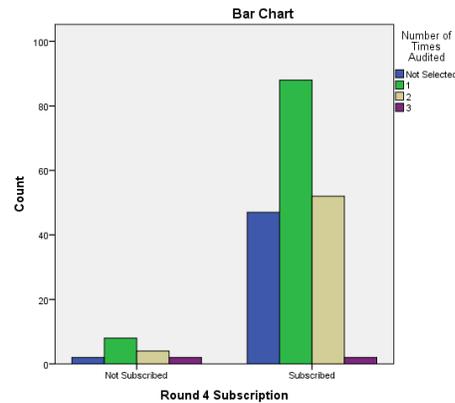


Figure 13: Round 4 Distributions between Enforcement Treatments and the Safe Harbor Provision Subscription Rate.

To determine the relationship enforcement treatments and subscription to a safe harbor provision in the fourth round, a Pearson's χ^2 was run. Since the dependent variable is categorical and not binary, the Fisher Exact test was not run for this hypothesis. The results are shown on

Table 21.

H_0 : Safe harbor subscription rate and number of enforcement treatments are independent in the fourth round.

H_a : Safe harbor subscription rate and number of enforcement treatments are not independent in the fourth round.

A Pearson's χ^2 was run and the p-value was 0.012, which shows that we can reject the null hypothesis confirming that there is association between enforcement treatments and subscription to the safe harbor provision. This confirms that enforcement treatments and subscription to the safe harbor provision are related at the community-standard 5 percent alpha protection level. These results confirm that a higher number of

enforcement treatments may lead to higher subscription to the safe harbor provision in the fourth round.

Table 21: Hypothesis 2a Round 4 Statistical Tests

Test	Value	df	Asymp. p-value (2-sided)
Pearson's χ^2	10.913 ^a	3	0.012
Likelihood Ratio	6.182	3	0.103
Linear-by-Linear Association	2.690	1	0.101

a. 4 cells (50.0%) have expected count less than 5. The minimum expected count is .31

As with the prior hypothesis, an effect size test was conducted. Table 22 shows the results from an effect size analysis using η^2 . The results found an η of 0.231 and an η^2 of 0.053. Using Cohen's interpretation guidelines (Ruscio, 2008), the results show a small effect size between enforcement treatments and subscription to the safe harbor provision. This implies that the enforcement treatments had a small effect on the safe harbor provision subscription rates the fourth round. These results are in line with the Phi and Cramer's V test results, which are statically significant at the community-standard 5 percent alpha protection level.

Table 22: Hypothesis 2a Round 4 Effect Size Tests

Test	Value	Approx. p-value
η	0.231	-
Phi	0.231	0.012
Cramer's V	0.231	0.012

IX.3.2.2 Round 5 Test Results

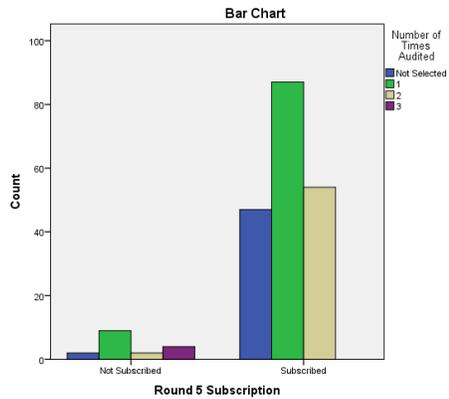


Figure 14 shows the distribution of subjects' decision to subscribe to a safe harbor provision given the number of enforcement treatments (1 thru 4 treatments). The results in this round are similar to those in round 4 and 6. Some of the similarities are larger number of participants who subscribed to the safe harbor provision, and similar distributions for within groups (subscribed and not subscribed). However, in this round, those individuals who experienced more than 3 enforcement treatments did not overly subscribe to the safe harbor provision, which contradicts expectations. The statistical tests for significance are shown below.

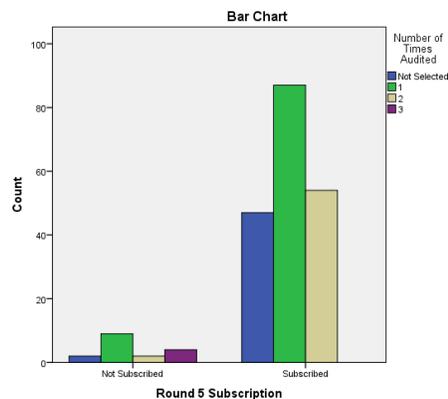


Figure 14: Round 5 Distributions between Enforcement Treatments and the Safe Harbor Provision Subscription Rate.

To determine the relationship enforcement treatments and subscription to a safe harbor provision in the fifth round, the Pearson's χ^2 was run, and results are show on [Table 23](#).

H_0 : The safe harbor subscription rate and number of enforcement treatments are independent in the fifth round.

H_a : The safe harbor subscription rate and number of enforcement treatments are not independent in the fifth round.

As mentioned before, The Fisher Exact test was not run. On the other hand, the Pearson's χ^2 was run and the p-value was 0.000, which shows that we must reject the null hypothesis confirming that there is association between enforcement treatments and subscription to the safe harbor provision. This confirms that enforcement treatments and subscription to the safe harbor provision are related at the community-standard 5 percent alpha protection level. The results confirm that a higher number of enforcement treatments may lead to higher subscription to the safe harbor provision in the fifth round.

Table 23: Hypothesis 2a Round 5 Statistical Tests

Test	Value	df	Asymp. p-value (2-sided)
Pearson's χ^2	47.167 ^a	3	0.000
Likelihood Ratio	23.497	3	0.000
Linear-by-Linear Association	4.981	1	0.026

a. 4 cells (50.0%) have expected count less than 5. The minimum expected count is .33

Furthermore, Table 24 shows the results from an effect size analysis using η^2 . The results found an η of 0.480 and an η^2 of 0.230. Using Cohen's interpretation guidelines (Ruscio, 2008), the results show a moderate effect size between enforcement treatments and subscription to the safe harbor provision. This implies that the enforcement treatments had a moderate effect on safe harbor subscription rates the fifth

round. These results are in line with the Phi and Cramer's V test results, which are statically significant at the community-standard 5 percent alpha protection level.

Table 24: Hypothesis 2a Round 5 Effect Size Tests

Test	Value	Approx. p-value
η	0.480	-
Phi	0.480	0.000
Cramer's V	0.480	0.000

IX.3.2.3 Round 6 Test Results

Figure 15 shows the distribution of subjects' decision to subscribe to a safe harbor provision given the number of enforcement treatments (1 thru 4 treatments). The results in this round are similar to those in round 4 and 5:

- Larger number of them subscribed to the safe harbor provision
- Similar distributions for both groups (subscribed and not subscribed). The statistical tests for significance are shown below.
- Higher subscription to the safe harbor provision for those who experienced higher enforcement treatments.

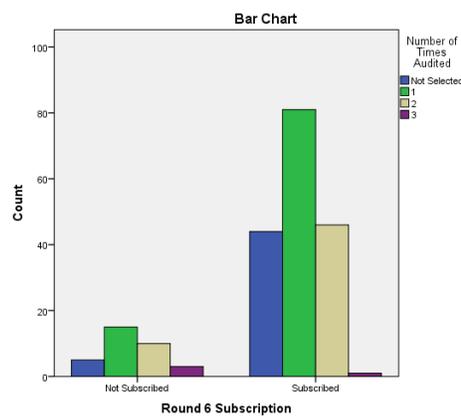


Figure 15: Round 6 Distributions between Enforcement Treatments and the Safe Harbor Provision Subscription Rate.

To determine the relationship enforcement treatments and subscription to a safe harbor provision in the sixth round, the Pearson's χ^2 was run and the results are shown on Table 25.

H_0 : The safe harbor subscription rate and number of enforcement treatments are independent in the sixth round.

H_a : The safe harbor subscription rate and number of enforcement treatments are not independent in the sixth round.

As mentioned before, the dependent variable is not a binary variable and the Fisher Exact test was not run. The results from the Pearson's χ^2 found a p-value was 0.009 in the sixth round, which shows that we can reject the null hypothesis confirming that there is association between enforcement treatments and subscription to the safe harbor provision. This confirms that enforcement treatments and subscription to the safe harbor provision are related at the community-standard 5 percent alpha protection level. The results confirm that a higher the number of enforcement treatments may lead to higher subscription to the safe harbor provision in the sixth round.

Table 25: Hypothesis 2a Round 6 Statistical Tests

Test	Value	df	Asymp. p-value (2-sided)
Pearson's χ^2	11.680 ^a	3	0.009
Likelihood Ratio	8.367	3	0.039
Linear-by-Linear Association	4.529	1	0.033

a. 2 cells (25.0%) have expected count less than 5. The minimum expected count is .64

Furthermore, [Table 26](#) shows the results from an effect size analysis using η^2 . The results found an η of 0.239 and an η^2 of 0.057. Using Cohen's interpretation guidelines (Ruscio, 2008), the results show a small effect size between enforcement treatments and subscription to the safe harbor provision. This implies that the enforcement treatments had a small effect on the safe harbor subscription rates in the sixth round. These results are in line with the Phi and Cramer's V test results, which are statically significant at the community-standard 5 percent alpha protection level.

Although the results were statistically significant at the 0.05 alpha scientific community protection level, we can only generalize these effect sizes to the population if the statistical power is above 0.8. In this case, the calculated post-hoc statistical power was 0.94. Therefore, we can generalize these results to the population.

Table 26: Hypothesis 2a Round 6 Effect Size Tests

Test	Value	Approx. p-value
η	0.239	-
Phi	0.239	0.009
Cramer's V	0.239	0.009

IX.3.2.4 Hypothesis 2a ANOVA Test Results

An ANOVA test was run as planned in the proposal. The ANOVA test evaluates the mean difference between the safe harbor subscription rate and number of enforcement treatments.

Table 27 shows, all the p-values were below the community-standard 5 percent alpha protection level. The results confirm a dependency between the safe harbor subscription rate and the number of enforcement treatments (i.e., 1- 4). Since the number of enforcements had an impact on the safe harbor subscription rate, this corroborates with the idea that subjects with higher enforcement treatments would be more likely to subscribe to the safe harbor provision.

Table 27: Hypothesis 2a ANOVA tests

	Sum of Squares	df	Mean Square	F	p-value
Round 4	0.785	3	0.262	3.767	0.012
Round 5	3.587	3	1.196	20.022	0.000
Round 6	1.577	3	0.526	4.048	0.008

IX.3.3 Hypothesis 2b

Hypothesis 2b sets out to determine whether business owners who experience enforcement will have higher payment compliance rate than those who do not experience enforcement. This section is also segmented into four areas of results based on each round of tests and ANOVA analysis.

The tests in this subsection will identify if there is any statistical differences in the employment tax payment compliance and the number of enforcement treatments. The dependent variable for Hypothesis 2b is employment tax payment compliance, and the independent variable is the number of enforcement treatments undergone during the experiment.

$$PMT = \alpha + \beta(EE) \quad (5)$$

Where PMT is employment tax payment compliance category

EE is the number of enforcement treatments

Formula 5 states that the employment tax payment compliance category is expected to improve as subject encounter higher levels of enforcement. Employment tax payment compliance category is a binary variable and it was calculated in two steps. First, the difference between expected payment and actual payment was calculated. Using the results from this step, a binary variable was compute suing the following formula:

If Actual payment – Expected payment $\geq 0 \rightarrow$ Compliant

If Actual payment – Expected payment $< 0 \rightarrow$ Non-Compliant

These computations were used for every round and throughout all analysis of payment compliance.

IX.3.3.1 Round 4 Test Results

Figure 16 shows the distribution of subjects' propensity to comply given the number of enforcement treatments (1 thru 4 treatments). As the picture shows, a larger number of participants subscribed to the safe harbor provision. The figure also shows similar distributions for both groups (subscribed and not subscribed), but the figure shows a higher rate of payment compliance for those who experience an enforcement encounter. The statistical tests for significance are shown below.

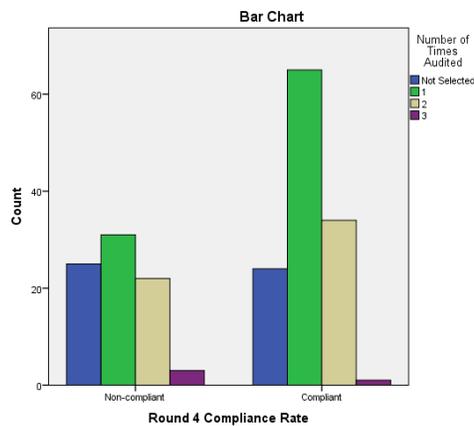


Figure 16: Round 4 Distributions between Enforcement Treatments and Payment Compliance Rate.

To determine the relationship enforcement treatments and payment compliance in the fourth round, the Pearson's χ^2 was run. Since the dependent variable is categorical and not binary, the Fisher Exact test was not run for this hypothesis.

H_0 : The employment tax payment compliance and enforcement treatments are independent in the fourth round.

H_a : The employment tax payment compliance and enforcement treatments are not independent in the fourth round.

The results are shown on Table 28. The Pearson's χ^2 was run and the p-value was 0.075, which shows that we can accept the null hypothesis confirming that there is no

association between enforcement treatments and employment tax payment compliance. This confirms that enforcement treatments and employment tax payment compliance are not related at the community-standard 5 percent alpha protection level. The results confirm that a higher the number of enforcement treatments may not lead to higher payment compliance in the fourth round.

Table 28: Hypothesis 2b Round 4 Statistical Tests

Test	Value	df	Asymp. p-value (2-sided)
Pearson's χ^2	6.918 ^a	3	0.075
Likelihood Ratio	6.878	3	0.076
Linear-by-Linear Association	0.297	1	0.585

a. 2 cells (25.0%) have expected count less than 5. The minimum expected count is 1.58

As with the prior hypothesis, an effect size test was conducted. Table 29 shows the results from an effect size analysis using η^2 . The results found an η of 0.184 and an η^2 of 0.033. Using Cohen's interpretation guidelines (Ruscio, 2008), the results show a small effect size between enforcement treatments and payment compliance. These results are also in line with the Phi and Cramer's V test results. The results would imply enforcement treatments had a small effect on payment compliance rate the fourth round. However, since the results were not statistically significant at the 0.05 alpha scientific community protection level, we can only generalize these effect sizes to the population if the statistical power is above 0.8. In this case, the calculated post-hoc statistical power was 0.76. Therefore, we cannot generalize these results to the population.

Table 29: Hypothesis 2b Round 4 Effect Size Tests

Test	Value	Approx. p-value
η	0.184	-
Phi	0.184	0.075
Cramer's V	0.184	0.075

IX.3.3.2 Round 5 Test Results

Figure 17 shows the distribution of subjects' propensity to comply given the number of enforcement treatments (1 thru 4 treatments). As the picture shows, a larger number of participants subscribed to the safe harbor provision. The figure also shows similar distributions for both groups (subscribed and not subscribed). Unlike the previous round, the figure shows a similar rate of payment compliance given the number of enforcement treatments. The statistical tests for significance are shown below.

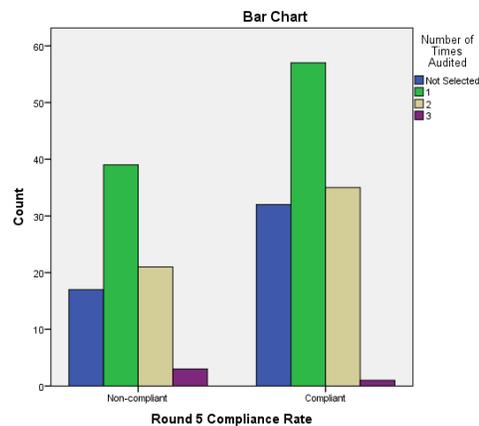


Figure 17: Round 5 Distributions between Enforcement Treatments and Payment Compliance Rate.

To determine the relationship enforcement treatments and payment compliance in the fifth round, the Pearson's χ^2 was run and the results are shown on Table 30. Since the dependent variable is categorical and not binary, the Fisher Exact test was not run.

H_0 : The employment tax payment compliance and enforcement treatments are independent in the fifth round.

H_a : The employment tax payment compliance and enforcement treatments are not independent in the fifth round.

The Pearson's χ^2 was run and the p-value was 0.437, which shows that we can accept the null hypothesis confirming that there is no association between enforcement

treatments and employment tax payment compliance. This confirms that enforcement treatments and payment compliance are not related at the community-standard 5 percent alpha protection level. The results confirm that a higher the number of enforcement treatments may not lead to higher payment compliance in the fifth round.

Table 30: Hypothesis 2b Round 5 Statistical Tests

Test	Value	df	Asymp. p-value (2-sided)
Pearson's χ^2	2.720 ^a	3	0.437
Likelihood Ratio	2.686	3	0.443
Linear-by-Linear Association	.600	1	0.439

a. 2 cells (25.0%) have expected count less than 5. The minimum expected count is 1.56

Table 31 shows the results from an effect size analysis using η^2 . The results found an η of 0.115 and an η^2 of 0.013. Using Cohen's interpretation guidelines (Ruscio, 2008), the results show a small effect size between enforcement encounter and payment compliance. These results are also in line with the Phi and Cramer's V test results. The results would imply enforcement treatments had a small effect on payment compliance rate the fifth round. However, since the results were not statistically significant at the 0.05 alpha scientific community protection levels, we can only generalize these effect sizes to the population if the statistical power is above 0.8. In this case, the calculated post-hoc statistical power was 0.38. Therefore, we cannot generalize these results to the population.

Table 31: Hypothesis 2b Round 5 Effect Size Tests

Test	Value	Approx. p-value
η	0.115	-
Phi	0.115	0.437
Cramer's V	0.115	0.437

IX.3.3.3 Round 6 Test Results

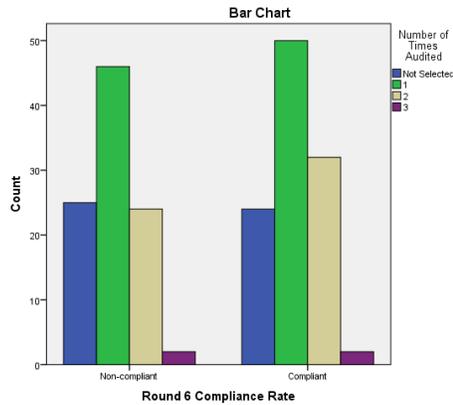


Figure 18 shows the distribution of subjects’ propensity to comply given the number of enforcement treatments (1 thru 4 treatments). As the picture shows, a larger number of participants subscribed to the safe harbor provision. The figure also shows similar distributions for both groups (subscribed and not subscribed). Unlike the previous round, the figure shows a similar payment compliance given the number of enforcement treatments. The statistical tests for significance are shown below.

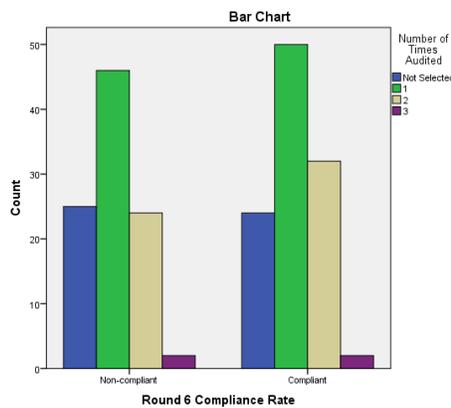


Figure 18: Round 6 Distributions between Enforcement Treatments and Payment Compliance Rate.

To determine the relationship enforcement treatments and payment compliance in the sixth round, the Pearson’s χ^2 was run. Again, the Fisher Exact test was not run.

H_0 : The employment tax payment compliance and enforcement treatments are independent in the fifth round.

H_a : The employment tax payment compliance and enforcement treatments are not independent in the fifth round.

The results are provided on [Table 32](#). The Pearson's χ^2 the p-value was 0.863, which shows that we can accept the null hypothesis confirming that there is no association between enforcement treatments and employment tax payment compliance. This confirms that that enforcement treatments and payment compliance are not related at the community-standard 5 percent alpha protection level. The results confirm that a higher the number of enforcement treatments may not lead to higher payment compliance in the sixth round.

Table 32: Hypothesis 2b Round 6 Statistical Tests

Test	Value	df	Asymp. p-value (2-sided)
Pearson's χ^2	0.742 ^a	3	0.863
Likelihood Ratio	0.743	3	0.863
Linear-by-Linear Association	0.559	1	0.455

a. 2 cells (25.0%) have expected count less than 5. The minimum expected count is 1.89

[Table 33](#) shows the results from an effect size analysis using η^2 . The results found an η of 0.060 and an η^2 of 0.003. Using Cohen's interpretation guidelines (Ruscio, 2008), the results show a small effect size between payment compliance and enforcement treatments. These results are also in line with the Phi and Cramer's V test results. The results would imply enforcement treatments had a small effect on payment compliance rate the sixth round. However, since the results were not statistically significant at the 0.05 alpha scientific community protection level, we can only generalize these effect sizes to the population if the statistical power is above 0.8. In this case, the calculated

post-hoc statistical power was 0.11. Therefore, we cannot generalize these results to the population.

Table 33: Hypothesis 2b Round 6 Effect Size Tests

Test	Value	Approx. p-value
η	0.060	-
Phi	0.060	0.863
Cramer's V	0.060	0.863

IX.3.3.4 Hypothesis 2b ANOVA Test Results

An ANOVA test was run for this hypothesis. The ANOVA test evaluates the mean difference between employment tax payment compliance and the number of enforcement treatments. As shown

Table 34, the p-values were greater than the community-standard 5 percent alpha protection level. Therefore, results confirm no dependency between the number of enforcement treatments and employment tax payment compliance.

Table 34: ANOVA tests for Hypothesis 2b

	Sum of Squares	df	Mean Square	F	p-value
Round 4	1.653	3	0.551	2.340	0.075
Round 5	0.647	3	0.216	0.901	0.442
Round 6	0.185	3	0.062	0.243	0.866

IX.3.4 Hypothesis 3

Hypothesis 3 sets out to determine whether business owners who chose a safe harbor provision will be more compliant, with respect to payment compliance. This section is segmented into three areas of results each one is focused on each round of testing.

The tests in this subsection will identify if there is any statistical differences in the employment tax payment compliance and the number of enforcement treatments. The

dependent variable for Hypothesis 3 is employment tax payment compliance, and the independent variable is subscription to the safe harbor provision.

$$PMT = \alpha + \beta(SR) \quad (6)$$

Where PMT is the employment tax payment compliance

SR is the subscription to the safe harbor provision

Formula 6 states that the employment tax payment compliance should be higher for those subjects who subscribe to the safe harbor provision than for subjects who do not.

Table 35 shows a sizable increase (about 10 percent) in the payment compliance rate once a safe harbor provision becomes an available choice. It also shows an improvement of approximately four percent in compliance as subjects undergo enforcement treatments (i.e. round 3). In the last round, we see a decrease in compliance when the income shock takes place.

Table 35: Compliance and Noncompliance Rate by Rounds

	Compliant	Non-Compliant
Round 1	49.3%	50.7%
Round 2	48.8%	51.2%
Round 3	51.2%	48.8%
Round 4	60.5%	39.5%
Round 5	61.0%	39.0%
Round 6	52.7%	47.3%

Table is similar to Table 45

IX.3.4.1 Round 4 Testing

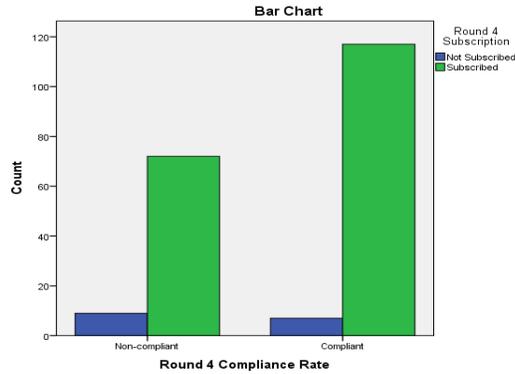


Figure 19 shows the distribution of payment compliance given the subscription to a safe harbor. As the picture shows, a larger number of subjects subscribed to the safe harbor provision. However, the figure shows slightly greater payment compliance for those subscribed to the safe harbor provision than for those who did not subscribe to the provision. The statistical tests for significance are shown below.

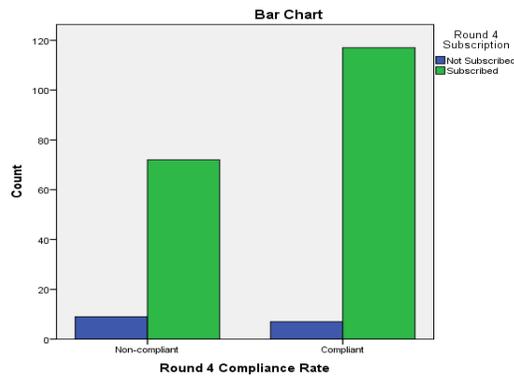


Figure 19: Round 4 Distributions between Payment Compliance and the Safe Harbor Provision Subscription Rate.

To determine the relationship payment compliance and subscription to a safe harbor provision in the fourth round, Fisher Exact test and Pearson's χ^2 were used. The results are show on.

H_0 : The employment tax payment compliance and the subscription to the safe harbor provision are independent in the fourth round.

H_a : The employment tax payment compliance and the subscription to the safe harbor provision are not independent in the fourth round.

The Fisher Exact test was not statistically significant (p-value of 0.124), so the null hypothesis is not rejected. The results confirm that there is no association between employment tax payment compliance and the subscription to the safe harbor provision. This conclusion is further corroborated by the Pearson's χ^2 (p-value of 0.154), which shows that we can accept the null hypothesis and conclude that compliance and subscription to the safe harbor provision are not related at the community-standard 5 percent alpha protection level. The results confirm that subscribing to safe harbor may not lead to higher employment tax payment compliance in the fourth round.

Table 36: Hypothesis 3 Round 4 Statistical Tests

Test	Value	df	Asymp. p-value (2-sided)	Exact p-value (2-sided)	Exact p-value (1-sided)
Pearson's χ^2	2.034 ^a	1	0.154		
Continuity Correction ^b	1.346	1	0.246		
Likelihood Ratio	1.982	1	0.159		
Fisher's Exact Test				0.186	0.124
Linear-by-Linear Association	2.024	1	0.155		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 6.32.

b. Computed only for a 2x2 table

Furthermore, [Table 37](#) shows the results from an effect size analysis using η^2 . The results found an η of 0.100 and an η^2 of 0.01. Using Cohen's interpretation guidelines (Ruscio, 2008), the results show a small effect size between compliance and subscription to the safe harbor provision. These results are also in line with the Phi and Cramer's V test results. The results would imply the safe harbor subscription had a small effect on payment compliance rate the fourth round. However, since the results were not

statistically significant at the 0.05 alpha scientific community protection level, we can only generalize these effect sizes to the population if the statistical power is above 0.8. In this case, the calculated post-hoc statistical power was 0.30. Therefore, we cannot generalize these results to the population.

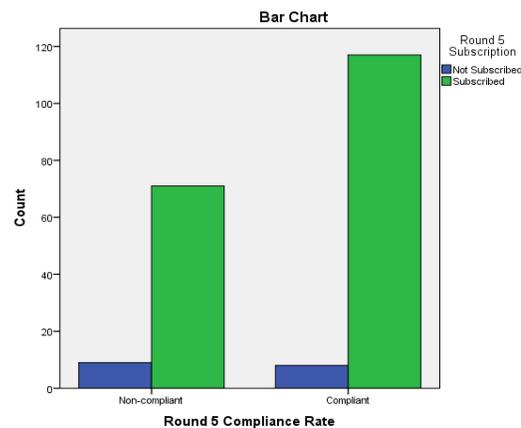
Table 37: Hypothesis 3 Round 4 Effect Size Tests

Test	Value	Approx. p-value
η	0.100	-
Phi	0.100	0.154
Cramer's V	0.100	0.154

IX.3.4.2 Round 5

Figure 20 shows the distribution of subjects' payment compliance given their subscription to the safe harbor provision. As the picture shows, a larger number of subjects subscribed to the safe harbor provision, and the figure shows slightly greater payment compliance for those subscribed to the safe harbor provision than those who did not subscribe to the provision (as in round 4). The statistical tests for significance are shown below.

Figure 20: Round 5 Distributions between Payment Compliance and the Safe Harbor Provision Subscription Rate.



Fisher Exact Test and Pearson's χ^2 tests were used to determine the relationship payment compliance and subscription to a safe harbor provision in the fifth round, and results are show on [Table 38](#).

H₀: The employment tax payment compliance and the subscription to the safe harbor provision are independent in the fifth round.

H_a: The employment tax payment compliance and the subscription to the safe harbor provision are not independent in the fifth round.

The Fisher Exact test p-value was 0.166, so the null hypothesis is not rejected. The results confirm that there is no association between employment tax payment compliance and the subscription to the safe harbor provision. This conclusion is further collaborated by the Pearson's χ^2 (p-value of 0.219), which shows that we accept the null hypothesis and conclude that employment tax payment compliance and subscription to the safe harbor provision are not related at the community-standard 5 percent alpha protection level. The results confirm that subscribing to safe harbor may not lead to higher employment tax payment compliance in the fifth round.

Table 38: Hypothesis 3 Round 5 Statistical Tests

Test	Value	df	Asymp. p-value (2-sided)	Exact p-value (2-sided)	Exact p-value (1-sided)
Pearson's χ^2	1.509 ^a	1	0.154		
Continuity Correction ^b	0.938	1	0.333		
Likelihood Ratio	1.470	1	0.225		
Fisher's Exact Test				0.299	0.166
Linear-by-Linear Association	1.501	1	0.220		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 6.63.

b. Computed only for a 2x2 table

Furthermore, [Table 39](#) shows the results from an effect size analysis using η^2 . The results found an η of 0.086 and an η^2 of 0.007. Using Cohen's interpretation guidelines (Ruscio, 2008), the results show a small effect size between compliance and subscription to the safe harbor provision. These results are also in line with the Phi and Cramer's V test results. The results would imply a safe harbor subscription had a small effect on payment compliance rate the fifth round. However, since the results were not statistically significant at the 0.05 alpha scientific community protection level, we can only generalize these effect sizes to the population if the statistical power is above 0.8. In this case, the calculated post-hoc statistical power was 0.22. Therefore, we cannot generalize these results to the population.

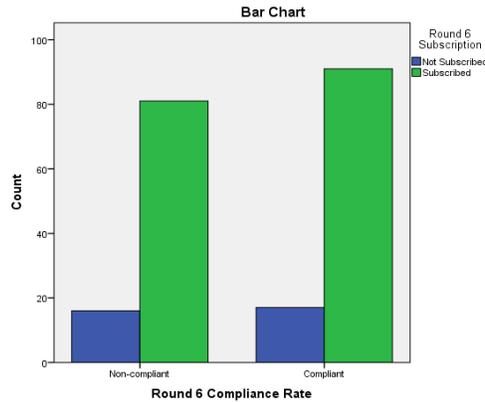
Table 39: Hypothesis 3 Round 5 Effect Size Tests

Test	Value	Approx. p-value
η	0.086	-
Phi	0.086	0.219
Cramer's V	0.086	0.219

IX.3.4.3 Round 6

[Figure 21](#) shows the distribution of subjects' decision to subscribe to a safe harbor and their payment compliance. As the picture shows, a larger number of subjects subscribed to the safe harbor provision. However, the figure does not shows much difference in payment compliance for those subscribed to the safe harbor provision than those subjects who did not subscribed to the safe harbor provision. This is a departure from rounds 4 and 5. The statistical tests for significance are shown below.

Figure 21: Round 6 Distributions between Compliance Payment and the Safe Harbor Provision Subscription Rate.



On [Table 40](#) the results for this round are provided and it tests the relationship between payment compliance and subscription to a safe harbor provision in the sixth round.

H_0 : The employment tax payment compliance and the subscription to the safe harbor provision are independent in the sixth round.

H_a : The employment tax payment compliance and the subscription to the safe harbor provision are not independent in the sixth round.

The Fisher Exact test p-value was 0.516 accepting the null hypothesis. Therefore, there is no association between employment tax payment compliance and subscription to the safe harbor provision in the sixth round. This conclusion is further collaborated by the Pearson’s χ^2 (p-value of 0.883), which shows that we accept the null hypothesis and conclude that compliance and subscription to the safe harbor provision are not related at the community-standard 5 percent alpha protection level. The results confirm that subscription to the safe harbor provision may not lead to higher employment tax payment compliance in the sixth round.

Table 40: Hypothesis 3 Round 6 Statistical Tests

Test	Value	df	Asymp. p-	Exact p-	Exact p-
------	-------	----	-----------	----------	----------

			value (2- sided)	value (2- sided)	value (1- sided)
Pearson's χ^2	0.022 ^a	1	0.883		
Continuity Correction ^b	0.000	1	1.000		
Likelihood Ratio	0.022	1	0.883		
Fisher's Exact Test				1.000	0.516
Linear-by-Linear Association	0.021	1	0.884		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 15.61.

b. Computed only for a 2x2 table

Furthermore, [Table 41](#) shows the results from an effect size analysis using η^2 . The results found an η of 0.010 and an η^2 of 0.000. Using Cohen's interpretation guidelines (Ruscio, 2008), the results show a small effect size between compliance and subscription to the safe harbor provision. These results are also in line with the Phi and Cramer's V test results. The results would imply the safe harbor subscription had a small effect on payment compliance rate the sixth round. However, since the results were not statistically significant at the 0.05 alpha scientific community protection level, we can only generalize these effect sizes to the population if the statistical power is above 0.8. In this case, the calculated post-hoc statistical power was 0.04. Therefore, we cannot generalize these results to the population.

Table 41: Hypothesis 3 Round 5 Effect Size Tests

Test	Value	Approx. p-value
η	0.010	-
Phi	0.010	0.883
Cramer's V	0.010	0.883

IX.3.4.4 Hypothesis 3 ANOVA test

An ANOVA test was run for this hypothesis. The ANOVA test evaluates the mean difference between employment tax payment compliance and the subscription to

the safe harbor provision. As shown on [Table 42](#), all the p-values were greater than the community-standard 5 percent alpha protection level. This confirms that there was no dependency between payment compliance and subscription to the safe harbor provision (i.e., subscribed and not subscribed). These results found not difference in the employment tax payment compliance for those who subscribe or did not subscribe. These results may be attributed to two events: 1) most of the subjects subscribed to the safe harbor and 2) overall employment tax payment compliance improved once subscription became available.⁴²

Table 42: Hypothesis 3 ANOVA Tests

	Sum of Squares	df	Mean Square	F	p-value
Round 4	0.486	1	0.486	2.035	0.155
Round 5	0.359	1	0.359	1.505	0.221
Round 6	0.005	1	0.005	0.021	0.884

IX.3.5 Hypothesis 4

Hypothesis 4 sets out to determine whether a business owner who chose a safe harbor provision will remain more compliant after an income shock⁴³ is experienced. In this section, we built up on the results obtained in the [Section 9.3.4: Hypothesis 3](#). However, in this section, a ratio variable was used as the dependent variable and binary variable remained as the independent variable.

The tests in this subsection will identify if there is any statistical differences in the employment tax payment compliance and the subscription to a safe harbor provision given an income shock. The dependent variable for Hypothesis 4 was employment tax

⁴² The safe harbor subscription rate is discussed in Section IX.3.1 and in Section IX.3.4. The improvement in employment tax payment compliance is discussed in Section IX.3.6

⁴³ Income shock is defined as a 50 percent drop in firm revenue

payment compliance and the independent variable is subscription to the safe harbor provision given the introduction of an income shock.

$$PMT = \alpha + \beta(SR | I) \tag{7}$$

Where PMT is the employment tax payment compliance category

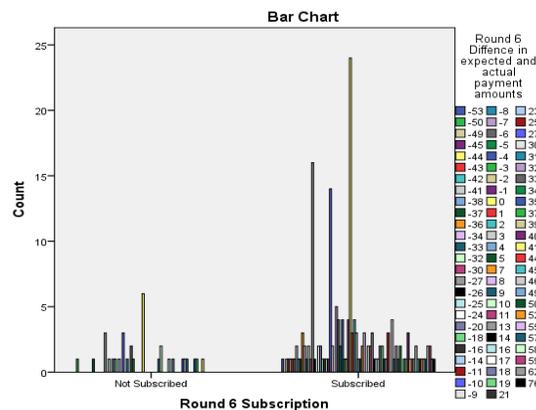
SR is the subscription to the safe harbor provision

I is the income shock imposed on round 6

Formula 7 states that the employment tax payment compliance should be higher for those subjects who subscribe to the safe harbor provision than for subjects who do not subscribe even during an income shock.

Figure 22 shows the distribution of subjects' decision to subscribe to a safe harbor and their payment compliance. As the picture shows, a larger number of subjects subscribed to the safe harbor provision. However, the figure does not shows much difference in payment compliance for those subscribed to the safe harbor provision and those subjects who did not subscribed to the safe harbor provision.

Figure 22: Round 6 Distributions between Payment Compliance and the Safe Harbor Provision Subscription Rate.



To determine the relationship between employment tax payment compliance and subscription to a safe harbor provision after the income shock is applied, a Pearson's χ^2

was run and the results are show on [Table 43](#). Because of the data classification, the Fisher Exact test was not run.

H_0 : The employment tax payment compliance and the subscription to the safe harbor provision given the income shock are independent in the sixth round.

H_a : The employment tax payment compliance and the subscription to the safe harbor provision given the income shock are not independent in the sixth round.

The Pearson's χ^2 the p-value was 0.883, which shows that we must accept the null hypothesis confirming that there is no association between employment tax payment compliance and subscription to the safe harbor provision once the income shock is applied. This confirms that enforcement treatments and payment compliance are not related at the community-standard 5 percent alpha protection level. The results confirm that subscribing to safe harbor during and income shock may not lead to higher payment compliance.

Table 43: Hypothesis 4 Statistical Tests

Test	Value	df	Asymp. p-value (2-sided)
Pearson's χ^2	68.543 ^a	76	0.716
Likelihood Ratio	69.431	76	0.690
Linear-by-Linear Association	0.177	1	0.674

a. 150 cells (97.4%) have expected count less than 5. The minimum expected count is 0.16.

b. Computed only for a 2x2 table

Furthermore, [Table 44](#) shows the results from an effect size analysis using η^2 . The results found an η of 0.578 and an η^2 of 0.334. Using Cohen's interpretation guidelines (Ruscio, 2008), the results show a moderate effect size between compliance and subscription to the safe harbor provision. These results are also in line with the Phi and Cramer's V test results. The results would imply the safe harbor subscription had a

moderate effect on payment compliance rate the sixth round. However, since the results were not statistically significant at the 0.05 alpha scientific community protection level, we can only generalize these effect sizes to the population if the statistical power is above 0.8. In this case, the calculated post-hoc statistical power was 1.00. Therefore, we can generalize these results to the population.

Table 44: Hypothesis 4 Effect Size Test

Test	Value	Approx. p-value
η	0.578	-
Phi	0.578	0.716
Cramer's V	0.578	0.716

IX.3.5.1 Hypothesis 4 ANOVA test

An ANOVA test was run for this hypothesis. The ANOVA test evaluates the mean difference between employment tax payment compliance and the subscription to the safe harbor provision given an income shock. As [Table 45](#) show, all the p-values were greater than the community-standard 5 percent alpha protection level. This confirms that there was no dependency between payment compliance and subscription to the safe harbor provision.

Table 45: Round 6 ANOVA Test

	Sum of Squares	df	Mean Square	F	p-value
Round 6	0.005	1	0.005	0.021	0.675

The results in hypothesis 3 and hypothesis 4 are similar, but required additional testing.

IX.3.6 Additional Payment Compliance Analysis

The payment compliance rate for rounds one through six are provided on [Table 46](#). The information for round one through three provides the propensity to underpay.

The payment compliance rate for rounds four through six provided a compliance propensity for subjects under the safe harbor provision. In general, the compliance increased for those rounds under the safe harbor provision.

Table 46: Payment Compliance frequency by Round

	Compliance	Non-compliance	Compliance Rate
Round 1	101	104	49.3
Round 2	100	105	48.8
Round 3	105	100	51.2
Round 4	124	81	60.5
Round 5	125	80	61.0
Round 6	108	97	52.7

Table similar to Table 34

IX.3.6.1 Payment Compliance ANOVA

An index was created to identify the compliance rate before and after the safe harbor subscription provision was permitted. Using ANOVA test, we tested to see if the propensity to comply had changed under the safe harbor provision. The results are provided on [Table 47](#) and they found statistical significance at 5%. This means that the p-value was less than the community-standard 5 percent alpha protection level. This confirms that there was dependency between payment compliance and subscription to the safe harbor provision.

Table 47: ANOVA Test between Compliance Rates under the Safe Harbor Provision

	Sum of Squares	df	Mean Square	F	p-value
Between groups	11.738	3	3.913	38.962	0.000
Within Groups	20.184	201	0.100		
Total	31.922	204			

Furthermore, [Table 48](#) shows the results from an effect size analysis using η^2 . The results found an η of 0.606 and an η^2 of 0.368. Using Cohen's interpretation guidelines (Ruscio, 2008), the results show a moderate effect size between compliance

and subscription to the safe harbor provision. These results are also in line with the Phi and Cramer's V test results. The results would imply the safe harbor provision had a moderate effect on payment compliance rate on the last three rounds. Since the results were statistically significant at the 0.05 alpha scientific community protection level and the statistical power is 1.00, we can generalize these results to the population.

Table 48: Compliance Rates under a Safe Harbor Provision Effect Size Test

Test	Value	Approx. p-value
η	0.606	
Phi	0.647	0.000
Cramer's V	0.373	0.000

IX.3.6.2 Logistic Regression Control Compliance

Using FA score obtained from both the Risk Components (see [Table 10](#)) a logistic regression was ran to demine if risk control or speculative risk had any impact on compliance or subscription rate. This is possible because our constructs were valid and reliable. The four tables below show the results.

IX.3.6.2.1 Employment Tax Payment Compliance

[Table 49](#) shows the payment compliance results obtained by running a logistic regression model using the FA analysis scores as independent variables (risk control, speculative risk and the interaction of risk control and speculative risk). [Table 50](#) Provides the Nagelkerke R^2 and the final omnibus model p-value. The model is

$$\log \frac{P(\text{Employment Tax Payment Compliance})}{1-P(\text{Employment Tax Payment Compliance})} = \alpha + \beta_1 (\text{risk control}) + \beta_2 (\text{speculative risk}) + \beta_3 (\text{risk control} \times \text{speculative risk})$$

Table 49: Payment Compliance logistic Regression Results Given Risk control

Compliance					
	B	S.E.	Wald	p-value	Odd Ratio
Round 1	-0.029	0.140	0.044	0.364	0.971
Round 2	-0.049	0.140	0.122	0.727	0.952
Round 3	0.049	0.140	0.122	0.727	1.050
Round 4	0.426	0.143	8.884	0.003 ^a	1.531
Round 5	0.446	0.143	9.716	0.002 ^a	1.562
Round 6	0.107	0.140	0.590	0.443	1.113

a. Model obtained statistically significant results for the constant (or y-intercept).

Table 50: Goodness of fit analysis

Test	p-value	Nagelkerke R ²
Round 1	0.746	0.000
Round 2	0.875	0.000
Round 3	0.938	0.000
Round 4	0.322	0.000
Round 5	0.477	0.000
Round 6	0.126	0.000

As the results from

Table 49 and Table 50 show, none of the dependent variables were statically significant. Furthermore, the goodness of fit test showed that this model was not able to explain any variation (R2 of 0.000). Therefore, risk controls (risk control or speculative risk) did not have any influence payment compliance. On the other hand, the fourth and fifth rounds were statistically significant and the odds ratio shows some effect on payment compliance. However, since the results were not statistically significant at the 0.05 alpha scientific community protection level, we can only generalize these effect sizes to the population if the statistical power is above 0.8. In this case, the calculated

post-hoc statistical power was 0.04. Therefore, we cannot generalize these results to the population.

In addition to risk controls, further exploration was conducted to determine if in fact payment compliance had improved by the implementation of the safe harbor provision. Other logistic regression models were run to explore this premise.

Table 51 provides the results from the final logistic regression model:

$$\begin{aligned} \log \frac{P(\text{Employment Tax Payment Compliance})}{1-P(\text{Employment Tax Payment Compliance})} &= \alpha + \beta_1 (\text{Prior round compliance}) + \beta_2 (\text{Tax liability threshold}) \\ &+ \beta_3 (\text{Availability to a safe harbor provision}) + \beta_4 (\text{Income shock}) + \beta_5 (\text{enforcement treatments in the prior round}) \\ &= -586 + 2.294(\text{Prior round compliance}) - 1.289(\text{Tax liability threshold}) + 0.460 (\text{Availability to a safe harbor provision}) - \\ &0.603(\text{Income shock}) + 0.412(\text{enforcement treatments in prior round}) \end{aligned}$$

Table 51: Payment Compliance logistic Regression Results

	Payment Compliance				
	B	S.E.	Wald	p-value	Odd Ratio
Prior round compliance	2.294	.166	191.727	0.000	9.919
Tax liability threshold	-1.289	.162	63.231	0.000	0.276
Availability to a safe harbor provision	.460	.177	6.732	0.009	1.584
Income Shock	-.603	.232	6.789	0.009	0.547
Enforcement treatments in prior round	.412	.200	4.249	0.039	1.510
Constant	-.586	.181	10.519	0.001	0.557

Table 52: Goodness of fit analysis

	Cox & Snell R ²	Nagelkerke R ²
Pseudo R ²	0.343	0.458

As the results from the

Table 51 and Table 52 show, subjects who were compliant in prior round would continue to maintain their compliant behavior. It also showed that payment compliance decreases when the subject was assigned to the high tax liability threshold. This is in accordance with the literature. The model also showed that when the safe harbor provision is available, subjects are more likely comply with their payment obligations. The model found that income shock and prior enforcement experience would reduce the payment compliance of participants. Furthermore, the goodness of fit test showed that this model was able to explain between 34 and 46 percent of the variation. The classification table did show a classification prediction accuracy of 79.2 percent.

Since the results were statistically significant at the 0.05 alpha scientific community protection level, we can generalize these effect sizes to the population if the statistical power is above 0.8. In this case, the calculated post-hoc statistical power was 1.00. Therefore, we can generalize these results to the population.

IX.3.6.2.2 Subscription Rate

Table 53 shows the subscription results obtained from the logistic regression model using FA analysis scores as independent variables (risk control, speculative risk and the interaction of risk control and speculative risk). Table 54 Provides the Nagelkerke R² and the final omnibus model p-value.

$$\log \frac{P(\text{Subscription Rate})}{1-P(\text{Subscription Rate})} = \alpha + \beta_1 (\text{risk control}) + \beta_2 (\text{speculative risk}) + \beta_3 (\text{risk control} \times \text{speculative risk})$$

Table 53: Payment Compliance logistic Regression Results Given Risk control

		Compliance				
		B	S.E.	Wald	p-value	Odd Ratio
Round 4	Risk Control ^a	-0.490	0.241	4.119	0.042	0.613
	Y-intercept	2.574	0.283	82.534	0.000	13.125

Round 5	Risk Control	-0.460	0.236	3.819	0.051	0.631
	Y-intercept	2.495	0.273	83.637	0.000	12.120
Round 6	Risk Control	-0.266	0.184	74.637	0.147	0.766
	Y-intercept	1.651	0.190	75.470	0.000	5.212

a. speculative risk and the interaction of risk control and speculative risk were not statistically significant.

Table 54: Goodness of fit analysis

Test	p-value	Nagelkerke R ²
Round 4	0.046	0.046
Round 5	0.054	0.041
Round 6	0.000	0.000

As the results from [Table 53](#) and [Table 54](#) show, one of the dependent variables (risk control) was statically significant. However, it was only significant at predicting subscriptions rate in round 4, but not in the other two rounds. The goodness of fit test showed that this model was not able to explain much of the variation, with an explain variation of around 4 percent. The classification table did show a classification prediction accuracy of 91.7 percent.

Although some of the results were statistically significant at the 0.05 alpha scientific community protection level, we can still generalize these effect sizes to the population if the statistical power is above 0.8. In this case, the calculated post-hoc statistical power was 0.742. Therefore, we cannot generalize these results to the population.

As with payment compliance, further logistics regressions were conducted to determine which factors drove the subjects' decision to subscribe to the safe harbor provision. The final model was:

$$\log \frac{P(\text{Subscription to Safe Harbor Provision})}{1-P(\text{Subscription to Safe Harbor Provision})} = \alpha + \beta_1 (\text{Prior round compliance}) + \beta_2 (\text{prior round enforcement encounter} \times \text{availability of safe harbor}) + \beta_3 (\text{Availability to a safe harbor provision} \times \text{tax liability threshold}) + \beta_4 (\text{prior round compliance} \times \text{tax liability threshold})$$

$$-1.695 + 1.970 (\text{Prior round compliance}) + 1.748 (\text{prior round enforcement encounter} \times \text{availability of safe harbor}) + 3.512 (\text{Availability to a safe harbor provision} \times \text{tax liability threshold}) - 2.564 (\text{prior round compliance} \times \text{tax liability threshold})$$

Table 55: Subscription to Safe Harbor Regression Results

	Payment Compliance				
	B	S.E.	Wald	p-value	Odd Ratio
Prior round compliance	1.970	.190	107.428	0.000	9.919
Interaction of prior round enforcement encounter and availability of a safe harbor	1.748	.319	29.942	0.000	0.276
Interaction of availability of a safe harbor and tax liability threshold	3.512	.243	208.095	0.009	1.584
Interaction of prior round compliance and tax liability threshold	-2.564	.306	70.058	0.009	0.547
Constant	-1.695	.161	111.157	0.039	1.510

Table 56: Goodness of fit analysis

	Cox & Snell R ²	Nagelkerke R ²
Pseudo R ²	0.320	0.427

Table 55 and

Table 56 provide the results from the model. The results found subjects who were compliant in prior rounds were almost 10 times more likely to subscribe to a safe harbor provision. On the other, the results found that subjects who encounter enforcement in the prior round were less likely to subscribe to a safe harbor provision once the provision became available. The model results also found that higher tax liability threshold subjects were more likely to subscribe to the safe harbor provision once it became available. Lastly, the model showed that subject who were compliant and had a high tax liability threshold were more likely to subscribe to the safe harbor provision.

The goodness of fit test showed that this model was able to explain between 32 and 42 percent of the variation. The classification table did show a classification prediction accuracy of 77.0 percent.

Since the results were statistically significant at the 0.05 alpha scientific community protection level, we can generalize these effect sizes to the population if the statistical power is above 0.8. In this case, the calculated post-hoc statistical power was 1.00. Therefore, we can generalize these results to the population.

X DISCUSSION

X.1 Contribution to Theory

The theoretical motivation of this study is to enhance our understanding of taxpayers' antisocial behavior using General Deterrence Theory. Deterrence is the impact that punishment has on those who have not committed a criminal act (Blumstein et al., 1978). In terms of tax enforcement, deterrence is the impact tax audits and their accompanying punishments, such as interest, penalties, and imprisonment, have on those individuals who have failed to report and pay taxes their tax requirements.

The SMORC states that there are two methods to increase individuals' marginal cost: (1) higher probability of being apprehended and (2) increased magnitude of punishment (Allingham & Sandmo, 1972; Ariely, 2012; Upadhyay, 2013). This study has shown similar results. It showed that increased probability of apprehension leads to increased compliance. Most importantly, this study has shown that individuals were willing to subscribe in large numbers to a safe harbor provision that eliminate the probability of apprehension in order to reduce their burden and avoid punishment (audit/enforcement encounter and penalties related to audit results). Furthermore, this study was able to show that individuals were willing to improve their compliance rate when enforcement was increased as the literature had prescribed in other circumstances (Alm et al., 1992; Bloomquist, 2009; Hasseldine et al., 2007; Kalplowitz, 1973; Lijiao et al., 2014).

General Deterrence Theory provides an explanation as to how different levels of punishment influence antisocial behavior. The general measurement of deterrence can be achieved by calculating the change in criminal rates at different levels of punishment (Blumstein et al., 1978). In this study a 10 percent improvement in compliance was

achieved. This provides some evidence that General Deterrence Theory in fact reduced the antisocial behavior associated with the employment tax payment compliance.

X.2 Contribution to Practice

The practical motivation was to improve the understanding of voluntary reporting and payment of employment tax in the hopes of influence related IRS processing. The results from this study suggest that provisions such as a safe harbor can be a method of reducing filing costs and audit costs and ultimately taxpayer burden. On the other hand, the results of this study were inconclusive in determining if such provisions can improve payment compliance. Nevertheless, the outcome of this study can improve timing and accuracy of employment taxes and it may improve the accuracy of employment tax payment.

X.2.1 Employment Tax Motivation

The practical focus on employment tax was driven by the impact it has on the tax gap, which accounts for an approximately 72 billion dollars (IRS, 2012b) loss to the US government every year. The practical focus for using a safe harbor was based on the IRS believes that improving voluntary compliance is necessary to reduce the tax gap (IRS, 2009). Given the complexity of the IRC and the IRS responsibility for ensuring that taxpayers understand and meet their tax obligations, including filing and payment of employment tax, this research results can assist the Service in improving their client's tax obligations and reducing taxpayer burden. This study set out to determine if a safe harbor provision could minimize the employment tax filing burden and improve employment tax payment compliance. As the results show, a safe harbor provisions can be a method for reducing filing burden and potentially improve payment compliance.

As the data in Table 1 suggested, third party matching already exists and filing compliance is high. However, that does not reduce enforcement activities or the time it takes to file the necessary Form 94X to comply with the payment of employment tax. The results of this study show that a safe harbor provision would reduce the filing burden and some enforcement activities related to employment tax filing. Furthermore, this research shows potential opportunities in improving employment tax payment compliance. The results show an increase of approximately 10 percent in payment compliance during a safe harbor periods that did not encounter an income shock. This shows that while a safe harbor provision will not prevent a subscriber from not complying during difficult financial epochs, it does show that subjects are willing to comply under normal circumstances. Therefore, this study has shown that tax authorities can reduce taxpayer burden and can use such provisions to reduce failure-to-pay employment tax.

X.2.2 Compliance Motivation

The practical focus on compliance is based on the premise that random audits lead to better predictions of noncompliance (Bloomquist, 2009). However, the current IRS state of smaller corporate audit coverage, current compliance identification methods, and shrinking enforcement personnel budgets, there is a need to improve voluntary payment, filing, and reporting compliance for employment tax. Furthermore, increasing deficits put pressure on the IRS to enhance compliance.

Although this study would not improve the NRP model used for predicting compliance, future studies could improve how or when the IRS conducts physical audit. Since subscribers of a safe harbor provision are more likely to be compliant and a large

number of subjects were willing to subscribe to a safe harbor provision, the IRS could substantially reduce the number of Revenue Agents (RAs) it uses to conduct field audits. The importance of this narrower focus study is because physical audits are costly and IRS budgets are shrinking.

Because of long running deficits, the IRS began identifying ways to reduce the tax gap (Brown & Mazur, 2003; Holland, 1958; IRS, 1996). After the great recession of 2008 the IRS instituted initiatives to curtail the \$72 billion tax gap from employment tax (IRS, 2009). This study is in line with the IRS's initiative to revising the collection due process for employment taxes and studying employment tax reporting and filing compliance (IRS, 2009).

This research study provides ways to simplify the filing and payment process associated with the employment tax and it shows a statistical improvement on the current process. This study builds on the simplification of employment tax payment, which Bloomquist (2003b) states could lead to increased compliance.

XI LIMITATIONS

The design used has great internal validity. However, the external validity for this type of design is not sufficient to infer to the greater population. One of the major reasons is the inability to determine the presence of non-response bias and the ability to randomly select subjects from the population. Therefore, inference to the subjects who undertook this experiment is possible, but not to the population. In future research, obtaining from the IRS a sample from the population of employers who actually pay employment tax to randomly select participants would be helpful in overcoming this limitation, and this could result in different conclusions.

Another limitation of this design can be the lack of realism that the web-based experiment may exhibit, which may lead to behaviors not observed in the naturally occurring world. Lastly, this study was a web-based experiment and there is a valid criticism as to its rigor. A major criticism is the ability for subjects to drop out of the experiment because it is easier to end a web based experiment than walking out of a laboratory experiment (Wade & Tingling, 2005). Another critic is that only those with web access are able to access the experiment. Nevertheless, Wade and Tingling (2005) do suggest that web experiments can be effective with “internal validity, volunteer bias, demand characteristics, external validity, financial cost and time” (pg. 81). To overcome this design limitation, a similar experiment could be augmented using lab and web subjects to compare and verify results. A more robust study would include a pilot using a sample from the actual population, where results could be compared to the control group not selected in the pilot.

During the completion of the experiment, subjects were instructed to identify how much employment tax withholding they needed to pay. However, about 20 percent of all

participants paid only the employer's (or employee) portion of the employment tax payment. It is possible that even though the instructions stated they needed to pay both employer and employee's portion, some subjects only paid the employer's portion without noticing that they had to pay both. Since there was not interaction between the student PI and the subjects, it is impossible to identify if this was done on purpose or if it was an oversight. On the other hand, if individuals who encounter enforcement had been subject to some additional educational information, it is possible that the hypothesis 3 and 4 results would have been quite different. This limitation suggests future research design changes, naturally.

Carryover effect (learning effect, or order effect) was another design limitation. Carryover effect occurs when experiment subject have undergone the same task or scenario a number of times, and subjects either become better through practice or become worse through fatigue. Since the experiment did not randomly assign the income shock through different rounds of the experiment, the study results cannot separate if the results are due to the introduction of the treatment (income shock) or if subjects were simply learning how to respond to the experiment. To resolve this limitation, future studies could use counterbalancing techniques. Future studies could implement randomize order of the income shock treatment variables as used in Kalambokidis et al. (2012).

Lastly, payment of participants is a limitation on the design. Since the participants were all paid the same amount regardless of their selection, participants did not have an economic incentive to maximize their payoffs. This limits the ability to identify whether or not subjects reacted to economic incentive. Future studies could use true economic penalties for incorrectly paying or filing their taxes during the experiment

by reducing economic payout (i.e., pay less to those who are found noncompliant). Furthermore, the experiment could implement a random payout selection. This rule would randomly identify the results of a round to compensate the subjects (Kalambokidis et al., 2012). The enhancements identified could be used to better mimic a real world environment and assess true taxpayer behavior.

APPENDICES

XI.1 Appendix A: Test Questionnaire

You have been asked to play the role of the typical small business owner with four fulltime employees. Your business gross receipts are \$350,000 dollars for this round. Your industry operating expenses are estimated to be around 20% of gross receipts and other expenses (such as cost of goods sold) are approximately 40% of gross receipts.

During this round, you should expect to operate as if it were a normal business environment. This will include paying your accounts payable, overhead, and taxes for your business. You will also be required to finalize a tax form, and should expect for all current laws to apply as you file and pay employment tax. In addition, you should expect the normal application of the law and that legal system will work its way through as expected.

Over the past several years, your business has experienced normal economic conditions and you have experienced average sales numbers, which you expect will grow at a normal rate in the future. In addition, your firm has maintained strong cash flow and your accounts receivables and accounts payables are up to date. With respect to your employees, you have had no issues paying their salaries and have made all the quarterly employment tax withholding deposits.

During this round, the Experiment Authorities will not conduct any audits.

The following information will help in assessing your company's financial situation for this questionnaire. You must use the information provided to determine the employment tax obligation.

1. Your company gross receipts were \$350,000 this quarter.
2. Your firm employs four (4) full time fulltime employees.
 - a. Their average income was \$15,250 for the quarter, a calculation which totals \$61,000 or (4 x \$15,250) in labor costs.
 - b. Your employment tax per employee should be relatively easy to pay at \$2,500 per employee. This totals \$10,000 per quarter for all employees.
3. Your employees' federal withholding for employment tax is \$10,000 (This is money withheld from employees' income to be deposited to the Enforcement Authorities by the employer).

Simplified Balance Sheet		
Accounts Receivables	Debits	Credits
Gross Receipts (Revenue)		\$350,000
Wages, tips, and other compensation	\$61,000	
Employment tax (Employer's Portion)	\$10,000	
Operating expenses (e.g., rent, utility, etc.)	\$70,000	
Other Expenses (e.g., COGS, inventory, etc.)	\$140,000	
Total Expenses		(\$281,000)
Net Income		\$61,000

XI.2 Appendix B: IRS Tax Form 944

Form **944 for 2014: Employer's ANNUAL Federal Tax Return**

Department of the Treasury — Internal Revenue Service

OMB No. 1545-2007

Employer identification number (EIN) -

Name (not your trade name)

Trade name (if any)

Address

Number Street Suite or room number

City State ZIP code

Foreign country name Foreign province/county Foreign postal code

Who Must File Form 944
 You must file annual Form 944 instead of filing quarterly Forms 941 only if the IRS notified you in writing. Instructions and prior-year forms are available at www.irs.gov/form944.

Read the separate instructions before you complete Form 944. Type or print within the boxes.

Part 1: Answer these questions for this year. Employers in American Samoa, Guam, the Commonwealth of the Northern Mariana Islands, the U.S. Virgin Islands, and Puerto Rico can skip lines 1 and 2.

1 Wages, tips, and other compensation	1	<input style="width: 90%;" type="text"/>
2 Federal income tax withheld from wages, tips, and other compensation	2	<input style="width: 90%;" type="text"/>
3 If no wages, tips, and other compensation are subject to social security or Medicare tax	3	<input type="checkbox"/> Check and go to line 5.
4 Taxable social security and Medicare wages and tips:		
	<i>Column 1</i>	<i>Column 2</i>
4a Taxable social security wages	<input style="width: 80%;" type="text"/> .	× .124 = <input style="width: 80%;" type="text"/> .
4b Taxable social security tips	<input style="width: 80%;" type="text"/> .	× .124 = <input style="width: 80%;" type="text"/> .
4c Taxable Medicare wages & tips	<input style="width: 80%;" type="text"/> .	× .029 = <input style="width: 80%;" type="text"/> .
4d Taxable wages & tips subject to Additional Medicare Tax withholding	<input style="width: 80%;" type="text"/> .	× .009 = <input style="width: 80%;" type="text"/> .
4e Add Column 2 from lines 4a, 4b, 4c, and 4d	4e	<input style="width: 90%;" type="text"/> .
5 Total taxes before adjustments. Add lines 2 and 4e	5	<input style="width: 90%;" type="text"/> .
6 Current year's adjustments (see instructions)	6	<input style="width: 90%;" type="text"/> .
7 Total taxes after adjustments. Combine lines 5 and 6	7	<input style="width: 90%;" type="text"/> .
8 Total deposits for this year, including overpayment applied from a prior year and overpayments applied from Form 944-X, 944-X (PR), 944-X (SP), 941-X, or 941-X (PR)	8	<input style="width: 90%;" type="text"/> .
9a Reserved		
9b Reserved		
10 Reserved		
11 Balance due. If line 7 is more than line 8, enter the difference and see instructions	11	<input style="width: 90%;" type="text"/> .
12 Overpayment. If line 8 is more than line 7, enter the difference <input style="width: 80%;" type="text"/> .		Check one: <input type="checkbox"/> Apply to next return. <input type="checkbox"/> Send a refund.

▶ You MUST complete both pages of Form 944 and SIGN it.

Next ➡

Name (not your trade name)	Employer identification number (EIN)
----------------------------	--------------------------------------

Part 2: Tell us about your deposit schedule and tax liability for this year.

13 Check one: Line 7 is less than \$2,500. Go to Part 3.
 Line 7 is \$2,500 or more. Enter your tax liability for each month. If you are a semiweekly depositor or you accumulate \$100,000 or more of liability on any day during a deposit period, you must complete Form 945-A instead of the boxes below.

Jan.	Apr.	Jul.	Oct.
13a <input style="width:100%;" type="text"/>	13d <input style="width:100%;" type="text"/>	13g <input style="width:100%;" type="text"/>	13j <input style="width:100%;" type="text"/>
Feb.	May	Aug.	Nov.
13b <input style="width:100%;" type="text"/>	13e <input style="width:100%;" type="text"/>	13h <input style="width:100%;" type="text"/>	13k <input style="width:100%;" type="text"/>
Mar.	Jun.	Sep.	Dec.
13c <input style="width:100%;" type="text"/>	13f <input style="width:100%;" type="text"/>	13i <input style="width:100%;" type="text"/>	13l <input style="width:100%;" type="text"/>

Total liability for year. Add lines 13a through 13l. Total must equal line 7. 13m

Part 3: Tell us about your business. If question 14 does NOT apply to your business, leave it blank.

14 If your business has closed or you stopped paying wages...
 Check here and enter the final date you paid wages.

Part 4: May we speak with your third-party designee?

Do you want to allow an employee, a paid tax preparer, or another person to discuss this return with the IRS? See the instructions for details.

Yes. Designee's name and phone number

Select a 5-digit Personal Identification Number (PIN) to use when talking to IRS.

No.

Part 5: Sign Here. You MUST complete both pages of Form 944 and SIGN it.

Under penalties of perjury, I declare that I have examined this return, including accompanying schedules and statements, and to the best of my knowledge and belief, it is true, correct, and complete. Declaration of preparer (other than taxpayer) is based on all information of which preparer has any knowledge.

X Sign your name here

Date

Print your name here

Print your title here

Best daytime phone

Paid Preparer Use Only

Check if you are self-employed

Preparer's name PTIN

Preparer's signature Date

Firm's name (or yours if self-employed) EIN

Address Phone

City State ZIP code

Form 944-V, Payment Voucher

Purpose of Form

Complete Form 944-V, Payment Voucher, if you are making a payment with Form 944, Employer's ANNUAL Federal Tax Return. We will use the completed voucher to credit your payment more promptly and accurately, and to improve our service to you.

Making Payments With Form 944

To avoid a penalty, make your payment with your 2014 Form 944 **only** if one of the following applies.

- Your net taxes for the year (Form 944, line 7) are less than \$2,500 and you are paying in full with a timely filed return.
- You already deposited the taxes you owed for the first, second, and third quarters of 2014, and the tax you owe for the fourth quarter of 2014 is less than \$2,500, and you are paying, in full, the tax you owe for the fourth quarter of 2014 with a timely filed return.
- Your net taxes for the third quarter are \$2,500 or more, net taxes for the fourth quarter are less than \$2,500, and you did not incur a \$100,000 next-day deposit obligation during the fourth quarter.
- You are a monthly schedule depositor making a payment in accordance with the Accuracy of Deposits Rule. See section 11 of Pub. 15 (Circular E), Employer's Tax Guide, for details. In this case, the amount of your payment may be \$2,500 or more.

Otherwise, you must make deposits by electronic funds transfer. See section 11 of Pub. 15 (Circular E) for deposit

instructions. Do not use Form 944-V to make federal tax deposits.

Caution. Use Form 944-V when making any payment with Form 944. However, if you pay an amount with Form 944 that should have been deposited, you may be subject to a penalty. See Deposit Penalties in section 11 of Pub. 15 (Circular E).

Specific Instructions

Box 1—Employer identification number (EIN). If you do not have an EIN, you may apply for one online. Go to IRS.gov and type "EIN" in the search box. You may also apply for an EIN by faxing or mailing Form SS-4, Application for Employer Identification Number, to the IRS. If you have not received your EIN by the due date of Form 944, write "Applied For" and the date you applied in this entry space.

Box 2—Amount paid. Enter the amount paid with Form 944.

Box 3—Name and address. Enter your name and address as shown on Form 944.

- Enclose your check or money order made payable to the "United States Treasury" and write your EIN, "Form 944," and "2014" on your check or money order. Do not send cash. Do not staple Form 944-V or your payment to Form 944 (or to each other).

- Detach Form 944-V and send it with your payment and Form 944 to the address provided in the Instructions for Form 944. Do not send a photocopy of Form 944-V because your payment may be misapplied or delayed.

Note. You must also complete the entity information above Part 1 on Form 944.

Detach Here and Mail With Your Payment and Form 944.

							
Form 944-V Department of the Treasury Internal Revenue Service		Payment Voucher				OMB No. 1545-2007	
		▶ Do not staple this voucher or your payment to Form 944.				2014	
1 Enter your employer identification number (EIN).		2 Enter the amount of your payment. ▶ <small>Make your check or money order payable to "United States Treasury"</small>		Dollars		Cents	
		3 Enter your business name (individual name if sole proprietor).					
		Enter your address.					
						Enter your city, state, and ZIP code or your city, foreign country name, foreign province/county, and foreign postal code.	

XI.3 Appendix C: Recruitment Script and Hand-out

Recruiting Email: Pre-Test

Subject line: Recruiting for pre-test on the filing an payment of employment tax study

Hello to all EDB candidates and alumni,

My name is Rafael Dacal. I am the student principal investigator within our program. The study title is “The Effects of Employment Tax Reporting Regimens on US Small Businesses and Proprietors’ Payment Compliance.” I am currently eliciting the help of 10 present or past EDB students to partake in a 30 minute structure interview on the topic of employment tax.

The purpose of this research is to study the tax behavior exhibited by sole-proprietors and small business owners. In particular, the study will test changes in the payment requirements of employment tax. This study does not require any personal information.

The purpose of the structure interviews is to refine the data collection instrument (DCI). As all of you know, participation in this research is voluntary. You are not required to partake in this study. If you decide to participate in the study and change your mind once you have started, you have the right to drop out at any time.

If interested, please contact me via this email or call (305) 582.8811.

Sincerely,

Rafael Dacal

Recruiting Email: Test

Subject line: Recruiting for experiment on the filing an payment of employment tax study

Hello to all EDB candidates and alumni,

My name is Rafael Dacal. I am the student principal investigator within our program. The study title is “The Effects of Employment Tax Reporting Regimens on US Small Businesses and Proprietors’ Payment Compliance.” I am currently eliciting the help of present or past EDB students to partake in a 30 minute experiment on the topic of employment tax.

The purpose of the web-based research is to study the tax behavior exhibited by sole-proprietors and small business owners. In particular, the study will test changes payment requirements of employment tax. This study does not require any personal information.

As all of you know, participation in this research is voluntary. You are not required to partake in this study.

If you interested in participating, please contact me via this email or call (305) 582.8811.

Sincerely,

Rafael Dacal



Now that you have agreed to participate in a business study



- Thank you for participating
- You will be able to make up to \$10
 - As a token of appreciation, and
 - To compensate for your time
- Please go to <http://www.qualtrics.com/XXXXXX> to complete the experiment
- This experiment should take approximately 30 minutes.

XI.4 Appendix D: Additional literature review

XI.4.1 Corporate responsibility

The corporate responsibility literature has discussed such constructs as brand reputation and social responsibility as they are associated with the filing and payment of taxes. At the center of the work on brand reputation is how firms manage their public perception as a social steward when designing tax strategies. Currently companies are managing their tax strategies by assessing how aggressive the public perceives the company tax strategies (Donohoe et al., 2014; Harvey Jr, 2014). If a firm is perceived as too aggressive by the public, the public perception may impact the company's filing strategy (Hanlon & Slemrod, 2009). Moreover, companies that demonstrate higher levels of corporate social responsibility were more likely to engage in lower levels of tax avoidance (Dowling, 2014; Lanis & Richardson, 2015) and, therefore, choose less aggressive tax strategies.

The link being made in with this literature review is how subjective norms effect individual actions. While individual filings are private and the exerted behavior is not a public, the subjective norm link could be considers tenuous. However, corporate filings are not private and connection on reputation and subjective norms could be derived. Since corporate tax literature was used as a proxy for the segment studied (small business), it could be suggested that corporate responsibility can affect employment tax behaviors in small business in similar fashion – not identical fashion.

XI.4.2 Employment Tax and Tax Incidence

In economics, the term-of-art tax incidence is used as a framing for who ultimately is responsible for paying the tax. The general perception is that the individual

consuming the product pays the taxes. However, that is not always the case. To illustrate this concept, let us discuss the consumption of labor. Firms consume labor. In times of economic decline, the number of employee vis-à-vis jobs is higher (i.e., an excess of labor). In such time, employers do not have increase wage to acquire new labors, and thus any increase in employment tax will be passed to employees because employees will accept less favorable job in such economic circumstances. In times of economic expansion, the number of employee vis-à-vis jobs is lower (i.e., a shortage of labor). In such time, employers would have increase wage to acquire new labors, and thus any increase in employment tax will be borne by employer.

In terms of labor markets, the party (either employer or employee) with the highest price elasticity of labor will pay a smaller portion of the tax. The literature has shown that this tax is more commonly “borne by labour” (pg. 189) (Hamermesh, 1979; Holmlund, 1981; Vroman, 1974). The literature has also shows that employment tax incidence has a bigger impact on the wage earning class than the investment class. This impact due because the investment class’s effective income is lower than the wage earning class (Sugin, 2014) and because the payroll tax has increased in the US from 4.2 percent to 6.2 percent in 2013 “as part of the fiscal cliff resolution” (pg. 94) (Bracha & Cooper, 2013).⁴⁴ The increase of employment tax has an impact on effective wages, which may lead to increase noncompliance of self-employed taxpayers who may be trying to increase their real wages.

Other studies have analyzed the impact of different taxes on income (e.g., income tax versus payroll tax). Some studies have shown that payroll taxes are a greater

⁴⁴ The changes required the payroll tax to revert to the previous level on January 1, 2013.

proportion of the tax incidence for the majority of US households because the average household has a larger proportion of wage income than investment income (Heim et al., 2014; Mitrusi & Poterba, 2001). Studies concluded that payroll taxes are becoming a greater portion of the tax incidence because of the introduction of the EITC, which reduces taxable income. On the other hand, there have been no significant changes in the employment tax law. Moreover, Heim et al. (2014) explains that the positive payment of employment tax has decreased from 79 percent to 75 percent between 2001 and 2011.

From these aforementioned studies, we can infer that employment tax is becoming a larger portion of the tax incidence for wage earners and that the tax incidence has decreased over the last decade. Nevertheless, the authors do not articulate a causal reason as to why these changes are taking place. They do attribute some of these changes to the great recession and changes in the code,⁴⁵ but these attributions are not based on research.

To be able to assess the employment tax incident, we also must understand how employment tax affects the supply of labor. Some research studies have discussed how decreases in payroll taxes increase labor participation due to a reduction in employment tax (Creedy, 2010; Engelhardt & Kumar, 2014; Kugler & Kugler, 2009). In other words, a decrease in payroll tax will increase labor supply because of an increase in real wages and a reduction in labor cost (Goerke, 2002; Månsson & Quoreshi, 2015; Saez et al., 2012b). This suggested increase in labor supply the tax incident on employees, which would have little impact on employment tax compliance. On the other hand, a reduction

⁴⁵ The change in the Code refers to the decrease in the payroll tax from 6.2 to 4.2 in 2010 to 2013.

on the supply of labor could affect employment tax compliance because employers have to make the payments.

What is the challenge to understanding the impact of employment tax compliance? MacRae and MacRae (1976) show that the supply of labor would not change with an increase in the payroll tax because there would be an offset between elastic and inelastic workers. Moffitt (1977) clarifies that the offsets occur when the budget constraint is linear, but a non-linear budget constraint would lead to a reduction in labor supply⁴⁶ anytime a payroll tax increase. In general, the literature shows that there is an inverse relation between labor supply and tax changes, but with more complex models with additional dimensions of taxpayer behavior, the relationship between labor supply and tax changes is ambiguous or negligible (Bauer & Riphahn, 2002; Collins & Plumlee, 1991; Cruces, Galiani, & Kidyba, 2010; Moffitt, 1977). Because of mixed results in the empirical studies, there is no clear way to identify how employment tax incident influences its compliance.

Beyond the tax incidence and its impact on labor and wages, a paper defined payroll tax dilemma. The payroll tax dilemma is when a firm cannot meet its operational liabilities and it uses employment tax withholdings to pay for these liabilities (Godfrey, 2004; Mauldin & Wilder, 1997). Companies normally face a liquidity crisis during economic downturns or economic distress, such as the great recession (Grady, 2013). Grady (2013) indicates that many of these businesses continue to illegally borrow from withheld taxes for expanded periods of time, and found that the majority of those businesses will eventually fail. Bloomquist (2003a) also argues that a financial strain is

⁴⁶ Most studies focus on the labor supply (instead of labor demand) because employees had borne the cost of an increase in employment tax.

one of the determinants of noncompliance for all taxpayers. One of the arguments of this study is that some of the authorized individual will succumb to the payroll tax dilemma when confronting a liquidity crisis (or strain). The argument is that individuals decide to use employment tax withholding as a short-term loan without clearly understanding the negative implications.

XI.4.3 Characteristics of non-filers

Much of the compliance literature focuses on non-filers because they introduce biases on the outcome of all predictions. The consequence of non-filers is that the absence of their data makes all noncompliance appear less fraudulent. Therefore, understanding non-filers is essential to minimize noncompliance (Feinstein, 1991). The inability to adjust the data (e.g., using weights) for non-filers increases the tax gap, reduces our understanding of tax compliance behavior, and impedes proper compliance enforcement.

Erard and Ho (2001) identified non-filers as “ghosts” (pg.26) because very little is known about this segment of the tax base. More explicitly, they labeled these taxpayers as ghosts because they do not comply with any filing requirements and it is very difficult for authorities or academic to identify or locate them. The non-filers in Erard and Ho’s (2001) study were located by the IRS through “intensive search by IRS agents” (pg. 30). They explain that if an individual is hard to find it would increase the taxpayer’s willingness not to file.

Erard and Ho’s (2001) also found that variables such as prior filing, marital status, and age could be used to identify known filer. For instance, older taxpayers are less likely to be noncompliance. The general consensus is that they are more risk adverse.

On the other hand, married couples are more likely to be noncompliant because their taxes are more complex and lead to more evasion. This view is consistent with Andreoni et al. (1998), Clotfelter (1983), and Feinstein (1991) studies. These studies found that the complexity of tax filing has an impact on taxpayers. For example, Erard and Ho's (2001) found that taxpayers who are at the margin of the filing requirement may be deterred from complying because of the cost associated with filing taxes. They also found that non-filers are more likely to be recipients of blue-collar business income, and are more likely to have their source of income from business or capital gains. Again, suggesting that lack of detection increases noncompliance.

The most interesting characteristic of ghosts is that income and taxable income are substantially smaller than other filers, and ghosts use less adjustment than other filers (Erard & Ho, 2001). This is interesting because one of their suggestions associated with this behavior was that the non-filer would appear as an accidentally noncompliant taxpayer, instead of deliberate act of evasion. This act of evasion draws attention to the issues of normative beliefs and subjective norms (Bloomquist, 2003b; Fischer et al., 1992). Normative beliefs are defined as the established behaviors, and it is assumed that all individuals will behave according to expectations (Ajzen, 1991; CSLI, 2011). Subjective norm is the peer pressure to behave in a certain way (Chatzisarantis & Biddle, 1998). The collective of normative beliefs of an individual is the foundation of his\her subjective norms (Ajzen, 1991). Therefore, it could be derived that these ghosts believe the act of non-compliance is acceptable at the individual and group level.

As with tax literature, compliance literature has also has shown that individuals' compliance will be affected similar belief and norms, and provides three plausible causes:

- Embracement of apprehension
- Satisfaction with the government or fiscal equity
- Fairness of the tax system

Gordon (1989) study examined the concept of morality as deterrence of tax evasion. Morality is affected by two forces: internal and external. The internal force could be described as an individual's ethics. He proposes that individuals with high ethical standards will continue to properly and voluntarily comply even when the cost of evading are substantially lower than the gains. The external forces can simply be explained as peer influence. He specifies reputation cost of being apprehended. There are two alternative outcomes. If there are many people being apprehended, noncompliance becomes a badge of honor and it may increase. On the other hand, if within a group, the evader is alone. He/she may feel a sense of shame because of the reputational cost, leading to improved compliance.

In his study, Gordon (1989) also found that "small evaders" (pg. 804) were more likely to become noncompliant as tax rate increased. Small evaders were defined as taxpayers who were already underreporting their taxes. Whereby, "non-evaders" (pg. 800) were less likely to change their filing behavior because the decision was not a monetary decision, but ethical. The ethical decision was based on personal cost of breaking the law, and the reputational cost associated with being apprehended.

As mentioned above, there appears to be a relationship between perceptions of fiscal equity tax compliance. Fiscal equity could be defined as perception on treatment of taxpayer, the level of taxpayer burden, and complexity of the code (Bloomquist, 2003a). An example of inequity could be a progressive tax, such as the income tax, where

taxpayers in the upper income brackets are required to pay more as a proportion of their income. Conversely, a regressive tax, such as sales tax, could be considered inequitable for taxpayers in the lower income brackets because they would be required to pay more taxes as a proportion of their income.

Another source of inequity was identified by Piketty and Saez (2003). In their paper, they provide IRS time series data on income, and it showed that income inequality over the last few decades has increased. Bloomquist (2003a) explains how income inequality influences overall reporting compliance, basing most of her argument on the concealment sources of income. On the other hand, she emphasizes how the lack of taxpayers' financial stability due to income inequality leads to noncompliance. She argues that because individuals' future expected cost of detection is lower than the present need for money. In other words, individuals' calculations of potential outcomes are biased due to temporal distance and individuals many times believe that the rosy outcome will prevail (Pennington & Roese, 2003).⁴⁷ This is due to people's inability to accurately and consistently predict future outcomes (Kahneman, 2011). Therefore, taxpayers' expected utility maximization formulation is incorrect and it results in noncompliance.

Spicer and Becker (1980) also provide empirical evidence that the relationship does exist between equity and compliance. Their theoretical argument is based on expected utility theory and "inequity theory" (pg. 174), and it stresses that individuals will use tax evasion as a vehicle to adjust inequities. They propose that if tax equity is perceived favorable to the taxpayer, he/she will reduce his/her noncompliance level. For

⁴⁷ This optimistic outlook could also be due to insensitivity to predictability (Tversky & Kahneman, 1974)

example, Bloomquist (2003a) provides evidence of this phenomenon using data following the Tax Reform Act of 1986. In her findings, lower effective taxes on the wealthiest American lead to better compliance. On the other hand, if tax equity hinders taxpayer, he/she will increase the tax evasion level.

Spicer and Becker (1980) results showed that “victims” (pg. 171) of inequity would increase their noncompliance behavior. This finding is also collaborated by Alm et al. (1992), which explained that to increase compliance the government should make individuals aware of how their taxes are benefiting those who pay. Examples of this could be Washington State DOT, Miami-Dade County (FL) transportation, and Forsyth County (GA). All of these governmental bodies provide some source of communiqué near public work sites to show taxpayers’ how their taxes are being used.

They too study the relationship between equity and compliance, and found a causal effect between the two. Studies claim that level of satisfaction with government services and will have an impact on compliance level (Alm, 1991; Alm et al., 1992). Lastly, fairness is based on the individuals’ perception of how many people are getting away with evading taxes (Andreoni et al., 1998; Erard, 1997; Erard & Ho, 2001). If individuals perceive that a large number of people are evading taxes, they too are likely to underreport.

REFERENCES

- Ajzen, I. 1991. The Theory of Planned Behavior. *Organizational Behavior & Human Decision Processes*, 50(2): 179.
- Allingham, M. G., & Sandmo, A. 1972. Income Tax Evasion: A Theoretical Analysis. *Journal of Public Economics*, 1(3-4): 281-400.
- Alm, J. 1991. A Perspective on the Experimental Analysis of Taxpayer Reporting: 577: American Accounting Association.
- Alm, J., Jackson, B. R., & McKee, M. 1992. Estimating the determinants of taxpayer compliance with experimental data. *National Tax Journal*, 45(1): 107-114.
- Alm, J., Jackson, B. R., & McKee, M. 2004. The Effects of Communication Among Taxpayers on Compliance, *2004 IRS Research Conference*. Washington, D.C.: IRS.
- Alm, J., & Yunus, M. 2009. Spatiality and Persistence in U.S. Individual Income Tax Compliance. *National Tax Journal*, 62(1): 101-124.
- Alreck, P. L., & Settle, R. B. 1995. Planning the Project. In N. McGuffin (Ed.), *The Survey Research Handbook*, 2nd ed.: 25-53. USA: Richard D. Irwin.
- Andreoni, J., Erard, B., & Feinstein, J. 1998. Tax compliance. *Journal of Economic Literature*, 36(2): 818.
- Apel, R. 2013. Sanctions, Perceptions, and Crime: Implications for Criminal Deterrence. *Journal of Quantitative Criminology*, 29(1): 67.
- Ariely, D. 2012. *The (honest) truth about dishonesty : how we lie to everyone-- especially ourselves / Dan Ariely*: New York : Harper, [2012]
First edition.

- Arno, P. S., Sohler, N., Viola, D., & Schechter, C. 2009. Bringing health and social policy together: the case of the earned income tax credit. *Journal Of Public Health Policy*, 30(2): 198-207.
- Athreya, K. B., Reilly, D., & Simpson, N. B. 2010. Earned Income Tax Credit Recipients: Income, Marginal Tax Rates, Wealth, and Credit Constraints. *Federal Reserve Bank of Richmond Economic Quarterly*, 96(3): 229-258.
- Averett, S., & Wang, Y. 2013. The Effects of Earned Income Tax Credit Payment Expansion on Maternal Smoking. *Health Economics*, 22(11): 1344-1359.
- Barney, D., Tschopp, D., & Wells, S. 2012. Tax Simplification through Readability. *CPA Journal*, 82(12): 6-10.
- Bauer, T., & Riphahn, R. T. 2002. Employment effects of payroll taxes – an empirical test for Germany. *Applied Economics*, 34(7): 865-876.
- Beck, P. J., Davis, J. S., & Jung, W.-O. 1991. Experimental Evidence on Taxpayer Reporting under Uncertainty: 535: American Accounting Association.
- Beck, P. J., Davis, J. S., & Woon-Oh Jung, W.-O. 1992. Experimental evidence on an economic model of taxpayer aggression under strategic and nonstrategic audits. *Contemporary Accounting Research*, 9(1): 86-112.
- Becker, G. S. 1968. Crime and Punishment: An Economic Approach. *Journal of Political Economy*, 76(2): 169.
- Berloff, G., & Modena, F. 2013. Income shocks, coping strategies, and consumption smoothing: An application to Indonesian data. *Journal of Asian Economics*, 24: 158-171.

- Bhattacharjee, A. 2012. Research Design, *Social Science Research: Principles, Methods, and Practices*, 2nd ed.: 159. USA: USF Tampa Bay Open Access Textbooks Collection.
- Bloomquist, K. M. 2003a. Tax Evasion, Income Inequality and Opportunity Costs of Compliance: 91-104. Washington: National Tax Association.
- Bloomquist, K. M. 2003b. Trends as Changes in Variance: The Case of Tax Noncompliance, *the 2003 IRS Research Conference*: 9. Washington DC: IRS.
- Bloomquist, K. M. 2009. A Comparative Analysis of Reporting Compliance Behavior in Laboratory Experiments and Random Taxpayer Audits. *Proceedings of the Annual Conference on Taxation*: 113-122.
- Blumenthal, M., Christian, C., & Slemrod, J. 1998. The Determinants of Income Tax Compliance: Evidence from a Controlled Experiment in Minnesota: 6575. Cambridge: National Bureau of Economic Research, Inc.
- Blumenthal, M., Erard, B., & Chih-Chin, H. 2005. Participation and Compliance with the Earned Income Tax Credit. *National Tax Journal*, 58(2): 189-213.
- Blumenthal, M., Kalambokidis, L., & Turk, A. 2012. Subsidizing Charitable Contributions With a Match Instead of a Deduction: What Happens to Donations and Compliance? *National Tax Journal*, 65(1): 91-116.
- Blumstein, A., Cohen, J., & Nagin, D. 1978. *Deterrence and incapacitation : estimating the effects of criminal sanctions on crime rates / Alfred Blumstein, Jacqueline Cohen, and Daniel Nagin, editors ; Panel on Research on Deterrent and Incapacitative Effects, Committee on Research on Law Enforcement and*

- Criminal Justice, Assembly of Behavioral and Social Sciences, National Research Council*: Washington : National Academy of Sciences, 1978.
- Bracha, A., & Cooper, D. 2013. Asymmetric Responses to Tax-Induced Changes in Personal Income: The 2013 Payroll Tax Hike versus Anticipated 2012 Tax Refunds. *Research Review*(20): 93-95.
- Brown, R. E., & Mazur, M. J. 2003. IRS's Comprehensive Approach to Compliance Measurement. *National Tax Journal*, 56(3): 689-700.
- Buettner, T., & Wamser, G. 2013. Internal Debt and Multinational Profit Shifting: Empirical Evidence from Firm-Level Panel Data. *National Tax Journal*, 66(1): 63.
- Burton-Jones, A. 2009. MINIMIZING METHOD BIAS THROUGH PROGRAMMATIC RESEARCH. *MIS Quarterly*, 33(3): 445-A447.
- Carroll, C. D. 2009. Precautionary saving and the marginal propensity to consume out of permanent income. *Journal of Monetary Economics*, 56: 780-790.
- Chatzisarantis, N. L. D., & Biddle, S. J. H. 1998. Functional significance of psychological variables that are included in the Theory of Planned Behaviour: a Self-Determination Theory approach to the study of attitudes, subjective norms, perceptions of control and intentions. *European Journal of Social Psychology*, 28(3): 303.
- Chen, Y., Ramamurthy, K., & Wen, K.-W. 2012. Organizations' Information Security Policy Compliance: Stick or Carrot Approach? *Journal of Management Information Systems*, 29(3): 157-188.

- Chittenden, F., Kauser, S., & Panikkos, P. 2003. Tax regulation and small business in the USA, UK, Australia and New Zealand. *International Small Business Journal*, 21(1): 93-115.
- Clark, J., Friesen, L., & Muller, A. 2004. The Good, the Bas, and the Regulator: An Experimental Test of Two Conditional Audit Schemes. *Economic Inquiry*, 42(1): 69-87.
- Clotfelter, C. T. 1983. Tax Evasion and Tax Rates: An Analysis of Individual Returns: 363: North-Holland Publishing Company.
- Cohen, J. 1988. *Statistical power analysis for the behavioral sciences*: Hillsdale, N.J. : L. Erlbaum Associates, 1988.
- 2nd ed.
- Collins, J. H., & Plumlee, R. D. 1991. The Taxpayer's Labor and Reporting Decision: The Effect of Audit Schemes: 559: American Accounting Association.
- Creedy, J. 2010. Elasticity of Taxable Income: An Introduction and Some Basic Analytics. *Public Finance & Management*, 10(4): 556-589.
- Crowne, D. P., & Marlowe, D. 1960. A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24(4): 349-354.
- Cruces, G., Galiani, S., & Kidyba, S. 2010. Payroll taxes, wages and employment: Identification through policy changes. *Labour Economics*, 17(4): 743-749.
- CSLI. 2011. Social Norms. In C. Bicchieri, & R. Muldoon (Eds.), *Stanford Encyclopedia of Philosophy*, Vol. 2015. California, USA: Center for the Study of Language and Information (CSLI), Stanford University.
- de la Vega, E. 2013. An Effective Anti-Poverty Program. *Social Policy*, 43: 8-10.

- Dengfeng, Y. A. N., & Sengupta, J. 2013. The Influence of Base Rate and Case Information on Health-Risk Perceptions: A Unified Model of Self-Positivity and Self-Negativity. *Journal of Consumer Research*, 39(5): 931-946.
- Donohoe, M. P. 2015. Financial Derivatives in Corporate Tax Avoidance: A Conceptual Perspective. *Journal of the American Taxation Association*, 37(1): 37-68.
- Donohoe, M. P., McGill, G. A., & Outslay, E. 2013. Back to the Drawing Board: The Structural and Accounting Consequences of a Switch to a Territorial Tax System. *National Tax Journal*, 66(3): 713.
- Donohoe, M. P., McGill, G. A., & Outslay, E. 2014. Risky Business: The Prosopography of Corporate Tax Planning. *National Tax Journal*, 67(4): 851.
- Dowling, G. G. D. u. e. a. 2014. The Curious Case of Corporate Tax Avoidance: Is it Socially Irresponsible? *Journal of Business Ethics*, 124(1): 173-184.
- Dwenger, N., & Steiner, V. 2012. Profit Taxation and the Elasticity of the Corporate Income Tax Base: Evidence from German Corporate Tax Return Data. *National Tax Journal*, 65(1): 117-150.
- Edwards, R. D. 2004. Macroeconomic Implications of the Earned Income Tax Credit. *National Tax Journal*, 57(1): 45-65.
- Engelhardt, G. V., & Kumar, A. 2014. Taxes and the Labor Supply of Older Americans: Recent Evidence from the Social Security Earnings Test. *National Tax Journal*, 67(2): 443.
- Erard, B. 1993. Taxation with representation. *Journal of Public Economics*, 52(2): 163.

- Erard, B. 1997. Self-selection with measurement errors A microeconomic analysis of the decision to seek tax assistance and its implications for tax compliance. *Journal of Econometrics*, 81(2): 319-356.
- Erard, B. 2002. Compliance Measurement and Workload Selection with Operational Audit Data, *The 2002 Internal Revenue Service Research Conference*: 28. George Washington University: IRS.
- Erard, B., & Ho, C.-C. 2001. Searching for Ghosts: Who Are the Nonfilers and How Much Tax Do They Owe? *Journal of Public Economics*, 81(1): 25-50.
- Erard, B., & Ho, C.-C. 2002. Developong an Econometric Model for MEasuring Tax Noncompliance Using Operational Audi Data *The 2002 American Statistical Association Conference*: 5. New York: American Statistical Association.
- Feinstein, J. S. 1991. An Econometric Analysis of Income Tax Evasion and Its Detection: 14: Rand.
- Feltham, G. D., & Paquette, S. M. 2002. The Interrelationship between Estimated Tax Payments and Taxpayer Compliance. *Journal of the American Taxation Association*, 24(2): 27.
- Fischer, C. M., Wartick, M., & Mark, M. M. 1992. Detection probability and taxpayer compliance: A review of the literature. *Journal of Accounting Literature*, 11: 1.
- Fisher, R. J. 1993. Social Desirability Bias and the Validity of Indirect Questioning. *Journal of Consumer Research*, 20(2): 303-315.
- Gallemore, J., Maydew, E. L., & Thornock, J. R. 2014. The Reputational Costs of Tax Avoidance. *Contemporary Accounting Research*, 31(4): 1103.

- Gittelman, S., Lange, V., Cook, W. A., Frede, S. M., Lavrakas, P. J., Pierce, C., & Thomas, R. K. 2015. Accounting for Social-Desirability Bias In Survey Sampling. *Journal of Advertising Research*, 55(3): 242-254.
- Godfrey, H. 2004. Avoid the Payroll Tax Trap. *Journal of Accountancy*, 198(5): 67-73.
- Goerke, L. 2002. Statutory and economic incidence of labour taxes. *Applied Economics Letters*, 9(1): 17-20.
- Gordon, J. P. F. 1989. INDIVIDUAL MORALITY AND REPUTATION COSTS AS DETERRENTS TO TAX EVASION. *European Economic Review*, 33(4): 797-805.
- Grady, T. A. 2013. The Criminalization of Payroll Tax Liabilities: Why This Harsh Direction? *Journal of Tax Practice & Procedure*, 15(5): 29-51.
- Grasmick, H. G., & Green, D. E. 1980. LEGAL PUNISHMENT, SOCIAL DISAPPROVAL AND INTERNALIZATION AS INHIBITORS OF ILLEGAL BEHAVIOR. *Journal of Criminal Law & Criminology*, 71(3): 325-335.
- Hair, J. F. 1998. *Multivariate data analysis with readings*: New York : Macmillan ; Toronto : Maxwell Macmillan Canada ; New York : Maxwell Macmillan International, c1998.
- 5th ed.
- Hair, J. F. 2014. *A primer on partial least squares structural equations modeling (PLS-SEM)*: Los Angeles : SAGE, c2014.
- Hamermesh, D. S. 1979. New Estimate of the Incidence of the Payroll Tax. *Southern Economic Journal*, 45(4): 1208.

- Hanlon, M., & Heitzman, S. 2010. A review of tax research. *Journal of Accounting & Economics*, 50(2/3): 127-178.
- Hanlon, M., & Slemrod, J. 2009. What Does Tax Aggressiveness Signal? Evidence from Stock Price Reactions to News about Tax Shelter Involvement. *Journal of Public Economics*, 93(1-2): 126-141.
- Harvey Jr, J. R. 2014. CORPORATE TAX AGGRESSIVENESS--RECENT HISTORY AND POLICY OPTIONS. *National Tax Journal*, 67(4): 831.
- Hasseldine, J., Hite, P., & Toumi, M. 2007. Persuasive Communications: Tax Compliance Enforcement Strategies for Sole Proprietors. *Contemporary Accounting Research*, 24(1): 171-194.
- Heim, B. T., Lurie, I. Z., & Pearce, J. 2014. WHO PAYS TAXES? A DYNAMIC PERSPECTIVE. *National Tax Journal*, 67(4): 755-778.
- Heinsman, D. T., & Shadish, W. R. 1996. Assignment methods in experimentation: When do nonrandomized experiments approximate answers from randomized experiments? *Psychological Methods*, 1(2): 154-169.
- Holland, D. M. 1958. DIVIDEND UNDERREPORTING ON TAX RETURNS. *Journal of Finance*, 13(2): 238-260.
- Holmlund, B. 1981. A NOTE ON CHANGES IN PAYROLL TAXES-DOES LEGAL INCIDENCE MATTER? *National Tax Journal*, 34(4): 479-482.
- Homburg, C., Fürst, A., Ehrmann, T., & Scheinker, E. 2013. Incumbents' defense strategies: a comparison of deterrence and shakeout strategy based on evolutionary game theory. *Journal of the Academy of Marketing Science*, 41(2): 185-205.

Hryshko, D. 2014. Correlated income shocks and excess smoothness of consumption.

Journal of Economic Dynamics and Control, 48: 41-62.

IBM. 2017. IBM Knowledge Center. In K. a. B. s. Test (Ed.), Vol. 2017: Web.

http://www.ibm.com/support/knowledgecenter/SSLVMB_20.0.0/com.ibm.spss.statistics.cs/fac_telco_kmo_01.htm: IBM.

IRS. 1996. Federal Tax Compliance Research: Individual Income Tax Gap Estimates for 1985, 1988, and 1992. In D. o. t. Treasury (Ed.): 71. Washington, DC: Internal Revenue Service.

IRS. 2006. The Examination (Audit) Process, Vol. 2015: provides DIF case scoring information. IRS.gov: IRS.

IRS. 2009. Update on Reducing the Federal Tax Gap and Improving Voluntary Compliance. In U. D. o. Treasury (Ed.): 35. Washington D.C.: IRS.

IRS. 2012a. IRS Releases New Tax Gap Estimates; Compliance Rates Remain Statistically Unchanged From Previous Study, Vol. 2016. Washington DC: IRS.

IRS. 2012b. Tax Gap for Tax Year 2006 Overview. In T. D. o. Treasury (Ed.): 3. Washington D.C.: Us Government.

IRS. 2013. Publication 15 (Circular E), Employer's Tax Guide: For use in 2014. In D. o. t. Treasury (Ed.): 67. US: US Government.

IRS. 2014a. Fiscal Year 2013 Enforcement and Services Results. In U. S. D. o. t. Treasury (Ed.). USA: IRS.

IRS. 2014b. Publication 505, Tax Withholding and Estimated Tax, For use in 2014. In D. o. t. Treasury (Ed.): 61. US: US Government.

IRS. 2015. IRS Audits, Vol. 2015. IRS.gov: IRS.

- Jay, J. R. 1815. For the North-American Journal. Tax Return in England: 24: Wells and Lilly.
- Johns, A., & Slemrod, J. 2010. THE DISTRIBUTION OF INCOME TAX NONCOMPLIANCE. *National Tax Journal*, 63(3): 397-418.
- Johnston, A. C., Warkentin, M., & Siponen, M. 2015. AN ENHANCED FEAR APPEAL RHETORICAL FRAMEWORK: LEVERAGING THREATS TO THE HUMAN ASSET THROUGH SANCTIONING RHETORIC. *MIS Quarterly*, 39(1): 113-A117.
- Jones, R. C., Thomas, J. F., & Lang, T. K. 2012. INCOME TAX REFORM AND THE FLAT TAX. *International Journal of Business, Accounting, & Finance*, 6(2): 1-12.
- Joost, M. E. P., & Smidts, A. 2000. Assessing the Construct Validity of Risk Attitude: 1337: Institute for Operations Research and the Management Sciences.
- Joulfaian, D., & Rider, M. 1998. DIFFERENTIAL TAXATION AND TAX EVASION BY SMALL BUSINESS. *National Tax Journal*, 51(4): 675-687.
- Kahneman, D. 2011. *Think Fast and Slow*. United States: Cornell University Press.
- Kaiser, H. 1974. An index of factorial simplicity. *Psychometrika*, 39(1): 31.
- Kalambokidis, L., Turk, A., & Blumenthal, M. 2012. Behavioral Experiments of Alternative Reporting Regimes: Transparency vs. Burden. In T. D. o. Treasury (Ed.): 45-66. Washington DC: IRS.
- Kalplowitz, S. A. 1973. An Experimental Test of a Rationalistic Theory of Deterrence. *Journal of Conflict Resolution*, 17(3): 535-572.

- Kenkel, D. S., & Schmeiser, M. D. 2014. Is Smoking Inferior?: Evidence from Variation in the Earned Income Tax Credit. *Journal of Human Resources*(4): 1094.
- Koskinen, J. 2015. Key Note Address to How Do IRS Budget Cuts Affect Taxpayers and the Tax System? In U.-B. T. P. Center (Ed.), *April 8, 2015 Urban-Brookings Tax Policy Center Proceedings*, Key Note Address ed.: 57. Alexandria, VA.
- Krumpal, I. 2013. Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & Quantity*, 47(4): 2025-2047.
- Kubick, T. R., Lynch, D. P., Mayberry, M. A., & Omer, T. C. 2015. Product Market Power and Tax Avoidance: Market Leaders, Mimicking Strategies, and Stock Returns. *Accounting Review*, 90(2): 675.
- Kugler, A., & Kugler, M. 2009. Labor Market Effects of Payroll Taxes in Developing Countries: Evidence from Colombia. *Economic Development & Cultural Change*, 57(2): 335-358.
- Kukk, M., & Staehr, K. 2013. Income Underreporting by Households with Business Income. Evidence from Estonia. *Working Papers of Eesti Pank*(6): 1-35.
- Lampenius, N., & Zickar, M. J. 2005. Development and Validation of a Model and Measure of Financial Risk-Taking. *Journal of Behavioral Finance*, 6(3): 129.
- Lanis, R. r. l. u. e. a., & Richardson, G. g. r. a. e. a. 2015. Is Corporate Social Responsibility Performance Associated with Tax Avoidance? *Journal of Business Ethics*, 127(2): 439-457.
- Lawes, L. E. 1927, *Congressional Digest*, Vol. 6: 231: Congressional Digest.
- Lee, K. 2001. Tax Evasion and Self-Insurance. *Journal of Public Economics*, 81(1): 73-81.

- Lijiao, C., Wenli, L., Qingguo, Z., & Smyth, R. 2014. Understanding personal use of the Internet at work: An integrated model of neutralization techniques and general deterrence theory. *Computers in Human Behavior*, 38: 220-228.
- Lunney, G. H. 1970. Using analysis of variance with a dichotomous dependent variable: an empirical study. *Journal of Educational Measurement*, 7: 263-269.
- MacRae, C. D., & MacRae, E. C. 1976. Labor Supply and the Payroll Tax. *American Economic Review*, 66(3): 408.
- Månsson, J., & Quoreshi, A. 2015. Evaluating regional cuts in the payroll tax from a firm perspective. *Annals of Regional Science*, 54(2): 323-347.
- Manzon, J. G. B., Sharp, D. J., & Travlos, N. G. 1994. An Empirical Study of the Consequences of U.S. Tax Rules for International Acquisitions by U.S. Firms. *Journal of Finance*, 49(5): 1893.
- Marcuss, R., Contos, G., Guyton, J., Langetieg, P., Lerman, A., Nelson, S., Schafer, B., & Vigil, M. 2013. INCOME TAXES AND COMPLIANCE COSTS: HOW ARE THEY RELATED? *National Tax Journal*, 66(4): 833.
- Markie, K. S., & Shackelford, D. A. 2012. CROSS-COUNTRY COMPARISONS OF CORPORATE INCOME TAXES. *National Tax Journal*, 65(3): 493.
- Martinez-Vazquez, J., & Rider, M. 2005. Multiple Modes of Tax Evasion: Theory and Evidence. *National Tax Journal*, 58(1): 51-76.
- Mauldin, S., & Wilder, W. M. 1997. Personal liability for employee withholding taxes. *CPA Journal*, 67(2): 30.
- Mazar, N., & Ariely, D. 2006. Dishonesty in Everyday Life and Its Policy Implications. *Journal of Public Policy & Marketing*, 25(1): 117-126.

- Merriam-Webster.com. 2014. m-w.com. In I. Merriam-Webster (Ed.), Vol. 2014.
- Mezulis, A. H., Abramson, L. Y., Hyde, J. S., & Hankin, B. L. 2004. Is There a Universal Positivity Bias in Attributions? A Meta-Analytic Review of Individual, Developmental, and Cultural Differences in the Self-Serving Attributional Bias. *Psychological Bulletin*, 130(5): 711-747.
- Mitchell, D. J. 2005. Beware the Value-Added Tax, Vol. 2015: The Heritage Foundation.
- Mitrusi, A., & Poterba, J. 2001. THE CHANGING IMPORTANCE OF INCOME AND PAYROLL TAXES ON U.S. FAMILIES. *NBER/Tax Policy & the Economy (MIT Press)*, 15(1): 95-119.
- Moffitt, R. A. 1977. Labor Supply and the Payroll Tax: Note. *American Economic Review*, 67(5): 1004-1005.
- Morris, D. 2010. Tax Penalties and Deterrence. *CPA Journal*, 80(9): 28-33.
- Mosberger, Á. K.-P. 2011. The elasticity of taxable income of high earners: Evidence from Hungary. *MNB Working Papers*, 2011(11): 1-34.
- Murthy, V. 2014. The Consequences of Willful Failure to Pay Payroll Taxes. *Journal of Accountancy*, 217(6): 54-58.
- Myung-Soo, J., Nelson, J. E., & Kiecker, P. 1997. A Model for Controlling Social Desirability Bias by Direct and Indirect Questioning. *Marketing Letters*, 8(4): 429-437.
- Nagin, D. S., & Pogarsky, G. 2001. Integrating celerity, impulsivity, and extralegal sanction threats into a model of general deterrence: Theory and evidence. *Criminology*, 39(4): 865-892.

- Neumark, D., & Wascher, W. 2011. DOES A HIGHER MINIMUM WAGE ENHANCE THE EFFECTIVENESS OF THE EARNED INCOME TAX CREDIT? *Industrial & Labor Relations Review*, 64(4): 712-746.
- Nor, J. M., Ahmad, N., & Saleh, N. M. 2010. Fraudulent financial reporting and company characteristics: tax audit evidence. *Journal of Financial Reporting and Accounting*, 8(2): 128-142.
- Nunnally, J. C., Bernstein, I. H. 1994. *Psychometric theory*: New York : McGraw-Hill, c1994.
3rd ed.
- Ong-Dean, C., Huie Hofstetter, C., & Strick, B. R. 2011. Challenges and Dilemmas in Implementing Random Assignment in Educational Research. *American Journal of Evaluation*, 32(1): 29-49.
- Ordover, H. 2014. SCHEDULARITY IN U.S. INCOME TAXATION AND ITS EFFECT ON TAX DISTRIBUTION. *Northwestern University Law Review*, 108(3): 905.
- Oxford_Dictionaries. 2015. Severity. In O. Dictionaries (Ed.), Vol. 2015. UK: Oxford University Press.
- Pennington, G. L., & Roese, N. J. 2003. Regulatory focus and temporal distance. *Journal of Experimental Social Psychology*, 39(6): 563.
- Peter, K. S. k. u. e., Buttrick, S. g. g. c., & Duncan, D. d. i. e. 2010. GLOBAL REFORM OF PERSONAL INCOME TAXATION, 1981-2005: EVIDENCE FROM 189 COUNTRIES. *National Tax Journal*, 63(3): 447-478.

- Phillips, M. D. m. p. u. e. 2014. INDIVIDUAL INCOME TAX COMPLIANCE AND INFORMATION REPORTING: WHAT DO THE U.S. DATA SHOW? *National Tax Journal*, 67(3): 531-567.
- Piketty, T., & Saez, E. 2003. INCOME INEQUALITY IN THE UNITED STATES, 1913-1998. *Quarterly Journal of Economics*, 118(1): 1.
- Pissarides, C. A., & Weber, G. 1989. An expenditure-based estimate of Britain's black economy. *Journal of Public Economics*, 39: 17-32.
- Planita, C. M. 2015. Understanding the Benefits of Individual Net Operating Losses in Our Current Business Environment. *Journal of Financial Service Professionals*, 69(1): 12.
- Plumley, A. 2006. **PRELIMINARY UPDATE OF THE TAX YEAR 2001 INDIVIDUAL INCOME TAX UNDERREPORTING GAP ESTIMATES.**
- Pressman, S. 2014. A Tax Reform That Falls Flat. *Challenge (05775132)*, 57(4): 82.
- Quackenbush, S. L. 2010. General Deterrence and International Conflict: Testing Perfect Deterrence Theory. *International Interactions*, 36(1): 60-85.
- Quackenbush, S. L. 2011a. *Understanding general deterrence : theory and application* (1st ed. ed.). New York: Palgrave Macmillan.
- Quackenbush, S. L. 2011b. *Understanding general deterrence [electronic resource] : theory and application / Stephen L. Quackenbush*: New York : Palgrave Macmillan, 2011.
- 1st ed.

- Rasholnikov, A. 2006. CRIME AND PUNISHMENT IN TAXATION: DECEIT, DETERRENCE, AND THE SELF-ADJUSTING PENALTY. *Columbia Law Review*, 106(3): 569.
- Rhee, R. J. 2012. A Financial Economic Theory of Punitive Damages. *Michigan Law Review*, 111(1): 33-87.
- Robben, H. S. J., Webley, P., Weigel, R. H., Warneryd, K.-E., Kinsey, K. A., Hessing, D. J., Martin, F. A., Elffers, H., Wahlund, R., Langenhove, L. V., Long, S. B., & Scholz, J. T. 1990. DECISION FRAME AND OPPORTUNITY AS DETERMINANTS OF TAX CHEATING: An international experimental study. *Journal of Economic Psychology*, 11(3): 341.
- Roth, J. A., Scholz, J. T., & Witte, A. D. 1989. *Taxpayer compliance / Jeffrey A. Roth, John T. Scholz, and Ann Dryden Witte, editors*: Philadelphia : University of Pennsylvania Press, c1989-.
- Ruscio, J. 2008. A probability-based measure of effect size: Robustness to base rates and other factors. *Psychological Methods*, 13(1): 19-30.
- Saez, E., Matsaganis, M., & Tsakloglou, P. 2012a. Earnings Determination and Taxes: Evidence From a Cohort-Based Payroll Tax Reform in Greece*. *Quarterly Journal of Economics*, 127(1): 493.
- Saez, E., Slemrod, J., & Giertz, S. H. 2012b. The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review. *Journal of Economic Literature*, 50(1): 3-50.

- Salihu, I. A., Sheikh Obid, S. N., & Annuar, H. A. 2013. MEASURES OF CORPORATE TAX AVOIDANCE: EMPIRICAL EVIDENCE FROM AN EMERGING ECONOMY. *International Journal of Business & Society*, 14(3): 412-427.
- Schelling, T. C. 1958. The strategy of conflict Prospectus for a reorientation of game theory. *Journal of Conflict Resolution*, 2(3): 203.
- Scholz, J. K. 1994. THE EARNED INCOME TAX CREDIT: PARTICIPATION, COMPLIANCE, AND ANTIPOVERTY EFFECTIVENESS. *National Tax Journal*, 47(1): 63-87.
- Scholz, J. T. 2003. Contractual Compliance and The Federal Income Tax System. *Washington University Journal of Law & Policy*, 13: 139.
- Schulz, S. 2014. Individual Differences in the Deterrence Process: Which Individuals Learn (Most) from Their Offending Experiences? *Journal of Quantitative Criminology*, 30(2): 215-236.
- SCOTUS. 1935. Gregory v. Helvering 293 U.S. 465. In U. S. Reports (Ed.), *No. 127*. Washington DC: SUPREME COURT OF THE UNITED STATES.
- Shadish, W. R., Clark, M. H., & Steiner, P. M. 2008. Can Nonrandomized Experiments Yield Accurate Answers? A Randomized Experiment Comparing Random and Nonrandom Assignments. *Journal of the American Statistical Association*, 103(484): 1334-1343.
- Slemrod, J. 1985. AN EMPIRICAL TEST FOR TAX EVASION. *Review of Economics & Statistics*, 67(2): 232.
- Spicer, M. W., & Becker, L. A. 1980. FISCAL INEQUITY AND TAX EVASION: AN EXPERIMENTAL APPROACH. *National Tax Journal*, 33(2): 171-175.

- St. Louis Federal Reserve Bank, F. 2014. Federal Surplus or Deficit [-] as Percent of Gross Domestic Product. In F. R. Bank (Ed.), *Debt to Gross Domestic Product Ratios* Vol. 2014: Economic Data. St. Louis, MO Federal Reserve Bank.
- Stewart, D. W. 1981. The Application and Misapplication of Factor Analysis in Marketing Research. *Journal of Marketing Research (JMR)*, 18(1): 51-62.
- Stiglitz, J. E. 1969. THE EFFECTS OF INCOME, WEALTH, AND CAPITAL GAINS TAXATION ON RISK-TAKING. *Quarterly Journal of Economics*, 83(2): 263.
- Straub, D. W., & Welke, R. J. 1998. Coping With Systems Risk: Security Planning Models for Management Decision Making. *MIS Quarterly*, 22(4): 441-469.
- Strully, K. W., Rehkopf, D. H., & Ziming, X. 2010. Effects of Prenatal Poverty on Infant Health: State Earned Income Tax Credits and Birth Weight. *American Sociological Review*, 75(4): 534.
- Sugin, L. 2014. CLASS IN AMERICA: PAYROLL TAXES, MYTHOLOGY, AND FAIRNESS. *Harvard Journal on Legislation*, 51: 113.
- TIGTA. 2011. LIMITATIONS IN THE SAMPLE SIZE FOR THE INTERNAL REVENUE SERVICE'S EMPLOYMENT TAX STUDY MAY IMPACT THE ABILITY TO DETERMINE COMPLIANCE LEVELS. In Treasury (Ed.): 1. Washington DC: The Treasury Department.
- TIGTA. 2014a. IMPROVEMENTS ARE NEEDED IN DOCUMENTATION AND DATA ACCURACY FOR THE EMPLOYMENT TAX STUDY. In Treasury (Ed.): 1. Washington DC: The Treasury Department.

- TIGTA. 2014b. A SERVICE-WIDE STRATEGY IS NEEDED TO INCREASE BUSINESS TAX RETURN ELECTRONIC FILING. In Treasury (Ed.): 1. Washington DC: The Treasury Department.
- Tittle, C., Botchkovar, E., & Antonaccio, O. 2011. Criminal Contemplation, National Context, and Deterrence. *Journal of Quantitative Criminology*, 27(2): 225-249.
- Trochim, W., & Donnelly, J. 2007a. *The Research Methods Knowledge Base* (3rd ed.). Cincinnati, OH: Atomic Dog Publishing (atomicdogpublishing.com).
- Trochim, W. M. K., & Donnelly, J. P. 2007b. Design. In M. Baird (Ed.), *The Research Methods Knowledge Base*, 3rd ed.: 157-175. USA: CENGAGE Learning.
- Trochim, W. M. K., & Donnelly, J. P. 2007c. Experimental Design. In M. Baird (Ed.), *The Research Methods Knowledge Base*, 3rd ed.: 185-207. USA: CENGAGE Learning.
- Turk, A., Muzikir, M., Blumenthal, M., & Kalambokidis, L. 2007. Charitable Contributions in a Voluntary Compliance Income Tax System: Itemized Deductions versus Matching Subsidies. In J. Dalton, & M. Gangi (Eds.), *The 2007 IRS Research Conference*, : 51–74. Washington, DC.: IRS.
- Tversky, A., & Kahneman, D. 1974. Judgment under Uncertainty: Heuristics and Biases: 1124: American Association for the Advancement of Science.
- Upadhyay, D. 2013. Book Review: The (Honest) Truth about Dishonesty: How We Lie to Everyone—Especially Ourselves. *Vision (09722629)*, 17(2): 194-195.
- Vitell, S. J. 2003. Consumer Ethics Research: Review, Synthesis and Suggestions for the Future. *Journal of Business Ethics*, 43(1/2): 33-47.

- Von Neumann, J., & Morgenstern, O. 1974. *Theory of games and economic behavior*: Princeton, Princeton University Press, 1953 [reprinted 1974, c1944] [3d ed.].
- Vroman, W. 1974. Employer payroll taxes and money wage behaviour. *Applied Economics*, 6(3): 189.
- Wade, M. R., & Tingling, P. 2005. A new twist on an old method. *Database for Advances in Information Systems*, 36(3): 69.
- Warren, N. 2014. A politically viable strategy for limiting personal income tax deductions: The case for a global cap. *Australian Tax Forum*, 29(3): 357-389.
- White, A. D., & Woodbury, D. F. 1985. THE EFFECT OF TAX LAWS AND TAX ADMINISTRATION ON TAX COMPLIANCE: THE CASE OF THE U.S. INDIVIDUAL INCOME TAX. *National Tax Journal*, 38(1): 1-13.
- Willison, R., & Warkentin, M. 2013. BEYOND DETERRENCE: AN EXPANDED VIEW OF EMPLOYEE COMPUTER ABUSE. *MIS Quarterly*, 37(1): 1-20.
- Winchester, R. 2009. THE GAP IN THE EMPLOYMENT TAX GAP. *Stanford Law & Policy Review*, 20(1): 127-151.
- Yitzhaki, S. 1974. INCOME TAX EVASION: A THEORETICAL ANALYSIS. *Journal of Public Economics*, 3(2): 201.
- Zagare, F. C. 2007. Toward a Unified Theory of Interstate Conflict. *International Interactions*, 33(3): 305-327.

VITA

Rafael Dacal currently works as a Senior Strategist for Analyst for the Internal Revenue Service (IRS). He works at the Small Business/Self-Employee (SB/SE) division. The SB/SE division services approximately 57 million taxpayers each year including proprietorships, partnerships, start-ups, and S Corporations. The teams Rafael leads focus on developing and maintaining the organization's strategic portfolio, which is critical for identifying and prioritizing investment grades activities. He has developed and implemented the performance measure process for SB/SE necessary to determine if strategic activities are accomplishing the organizational goals and objective.

Prior to working at the Strategy Office, Rafael worked as a Senior Operations Research Analyst for the Internal Revenue Service. The teams Rafael led conduct research using different techniques to identify trends, taxpayer behavior, and compliance issues. Some of the techniques he has employed in his research work include trend analysis, market research, impact measurement, surveys, profiling, and optimization. Some of the key projects Rafael worked on included the Examination workload selection, which resulted in the IRS ability to improve fraud detection. He also worked on the Filing Extension Simplification research project, which resulted in streamlining the request to file an extension. The outcome of this project affects all taxpayers who request an extension to file. In addition to his research work, Rafael is part of the Quality Assurance Team. This team ensures that all IRS research methodologies are robust and that the quality meets standards prior to releasing results or reports to clients. He has also preformed ad hoc research for Deputy Commissioners and the Director of Research.

After graduating from high school, Rafael served three years in the US Army (2nd Armor and 1st Cavalry Divisions). Once he completed his military service, Rafael

attended Florida International University (FIU) located in Miami. FIU is the fourth largest research university in the United States with a Carnegie Foundation top-tier rating. He received a Bachelor's of Arts in Economics, a Master's of Arts in Economics, and a Master's of Science in Finance. Additionally, he has taught for over twelve years in the field of economics, business statistics, and finance.

Rafael was born in Cuba and lived five years in Venezuela before he moved to Miami, Florida. Rafael is married to María. They moved from Miami to the greater Atlanta area in 2011. Rafael and María have two children: Nicole and Joshua. The family enjoys traveling around the world, and tries to visit a different country and state every year.

While living in Florida, Rafael volunteered for many years at Miami Shake-a-Leg, a program established to assist disadvantaged and disabled children. Over the years, he has assisted the organization with fundraising and has taught disabled children how to sail.