NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

MBA PROFESSIONAL PROJECT

FRAUD CLASSIFICATION DETECTION IN DOD CONTRACTS USING A STATISTICAL ANALYSIS PERSPECTIVE

December 2018

By: **Joshua M. Kidwell**

Advisor: Rene G. Rendon
Co-Advisor: William A. Muir
Co-Advisor: Juanita M. Rendon

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704–0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.			
1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** December 2018	**3. REPORT TYPE AND DATES COVERED** MBA Professional Project	
4. TITLE AND SUBTITLE FRAUD CLASSIFICATION DETECTION IN DOD CONTRACTS USING A STATISTICAL ANALYSIS PERSPECTIVE			**5. FUNDING NUMBERS**
6. AUTHOR(S) Joshua M. Kidwell			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.			**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

The purpose of this research was to propose data analysis methods to discover statistical relationships of contract data contained in the Federal Procurement Data System (FPDS) and fraud indicators. The primary means by which this research was conducted was by analyzing a sample of contracts associated with fraud investigations in the FPDS and collecting any statistical data related to types of contracts, the fraud statute investigated, and other contract variables attributable to FPDS. The objective was to determine if there is a statistical relationship between contracts associated with fraud investigations and contract variables from a contracting officer, auditor, or investigating official's perspective. The research findings were based on examples and illustrations of statistical relationships discovered and analyses explored as well as proposed applications of use for the data in detecting fraud in Department of Defense (DoD) contracts.

14. SUBJECT TERMS procurement, contract, fraud, FPDS, DoD, data, analytics, analysis, detection, statistical, contracting, auditing, investigation, audit			**15. NUMBER OF PAGES** 77
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540–01-280-5500 Standard Form 298 (Rev. 2–89) Prescribed by ANSI Std. 239–18

THIS PAGE INTENTIONALLY LEFT BLANK

FRAUD CLASSIFICATION DETECTION IN DOD CONTRACTS USING A STATISTICAL ANALYSIS PERSPECTIVE

Joshua M. Kidwell, Master Sergeant, United States Air Force

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF BUSINESS ADMINISTRATION

from the

NAVAL POSTGRADUATE SCHOOL
December 2018

Approved by: Rene G. Rendon
Advisor

William A. Muir
Co-Advisor

Juanita M. Rendon
Co-Advisor

Rene G. Rendon
Academic Associate
Graduate School of Business and Public Policy

THIS PAGE INTENTIONALLY LEFT BLANK

FRAUD CLASSIFICATION DETECTION IN DOD CONTRACTS USING A STATISTICAL ANALYSIS PERSPECTIVE

ABSTRACT

The purpose of this research was to propose data analysis methods to discover statistical relationships of contract data contained in the Federal Procurement Data System (FPDS) and fraud indicators. The primary means by which this research was conducted was by analyzing a sample of contracts associated with fraud investigations in the FPDS and collecting any statistical data related to types of contracts, the fraud statute investigated, and other contract variables attributable to FPDS. The objective was to determine if there is a statistical relationship between contracts associated with fraud investigations and contract variables from a contracting officer, auditor, or investigating official's perspective. The research findings were based on examples and illustrations of statistical relationships discovered and analyses explored as well as proposed applications of use for the data in detecting fraud in Department of Defense (DoD) contracts.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

ACFE	Association of Certified Fraud Examiners
AFOSI	Air Force Office of Special Investigations
CAAT	computer assisted audit technique
CPA	certified public accountant
DCAA	Defense Contract Audit Agency
DoD	Department of Defense
DoJ	Department of Justice
EDI	electronic data interchange
FY	fiscal year
FPDS	Federal Procurement Data System
GIGO	garbage in garbage out
IDIQ	indefinite delivery indefinite quantity
MCIO	military criminal investigative office
NDAA	National Defense Authorization Act
NCIS	Naval Criminal Investigative Service
OU	multi-subset observation undersampling
PII	personally identifiable information
PIIN	procurement instrument identification number
PSC	product service code
SQL	Structured Query Language
T&M	time and materials
USD(AT&L)	Under Secretary of Defense, Acquisition, Technology, and Logistics
VU	multi-subset variable undersampling

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I would like to thank Dr. Rene Rendon and Dr. Juanita Rendon. Their teachings throughout the courses at Naval Postgraduate School and guidance throughout this research study have given me the foundation necessary to present these findings and apply it to future practice. Furthermore, Dr. Juanita Rendon’s careful and meticulous review helped me make this report a much better product. I would also like to thank Maj William Muir, who not only provided me the idea for the basis behind this research but also contributed his much needed statistical analysis expertise. His guidance and enthusiasm encouraged me to make this research study a worthwhile learning experience. I also want to give a special thanks to the Air Force Office of Special Investigations Office of Procurement Fraud because it would not have been possible to conduct this research without their support. I want to thank my whole family for enduring the long hours I have put into this effort, especially my wife, Chibi, who I have a hard time keeping up with. Finally, I want to thank the Air Force leaders for providing the opportunity to attend the Naval Postgraduate School. It was a valuable and humbling experience shared with the few enlisted and many commissioned officers from multiple domestic and international military service departments at home and abroad.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND

Within the Department of Defense (DoD), procurement fraud has been an issue of concern that has captured the attention of acquisition professionals, contracting agencies, investigative agencies, auditing agencies, and even Congress. The means by which procurement fraud has been addressed in the last decade has, by and large, included initiatives such as increasing military investigative services efforts in procurement fraud within Military Criminal Investigative Organizations (MCIO). These MCIOs include agencies such as the Air Force Office of Special Investigations (AFOSI) and the Naval Criminal Investigative Service (NCIS). Other efforts have included increasing auditing personnel within DoD agencies such as the Defense Contract Audit Agency (DCAA). These efforts are based on personnel increases to strengthen the detection and prosecution of fraud. However, prevention, a most critical effort to fraud deterrence, is an area that is only just beginning to be emphasized. Juanita Rendon and Rene Rendon (2016) argued that prevention is critical to fighting procurement fraud. They observe that using first line of defense prevention methods such as training procurement professionals in contract management on internal controls to deter procurement fraud schemes will not only prevent potential losses, but will also reduce the strain on DoD auditing and investigative agency resources (Rendon & Rendon, 2016).

In 2011, Office of the Undersecretary of Defense for Acquisition, Technology, and Logistics (OUSD(AT&L)) produced a report to Congress on contracting fraud in the DoD. The DoD report to Congress included data on recoveries, indictments, and convictions resulting from procurement fraud cases and utilized data collected from the Department of Justice (DoJ), Defense Criminal Investigative Organizations (DCIO), and the Federal Procurement Data System (FPDS). Their findings, although informative, were very difficult for the researchers to compile as they described in their report. The data was difficult to compile because records had vastly different classification standards among the different agencies. For example, the Department of Justice (DoJ) and DCIO's records classified fraud cases with the lead charge and not the fraud scheme which made it very difficult to collect data specific to fraud incidents. In order to collect all of the actual fraud schemes that resulted in recoveries, indictments, or convictions in court, it would have involved analyzing thousands of case files and would have required an exorbitant amount of time and resources (Office of the Undersecretary of Defense for Acquisition and, Technology, and Logistics [USD(AT&L), 2011]). For this same reason, the data collected for this research follows a similar method of using the lead fraud statute investigated to categorize the contract data under a fraud classification in an efficient manner.

B. PURPOSE OF RESEARCH

The purpose of this research is to determine if there are any statistical relationships between contracts associated with fraud investigations and their associated contract variables. If there are statistical relationships discovered, the findings will be used to determine potential applications for measuring the risk of fraud in different types of contracts. The audience perspective will include contracting officers, auditors, investigative officials, and other officials within the DoD acquisition community. The goal of the research is for any contracting officer, auditor, or investigative official to be able to look at data within FPDS and apply fraud risk analysis to particular types of contracts. The findings may prove to be useful by increasing awareness of fraud risk and mitigating a higher risk of fraud in contracts by arming contracting professionals, auditors, and investigators with a data driven tool on their tool belt. Additionally, it could help to identify types of contracts with a higher risk for fraud.

C. RESEARCH QUESTIONS

The four research questions for this research study include the following:

1. What is the statistical relationship between contracts associated with a class of fraud investigation and contract type? (e.g., fixed price, cost plus incentive, etc.)

2. What is the statistical relationship between contracts associated with a class of fraud investigation and competition type involved? (e.g., full and open competition, sole source, etc.)

3. What is the statistical relationship between contracts associated with a class of fraud investigation and business size status?

4. What is the statistical relationship between contracts associated with a class of fraud investigation and the contracts' associated product service codes?

D. METHODOLOGY

This research study involved a literature review, data collection, and statistical data analysis. The literature review includes sources which emphasized data analysis approaches toward detecting and preventing fraud. The primary means of how this research was conducted was by analyzing samples of contracts investigated for fraud or associated with fraud investigations in the FPDS and by collecting any statistical data related to types of contracts, the fraud statute investigated, and other data attributable to FPDS. The data collected included Procurement Instrument Identification Numbers (PIIN) of DoD acquisition and procurement contracts associated with fraud cases and/or investigated for fraud as well as the associated fraud statute. This dataset was provided by AFOSI procurement fraud analysts for the purpose of this research with express permission from AFOSI Procurement Fraud Investigations Program Management. The data was extracted from final closed fraud cases from FY12–FY17. While some contracts included in this dataset were from FY12–FY17, much of the data from FPDS was derived from contracts that were from years prior to FY12 because many frauds are not reported resulting in an investigation until a few years after the contract is awarded.

Furthermore, the federal statute of limitations for fraud is five years for criminal charges, and investigations may be opened a couple of years into a contract's life cycle and could take two or more years to close the case. Due to this nature of the dataset, award dates for specific contracts were not used. The primary data source utilized was data collected from FPDS, a public database located online at www.fpds.gov. This activity was accomplished by taking the PIINs associated with contracts involving fraud investigations and cross-referencing those contracts within the FPDS. The contract variables contained in the FPDS were collected for statistical analysis to find any significant consistencies or patterns to determine if there are any fraud risk models that could prove useful in data analysis applications. The contract variables used in the analysis include contract type (fixed price, cost reimbursable, or incentivized), type of competition (full and open or sole source), whether the contract was a small business set aside or not, contractor business size status, and associated product service codes.

Data was analyzed using discriminant analysis within Stata, a statistical analysis software program (StataCorp, 2017). The way the data findings are presented in this report is in a manner in which it is not possible to identify any specific contractors or individuals associated with the contract data analyzed. Contract PIINs were used for cross-reference in FPDS and subsequent data analysis, but were not cited in this report. No personally identifiable information (PII) was utilized. No data collection involving human interaction, interviews, or surveys were utilized for this research.

For the purposes of this research, each contract was treated as an occurrence or reporting of a fraud allegation and not proven fraud. It is important for the reader of this report to note that although the data used for this research project were from contracts associated with fraud investigations, this does not necessarily mean these instances resulted in convictions for fraud in a court of law. It is also important to note that even when there is no conviction in a court of law for an allegation of fraud, it does not mean that the risk for fraud was not there or that fraud did not occur. Many criminal investigations into fraud or other crimes, especially those lacking sufficient evidence, do not result in convictions, and it is no different with the dataset collected for this research. The dataset collected did not include data related to criminal prosecution or differentiate between contracts that had fraud resulting in criminal conviction and those that had alleged fraud that did not result in a criminal conviction. That data was not used and was not made available. Each individual instance in the quantified data presented in this report should be thought of as a fraud allegation or reported fraud for the purposes of identifying fraud risk by analyzing contract variables for potential prevention, detection, and prediction of procurement fraud.

E. IMPORTANCE OF RESEARCH

According to the Association of Certified Fraud Examiners (ACFE) (2018), the total loss due to occupational fraud occurring around the world from January 2016 to October 2017 amounted to $7.1 billion. Fifty-five percent of the cases included in ACFE's report amounted to financial losses of less than $200,000, and more than one fifth amounted to financial losses of at least $1 million (Association of Certified Fraud Examiners [ACFE], 2018). This shows how prevalent the fraud threat is today around the world, and procurement fraud in DoD contracts should not be excluded. With the threat of fraud being just as big today as it has ever been, a growing factor in fighting fraud is the focus on fraud detection. ACFE (2018) argued that the longer that fraud goes undetected, the larger that fraud schemes tend to grow. Fraud schemes that go undetected for over 60 months are more than 20 times as costly as those that are caught in the first six months (ACFE, 2018). Furthermore, fraudsters who are not caught within the first three years of initiating their fraudulent behaviors tend to rapidly increase their frauds (ACFE, 2018). This is why it is important for organizations in government and private sectors to implement proactive fraud detection methods to catch fraudulent behavior early and prevent or minimize the damages (ACFE, 2018). One such method of proactive fraud detection identified in the ACFE's report was proactive data monitoring and analysis which was associated with a more than a 50% reduction in fraud losses. This is significant even though this method of analysis only accounts for 37% of the most common anti-fraud controls (ACFE, 2018).

One of the initiatives in the latest National Defense Authorization Act (NDAA) for Fiscal Year (FY) 2018 included emphasizing data analytics within DoD acquisition programs to assist in streamlining processes and improving acquisition program outcomes (National Defense Authorization Act, 2017). In line with the NDAA's focus in applying data analytics toward improving acquisition program outcomes, it is argued that a data analysis method to detect the risk of fraud in specific types of contracts associated with DoD acquisition programs using existing data systems may prove to be a valuable tool for contracting professionals, auditors, and investigators within the DoD. Within contracting, internal control method effectiveness in preventing fraud may be amplified by creating data analysis methods to detect the risk of fraud and focus those internal control efforts through more efficient and empirically based means. To address the need for these data analysis methods, this study involves research and data analysis to discover statistical relationships between contracts associated with fraud investigations and their associated contract variables contained in the FPDS.

F. ORGANIZATION OF REPORT

This report consists of four chapters. Chapter II is the literature review, which covers scholarly and peer-reviewed articles and sources from the government on data driven analytics and its importance to improving processes such as fraud prevention, detection, and investigations. Chapter III covers the methodology used in this research and will define the nature of the FPDS data relevant to this analysis. Chapter IV consists of the analysis of the FPDS data collected and will introduce and present the strength of the statistical relationships between the contracts' alleged fraud statutes involved and their associated contract variables. Chapter V includes a summary, conclusion, and areas for further research.

G. SUMMARY

This chapter introduced the background, research questions, purpose of the research, and a brief discussion of the methodology. Also provided was the importance of the research and organization of the report. The following chapter sets the foundation for this research by providing a literature review.

II. LITERATURE REVIEW

A. INTRODUCTION

Before addressing the primary purpose of this research within this literature review, it is important to provide an overview of procurement fraud resolution efforts and the definition of fraud. A basic definition of fraud provided by Henry Campbell Black in *Black's Law Dictionary* is "an intentional perversion of truth for the purpose of inducing another, in reliance upon it, to part with some valuable thing or to surrender legal right" (as cited in Caulfield, 2014, p. 53). Basically, fraud is an intentional act to deceive in order to gain something of value that does not legally belong to the fraudster. Fraud occurs in many arenas, big and small. There are corporate frauds, insurance frauds, healthcare frauds, and credit card frauds, to name a few. One of the most damaging types of fraud is procurement fraud. Within the DoD, procurement fraud occurs in the field of acquisition and defense contracts. This type of fraud can cause as little as a few thousand dollars of damages to up to and beyond several million dollars and could even cost lives. Several common fraud schemes committed in procurement fraud include cost mischarging, bribery, false claims, bid rigging, kickbacks, and product substitution.

There are many elements to fighting procurement fraud including reactionary measures such as criminal investigations, prosecution, civil remedies, and recoveries, and administrative remedies such as contractor debarment and suspension. In many cases, fraud is not discovered to have been in its planning or development stages until it has already occurred. As a result, there may be victims of that fraud, whether it be victims such as taxpayers or U.S. Army soldiers hurt by a fraudulent and defective product. This is why fighting fraud is, the majority of the time, a reactionary effort. The problem with the focus being on reactionary measures is that it overshadows where the majority of efforts should be placed, which is in prevention.

One of the most critical elements of combating procurement fraud, within as well as outside the DoD, is prevention. There are many preventative efforts that are applicable to deterring fraud including proper monitoring, internal controls, fraud awareness training, and effective internal policy. While these preventative measures are very common and have been around a long time, there is a type of preventative measure that is still fairly new in this digital age and enables a variety of approaches to tackle the issue of fraud prevention. The argument presented here is that data analytics could serve as an effective fraud prevention measure in the DoD acquisition and procurement environment. This chapter will cover the elements of procurement fraud and applicable criminal statutes. Moreover, this chapter will cover literature addressing the potential use for data analytics in fraud prevention, detection, and to a degree, possible prediction.

B. PROCUREMENT FRAUD

The following section covers literature on the fraud and auditability theories. This section also covers common procurement fraud schemes as well as fraud prevention and detection factors. Finally, this section will also cover literature identifying fraud exposure areas and preventative techniques.

1. Fraud Triangle

This research analyzes characteristics of DoD contracts associated with investigations in procurement fraud. Before going into research related to procurement fraud in great detail, it is important to understand how the phenomenon of fraud occurs. One of the most common analyses of the origin of fraud is the concept of the fraud triangle. Dr. Donald Cressey conducted research on circumstances that attracted fraudsters to commit ethical violations and commit their first fraudulent act. Dr. Cressey's research findings later became known as the fraud triangle (Dorminey, Fleming, Kranacher, Riley, & Richard, 2010).

As shown in Figure 1, the fraud triangle is made up of three elements, to include pressure, opportunity, and rationalization. The first element, pressure, is often associated with financial hardship including living beyond one's means, greed, poor credit, unexpected medical expenses, gambling losses and addiction, or alcohol and substance abuse (Dorminey et al., 2010). The second element, opportunity, is present with poor internal controls, poor supervision, weak culture of ethics, poor training, lack of prosecution of perpetrators, and ineffective fraud programs (Dorminey et al., 2010). The third element, rationalization, is the psychological reasoning that makes fraudsters believe that what they are doing is just and for their own good. Dr. Cressey contended that moral rationalization must be present in order for the crime to take place (Dorminey et al., 2010). The criminal does not see himself as a criminal so he must first be able to rationalize his planned action before he commits his crime. For example, a fraudster may justify his actions by telling himself that the fruits of his actions are just a loan to be paid back when he receives his next paycheck. Of course the payback does not typically come to fruition when that act was fraudulent in the first place (Dorminey et al., 2010).

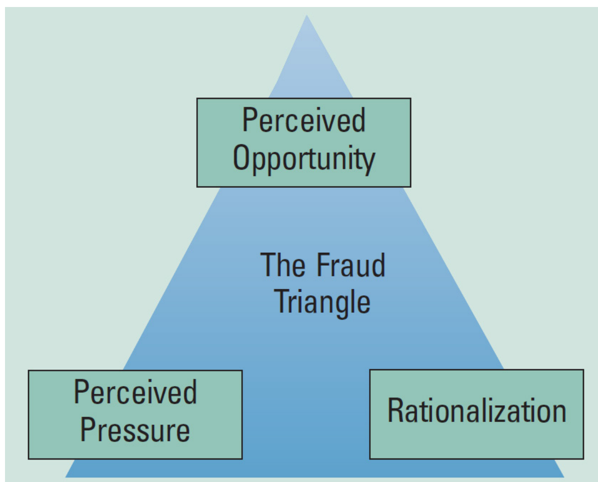

Figure 1. Fraud Triangle. Source: Dorminey et al. (2010).

2. Procurement Fraud Schemes

There are various procurement fraud schemes that are committed under the conditions of the fraud triangle previously mentioned. A key factor for procurement fraud prevention highlighted by Rendon and Rendon (2015) is procurement contracting officers understanding the *where*, *why*, and *how* procurement fraud occurs. First, the *where* represents the contract management process area or phase in the contracting process. Secondly, the *why* represents the internal control area that prevents a successful fraud scheme. Thirdly, the *how* is the procurement fraud scheme itself (Rendon & Rendon, 2015). The contracting officer, auditor, and investigator should understand the relationships between the procurement fraud scheme, the associated contract management process phase in which those schemes could occur, and the associated internal controls that need to be present and effective to prevent the fraud scheme from happening (Rendon & Rendon, 2015). The "Procurement Fraud Matrix" provides a visual representation showing the focus areas for training to become effective at procurement fraud prevention (see Figure 2).

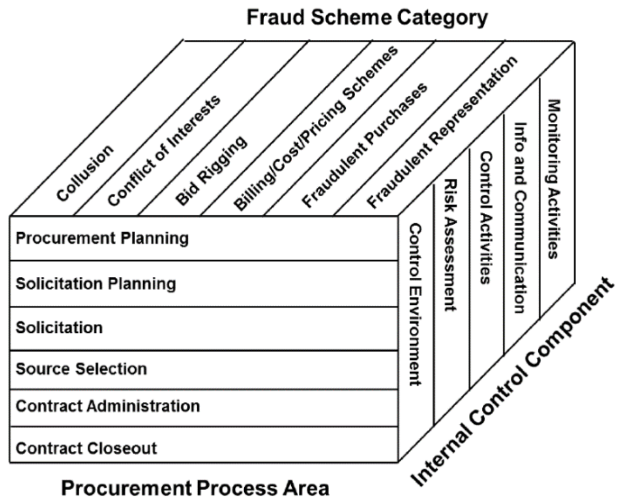

Figure 2. Procurement Fraud Matrix. Source: Rendon and Rendon (2015).

In relation to the procurement fraud matrix, Juanita Rendon & Rene Rendon's (2016) research into the relationships between procurement fraud schemes, contract management process phases, and internal controls uncovered notable findings. Juanita Rendon and Rene Rendon (2016) analyzed incidents related to procurement fraud reported by the DoD Office of General Counsel. Findings from their research identified 50% of the fraud occurrences as contractor and government collusion schemes. Bribery and billing, cost, and pricing schemes each accounted for 25% and 35% of the fraud occurrences, respectively (Rendon & Rendon, 2016). Juanita Rendon and Rene Rendon's (2016) research also highlighted internal control component areas identified as weak and contributing to occurrences of fraud to include risk assessment, control environment, monitoring, and control activities. A critical finding in Juanita Rendon and Rene Rendon's (2016) research showed that control activity material weaknesses accounted for 100% of the fraud occurrences. The monitoring activities and control environment components accounted for weaknesses in 95% and 80% of the fraud incidents, respectively (Rendon & Rendon, 2016). Fraud occurrences were identified throughout the contract management process phases, including procurement planning, solicitation planning, solicitation, source selection, and contract administration (Rendon & Rendon, 2016). Of all of the fraud occurrences, 80% occurred during the source selection phase, and 40% occurred during the contract administration phase (Rendon & Rendon, 2016).

The implications of Juanita Rendon and Rene Rendon's (2016) research illustrated that it is critical for contracting officers, auditors, and investigators to understand these procurement fraud schemes. These schemes identified in Rene Rendon & Juanita Rendon's (2015) Procurement Fraud Matrix included collusion, conflicts of interest, bid rigging, billing, cost, and pricing schemes, fraudulent purchases, and fraudulent representation. These fraud schemes can cause serious damage to the integrity, accountability, and transparency of DoD's acquisition, procurement, and contracting functions (Rendon & Rendon, 2015). Each of these fraud schemes can be criminally charged under applicable fraud statutes listed in the U.S. Code (U.S.C.).

3. Fraud Statutes

These procurement fraud schemes are criminally chargeable under applicable criminal fraud statutes contained in the U.S.C. The following are brief descriptions and excerpts of these statutes from the U.S.C. The following excerpts are titled with the common fraud statute name as used in this study's dataset in italics followed by the specific U.S.C. statute title that it falls under.

a. Procurement Integrity Act

Title 41 U.S.C. §2102. Prohibitions on Disclosing and Obtaining Procurement Information

A person "shall not knowingly disclose contractor proposal, bid, or source selection information before the award of a Federal agency procurement contract" (Prohibitions on Disclosing and Obtaining Procurement Information, 2011).

b. Acts Affecting a Personal Financial Interest

Title 18 U.S.C. §208. Acts Affecting a Personal Financial Interest

Acts affecting a personal financial interest are officers and employees of certain U.S. Government offices who commit an act in that capacity to pursue a financial interest including prospective employment (Acts Affecting a Personal Financial Interest, 1962).

c. Anti-Kickback Act of 1986

Title 41 U.S.C. §8701. Public Contracts–Kickbacks

Kickbacks are "any money, fee, commission, credit, gift, gratuity, thing of value or compensation of any kind" that is given to a contractor or contractor employee to wrongfully gain or reward favorable treatment related to a prime contract (Public Contracts–Kickbacks, 2011).

d. Bribery

Title 18 U.S.C. §201 Bribery of Public Officials and Witnesses

Bribery is "when a person directly or indirectly, corruptly gives, offers or promises anything of value to any public official or person who has been selected to be a public official, or offers or promises any public official…with intent to influence an official act… influence that official to engage in fraud…or act unlawfully" (Bribery of Public Officials and Witnesses, 1994).

e. Civil False Claims Act

Title 31 U.S.C. §3729 False Claims

Civil false claims primarily include acts committed when a person presents or makes a fraudulent claim for approval to be paid, under false pretenses (False Claims, 2009).

f. Conspiracy to Commit Offense or to Defraud the Government

Title 18 U.S.C. §371 Conspiracy to Commit Offense or to Defraud United States

"If two or more persons conspire either to commit any offense against the United States, or to defraud the United States, or any agency thereof in any manner or for any purpose, and one or more of such persons do any act to effect the object of the conspiracy, each shall be fined under this title or imprisoned not more than five years, or both" (Conspiracy to Commit Offense or to Defraud United States, 1994).

g. Conspiracy to Defraud with Respect to Claims

Title 18 U.S.C. §286 Conspiracy to Defraud the Government with Respect to Claims

Conspiracy to defraud with respect to claims occurs when "whoever enters into an agreement or conspiracy to defraud the United States" or its agencies "by obtaining or aiding to obtain the payment of any false, fictitious, or fraudulent claim" (Conspiracy to Defraud the Government with Respect to Claims, 1994).

h. Counterfeiting

Title 18 U.S.C. §470–514 Counterfeiting and Forgery

Counterfeiting is when someone produces, markets, obtains, exchanges, or provides any forged, faked, or altered obligation or other U.S. security with the intent of passing that instrument as true and genuine (Dealing in Counterfeit Obligation or Securities, 2001).

i. False Official Statements

Title 10 U.S.C. §907 Art. 107 False Official Statements

False official statements are made when "any person subject to the U.S. military's Uniform Code of Military Justice who, with intent to deceive, signs any false record, return, regulation, order, or other official document, knowing it to be false, or makes any other false official statement knowing it to be false" (False Official Statements, 2016).

j. False Statements

Title 18 U.S.C. §1001. Statements or Entries Generally

False statements are committed by whoever, "within the jurisdiction of the executive, legislative, or judicial branch of the Government of the United States, knowingly and willfully falsifies, conceals, or covers up by any trick, scheme, or device a material fact; makes any materially false, fictitious, or fraudulent statement or representation; or makes or uses any false writing or document knowing the same to contain any materially false, fictitious, or fraudulent statement or entry" (Statements or Entries Generally, 2006).

k. False, Fictitious, or Fraudulent Claims

Title 18 U.S.C. §287 False, Fictitious or Fraudulent Claims

Fraudulent claims are submitted when a person knowingly makes or presents to anyone in the U.S. military, naval, civil service, or other U.S. agencies, any fraudulent claim against the U.S. (False, Fictitious, or Fraudulent Claims, 1986).

l. Forgery: Making and Altering

Title 18 U.S.C. §470–514 Counterfeiting and Forgery

Forgery occurs when "whoever falsely makes, forges, alters, or counterfeits any bid, bond, contract, proposal, security, public record, or other writing for the purpose of defrauding the United States" (Contractors' Bonds, Bids, and Public Records, 1994).

m. Fraud Offenses (False Pretenses/Swindle/Confidence Game) and Mail Fraud

Title 18 U.S.C. §1341 Frauds and Swindles

Mail fraud occurs when a person has planned a scheme to commit fraud and utilizes any post office in furtherance of that fraudulent scheme (Frauds and Swindles, 2008).

n. Public Money, Property, or Records, Larceny and Theft

Title 18 U.S.C. §641 Public Money, Property, or Records

Public money, property, or records, larceny and theft is when someone “embezzles, steals, purloins, or knowingly converts to his use or the use of another, or without authority, sells, conveys or disposes of any record, voucher, money, or thing of value of the United States or of any department or agency thereof” (Public Money, Property, or Records, 2004).

o. Major Fraud against the United States

Title 18 U.S.C. §1031 Major Fraud Against the United States

Major fraud against the United States is committed when someone “knowingly executes, or attempts to execute, any scheme or artifice with the intent to defraud the United States; or to obtain money or property by means of false or fraudulent pretenses…if the value of such grant, contract, subcontract, subsidy, loan, guarantee, insurance, or other form of Federal assistance, or any constituent part thereof, is $1,000,000 or more” (Major Fraud Against the United States, 2009).

p. Possession of False Papers to Defraud United States

Title 18 U.S.C. §1002. Possession of False Papers to Defraud United States

This offense occurs when a person “knowingly and with intent to defraud the United States, or any agency thereof, possesses any false, altered, forged, or counterfeited writing or document for the purpose of enabling another to obtain from the United States, or from any agency, officer or agent thereof, any sum of money” (Possession of False Papers to Defraud United States, 1994).

q. Wire Fraud

Title 18 U.S.C. §1343 Fraud by Wire, Radio, or Television

Wire fraud is committed when a person “having devised or intending to devise any scheme or artifice to defraud, or for obtaining money or property by means of false or fraudulent pretenses, representations, or promises, transmits or causes to be transmitted by means of wire, radio, or television communication in interstate or foreign commerce, any writings, signs, signals, pictures, or sounds for the purpose of executing such scheme or artifice” (Fraud by Wire, Radio, or Television, 2008).

4. Fraud Prevention and Detection Factors

As previously detailed, procurement fraud can be charged under a number of statutes. In order to effectively conduct a meaningful investigative analysis according to these legal statutes, it is first important to understand where procurement fraud occurs within the organization. According to Caulfield (2014), procurement fraud is more likely to be discovered within organizations that have weak internal controls. Having an understanding of the internal controls within an organization is a key point in effective fraud prevention and detection (Caulfield, 2014).

Another factor that must be understood for an acquisition professional, auditor, or investigator to effectively prevent or detect fraud is the perspective of the fraudster. The fraudster usually has some sort of connection with the organization in which the fraud is discovered. The fraudster is often the "trusted agent" (Caulfield, 2014). This can be a great challenge for employees within that organization to recognize the fraudster who is hiding his fraudulent behavior behind his trusted agent duties. Knowledgeable employees within an organization as well as knowledgeable auditors, acquisition professionals, and investigators are valuable weapons against the hurdles put forth in discovering the fraudster within any organization (Caulfield, 2014).

Another factor that is important to consider in preventing and detecting fraud is utilizing the perspective knowledge of the auditor. From the auditor's perspective, in order to be able to easily recognize and detect fraud, an organization needs to be auditable. According to Rendon and Rendon (2015), procurement agencies need auditability in order to maintain transparency, accountability, and integrity in their programs and is the first line of defense in fighting procurement fraud. Auditability theory encompasses various factors including governance which focuses on internal control effectiveness, capable processes, and competent personnel (Rendon & Rendon, 2015). The auditability triangle contains the elements; internal controls, processes, and personnel, each of which contain smaller components (see Figure 3). The internal control components include internal controls which are enforced, monitored, and reported. The processes components include processes which are institutionalized, measured, and improved. The personnel components include personnel who are educated, trained, and experienced (Rendon & Rendon, 2015).

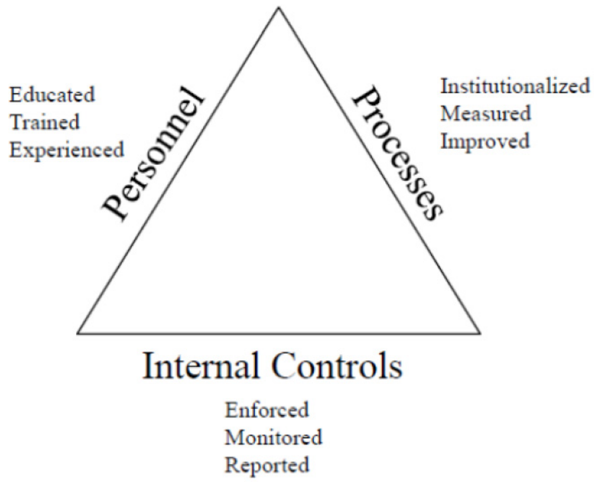

Figure 3. Auditability Triangle. Source: Rendon and Rendon. (2016).

In addition to being knowledgeable, what is often overlooked are risk management initiatives focused on reducing fraud, waste, and mismanagement. While the focus of this research is on the viability of a potential data analysis technique designed to detect procurement fraud in contracts, it is also important to understand the concerns of procurement fraud from the government's perspective and how these concerns were addressed prior to our current state of technological and data handling capability. Although the suggestion of leveraging data to reduce fraud was not as emphasized prior to the 21st century, fraud, waste, and mismanagement risk reduction was a growing concern. The government addressed these concerns using several methods. As identified by Weisman (1987), in the late 1980s, risk management program initiatives had been pursued by the government based on increased concern over fraud and mismanagement risk. These fraud risk management program initiatives were broken down into specific areas. These focus areas included identifying exposure areas and preventative techniques (Weisman, 1987).

5. Identifying Exposure Areas

A major focus of this research is how to effectively prevent fraud. Effectively preventing fraud is heavily dependent on identifying and focusing on exposure areas within the organization where fraud allegations occur the most. One of the major challenges in focusing on those areas includes recognizing the difference between errors and fraud. It is sometimes easy to misconstrue the two, and it is critical to not only be able to identify the error, but also to recognize and discern when that error may be a product of intentional behavior indicative of fraud or negligent acts indicative of complacency. When employees exhibit these conditions of behavior, it can be harmful to other employees, the organization itself, and in relation to DoD contracts and acquisition, it can be harmful to the taxpayers (Weisman, 1987).

There are several areas found within contracting functions that should be focused on in order to effectively identify and recognize error or fraud. Labor charging is one of these areas and, according to Weisman (1987), is the highest cost element within DoD contracts. A critical element in document related internal controls and verification is the use of third party documentation. Labor costs are typically lacking in this area. Third party documentation typically does not support labor costs which results in fraud vulnerabilities (Weisman, 1987). Errors may occur in labor charging systems, and it is critical to have strong internal controls to prevent errors. Sometimes when these errors are discovered, they can be interpreted as fraud. Conclusions of fraudulent behavior can be more convincing when the error is benefitting the contractor than if the error was not benefitting the contractor (Weisman, 1987).

Defective pricing is another exposure area in which the government focuses. Contractors are required to provide current, accurate, and complete pricing data to the government as of the date of certification, and if the pricing data is overstated, they violate the Truth in Negotiations Act. If the government discovered the overstated pricing data, it is entitled to an adjustment (Weisman, 1987). This could be considered fraud if the contractor knew at the time he certified the pricing data that it was not accurate, current, or complete, as in the case of overstated pricing data. Even if the contractor made an error in the pricing data, it could be construed as fraud, especially if the error gave the contractor a financial advantage (Weisman, 1987).

Another area of serious concern to the government is the occurrence of gratuities, kickbacks, briberies, and conflicts of interest between government employees and contractors. Weisman (1987) detailed the occurrence of these fraud schemes as an area of emphasis for the government because of the “revolving door.” Weisman (1987) defines the "revolving door" as "the movement of people from DoD to industry and vice versa" (p. 30). This "revolving door" condition includes the symptoms of private contractors with prior DoD personnel pensions and DoD personnel with prior contractor company pensions. These conditions generate an area of risk to kickbacks, bribery, and conflicts of interest because the relationships necessary for these transactions to cultivate already exist via the "revolving door." Some of these transactions have been investigated through sting operations to catch the government employees and contractors engaging in these fraudulent behaviors (Weisman, 1987).

Product substitution is a common fraud scheme which is the intentional substitution of a product detailed in the contract without disclosure resulting in a probable cheaper cost to the contractor and potentially inferior product which is not in accordance with the contract specifications. This was a focus area for the government in the late 1980s just as it is today. In addition to product deliverables, other areas relevant to this scheme include materials, production processes, and testing (Weisman, 1987). Investigations in 1984–1985 uncovered product substitutions related to a DoD moratorium which included shipments of weapon systems. These weapon systems included semi-conductors that were not tested in accordance with specifications stipulated in the contract. These semi-conductors were manufactured by Texas Instruments, who in defense, argued over excessive contract specifications. In order to avoid fraud allegations of product substitution, it is important for contractors to be clear in their intent in accordance with the contract specifications (Weisman, 1987).

6. Preventative Techniques

Weisman (1987) asserts that the primary objective of preventative techniques for government fraud, waste, and mismanagement is keeping the disputes within identified exposure areas, such as in the mechanical and accounting error arena, and keeping the disputes out of the litigation arena. It is simpler, cheaper, and much less risky to deal with an accounting issue than to deal with a matter that will end up in litigation (Weisman, 1987). Weisman (1987) illustrated that written company policies and procedures are effective preventative techniques to procurement fraud. These written policies should cover all areas of checks and balances regarding how to accomplish procedures, how systems work, and how error prevention is accomplished. They should be meticulous in these documentation efforts (Weisman, 1987).

Another valuable preventative technique is internal audit. It is essential to monitor policies and procedures to make sure they are followed and are effective. Strong internal audit functions are one of the best ways to monitor effectiveness and compliance (Weisman, 1987). A defense contractor's internal audit function should perform similarly to DCAA. In labor charging activities, internal audit procedures should include payouts, door checks, and floor checks. Internal auditors should perform post award cost audits in activities conducive to defective pricing. Fraud prevention programs should include monitoring by functions such as internal auditing to be effective (Weisman, 1987).

C. PROCUREMENT FRAUD DETECTION

A key element of antifraud programs is risk evaluation, which involves taking steps to mitigate the risks that are discovered in the evaluation and leveraging certified public accountants' (CPA) auditing expertise ("Detecting procurement fraud," 2003). Considering Weisman's (1987) research was over 30 years ago, fraud prevention efforts mentioned in his research did not include efforts related to data analysis. Although data analysis related prevention efforts were excluded, many of the efforts detailed in 1987 by Weisman are still relevant today. These areas include internal control functions, written company policies and procedures, code of ethics implementation and enforcement, internal audit, training programs, competent people development, communication, ombudsman functions, voluntary reporting and whistleblower support, automated time charging systems, and periodic checkups and monitoring, to name a few (Weisman, 1987).

Procurement fraud schemes, fraud statutes, and prevention efforts have been discussed. The following section will address the consequences of fraud and data analysis approaches for detecting procurement fraud in DoD contracts.

1. Consequences of Fraud

Some of the forms of procurement fraud and prevention methods as detailed by Weisman (1987) have been previously covered, but it is important to also briefly cover some of the potential consequences to those engaging in procurement fraud before covering the data analysis approaches. Detected procurement fraud can initiate investigations uncovering evidence and resulting in guilty pleas, acquittals, and civil case settlements. Detecting procurement fraud can also result in administrative remedies, such as suspension and debarments temporarily or permanently barring an entity from doing business with the government. Usually administrative remedies such as these are sought when there is not enough evidence to prove in a court of law that fraud occurred. Debarments can coincide with guilty pleas in criminal cases so that continued risk to the government will not continue with the same entity exposed to fraud (Karpoff, Lee, & Vendrzyk, 1999).

2. Data Analysis Approaches in Procurement Fraud Detection

Even though Karpoff et al.'s (1999) research does not detail a data analysis method related to fraud detection, it is important to mention that it does apply a data analysis method to evaluate the effectiveness of penalties. Even back in the late 1990s, the idea that data analytics can be useful for multiple approaches toward procurement fraud was useful and supported. Karpoff et al. (1999) collected data related to procurement fraud penalties from categories including both enforced penalties and secondary effect consequences. These two categories included fines, civil claims paid, restitution and damages paid, repaid investigation costs, and in the form of consequences, revenue and stock value changes (Karpoff et al., 1999). This data was collected from penalties enforced on various contractors big and small. Based on the data analysis conducted, Karpoff et al. (1999) concluded that contractors with greater influence are typically sanctioned much lighter for procurement frauds in comparison to lesser influential contractors.

a. Data Pools and Fraud Risk

This literature review includes research related to improving a government auditor's procurement fraud detection abilities with information technology and data analytics. One method of procurement fraud detection through the use of data analytics is by computer assisted audit techniques (CAAT). This technique is applied to payment processes and is one of the most efficient procurement fraud detection methods. One of the reasons this method is so efficient is because it can be used for large and small amounts of data (Ramamoorti & Curtis, 2003). Ramamoorti and Curtis (2003) also argue that out of all of the procurement phase processes, the payment process had the best data availability. Their research focused on the payment function in the procurement process, and the prospects of data analysis to root out fraud (Ramamoorti & Curtis, 2003).

Ramamoorti and Curtis (2003) highlighted the fraud risk hypothesis approach, also known as fraud theory approach, developed by experienced auditors through regular practice. The fraud risk hypothesis approach is utilized when auditors or investigators develop a theory early in their investigation regarding whether or not a finding was an unintentional error or deliberate fraud. They also develop their theory by determining what the potential fraud scheme might be and how the system being audited may be exploited by the fraudster (Ramamoorti & Curtis, 2003). In tandem with implementing the fraud risk hypothesis, the auditors or investigators should be aware of the organization's Electronic Data Interchange (EDI). EDI is "the movement of business data electronically between or within firms (including their agents or intermediaries) in a structured, "computer-processable" [sic] data format that permits data to be transferred without rekeying from a computer-supported business application in one location to a computer-supported business application in another location" (Hill & Ferguson, 1989, para. 21). This is important because it arms the auditors and investigators with a data pool to use for potential analysis that may uncover anomalies indicative of fraud risk or the fraud incident itself (Ramamoorti & Curtis, 2003).

In order to effectively leverage data analytics toward procurement fraud identification, it is important to understand the principles of effective procurement fraud detection and investigation. First, it is important to know your data. "Garbage in, garbage out" (GIGO) is the concept that should be applied when the auditors or investigators are getting familiar with the data pool that is in front of them. Although defects and impurities in the data must be recognized, it may not necessarily be feasible or possible to cleanse the data pool of these anomalies (Ramamoorti & Curtis, 2003). Understanding the entry fields, such as invoice date and payment date, and what the fields truly imply can help the auditors and investigators conduct the data mining process effectively (Ramamoorti & Curtis, 2003).

Ramamoorti and Curtis' (2003) research also emphasized the importance of knowing the purpose, scope, and nature of the tests applied as well as the software to be used to analyze the data and apply those tests for fraud detection. Audit tests should be prioritized according to effectiveness and cost. Due to budget constraints and personnel availability and limitations, the selected audit test for procurement fraud should be chosen based on the appropriate trade-off between effectiveness and cost of running the test (Ramamoorti & Curtis, 2003). Methods, such as Benford's law, can be applied in these tests effectively. The use of methods utilizing Benford's law detects anomalies in frequencies expected in number listings (Ramamoorti & Curtis, 2003). Additional tools proven valuable when utilized for these audit tests include software such as Microsoft's SQL Server or CaseWare's IDEA®, but just like knowing the appropriate test to use, it is important to know the appropriate software to use for the data pool being analyzed as well as understanding the software itself (Ramamoorti & Curtis, 2003). Although software can be a powerful tool for applying data analysis to a fraud investigation, it should not be considered a substitute for appropriate training. The successful auditors are those that pursue training and other learning opportunities to improve and augment their skills (Ramamoorti & Curtis, 2003). Using these software tools along with the training and experience obtained by successful auditors or investigators can more effectively root out fraud and prevent false positives for fraud identification (Ramamoorti & Curtis, 2003).

The risk of false positive fraud determinations is a reality for auditors when conducting their audits as well as investigators when they are conducting criminal investigations. This risk increases with organizational complexity and size (Ramamoorti & Curtis, 2003). To address this problem, Ramamoorti and Curtis (2003) illustrated how Type I (false positive) and Type II (false negative or miss) error analysis can be applied to an auditor's data analysis test for detecting fraud. Type II errors are of great concern in that it is not uncommon for an auditor to make a Type II error for an actual fraud occurrence. These misses and false positives can be reduced significantly by developing tests with greater detail and specificity to appropriately target specific risk areas for fraud in an organization. To illustrate, a specific method of tailoring data analysis tests to better target problem areas is to change the thresholds of the test. For example, targeting a lower threshold that produces too many hits and an upper threshold that produces too few hits for proper analysis may call for a decrease in the upper threshold to a level that facilitates a more effective analysis and meaningful test to meet the auditor's objective (Ramamoorti & Curtis, 2003).

b. Proactive Data Monitoring and Analysis for Fraud Detection

In more recent years, data analytics has become even more prominent and important to preventing and detecting fraud. Bănărescu (2015) pointed out one that of the most effective tools in anti-fraud efforts identified in the 2014 ACFE report was proactive approaches toward data monitoring. These approaches included data analyses applied toward reducing fraud losses and duration of ongoing fraud schemes (as cited in Bănărescu, 2015). Nearly 35% of organizations experiencing fraud utilized proactive approaches toward data analysis and monitoring in their fraud detection and prevention efforts (as cited in Bănărescu, 2015). Those organizations that utilized data analysis tools experienced nearly a 59.7% reduction in median loss and a 50% reduction in fraud scheme duration (as cited in Bănărescu, 2015). While there appeared to be positive benefits to data, there was also a study identifying the barriers to implementing data analytical systems during audit activities. Some of these barriers included a lack of CAAT tools and data analytic software, gaps in audit methods applied toward data analytics support, and the perception that audit activity efficiency would not increase with the use of data analytics and CAAT tools (as cited Bănărescu, 2015).

Data mining, for the purpose of analytical processing, is designed to extract and explore data within datasets to determine relationships and patterns (Bănărescu, 2015). Data mining can also be defined as, "the nontrivial extraction of implicit, previously unknown, and potentially useful information from data" (as cited in Bănărescu, 2015). There were several methods of data mining to augment an auditor's approach toward detecting fraud including text data mining or "text analytics" (Bănărescu, 2015, p. 1832).

There are two categories of text data mining products, including one that optimizes activities of organizations which involved navigation and document viewing, and the other one that provides text analysis functions including the data summarization, extraction, and classification (as cited in Bănărescu, 2015).

In addition to text data mining, geospatial analysis is another data analysis method for auditor activities. Geospatial analysis is a visual means of analyzing events and their relevance to the location where those events occurred in order to discover fraud behavior patterns (Bănărescu, 2015). Just as in previous examples, the limitations of these data analysis methods are that the implementation of these analytical systems is reliant upon retrieving data from a variety of sources. Additionally, these systems are not autonomous because they rely on some sort of human interaction from auditors and other data miners to be effective. They combine both human and technical factors that will not function alone effectively without the other (Bănărescu, 2015).

As highlighted by Bănărescu (2015), integration of data analytical methods such as text data mining and geospatial analysis exhibit various limitations and benefits. Some of the limitations of data analytical software and tools identified included the cost of prevention and detection software, large portions of data not being included in datasets for analysis, and the requirement of human interaction with the software interface. Furthermore, due to the complexity of analytical data research, additional descriptive tools such as tables and graphs are likely required to adequately present the data. Some of the benefits identified included getting results in real time to fraud issue related questions, automatic data collection, elimination of erroneous duplicate records, data integrity improvement, high productivity substitute for manual work, increased rate of fraud detection, and rapid fraud detection and recovery from fraud activity consequences (Bănărescu, 2015).

Bănărescu (2015) concluded that a number of data analysis approaches and software tools can be effective in detecting and preventing fraud and that there is no perfect combination of tools to use. To organizations new to fraud detection efforts, Bănărescu (2015) recommends getting started using such tools and approaches to make those improvements in fraud detection and prevention and to build a system framework. This system framework can be built in a series of steps to include defining the intention of fraud detection and prevention and creating a function for that purpose, creating an infrastructure that facilitates deliberate internal and external data collection, enabling real time detection of irregularities through the use of predefined fraud scheme models, and creating a recovery system, to name a few (Bănărescu, 2015).

c. *Solving the Fraud Data Analytical Problem*

While the previous discussion provides some solutions to fraud prevention and detection, this section presents some problems in fraud detection. Two such problems presented were the "needle in a haystack problem" and the "curse of the data dimensionality problem" (Perols, Bowen, Zimmermann, & Samba, 2017, p. 222). The former problem termed formally as "relative rarity" manifests when identified fraud occurrences have a significantly smaller percentage relative to non-fraud occurrences. To illustrate this point, Perols et al. (2017) found that merely 0.6% of all financial statements audited in the U.S. were determined fraudulent. This presents a challenging problem because it mandates classifier algorithms to include an extremely big volume of possible patterns without a sufficient pool of fraud occurrences to identify patterns derived from immense and extraneous data. As a result, the risk of false positives increases for a false negative rate when a new sample has been applied (as cited in Perols et al., 2017). The authors further explained that algorithms applied to reduce erroneous classifications have a tendency toward bias in correctly applying observation classification (Perols et al., 2017). For example, if non-fraudulent observations accounted for 99% of the total observations, then all observations identified as non-fraudulent in a prediction model would be 99% accurate (Perols et al., 2017). This scenario would correctly classify virtually all non-fraudulent observations in totality. Contrastingly, this scenario would also correctly classify 0% of fraudulent observations (Perols et al., 2017). A solution to the "needle in a haystack problem" or "relative rarity" was a data analysis method termed "Multi-Subset Observation Undersampling" (as cited in Perols et al., 2017).

As explained by Perols' et al. (2017), Multi-Subset Observation Undersampling (OU) utilizes poly-subsets. All fraud observations are contained in each subset which also contain different subsamples of non-fraud observations. The OU method balances out the ratio of majority non-fraud observations and minority fraud observations, and as a result, improves the precision of algorithmic classification in fraud occurrences (Perols et al., 2017). By applying this method using multiple prediction models of non-overlapping majority observation subsets, each prediction model will differ from each other (see Figure 4). This is a significant effect because multiple subsets are more likely to present patterns predictive of fraud (Perols et al., 2017). An additional similar solution presented in the article was utilized for the "data dimensionality problem" (Perols et al., 2017).

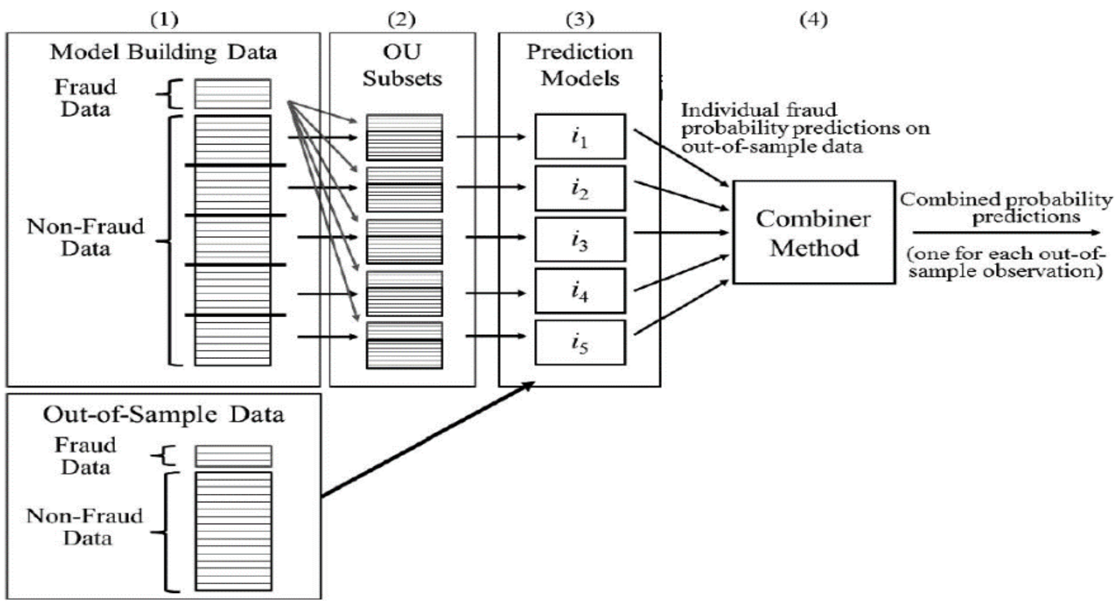

Figure 4. Multi-Subset Observation Undersampling (OU). Source: Perols et al. (2017).

Under the "curse of the data dimensionality problem," when the number of explanatory variables increase in the dataset, data requirements increase exponentially (as cited in Perols et al., 2017). In relation to the problem of relative rarity of fraud in the enormous volume of non-fraud occurrences presented before, the data dimensionality problem addresses a similar issue concerning a voluminous number of variables. The study of this problem found that detected fraud cases were small in number compared to the voluminous independent variable quantities found in other studies of fraud occurrence.

Only an insignificant number of fraud occurrences are present for the observer to see patterns in a large volume of independent variables and fraud (Perols et al., 2017). The solution for this problem was a method termed "stepwise backward variable selection" in which a "parsimonious fraud prediction model" is built (as cited in Perols et al., 2017). Although this method may be useful, it may discard variables that could potentially be useful. Perols et al. (2017) applied this method with the intention of remaining potentially useful explanatory variables. This approach was accomplished by designing an improved method of handling data rarity coined "Multi-Subset Variable Undersampling (VU)." With this method, the set of explanatory variables are randomly split into different subsets without replacement. Similarly, to OU, in conjunction to application of a classification algorithm, each subset is applied toward building a fraud prediction model which in turn is applied to non-sample data (see Figure 5). The intended result is combined model predictions which provide an overall fraud probability prediction (Perols et al., 2017).

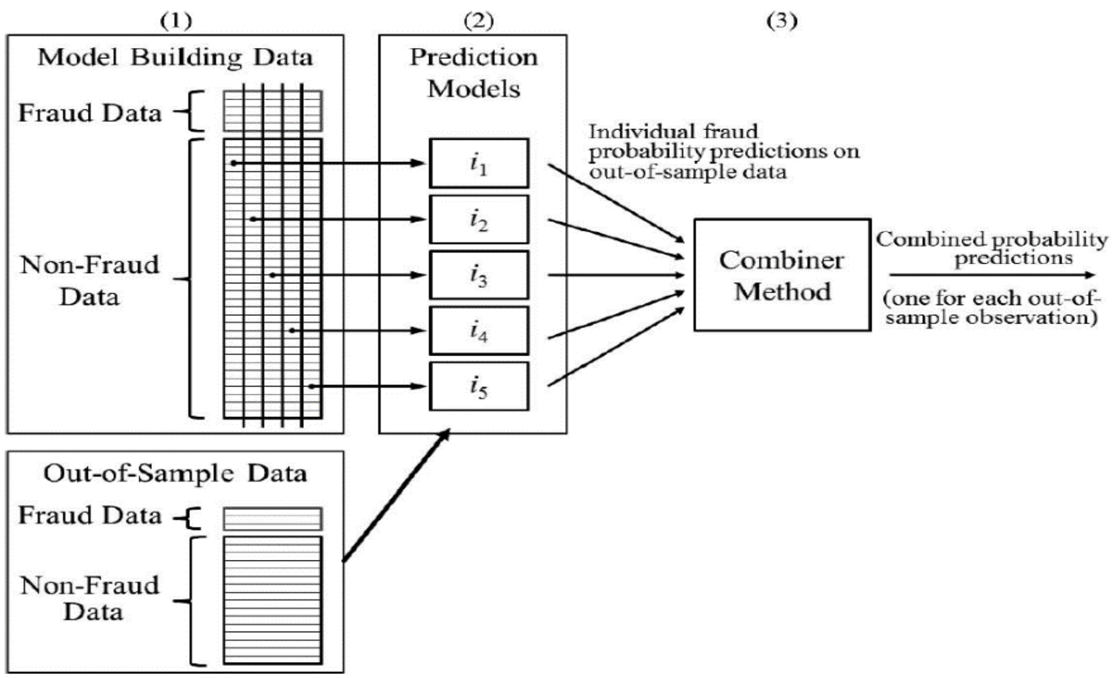

Figure 5. Multi-Subset Variable Undersampling (VU). Source: Perols et al. (2017).

In addition to presenting the two former examples of data analysis techniques to predicting fraud, Perols et al. (2017) concluded that their findings made several important contributions to prior research in this area. One of the most notable contributions was that they identified and addressed data rarity problems in financial statement fraud. They accomplished this by introducing and evaluating new multiple data pre-processing techniques applicable to the accounting fields. The OU and VU techniques were applied to experiments which resulted in findings that supported a significant reduction in expected misclassification costs by approximately 10% (Perols et al., 2017). Perols et al. (2017) argue that the results of using these methods show improvement in the quality of fraud prediction.

Eger, Juanita Rendon, and Rene Rendon (2014) contend that the currently employed fraud prevention process in the DoD is a system which is outdated due to its reactionary nature involving post-fraud detection based on whistle-blower reporting. Reliance on data analytics for proactive fraud detection is the driving force in the commercial business sector. The commercial business sector employs these methods to prevent fraudulent behavior by looking for data anomalies that indicate potential fraud (Eger et al., 2014). The government can learn from the commercial sector in fraud prevention and implement their own proactive data analytics tools. The development of these tools can begin with taking existing data holding systems, collaborating with commercial industry, and modeling potential data analytical tools based on detection of procurement fraud schemes (Eger et al., 2014). Eger et al. (2014) proposed a conceptual framework for a successful data analytic fraud prevention program (see Figure 6).

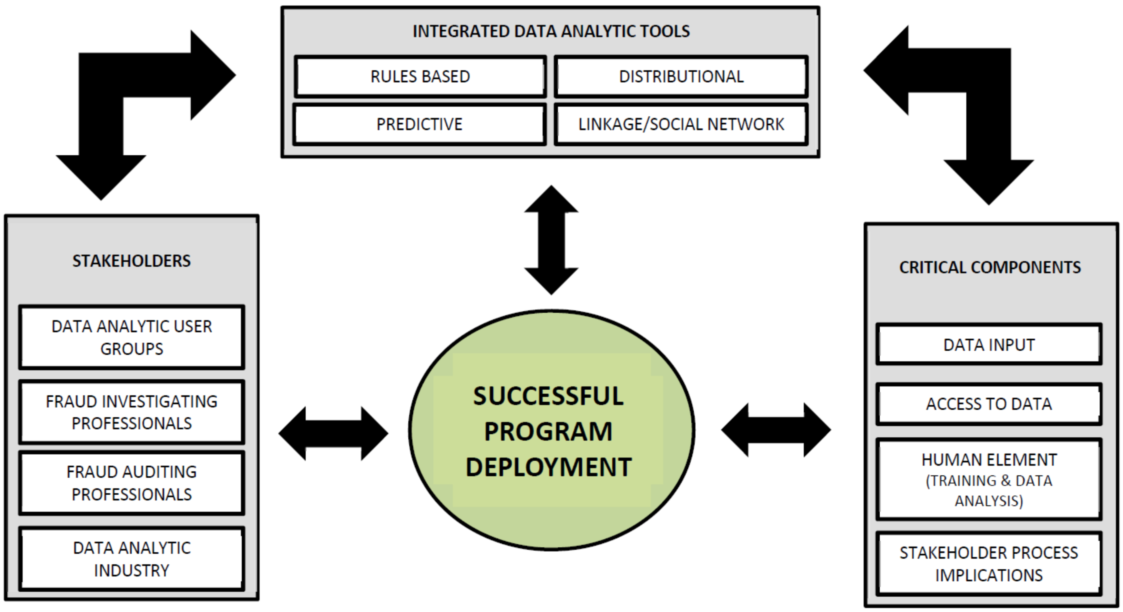

Figure 6. DoD Data Analytic Fraud Prevention Program Conceptual Framework. Source: Eger et al. (2014).

Stakeholders in government procurement including professionals from auditing and investigative fields were interviewed. Based on interviews of government procurement professionals, Eger et al. (2014) found, from their perspectives, that a data analytical fraud prevention program had more procedural implications than policy implications. Furthermore, government procurement professionals asserted that the data analytical fraud prevention program increases the confidence level of stakeholders within government procurement and does not relieve them of their management responsibilities in due diligence (Eger et al., 2014). Based on interviews of auditing and investigative professionals, Eger et al. (2014) discovered their opinion to be in favor of the use of a data analytics fraud prevention program as a valuable tool. They also indicated minimal policy implications in implementation of such a program. Although they predicted a significant workload increase due to the potential increase in proactively detected fraud incidents, the auditing and investigative professionals contended that the increase in workload is outweighed by the potential benefits (Eger et al., 2014).

D. SUMMARY

This chapter provided information gathered from the literature review. This literature review included the applicable fraud statutes that were included in the dataset for this research's analysis and accompanying legal statute descriptions. This provided for an understanding of the fraud statutes beyond their title alone. Additionally, this chapter included literature that provided information on procurement fraud in defense contracts and the associated challenges involved in identifying and resolving fraud. Finally, this chapter also included relevant data analysis concerns and techniques for detecting and predicting fraud. The following chapter discusses the methodology used for this research.

THIS PAGE INTENTIONALLY LEFT BLANK

III. METHODOLOGY

A. INTRODUCTION

The purpose of this research was to determine if there are statistical relationships and if so, what the statistical relationships are, between defense contracts associated with classes of fraud investigations and their associated contract variables. This chapter will provide an explanation of how the data was collected for the dataset. This chapter will also provide variable descriptions to illustrate the characteristics of the contract variables included in the dataset. Additionally, this chapter includes an explanation of the method of statistical analysis applied to the dataset. To answer the research questions for this study, the methodology for this research study applied multinomial logistic regression to determine the existence of and describe the statistical relationships between the contract variables belonging to contracts associated with alleged fraud investigations and applicable fraud statutes. In order to explain the methodology used for this research study, it is appropriate to first describe the sample dataset, then illustrate how the data was selected and collected.

B. SAMPLE

The dataset analyzed for this research included six variables identified as the following; fraud statute class, contract type, competition, small business set aside condition, small business status, and the Product Service Code (PSC) category. Of the six variables, the fraud statute class was classified as the dependent variable and the other five variables were classified as independent variables. This dataset was chosen for this research because the purpose was to determine if there are statistical relationships between these contract variables and fraud statute classes that could potentially indicate a higher or lower frequency of alleged fraud class based on the types of contracts related to these investigations. The results sought include determining likelihood of fraud statute class membership based on contract variable. These variables were collected from two sources.

The first of the two sources of data was Air Force Office of Special Investigations (AFOSI) procurement fraud analysts who provided the data containing the contract

Procurement Instrument Identification Number (PIIN) and applicable fraud statute of the alleged fraud investigation. The AFOSI procurement fraud analysts were not requested to apply any particular method to the data collection so as to prevent bias in the data collection efforts. The AFOSI procurement fraud analysts did not provide any information on how the data was collected other than case status and date range from where the data came. Contract PIINs and the applicable fraud statutes were provided from final closed fraud cases from FY12–FY17. No other parameters were specified or indicated. Beyond the contract PIINs and fraud statute variables, no other data from this source was included in the research analysis. The data contained no personally identifiable information (PII). The PIINs were used solely for the purpose of cross-referencing within the Federal Procurement Data System (FPDS), the second of the two sources, to collect the associated contract variables for analysis. The fraud statutes served as the dependent variable in the analysis and were further broken down into fraud statute classes to better serve the analysis.

The second of the two sources was the FPDS, a public database located at www.fpds.gov. The four contract variables, including contract type, contract competition type, small business set aside condition, contractor small business condition, and Product Service Code (PSC) were collected from FPDS for each cross-referenced contract PIIN. The PSCs were divided into broader classes to better serve the model by cross-referencing each PSC in the PSC Manual available at www.acquisition.gov/PSC_Manual. Once the contract variables were extracted from FPDS, the contract PIINs served their purpose and were no longer used. Once the data collection was complete, each set of the four contract variables and associated fraud statute were organized into separate individual observations. This was accomplished by stripping the PIINs from the dataset and replacing them with sequential observations numbers from 1 through 313, so as to remove any contract identification prior to analysis to remove any possible bias. There were a total of 313 observations containing variables from DoD contracts related to alleged fraud investigations in the sample.

C. OPERATIONALIZATION OF VARIABLES

1. Dependent Variable–Fraud Statute Class

The dependent variable in this research analysis was the fraud statute class. The fraud statutes, each detailed in the literature review, were selected from the fraud family of criminal statutes contained in the U.S.C. AFOSI investigates and classifies a criminal fraud investigation by its lead charge found among the listed criminal statutes. Each observation in the sample used for this project had a single fraud statute assigned to it. For the purpose of this research analysis, the fraud statute class was assigned as the dependent variable. Of the total sample, approximately 36% of the sample observations were related to alleged fraud investigations into False, Fictitious, or Fraudulent Claims, approximately 20% were related to Civil False Claims Act violations, and approximately 9% were related to False Statements. Each of the other fraud statutes had much lower frequencies from 7% and below of the total sample of observations (see Figure 7). To improve model functionality, the fraud statues with the bottom 17 frequencies in the observation sample were aggregated into a single class: all others. This aggregation was accomplished to better serve model functionality and avoid multi-collinearity problems. The fraud statute classes with the top three highest frequencies maintained their independent class identities.

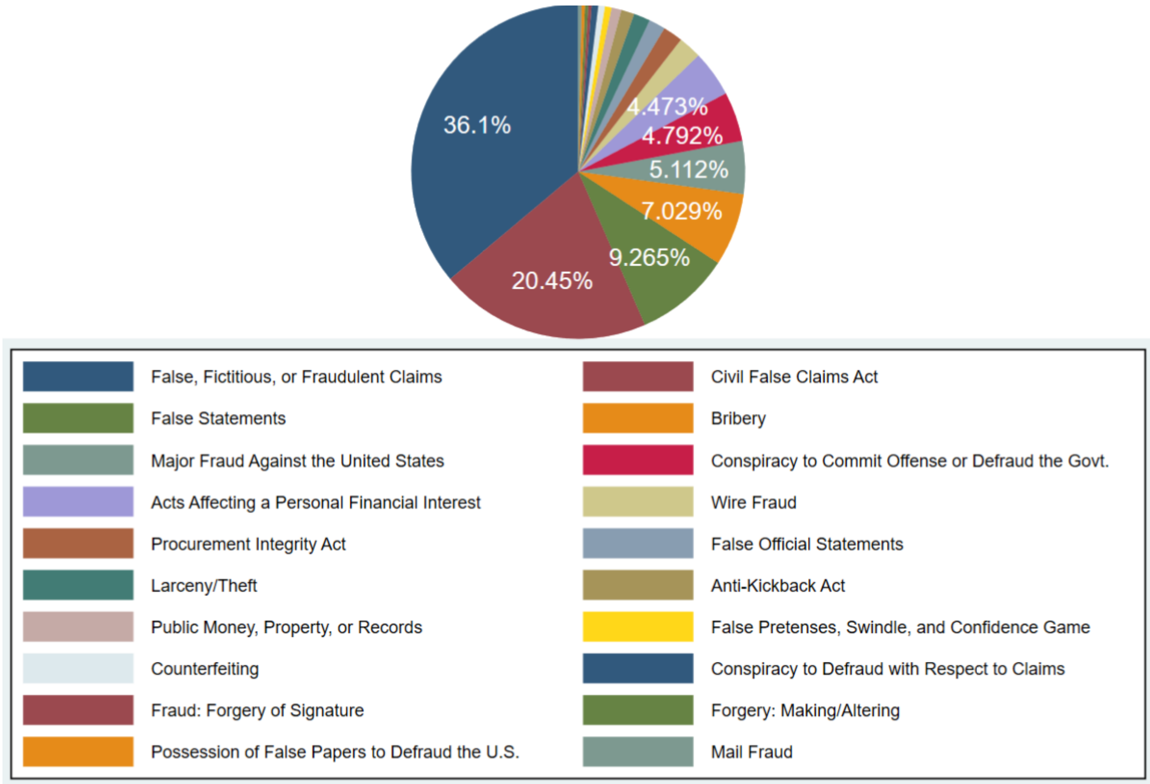

Figure 7. Percent of Observations in the Sample Related to Alleged Fraud Investigations by Statute

2. Independent Variables

The independent variables were the contract type, competition type, small business set aside condition, contractor small business condition, and PSC category. These variables, collected from FPDS, were organized in order to facilitate analysis to determine statistical relationships between the independent variables and the dependent variable, the fraud statute classes. These contract variables serve as the contract characteristics that could potentially reveal patterns related to susceptibility to a class of fraud. These variables directly support the purpose of this research. As data was collected, some challenges in one of the independent variables was encountered. There was one independent variable that had 166 different classes; the PSC. With a sample of only 313 observations, this made it very difficult to determine a statistical relationship between the PSC and other variables, so the PSC was categorized into 19 different classes to improve the model results. Due to these variables all being categorical, the multinomial logistic regression model was selected to conduct the analysis.

a. Contract Type

The following table depicts the frequency of contract types by fraud statute class within the sample of 311 observations. False, fictitious, or fraudulent claims had the highest frequency of the fraud statute classes accounting for 36.3% of the total observations. Fixed price contracts had the highest frequency among all of the differing contract types accounting for 76.9% of the total observations. 28.6% of all observations were both associated with false, fictitious, and fraudulent clams and fixed price type contracts. Among the lowest frequencies that occurred in the sample by fraud statute class included false statements and miscellaneous (Indefinite Delivery Indefinite Quantity [IDIQ], Time and Materials [T&M]) type contracts (see Table 1).

Table 1. Contract Type Frequency

Key
frequency *cell percentage*

	Type Category			
Statute Category	Fixed Pri	Cost	Misc. (ID	Total
False, Fictitious Or	89 28.62	20 6.43	4 1.29	113 36.33
All Others	83 26.69	16 5.14	8 2.57	107 34.41
Civil False Claims Ac	49 15.76	7 2.25	6 1.93	62 19.94
False Statements	18 5.79	8 2.57	3 0.96	29 9.32
Total	239 76.85	51 16.40	21 6.75	311 100.00

b. Competition

Table 2 illustrates the frequency of contract competition type by fraud statute class within the sample of 307 observations. Again, false, fictitious, and fraudulent claims had the highest frequency among the fraud statute classes accounting for 35.5% of the total observations in the sample. Full and open competition had the highest frequency among the contract competition types accounting for 74.6% of the total observations in the sample. Twenty-five and four tenths percent of all observations were associated with both false, fictitious, or fraudulent claims and full and open competition type contracts. Among the lowest frequencies of the fraud statute classes in the sample, false statements and contracts that were not competed were included (see Table 2).

Table 2. Competition Type Frequency

Key
frequency *cell percentage*

Statute Category	Competition Category Full/Open	 Not Compe	Total
False, Fictitious Or	78 25.41	31 10.10	109 35.50
All Others	86 28.01	21 6.84	107 34.85
Civil False Claims Ac	43 14.01	20 6.51	63 20.52
False Statements	22 7.17	6 1.95	28 9.12
Total	229 74.59	78 25.41	307 100.00

c. Small Business Set Aside and Small Business

Table 3 illustrates the frequency of small business set asides (SBSA) and small business (SB) type contracts by fraud statute class within samples of 312 and 313 observations, respectively. Again false, fictitious, and fraudulent claims had the highest frequency among the fraud statute classes accounting for 36.2% of the SBSA observations and 36.1% of the SB type contract observations. SBSA was a “no” more than it was a “yes” accounting for 61.5% of the observations in that sample. SB was a “yes” more than it was a “no” accounting for 53.4% of the observations in that sample (see Table 3).

Table 3. Small Business Set Aside and Small Business Frequency

Key
frequency *cell percentage*

Statute Category	SB Set Aside No	SB Set Aside Yes	Total
False, Fictitious Or	**62** **19.87**	**51** **16.35**	**113** **36.22**
All Others	**65** **20.83**	**42** **13.46**	**107** **34.29**
Civil False Claims Ac	**52** **16.67**	**11** **3.53**	**63** **20.19**
False Statements	**13** **4.17**	**16** **5.13**	**29** **9.29**
Total	**192** **61.54**	**120** **38.46**	**312** **100.00**

Key
frequency *cell percentage*

Statute Category	Small Business Yes	Small Business No	Total
False, Fictitious Or	**67** **21.41**	**46** **14.70**	**113** **36.10**
All Others	**59** **18.85**	**48** **15.34**	**107** **34.19**
Civil False Claims Ac	**23** **7.35**	**41** **13.10**	**64** **20.45**
False Statements	**18** **5.75**	**11** **3.51**	**29** **9.27**
Total	**167** **53.35**	**146** **46.65**	**313** **100.00**

d. PSC Category

Table 4 illustrates the four PSC categories with the highest frequency of occurrence. False, fictitious, and fraudulent claims had the highest frequency among the fraud statute classes accounting for 36.1%. The facilities and construction PSC had the highest frequency accounting for 18.5% of the observations in the sample. The next highest frequencies were Research and Development (R&D), professional services, and information technology (IT) accounting for 17.3%, 16%, and 13.7% of the observations, respectively (see Table 4). Among the lowest frequencies of the PSC categories that occurred, human capital, medical, and travel & lodging were included.

Table 4. PSC Category Frequencies

Key
frequency   *cell percentage*

Statute Category	PSC Category: Facilitie	Research	Professio	Informati	Sustainme	Equipment	Transport	Weapons &	Total
False, Fictitious Or	16 5.11	19 6.07	17 5.43	15 4.79	8 2.56	8 2.56	4 1.28	8 2.56	113 36.10
All Others	35 11.18	13 4.15	21 6.71	6 1.92	6 1.92	1 0.32	6 1.92	4 1.28	107 34.19
Civil False Claims Ac	4 1.28	8 2.56	9 2.88	21 6.71	4 1.28	6 1.92	2 0.64	1 0.32	64 20.45
False Statements	3 0.96	14 4.47	3 0.96	1 0.32	3 0.96	1 0.32	1 0.32	0 0.00	29 9.27
Total	58 18.53	54 17.25	50 15.97	43 13.74	21 6.71	16 5.11	13 4.15	13 4.15	313 100.00

Statute Category	PSC Category: Industria	Office Ma	Aircraft,	Electroni	Miscellan	Clothing,	Federal S	Security	Total
False, Fictitious Or	3 0.96	3 0.96	1 0.32	4 1.28	2 0.64	1 0.32	0 0.00	2 0.64	113 36.10
All Others	3 0.96	0 0.00	1 0.32	1 0.32	3 0.96	2 0.64	1 0.32	1 0.32	107 34.19
Civil False Claims Ac	2 0.64	4 1.28	3 0.96	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	64 20.45
False Statements	0 0.00	0 0.00	1 0.32	0 0.00	0 0.00	0 0.00	2 0.64	0 0.00	29 9.27
Total	8 2.56	7 2.24	6 1.92	5 1.60	5 1.60	3 0.96	3 0.96	3 0.96	313 100.00

Statute Category	PSC Category: Human Cap	Medical	Travel &	Total
False, Fictitious Or	1 0.32	1 0.32	0 0.00	113 36.10
All Others	1 0.32	1 0.32	1 0.32	107 34.19
Civil False Claims Ac	0 0.00	0 0.00	0 0.00	64 20.45
False Statements	0 0.00	0 0.00	0 0.00	29 9.27
Total	2 0.64	2 0.64	1 0.32	313 100.00

D. MODEL SPECIFICATION

Multinomial logistic regression analysis was applied as the primary model to answer the research questions and determine statistical relationships between each independent variable and the dependent variable class, which was the fraud statute class. The following is the equation that was applied for the analysis where P is the probability logistic regression function, X_n is the dependent variable fraud statute class, and b_1 through b_m are independent variable coefficients:

Log(P(FraudClass=Xn)/(P(FraudClass=FalseFictitiousFraudulentClaims))
=b1+b2(Type=2)+b3(Type=3)+b4Competition+b5SmallBusinessSetAside
+b6SmallBusiness+b7(PSCCategory=2)+b8(PSCCategory=3)+…+b23(P
SCCategory18)

E. SUMMARY

This chapter provided a description of the methodology used in this research. The sample dataset and the independent and dependent variables were explained. The following chapter will provide discussion of the analysis and findings of this research.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. ANALYSIS

This chapter provides the analysis and findings of this research and will include descriptive frequencies of the independent variables in relation to the dependent variable. The multinomial logistic regression will determine statistical significance of relationships between the dependent and independent variables. Finally, the analysis will determine appropriateness of model fit based on the explanatory power of the independent variables.

A. MULTINOMIAL LOGISTIC REGRESSION

The multinomial logistic regression function was applied to the statistical model of relationships between the dependent and independent variables. The regression function did not converge with all 20 original dependent variable classes. The model was applied progressively while combining the dependent variable classes with the lowest frequency into the all others category. The model successfully converged with four dependent variable classes, reduced to classes with the three highest frequencies and the fourth all others category (see Table 5).

Table 5. Multinomial Logistic Regression Model

```
Iteration 0:   log likelihood = -388.22993
Iteration 1:   log likelihood = -332.60041
Iteration 2:   log likelihood = -328.27523
Iteration 3:   log likelihood = -320.94723
Iteration 4:   log likelihood = -320.21679
Iteration 5:   log likelihood = -320.05516
Iteration 6:   log likelihood = -320.02128
Iteration 7:   log likelihood = -320.01569
Iteration 8:   log likelihood = -320.01435
Iteration 9:   log likelihood = -320.01406
Iteration 10:  log likelihood = -320.01399
Iteration 11:  log likelihood = -320.01398

Multinomial logistic regression                 Number of obs     =        304
                                                LR chi2(66)       =     136.43
                                                Prob > chi2       =     0.0000
Log likelihood = -320.01398                     Pseudo R2         =     0.1757
```

StatuteCategory_n	RRR	Std. Err.	z	P>|z|	[95% Conf.	Interval]
All_Others						
TypeCategory_n						
Fixed Price	1.194633	.5557882	0.38	0.702	.4799804	2.973346
Misc. (IDIQ, T&M)	4.242179	3.906398	1.57	0.117	.6978509	25.78786
CompetitionCategory_n						
Not Competed (Sole Source, etc.)	.5272464	.1981811	-1.70	0.089	.2523855	1.101445
SBSetAside_n						
Yes	.6793276	.3122884	-0.84	0.400	.275918	1.672548
SmallBusiness_n						
Yes	.9238008	.4579844	-0.16	0.873	.3496116	2.441017

PSCCategory_n						
Clothing, Textiles & Subsistence S&E	1.873527	3.61199	0.33	0.745	.0428174	81.97839
Electronic & Communication Equipment	.2805087	.5131138	-0.69	0.487	.0077787	10.11547
Equipment Related Services	.1096547	.1957473	-1.24	0.216	.0033153	3.626839
Facilities & Construction	2.199366	3.228898	0.54	0.591	.1237785	39.07959
Federal Supply Class (PSC Rescinded)	9.57e+08	2.71e+13	0.00	0.999	0	.
Human Capital	1.368786	2.836445	0.15	0.880	.0235748	79.47384
Industrial Products & Service	1.410386	2.373007	0.20	0.838	.05214	38.15089
Information Technology	.4559622	.6945903	-0.52	0.606	.0230277	9.028325
Medical	1.368786	2.836445	0.15	0.880	.0235748	79.47384
Miscellanious S&E	1.509612	2.595319	0.24	0.811	.0519378	43.87807
Office Management	6.22e-08	.0001391	-0.01	0.994	0	.
Professional Services	1.44817	2.158489	0.25	0.804	.0780025	26.88627
Research & Development	.8036785	1.231284	-0.14	0.887	.0399016	16.18729
Security & Protection	.4382632	.8313289	-0.43	0.664	.0106443	18.04491
Sustainment S&E	.7662988	1.192819	-0.17	0.864	.0362594	16.19482
Transportation & Logistics Services	2.771806	4.602882	0.61	0.539	.1069673	71.82483
Weapons & Ammunition	.6407493	.9965816	-0.29	0.775	.0303936	13.50809
_cons	1.106126	1.665653	0.07	0.947	.0578133	21.16321
Civil_False_Claims_Act						
TypeCategory_n						
Fixed Price	2.507846	1.4579	1.58	0.114	.8025403	7.836727
Misc. (IDIQ, T&M)	9.655936	9.676219	2.26	0.024	1.354581	68.83097
CompetitionCategory_n						
Not Competed (Sole Source, etc.)	.8483989	.3605233	-0.39	0.699	.3688806	1.951257
SBSetAside_n						
Yes	.4455462	.2596022	-1.39	0.165	.1422103	1.3959
SmallBusiness_n						
Yes	.4028426	.2292773	-1.60	0.110	.1320308	1.229123
PSCCategory_n						
Clothing, Textiles & Subsistence S&E	8.68e-08	.0003072	-0.00	0.996	0	.
Electronic & Communication Equipment	2.37e-08	.0000659	-0.01	0.995	0	.
Equipment Related Services	.3197715	.4163199	-0.88	0.381	.024925	4.102458
Facilities & Construction	.1455469	.1902043	-1.47	0.140	.0112365	1.885273
Federal Supply Class (PSC Rescinded)	.394929	18490.23	-0.00	1.000	0	.
Human Capital	8.06e-08	.0003352	-0.00	0.997	0	.
Industrial Products & Service	.7341421	1.132997	-0.20	0.841	.0356555	15.1159
Information Technology	.8920194	1.104362	-0.09	0.926	.0788037	10.09722
Medical	8.06e-08	.0003352	-0.00	0.997	0	.
Miscellanious S&E	5.91e-08	.0001494	-0.01	0.995	0	.
Office Management	1.437784	2.134654	0.24	0.807	.0783302	26.39113
Professional Services	.5215182	.6652611	-0.51	0.610	.0428012	6.354517
Research & Development	.6756775	.8992856	-0.29	0.768	.0497556	9.17565
Security & Protection	2.20e-08	.0000735	-0.01	0.996	0	.
Sustainment S&E	.3289231	.4591031	-0.80	0.426	.0213308	5.072027
Transportation & Logistics Services	.5377156	.8448301	-0.39	0.693	.0247281	11.69271
Weapons & Ammunition	.0636983	.100786	-1.74	0.082	.0028663	1.415572
_cons	1.162914	1.533588	0.11	0.909	.0877058	15.41936

False_Statements						
TypeCategory_n						
Fixed Price	1.39864	.8609421	0.55	0.586	.4185458	4.673785
Misc. (IDIQ, T&M)	10.96204	12.13241	2.16	0.031	1.252596	95.93383
CompetitionCategory_n						
Not Competed (Sole Source, etc.)	.9318953	.6094888	-0.11	0.914	.2586155	3.357992
SBSetAside_n						
Yes	2.286849	2.172161	0.87	0.384	.3554059	14.71466
SmallBusiness_n						
Yes	.461543	.4535484	-0.79	0.431	.0672597	3.167157
PSCCategory_n						
Clothing, Textiles & Subsistence S&E	7.27e-08	.0003686	-0.00	0.997	0	.
Electronic & Communication Equipment	1.61e-08	.0000715	-0.00	0.997	0	.
Equipment Related Services	.155938	.2844725	-1.02	0.308	.0043665	5.568913
Facilities & Construction	.2202719	.3556664	-0.94	0.349	.0093014	5.21637
Federal Supply Class (PSC Rescinded)	3.12e+09	8.84e+13	0.00	0.999	0	.
Human Capital	4.75e-08	.0002795	-0.00	0.998	0	.
Industrial Products & Service	7.69e-08	.0002103	-0.01	0.995	0	.
Information Technology	.1019587	.1849002	-1.26	0.208	.002916	3.565029
Medical	4.75e-08	.0002795	-0.00	0.998	0	.
Miscellanious S&E	4.04e-08	.0001589	-0.00	0.997	0	.
Office Management	3.62e-08	.0001509	-0.00	0.997	0	.
Professional Services	.2137652	.3512405	-0.94	0.348	.0085372	5.35251
Research & Development	1.101448	1.795618	0.06	0.953	.0451134	26.89193
Security & Protection	2.32e-08	.0001214	-0.00	0.997	0	.
Sustainment S&E	.3247242	.5627552	-0.65	0.516	.0108733	9.69769
Transportation & Logistics Services	.555835	1.081537	-0.30	0.763	.0122656	25.18858
Weapons & Ammunition	3.30e-08	.0000774	-0.01	0.994	0	.
_cons	.5514682	.8968713	-0.37	0.714	.0227606	13.36159
False__Fictitious_Or_Fraudulent	(base outcome)					

Note: _cons estimates baseline relative risk for each outcome.

The sample size was reduced from 313 to 304 observations. This was due to some variable data missing because it was not available in the Federal Procurement Data System (FPDS). The logistic regression included the four dependent variable classes, which are listed from highest to lowest frequency, false, fictitious, and fraudulent claims, all others, Civil False Claims Act, and false statements. All results in the multinomial logistic regression were relative to the base outcome, false fictitious, or fraudulent claims. False, fictitious, and fraudulent claims was selected as the base outcome in the function by default due to holding the highest frequency of occurrence among the dependent variable classes. The independent variables included contract type, binary outcome for competition type, binary outcome for small business set aside, binary outcome for small business, and PSC category.

Each of the resulting independent variable class coefficients were relative to the omitted dummy variable class assigned to each variable. By function design, the dummy variable classes were omitted by default based on alphabetical order, with the first alphabetically labeled class within the variable selected for omission. The dummy variable class assigned to contract type was cost reimbursable type. The dummy variable class assigned to competition type was full and open. The dummy variable class assigned to both small business set aside and to small business was the negative binary response, "no." Finally, the dummy variable assigned to the PSC category included aircraft, ships, submarines, and land vehicles. The following sections describe the overall and independent variable model fit as well as the regression findings.

B. OVERALL AND INDEPENDENT VARIABLE MODEL FIT

The multinomial regression model had a statistically significant overall fit with the included independent variables, (LR χ^2 (66) = 136.43, $p < .01$) as measured by pseudo-R^2 = 0.1757.

1. Contract Type

A joint test of contract type was not significant, Δ LR χ^2 (6) = 8.62, $p > .05$.

2. Competition

A joint test of competition was not significant, Δ LR χ^2 (3) = 3.16, $p > .05$.

3. Small Business Set Aside

A joint test of small business set aside was not significant, Δ LR χ^2 (3) = 3.87, $p > .05$.

4. Small Business

A joint test of small business was not significant, Δ LR χ^2 (3) = 3.30, $p > .05$.

5. PSC Category

A joint test of PSC category was significant, Δ LR χ^2 (51) = 101.64, $p < .01$.

C. DISCUSSION OF THE FINDINGS

The purpose of this research was to determine if there is a statistical relationship between contracts associated with fraud statute classes of alleged fraud investigations and their associated contract variables. The following is a discussion of related findings based on the analysis.

1. Findings

The hypothesis (H_a) being tested is that the dependent variable contract type has a statistically significant ($p < .05$) effect on fraud statute classification. The results of the logistic regression analysis rejected the null hypothesis (H_0) that there is no effect on fraud statute class based on contract type as there were a few results which held statistical significance. The multinomial logistic regression resulted in identification of the effect that miscellaneous (IDIQ, T&M) contracts, relative to cost reimbursable type contracts, multiplies the odds 9.66 times of a contract associated with an alleged fraud investigation being classified as Civil False Claims Act relative to false, fictitious, or fraudulent claims. The null hypothesis (H_0) that independent variable miscellaneous (IDIQ, T&M) type contracts did not have an effect on dependent variable classification was rejected. Although these results were statistically significant, the inclusion of the independent variable miscellaneous (IDIQ, T&M) type contracts in the model was determined to have no improvement in the overall model. In contrast, there were indications based on frequency, the multinomial logistic regression model, and the predicted probabilities that with more data, more statistically significant results may show a stronger relationship between dependent variable fraud statute class and independent variable contract type.

There were no statistically significant results to reject the null hypothesis (H_0) that there is no effect of independent variable competition type on the dependent variable fraud statute class. Furthermore, the model fit test determined that inclusion of this independent variable did not improve the overall model. A notable finding that may indicate more data could reveal a stronger relationship between these variables was discovered during the model fit test on the PSC category. When the PSC category was excluded from the logistic regression model, the *p-value* on competition type decreased from 0.089 to 0.059 improving the statistical significance of competition type. This may indicate that, with more data, it may be possible to determine a statistical relationship between these variables.

There were no statistically significant results to reject the null hypothesis (H_0) that there is no effect of independent variable small business set aside or small business conditions on the dependent variable fraud statute class. Furthermore, the model fit test determined that inclusion of these independent variables did not improve the overall model. A notable finding that may indicate more data could reveal a stronger relationship between these variables was discovered during the model fit test on PSC category. When the PSC category was excluded from the logistic regression model, the *p-value* on small business set aside effect on Civil False Claims Act classification decreased to 0.040 improving statistical significance. This may indicate that, with more data, it may be possible to determine a statistical relationship between these variables.

The results of the logistic regression analysis did not reject the null hypothesis (H_0) that there is no effect on fraud statute class based on the PSC category as there were no results which held statistical significance. In contrast, the model fit joint test concluded that the independent variable PSC category provided statistically significant improvement in the overall model fit. To explore this anomaly further, a multinomial logistic regression was conducted with PSC category isolated as the only independent variable excluding all other independent variables. The findings indicate that there are statistically significant results to conclude that the null hypothesis (H_0) that there is no effect of PSC category on fraud statute classification could be rejected. The incongruence between the primary multinomial logistic regression and the single independent variable regression indicated that inclusion of the other independent variables decreased reliability of PSC category results. It could be concluded that this may be due to high multi-collinearity. More data is needed in order to improve overall model fit, overcome high multi-collinearity, and determine individual PSC category statistical significance in relation to the other independent variables.

2. Implications Based on the Findings

Based on the findings discovered using the multinomial logistic regression analysis for this research, there appears to be a lot of promise in applying these methods to potentially apply fraud risk based on classes of fraud and contract variables. While this research was just a start to developing a method of detecting fraud statute classes within a sample of observations containing only fraud related observations, there could also be potential in applying this method across a larger observation data sample that includes non-fraud related data. Since the observation sample was limited to only 313, and there were multiple dependent and independent variables with varying classes, this significantly reduced the possibility of obtaining statistically significant results. This resulted in many of the dependent variable classes, and independent variable classes lacking evidence of statistical relationship; thus, not supporting a rejection of the null hypothesis (H_0).

The majority of the data produced by the multinomial logistical regression analysis was not supportive of a statistical relationship among the dependent and independent variables in most instances during this research study. This is not to say that multinomial logistical regression is not appropriate for this type of analysis. The implication is quite the opposite, as there were results that proved otherwise. Furthermore, there were indications based on frequency, model fit, predicted probabilities, and exclusion of independent variables that an increase in the number of observations in the sample could likely produce more statistically significant results in multinomial logistic regression application. Most of the statistically significant results that were found in this research were related to the dependent variable class Civil False Claims Act because it had the second highest frequency and seemed to have enough data to produce evidence of statistical relationship. Aside from the dependent variable class false, fictitious, or fraudulent claims, for the majority of the other dependent variable classes, there was no evidence of statistical relationships presumably due to lack of data. More data is needed to determine the statistical relationships of the other variables.

D. SUMMARY

This chapter provided the findings and analysis for this research. The chapter also provided the descriptive frequencies of the independent and dependent variables. The multinomial logistic regression was applied to the statistical model of relationships between the dependent variable fraud statute class and the independent variables. Finally, this chapter provided the results of the appropriateness model fit based on the explanatory power of the independent variables. The following chapter provides a summary, conclusion, and areas for further research.

V. SUMMARY, CONCLUSION, AND AREAS FOR FURTHER RESEARCH

A. SUMMARY

Within the DoD, procurement fraud has been an issue of concern that has captured the attention of acquisition professionals, contracting agencies, investigative agencies, auditing agencies, and Congress. Methods of detecting and predicting procurement fraud are more accessible and necessary than ever before. As the literature review illustrated, there are many methods of data analyses which could be applied to detecting fraud. This research presented multinomial logistic regression as one practical method that could be applied to detect or predict procurement fraud statute class membership. Using this method proved its applicability to the problem of detecting fraud statute class based on contract variables. Although there were few statistically significant results in the findings, there was no absolute absence of statistical findings. The few statistically significant findings that were present indicated potential for improved effectiveness with a larger, more robust dataset.

B. CONCLUSION

The purpose of this research was to determine the statistical relationship between fraud statute class of investigations and the contract variables of the associated contract. As detailed during the analysis, there were statistically significant relationships discovered. Although few statistically significant relationships were discovered, there was evidence that improved dataset robustness could potentially reveal more statistically significant relationships between the variables. With improved applicability of the multinomial logistic regression to detect and predict fraud in DoD contracts, acquisition, contracting, auditing, investigations professionals as well as tax payers could benefit in the fight against fraud. The following are the answers to the research questions based on these findings.

1. What is the statistical relationship between contracts associated with a class of fraud investigation and the type of contract?

Based on the research findings, an answer to this question was partially determined. The multinomial logistic regression resulted in identification of the effect that miscellaneous (IDIQ, T&M) contracts had a higher probability of being classified as Civil False Claims Act relative to false, fictitious, or fraudulent claims. Although these results were statistically significant, the inclusion of this independent variable in the model was determined to have no improvement in the overall model. In contrast, there were indications based on frequency, the multinomial logistical regression model, and predicted probabilities that with more data, more statistically significant results may show a stronger relationship between dependent variable fraud statute class and independent variable contract type.

2. What is the statistical relationship between contracts associated with a class of fraud investigation and type of competition involved?

According to the research findings, an answer to this question could not be determined. There were no statistically significant results to reject the null hypothesis (H_0) that there is no effect of independent variable competition type on the dependent variable fraud statute class. More data would be needed to determine further statistical significance of the relationship.

3. What is the statistical relationship between contracts associated with a class of fraud investigation and business size status?

According to the research findings, an answer to this question could not be determined. There were no statistically significant results to reject the null hypothesis (H_0) that there is no effect of independent variable small business set aside or small business conditions on the dependent variable fraud statute class. More data would be needed to determine further statistical significance of the relationship.

4. What is the statistical relationship between contracts associated with a class of fraud investigation and the contracts' associated product service codes?

Based on the research findings, an answer to this question was partially determined. The results of the logistic regression analysis determined there was no evidence that PSC category had an effect on fraud statute class. In contrast, the model fit test concluded that the independent variable PSC category provided statistically significant improvement in the overall model fit. More data is needed in order to improve overall model fit and determine individual PSC category statistical significance in relation to the other independent variables. The following section will provide areas for further research.

C. AREAS FOR FURTHER RESEARCH

There is a lot of promise in applying this model to future research related to the detection of procurement fraud. As previously stated, the dataset for this research was limited to only observations with alleged fraud. Future research could include non-fraud observations as a base outcome and model the logistic regression analysis to determine a statistical relationship between the occurrence of fraud itself and contract variables. A challenge regarding this suggestion to be cognizant of includes the reality that there would be no way of knowing with certainty that an observation is virtually free of fraud. It would be critical to determine the nature and appropriateness of the data used in order to effectively collect non-fraud observations. A potential answer to this issue would be to frame the nature of this data differently, such as observations not related to fraud investigations. An additional way to expand the dataset for future research would be to include federal contracts in addition to DoD specific contracts.

Finally, another application for further research would be to apply the answer to the relative rarity problem introduced in the literature review chapter. Perols' et al., (2017) Multi-Subset Observation Undersampling (OU) and Multi-Subset Variable Undersampling could both be applied to similar research in the future. These two methods would address the problems of relative rarity and excessive independent variables experienced with the limited quantity of observations in this research.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Acts Affecting a Financial Interest, 18 U.S.C. §208 (1962).

Association of Certified Fraud Examiners (ACFE) (2018). *Report to the nations: 2018 global study on occupational fraud and abuse*. Austin, TX, Author. Retrieved from https://s3-us-west-2.amazonaws.com/acfepublic/2018-report-to-the-nations.pdf

Bănărescu, A. (2015). Detecting and preventing fraud with data analytics. *Procedia Economics and Finance*, 32, 1827–1836. https://doi.org/10.1016/S2212-5671(15)01485-9

Bribery of Public Officials and Witnesses, 18 U.S.C. §201 (1994).

Caulfield, T. (2014). The anatomy of procurement fraud. *Contract Management, 54*, 52–55. Retrieved from http://libproxy.nps.edu/login?url=https://search-proquest-com.libproxy.nps.edu/docview/1508191293?accountid=12702

National Defense Authorization Act for Fiscal Year 2018. (Conference report). H.R. 2810, 115th Cong. (2017). (Conference report). Retrieved from https://www.congress.gov/115/crpt/hrpt404/CRPT-115hrpt404.pdf

Conspiracy to Commit Offense or to Defraud United States, 18 U.S.C. §371 (1994).

Conspiracy to Defraud the Government with Respect to Claims, 18 U.S.C. §286 (1994).

Contractors' Bonds, Bids, and Public Records, 18 U.S.C. §494 (1994).

Dealing in Counterfeit Obligations or Securities, 18 U.S.C. §473 (2001).

Detecting procurement fraud. (2003). *CPA Consultant, 17*(1), 11. Retrieved from http://libproxy.nps.edu/login?url=https://search-proquest-com.libproxy.nps.edu/docview/194694470?accountid=12702

Dorminey, J. W., Fleming, A. S., Kranacher, M., & Riley, Richard A., Jr. (2010). Beyond the fraud triangle. *The CPA Journal, 80*(7), 17–23,3. Retrieved from http://libproxy.nps.edu/ login?url=https://search-proquestcom.libproxy.nps.edu/docview/637276283? accountid=12702

Eger III, R. J., Rendon, J. M., Rendon, R. G. (2014). *DON online fraud protection riding on an integrated business intelligence foundation* (Technical report). Monterey, CA: Naval Postgraduate School.

False Claims, 31 U.S.C. §3729 (2009).

False, Fictitious or Fraudulent Claims, 18 U.S.C. §287 (1986).

False Official Statements, 10 U.S.C. §907 Art. 107 (2016).

Fraud by Wire, Radio, or Television, 18 U.S.C. §1343 (2008).

Frauds and Swindles, 18 U.S.C. §1341 (2008).

Hill, N. C., & Ferguson, D. M. (1989). Electronic data interchange: a definition and perspective. In EDI Forum: *The Journal of Electronic Data Interchange* (Vol. 1, No. 1, pp. 5–12).

Karpoff, J. M., Lee, D. S., & Vendrzyk, V. P. (1999). Defense procurement fraud, penalties, and contractor influence. *Journal Of Political Economy*, *107*(4), 809.

Major Fraud Against the United States, 18 U.S.C. §1031 (2009).

Perols, J. L., Bowen, R. M., Zimmermann, C., & Samba, B. (2017). Finding needles in a haystack: Using data analytics to improve fraud prediction. *Accounting Review*, *92*(2), 221–245. https://doi.org/10..2308/accr-51562

Possession of False Papers to Defraud United States, 18 U.S.C. §1002 (1994).

Prohibitions on Disclosing and Obtaining Procurement Information, 41 U.S.C. §2602 (2011).

Public Contracts–Kickbacks, 41 U.S.C. §8701 (2011).

Public Money, Property, or Records, 18 U.S.C. §641 (2004).

Ramamoorti, S., & Curtis, S. (2003). Procurement fraud & data analytics. *The Journal of Government Financial Management, 52*(4), 16–24. Retrieved from http://libproxy.nps.edu/login?url=https://search-proquest-com.libproxy.nps.edu/docview/222375295?accountid=12702

Rendon, J. M., and Rendon, R. G. (2016). Procurement fraud in the U.S. Department of Defense: Implications for contracting processes and internal controls. *Managing Auditing Journal, 31*(7), pp. 748–767

Rendon, R. G., and Rendon, J. M. (2015). Auditability in public procurement: an analysis of internal controls and fraud vulnerability. *Int. J. Procurement Management,* Vol. 8, No. 6, pp. 710–730

StataCorp. (2017). *Stata Statistical Software: Release 15*. College Station, TX: StataCorp LLC.

Statements or Entries Generally, 18 U.S.C. §1001 (2006).

Under Secretary of Defense for Acquisition, Technology, and Logistics (USD(AT&L)) (2011). *Report to Congress on contracting fraud*. Retrieved from https://www.sanders.senate.gov/imo/media/doc/102011%20-%20DOD%20Fraud%20Report.pdf

Weisman, A. R. (1987). The prevention of fraud, waste and mismanagement in defense contracting. *The CPA Journal, 57*(6), 26. Retrieved from http://libproxy.nps.edu/login?url=https://search-proquest-com.libproxy.nps.edu/docview/212262342?accountid=12702

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California