EVALUATION OF THE CAN SPAM ACT: TESTING DETERRENCE AND OTHER
INFLUENCES OF EMAIL SPAMMER BEHAVIOR OVER TIME

By

ALEX C. KIGERL

A dissertation submitted in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY

WASHINGTON STATE UNIVERSITY
Department of Criminal Justice and Criminology

MAY 2014

© Copyright by ALEX C. KIGERL, 2014
All Rights Reserved
To the Faculty of Washington State University:

The members of the Committee appointed to examine the dissertation of ALEX KIGERL find it satisfactory and recommend that it be accepted.

__________________________________
Zachary Hamilton, Ph.D., Chair

__________________________________
Jacqueline van Wormer, Ph.D.

__________________________________
David Makin, Ph.D.

__________________________________
Warren Harrison, Ph.D.
Both email and email spam have experienced a growth that has paralleled the similar growth in technology worldwide. Email spam is more than just a nuisance, such mass unsolicited messages may also be harmful or fraudulent. The Controlling the Assault of Non-Solicited Pornography and Marketing Act (CAN SPAM) is federal legislation that was passed and enforced in the United States starting in January 1, 2004; created in response to the growing spam problem. The Act was intended to regulate the methods and content of spam that could be transmitted, requiring spammers to comply with a number of ethical standards when sending spam.

A series of reports and evaluations by cybersecurity firms and researchers followed in response to the passing of the Act to assess its efficacy, most of which were not positive about the Act’s success. However, none of these evaluations used methods that were sufficiently rigorous, failing to capture the continuous nature of CAN SPAM Act’s enforcement, ignoring a variety of possible spurious influences, and only considering a relative few number of measures of spamming behavior.
This research proposes to address all of these limitations by analyzing a sample of 5,490,905 spam emails received in the United States from March, 1998 to November, 2013. A time series dataset was built from the spam sample by software which processed each spam message to build 17 measures of spammer behavior, falling under the categories of spam volume, spam compliance with the CAN SPAM Act, spam severity (malware and fraud), and spam locality. Each measure was incorporated into a multiple time series design and regressed on 11 measures of CAN SPAM Act enforcement, public opinion, and attention, all while controlling for multiple economic, technological, and related time series measures. The study is informative as to the causes of spamming behavior on a multitude of different dimensions relevant to illicit email spam that has policy implications for both the CAN SPAM Act and other possible anti-cybercrime legislation.
# Table of Contents

Abstract ................................................................................................................................. iii

Table of Contents .................................................................................................................... v

Chapter 1 .................................................................................................................................. 1

Introduction .............................................................................................................................. 1

The Present Study Proposal ...................................................................................................... 5

Research Questions ................................................................................................................... 6

Chapter Overviews ................................................................................................................... 7

Chapter 2: Crime in Cyberspace: Cybercrime and Unlawful Internet Email Spam .................. 9

Why Send Spam ....................................................................................................................... 10

Spam’s Place in the Cybercrime Hierarchy ............................................................................. 15

Harmful Consequences of Spam ............................................................................................ 21

- Spam Scams .......................................................................................................................... 22
- Botnets ................................................................................................................................. 27

Defenses Against Spam .......................................................................................................... 29

Chapter 3: The CAN SPAM Act: Legal Recourses to Spam and their Effectiveness ............... 33

Regulations of Spam ................................................................................................................. 36

- Regulations of Commercial Email ....................................................................................... 37
- Aggravating Factors of Illegal Spam ................................................................................... 39
- Sexually Explicit Material .................................................................................................... 40

Penalties for CAN SPAM Violations ....................................................................................... 41

- Enforcement by States ......................................................................................................... 42
- Enforcement by Internet Access Providers ......................................................................... 42
- Enforcement by the Federal Trade Commission ................................................................. 44
- Rewards for Reporting Violations ....................................................................................... 45
- Updates to the CAN SPAM Act ............................................................................................ 45

Efficacy and Evaluation of the CAN SPAM Act .................................................................... 47

Chapter 4: Theories of Cybercrime: Understanding Spamming Behavior and the Theoretical Foundations

Underpinning Anti-Spam Efforts ............................................................................................. 54
Chapter 6: Results .................................................................................................................. 134

Descriptives .......................................................................................................................... 134

Interrater Reliability Testing ................................................................................................. 137

Scale Reliability Testing ....................................................................................................... 139

Multivariate Analyses .......................................................................................................... 140

  Time Series Regression of Spam Count per Month ............................................................. 140
  Time Series Regression of Percent US IPAs Per Month ...................................................... 143
  Time Series Regression of Percent US TLDs Per Month .................................................... 144
  Time Series Regression of Percent Spam Opt-Out Compliance Per Month ..................... 147
  Time Series Regression of Percent Spam Address Compliance Per Month ..................... 148
  Time Series Regression of Percent Spam Meaningful Subject Compliance Per Month .... 151
  Time Series Regression of Percent Spam Advertisement Notice Compliance Per Month ... 153
  Time Series Regression of Percent Spam Header “To” Found Per Month ......................... 154
  Time Series Regression of Percent Duplicate Email From and To Fields Per Month ......... 156
  Time Series Regression of Percent of Emails with Sender Name Found Per Month ........ 157
  Time Series Regression of Percent of Emails with Header “From” Found Per Month ........ 158
  Time Series Regression of Percent of Emails with Duplicate “To” and Return-Path Fields Per Month .................................................................................................................. 159
  Time Series Regression of Percent of Emails with Duplicate “To” and Reply-To Fields Per Month .................................................................................................................. 161
  Time Series Regression of Average Fraud Probability Per Month .................................... 162
  Time Series Regression of Percent Fraudulent Emails Per Month ................................... 164
  Time Series Regression of Percent of Spam with Malicious Links Per Month .................. 165
  Time Series Regression of Percent of Spam with Embedded Scripts Per Month .............. 167

Chapter 7: Discussion ............................................................................................................. 169

Implications of Findings ....................................................................................................... 169

  Spam Volume ...................................................................................................................... 169
  Spam Locality .................................................................................................................... 171
  Spam Compliance with the CAN SPAM Act ...................................................................... 172
  Spam Severity and Illegality .............................................................................................. 175

Limitations .............................................................................................................................. 177
Chapter 1

Introduction

Cybercrime has grown in parallel with the similar expansion of the internet and information technology. As technology improves, so too do offenders who utilize that technology. Electronic spam has similarly kept pace with the digital age, with higher volumes of spam sent that can be both fraudulent and dangerous (Gudkova, 2013). Spam is the sending of unsolicited bulk electronic messages to multiple recipients. Spam can take many forms where electronic communications can be sent (text messages, forums, chat rooms), but its most common form is that of email spam (Rao & Reiley, 2012), of which 72% of all emails sent worldwide are in fact spam and not legitimate (Gudkova, 2013).

Additionally, almost 70% of all internet traffic, not just email related traffic on the internet, is spam of any kind; the rest being other internet activity (Lachhwani & Ghose, 2012). As spam improves, so too do spam filters, which block a majority of spam sent. To compensate, spammers have to invest more time and effort into sending even more spam to maintain existing profits. This race between spammers and cybersecurity developers results in the same amount of spam reaching inboxes, but contributes to a higher demand on internet servers (Jonsson, 2009).

Of the spam that does reach user inboxes, the barrage of solicitations can be more than just a nuisance, they can cost time and resources to personally filter and delete. Spam can waste employee time that is estimated to cost $22 billion in lost productivity (Lachhwani & Ghose, 2012). Spam costs both time and money for internet service providers who have to carry the extra load on their networks, business in terms of employee time, and individuals who are also affected by both email and other spam.
Spam costs more than just employee productivity and traffic bearing costs, however. Most email spam is also fraudulent in some way, from simple falsification of the sender’s identity or origin (Sanchez, Duan, & Dong, 2010) to outright criminal fraud (Nhan, Kinkade, & Burns, 2009). Phishing scams alone, which attempt to deceive the recipient into divulging personal information, usually through email, often net about a thousand dollars from individual victims (Brody, Mulig, & Kimball, 2007). Phishing scams that successfully target businesses, however, tend to cost over ten times that amount, and phishing scams tend to target businesses for this reason. The financial losses to the United States from such phishing attacks on the whole are almost $53 billion a year. Fraudulent spam can also seek to deceive recipients into wiring money directly, rather than just personal information such as credit cards. Called advance fee fraud, these scams cost individuals substantially more, on average victims lose $34,000 from such schemes.

Most email spam is sent from botnets, with as much as 88% of the spam reaching user inboxes originating from such malware (Stone-Gross, Holz, Stringhini, & Vigna, 2011). Botnets are malware installations on victim PCs remotely controlled by the botnet operator. The operator can choose to send spam from the botnet installations themselves, and not the offender’s own computer. Many botnets are installed for just such a purpose, as sending spam tends to be one of the more common reasons to install a botnet. Spam does not just cost money and time, it can also cost personal computing hardware when hacked and compromised. It is estimated that as much as 14% of internet-ready PCs are infected with at least one botnet (Kindsight, 2012).

In light of these problems, the United States Congress enacted the Controlling the Assault of Non-Solicited Pornography and Marketing Act of 2003, which was intended to regulate commercial electronic marketing efforts such as spam (CAN SPAM Act of 2003). The Act went
into effect and enforcement in January 1, 2004. Spam was defined as any commercial electronic mail message, excluding electronic mail messages that were not commercial (such as political promotion or solicitation). The statute did not ban sending spam messages outright, but instead created requirements to follow for those intending to promote a commercial good or service via bulk electronic communications.

In order to send spam messages, senders must comply with a set number of regulations. Among them includes the requirement of accurate message header information, especially among email spam. This prohibits falsifying sender email address or name information, as well as recipient address information. Senders of spam do not have to require users opt-in to receive spam email, but do have to honor requests by recipients to discontinue sending spam (opt-out). The sender must also include a valid physical mailing address in the body of the email, as well as indicate unambiguously that the message is an advertisement. The subject line of an email must also be meaningful, such that a recipient can guess at the contents of the message without having to open it from reading the subject title itself. Lastly, any sexually explicit material must be labeled as such in the subject field, unless the recipient has expressed prior opt-in consent.

Violators of these regulations can be sued for each electronic message sent that does not comply with the CAN SPAM Act. Considering the high volume of messages required to be successful in the spam business, the amount of the damages that can be levied against the spammer can accrue in the millions of US dollars. In addition, prison sentences can be given for messages that are sexually explicit. The authorities enforcing these regulations predominantly include the Federal Trade Commission, the state Attorney General, or internet access providers (IAP). IAPs are distinguished from internet service providers (ISPs) in that they can include both ISPs but also the providers of any internet networking service, such as social networking or
email services. IAPs are the closest the CAN SPAM Act permits to allowing private entities bring suit against spammers.

Following the passing of the Act, the FTC submitted a report on the Act’s efficacy. The FTC concluded that legitimate email marketers were now in compliance with the Act, and that spam rates overall in the United States had flattened out and ceased their upward trend shortly after the Act’s passing (Majoras, Leary, Harbour, & Leibowitz, 2005). Other commentators were not so positive about the success of the new law. Private cybersecurity firms reported that spam had in fact increased following the Act (Gross, 2004; Zeller, 2005). However, another study, that conducted a time series impact assessment, found no difference in spam volume following the CAN SPAM Act, neither an increase nor a decrease (Kigerl, 2009).

While the FTC reported that legitimate marketers were in compliance, many reports concluded that the remaining illegal spammers paid no attention to the CAN SPAM regulations. Compliance was said to be extremely low, from six percent of spam emails sampled to be fully compliant to as little as three percent in compliance (Grimes, 2007; Gross, 2004; Yu, 2011). However, none of these studies utilized a pretest measure of compliance prior to the CAN SPAM Act. The reports simply quantified the percent of spam emails in compliance and considered the absolute effect sizes to be low, without a relative comparison prior to the Act’s intervention. Kigerl (2009), however, did perform such a comparison and found no significant impact of the CAN SPAM Act on compliance.

Despite the mixture of findings, all assessments to date of the CAN SPAM Act only consider the impact of the legislation to be dichotomous. That is, comparisons are made prior to the CAN SPAM Acts passing and on or after the Act went into enforcement on January 1, 2004. The measures of the CAN SPAM Act are equivalent to a single binary measure, false prior and
true after the Act. Such analyses do not capture the true variation and theoretical influence a policy or law is intended to have on the target human behavior. This study seeks to remedy this limitation by including 11 different continuous measures of CAN SPAM Act activity, enforcement, attention, and public attitudes. Other research has found such measures of law enforcement have a significant impact on other forms of cybercrime (Guitton, 2012; Png & Wang, 2007), so it is questioned whether they also have an influence on spam as the cybercrime outcome.

In addition to prior research having only limited measures of law enforcement, no research to date has considered variables in addition to the CAN SPAM Act. Specifically, only the CAN SPAM Act is analyzed when assessing spam time series within the United States, ignoring other possible influences on spammer behavior. This study incorporates control variables into all analyses, intending to capture additional influences of spam.

If the CAN SPAM Act does not have an influence, then what might cause observed variation? Multiple economic and technological variables have been found to relate to spam rates in cross sectional studies (Kigerl, 2012); it is questioned whether those same metrics also influence time series estimates of spammer behavior. It is also additionally important to control for alternate influences of spam because much of the growth in spam is anticipated to be spuriously related to the similar and parallel growth in information technology worldwide. It is important to account for these relationships when assessing if measures of law enforcement interventions are to be evaluated for their influence on spammer behavior.

**The Present Study Proposal**

The present study will consider the impact of the CAN SPAM Act on email spam in the United States. The CAN SPAM Act will be operationalized in terms of law enforcement, public
awareness, and public attitudes towards the Act. The outcomes the Act is to be tested against will include a number of measures of spamming behavior, including four spam categories of spam volume, compliance with CAN SPAM law, spam severity, and locality. Additional economic and technological control measures possibly predictive of spam will also be included in time series models of spamming behavior. The study period begins in March, 1998 and ends in March, 2013.

The present research will expand and improve upon existing literature evaluating the CAN SPAM Act in a number of ways. The measures of CAN SPAM activity are continuous, whereas prior literature evaluates the Act purely as a dichotomous impact, that being before and after the passing of the Act. The new measures will not only capture the continuous nature of the Act, but will be operationalized to capture different aspects of the Act as yet untested in the literature. The measures of spamming outcomes also include a greater number of differing predictors, many of which have never been tested against the CAN SPAM Act specifically. Finally, the control time series metrics included in all models will rule out possible spurious influences, something prior literature has not yet done.

**Research Questions**

Each of the measures of spam falls under one of four categories. Each category is intended to relate to one of four research questions this study offers. They include spam volume, compliance, severity, and locality over time. It is questioned whether the CAN SPAM Act or possible deterrence strategies have a measurable impact on any of these four categories of spam activity.
Question 1. Has the CAN SPAM Act impacted spam volume over time? A possible measure of success of the CAN SPAM Act would be whether enforcement or other activity has reduced the amount of spam sent.

Question 2. Has the CAN SPAM Act impacted CAN SPAM compliance over time? It is unnecessary for spam volume to change for the Act to have had a successful impact, as sending spam is legal, given it complies with regulations. It may be that spam compliance with spam law increases with higher levels of CAN SPAM enforcement or attribution of spamming individuals in the news.

Question 3. Has the CAN SPAM Act impacted spam severity over time? Spam severity above and beyond the scope of the CAN SPAM Act will be measured as both malware distribution and fraud and tested against CAN SPAM Act enforcement.

Question 4. Has the CAN SPAM Act impacted spam locality over time? The CAN SPAM Act only has limited jurisdiction, even over spam that is sent within or to the United States. It is questioned whether the CAN SPAMAct displaces the geographic origins of which spam messages may be associated.

Chapter Overviews

Chapter 2 discusses spam and provides a general overview of cybercrime overall and where spam crimes fit in. The chapter discusses the reality that most defenses against spam are technological, as opposed to legal, suggesting possibilities for future legal development of better spam fighting policies. Chapter 3 introduces the CAN SPAM Act, which is federal anti-spam legislation created in response to the spam problem in the United States. Limitations of existing evaluations of the CAN SPAM Act’s effectiveness are discussed, suggesting the need for the present study. Chapter 4 introduces theories of cybercrime and spam, including deterrence and
routine activity theory. The present study draws on new elements of both theories to compare against spam crime that is yet untested in the literature. Chapter 5 contains the study methodology. The chapter explains in detail the software written to process the sample of emails that will be used to build a dataset for analysis. Finally, chapter six concludes on the possible outcomes of this evaluation and suggests possible policy implications depending on each potential finding.
Chapter 2: Crime in Cyberspace: Cybercrime and Unlawful Internet Email Spam

Spam is the sending of mass unsolicited electronic mail messages to many recipients, usually with commercial motivations. Based on this definition, there are many categories of spam; text message spam, comment spam, search engine spam, social networking spam, and of course, email spam. Likely the most concerning form of spam would be this last type. Email spam is one of the most effective means to reach as large a number of recipients as possible. More people regularly use email than forums, private messaging, and social networking sites for their electronic communications, and sending tremendous amounts of text message spam can be more difficult to automate cheaply. Email is the most effective method of reaching potential buyers and/or victims. Because of the ease in which a single message can be propagated to thousands of recipients, spam has caught on since its emergence. It is an industry that is growing at an accelerated rate.

Spam rates have been increasing over time with the continued spread and reach of the internet. As more individuals become connected to and users of the internet, the profitability of spam as a business grows. Spam activity had experienced continued growth over the years since its inception (Lee, 2005), but has since leveled off and remained at a steady, if still high, rate (Gudkova, 2013). In 2003, just 45% of all email sent over the internet was considered spam (McCain, 2003), but today spam amounts to 72% of all emails sent (Gudkova, 2013). Spam is so high today such that 68.8% of all internet traffic, not just email related traffic, is spam (Lachhwani & Ghose, 2012). Spam filters have become more effective, and, as a result, block more spam from reaching your inbox. As a consequence, spammers have had to send even more spam on average to compensate, keeping the amount of spam users receive about the same, but increasing the traffic load on the internet in the process (Jonsson, 2009). Even if all spam was
caught by existing spam filters and related anti-spam technology, one can argue that, the energy expenditure alone from processing so much network activity is wasteful.

**Why Send Spam**

The purpose for sending spam is the same for perpetrating most forms of serious cybercrime: profit (Kanich, Weaver, McCoy, Halvorson, Kreibich, Levchenko, Paxson, Voelker, & Savage, 2011; Potdar, Ridzuan, Hayati, Talevski, Yaganeh, Firuzeh, & Sarencheh, 2010). Modern day spam has been profit motivated since it emerged in 1994, when two lawyers from Arizona sent messages to 20 million recipients offering them immigration services (Rutenburg, 2011). Their ISP terminated their service, but not after netting themselves over $100,000 from their scheme. A theoretical foundation that can explain cybercrime and spam related violations is routine activity theory. Three preconditions are put forth by the theory which must be met before a crime is committed, be that crime in cybercrime or elsewhere: (1) there must be the absence of a capable guardian which might guard against committing the crime, (2) there must be a target the offender wants something from (credit card numbers, bank accounts), and (3) the offender must be sufficiently motivated to commit the crime, for whatever reason. The third condition, motivated offender, is clear when considering that the majority of cybercrimes are profit motivated. Cybercrime can easily be turned into a successful business model, with the offender safely and comfortably making such gains from his/her own home (Kshetri, 2010).

Despite this possibility of easy money, it does take a large amount of spam to make a profit from it. In 2008, the average daily volume of spam sent worldwide was 120 billion messages per day (Kleiner, 2008). The tremendous amount of spam that has to be transmitted requires us to ask: why send so much spam, especially when everyday internet users are savvier to the schemes inherent in most unsolicited messages of this type? Spam as a business remains
as lucrative as ever, however, with most spam emails used for advertising purposes (Types of Spam, 2009).

Sending spam has to be profitable enough to maintain a growing industry of career spammers. Yet due to the necessary secrecy the profession requires, it is difficult to determine how much the average spammer makes from a full time spam business. Due to the anonymous nature of the internet, and the near meaninglessness of geographic and international borders in cyberspace, not many spammers are apprehended.

There exist some cases where spammers have been caught and their illegal incomes analyzed. Those spammers who are apprehended can shed light on the details of their operations. During 2005, Levon Gillespie, was served a court summons from lawyers at Microsoft, claiming that Gillespie had violated state and federal law by sending spam on Microsoft’s email networks. Gillespie was accused of providing bullet proof hosting to bulk advertising, by hosting their illegal websites on offshore servers to isolate them from local anti-spam laws.

Gillespie failed to attend his court hearing, and was fined 1.4 million dollars for his spam related crimes due to a default judgment. A follow up phone call to Gillespie by journalists revealed that he was not even aware of the court decision and that no one had contacted him about it. Gillespie responded that he would continue his spam business because of the substantial profits that he could continue to make. Gillespie reported that his spam operations earned him a six figure income.

The money that can be made by sending spam can justify the possible risks of being prosecuted, fined, or punished, especially if these possible risks are very low. Gillespie chose to continue his illegal spam business even after being caught, and was even willing to be open
about it to reporters. Another such instance involved a spammer who actually chose to
discontinue his illegal spamming endeavors in 2004, but not because of the risks that the spam
business entailed. Instead, the social stigma of earning a living by sending spam was the final
deterrent to change his mind (Zink, 2008). The former spammer felt embarrassed revealing his
work to others asking him about what he does for a living, so opted to quit the spam business and
write a book about it instead. In his writings he explained how readers can become involved in
the spamming business themselves.

The former spammer presented about his former profession at the Spam Symposium in
Europe during 2007. He claimed that an average of 40 million spam emails were sent a week,
and that recipients followed the links in a given spam message .12% of the time. Among those
click-throughs, only 200 actually purchased a product, meaning only one in 166,000 sent
messages resulted in a buyer. However, this was enough for 240 transactions per week,
amounting to $336,000 a year after factoring in expenses (Zink, 2008).

This was prior to 2004, however, when spam rates were not quite as high. There is a
higher volume of spam sent today. This could be due to the fact that there are more internet
users today as well, which means more individuals with the skills necessary to conduct spam
operations. It could also be that sending far more spam is necessary today as users of email are
more keen to the types of strategies employed by spammers. One buyer among over 100,000
advertisements sent is a tough crowd to sell a product. Yet that one out of thousands who
actually makes a purchase or falls victim to a ploy is sufficient incentive to continue flooding
inboxes with more junk spam. It is not difficult to educate a particular email user about the
nature of illegal spam, but unless every last person among the 2.4 billion internet users across the
globe (Internet World Stats, 2012) is able to understand the subterfuge of every message, sending spam will continue to be profitable.

More recent investigations into the interworkings of spam profitability can be done without actually acquiring the spammer’s personal effects (records, computers). Most spam is sent from botnets, which are multiple malware infected PCs controlled by a single owner (in this case, the spammer) (Schiller, Binkley, Harley, Evron, Bradley, Willems & Cross, 2007). These decentralized computers can all work together in tandem to send more spam than a single PC could alone. Considering the high bulk of email recipients necessary to turn a profit, sending spam from one’s own computer is simply not enough. Instead, spam is sent from thousands of unsuspecting victim’s own computer resources. Not only does this allow a greater volume of spam to be sent, but it also means that the spam sending hosts do not reside on the spammer’s own property; making tracing the spammer more difficult.

Because this spam is sent from botnets installed on potentially anyone’s own computer, such spam sending malware is subject to inspection and research. A botnet client can be deliberately baited and installed for research purposes, in which case all incoming and outgoing messages to and from that botnet client can be analyzed. The botnet client usually communicates with the command and control server, operated by an actual human being called a bot master. An important example where this type of research was carried out was of the Storm Botnet, a botnet which was predominantly used to send spam (Kanish, Kreibich, Levchenko, Enright, Voelker, Paxson & Savage, 2008).

Kanish and colleagues set up multiple honeynet servers, computers intended to bait malicious installations such as that of a botnet for purposes of analyzing them. Eight such servers were created to install botnet proxies from the Storm Worm malware variant. Proxy bots
are used as relays between spambots (which send the actual spam) and the master servers, which
the spammer uses to send commands to the spambots with. Because these proxy servers were
maintained by the researchers, all electronic communications of the bot master to the spambot
clients on the network could be intercepted and analyzed.

A total of 75,869 spambots were found to be connected to the proxy bot servers (Kanish et al., 2008). The proxy bots would relay email lists (to be spammed) and email templates (the
spam message) to the spambot clients. The researchers intercepted these spam templates and
replaced them with innocuous ones with links to a false pharmaceutical website hosted on a
server under the control of the researchers. This was done because the original spam messages
linked to websites not under the researchers control, in which case traffic to those websites could
not be analyzed.

By analyzing traffic sent to the researchers’ own false pharmacy website, the number of
spam recipients of the Storm Worm botnet who followed the included spammed links (response
rate) could be measured. Additionally, of those who clicked through to the website, the number
of visitors who attempted to make a purchase (conversion rate) could also be monitored. If the
user attempted to purchase the advertised goods, the recipient would be redirected to a 404 page
before they could enter their credit card information, for legal reasons. However, this event was
recorded by the researchers as a purchase.

After 26 days of this experiment, over 200 million separate email addresses were sent
pharmaceutical spam messages. Among those spammed, 350 million individual email messages
were sent. Out of the millions of email recipients, only 28 purchases were made. Each purchase
amounted to about $100, totaling about $2800 a month in sales. Yet the researchers were only
tracking about 1.5 percent of the botnet. Based on actual estimates of Storm Worms total network size, actual profits would be close to $7,000 a month, tax free (Kanish et al., 2008).

The conversion rate, or the number of buyers per advertisement, comes down to one purchase for every 12.5 million spam messages. Despite what sounds like a bad advertising strategy, spam continues to be profitable. Yet while the United States tends to receive more spam than other nations around the world, it tends to have the lowest response rate (Kanish et al., 2008). Americans may have a low response rate because of better spam filtering technology, or perhaps because of more familiarity with spam tactics. Regardless, spammers are still incentivized to infect as many victim PCs as possible to acquire them as part of their botnet. The spam business requires a large number of recruited bots to become successful.

Without such profitability, it is doubtful spam would continue to be sent out in the torrents of electronic messages we see today. Yet the business aspects of spam are only a tiny fraction of the entire cybercrime marketplace. Spammers often do not act alone, and will often become part of business partnerships with other sorts of cybercriminals that specialize in their own areas. Cybercrime is organized and specialized, and spam is highly interconnected to the rest of the cybercrime marketplace as a whole.

**Spam’s Place in the Cybercrime Hierarchy**

Spam infiltrates electronic communications of every conceivable type, and one reason is how easy sending spam has become and how profitable it can be. Spam is not the only type of cybercrime with this appeal, and the anonymity of conducting illicit acts hidden behind a screen in cyberspace may be additional cause for the prevalence of cybercrimes. Combine this anonymity with the limitless number of targets that can be reached over the vastness of the internet, and any motivated offender can locate any kind of suitable target he/she wishes. These
different types of environmental conditions that characterize the internet are consistent with routine activity theory, which suggests that when the three conditions of a motivated offender discovering a suitable target without a capable guardian, crime will follow.

The nature of cyberspace can make perpetrating crime easier, and criminals of all types can find their place. Given the anonymous nature and impersonal distances and cover that cyberspace provides, spam crimes have grown well adapted to making money by integrating with other cybercrime groups and even legitimate businesses. A motivated spammer may lack all the required skills necessary to run a profitable spam business, but still can team up with or hire other individuals in the cybercrime community that together can successfully implement their spam operations (Anderson, Bohme, Clayton & Moore, 2008).

Spam is a type of white collar crime, where the spammer is just one player amongst a larger superstructure of specialized roles and groups that together compose a very organized business market. Cybercrimes, including spam, are a specialized type of white collar crime combining what Friedrichs (2009) calls contrepreneurial crime and technocrime. Contrepreneurial crime is conducted by legitimate businesses that engage in certain illegal dealings or misconduct through that business. Technocrime is any crime that utilizes advanced sources of technology. Some legal business may contract out to spammers to promote their products or their business (Saltzman, 2009). While sending spam is typically a technocrime, as they invariantly rely on technology to mass produce spammed messages, spammers who sell their services to legitimate businesses become part of contrepreneurial crime.

The network of cybercriminals that exists today composes individuals of different skillsets and statuses that can be part of large cybercrime gangs, legitimate businesses, or decentralized individuals acting in their own self-interest. Spam is just one component of an
organized cybercrime network, which is why, in order to understand spam, understanding of cybercrime as a whole is important. Spammers often conduct business with other cybercriminals, whether to hire out their services, be hired themselves, or trade goods amongst other participants in the cybercrime community (Anderson, Bohme, Clayton & Moore, 2008). Cybercrime today is similar to legitimate private markets, with a division of labor, goods and services, and supply and demand. It is highly profit motivated, with 85% of malware written with the intention of profit in mind (Paul, 2006). Not only is cybercrime profitable, but there are also low risks for being caught and prosecuted (Kshetri, 2009), which makes cybercrime a more enticing investment compared to other forms of crime. The anonymity and technical sophistication of cybercrime makes tracking down the offender and prosecuting them difficult and expensive. These low risk investments coupled with the high return that can be made by targeting potential victims in the millions simply by automated software anonymously orchestrated over the internet makes cybercrime a compelling business model.

Cybercrime has a distinct division of labor (Anderson et al., 2008). For example, a scammer may operate a phishing scheme that utilizes a fraudulent website intended to trick a visitor into entering personal information and submitting it to the server. However, the phishing scammer would need to find a means to lure a victim to his/her fraudulent website, in which case they might contract a spammer to mass mail solicitations to potential victims to visit the site, masquerading as their trusted bank. The spammer may send the spam via a botnet that he/she has rented out from a botnet master, purchasing thousands of bots from which to send spam. The spammer may also use the botnet to send his/her own spam schemes, to attempt to recruit mules, which are people who launder money for the spammer, often unaware of the illegal nature of the “work” they do. The mule may accept bank payments from the credential information
misappropriated from the phisher, with the mule wiring most of the money to another country via wiring services to obscure its source. The mule keeps a small percentage of the misappropriated funds, but takes on a majority of the risk, as the financial paper trail leads to his/her own personal bank account (Anderson et al., 2008).

The cybercriminals may misappropriate so many stolen credential goods (credit cards, bank logins) that they are not able to use them all themselves. Instead, they may sell many or even most of them over the internet to dispose of them in the many online communities (forums, chat groups) available to them. Stolen credit card information that can be used to make online purchases can be sold for as little as fifty cents to ten dollars each, depending on which country the original card holder was from (Giles, 2009).

Stolen credit cards which include more information associated with them (social security number, etc.) may be sold for even more, which can include enough information to access a bank account. Additional data associated with a credit card can also include a digital copy of the magnetic strip or smart chip embedded in the card, which can be used to not only make online purchases, but to also clone the card and make ATM withdrawals (Giles, 2009). Not all participants of the cybercrime community professing to sell these kinds of goods are to be trusted. Customers can post feedback and ratings of each traders who sells credential goods, rating them higher for better service, speedy responsiveness and delivery, and of course for actually delivering what was advertised. Sellers can eventually be designated as verified or trusted amongst the community or by the website moderator, which is a good title to have for such a business. Potential buyers should be confident that they can trust a seller to deliver if they bear such a title (Giles, 2009).
The credential goods sold can be bought by those looking to buy a few goods over the internet free of charge, but can also be used by other cybercriminals to further their own operations. Paying for goods and services to fund your illegal cybercrime business with your own personal credit card might not be sound judgment. A far better alternative would be to purchase them with someone else’s credit card, which can be easily and cheaply bought. They can be used to safely purchase servers which can be leased out to spammers who can use them as command and control nodes for a botnet that can send spam (Brody, Mulig, & Kimball, 2007). It is all part of the same cycle.

Most of the cybercrime roles mentioned here require certain skills that the job requires (writing malware, managing a botnet). Yet technical skills are not always required to participate in the cybercrime business world. Services provided by other cybercriminals can make it easy for any individual with little technical know-how to be a part of the cybercriminal enterprise. These services include crimeware (software) that those with money to spend can purchase. Given someone has the money to pay for such services, anyone can become a spammer, deploy a phishing scheme, or manage a botnet (Wiedrick-Kozlowski & Stinchombe, 2008).

While this is beneficial to those without the technical skills to deploy such crimeware on their own, those who write the malware can be equally appreciative of such changing markets (Wiedrick-Kozlowski & Stinchombe, 2008). Cyber “criminals” who write viruses and worms can sell them to others instead of launch them on unsuspecting victims themselves, which may not even be illegal. While using such software to infect computers or computer networks is illegal, distributing these products as is to those willing to pay for them does not make the distributor liable for what their customers do.
While these services are more readily available today, purchasing them does not come cheap. A particular buyer can use their money to rent bullet proof website hosting in offshore countries to operate an illegal online business. Servers that are “bullet proof” are websites that are hosted in nations with fewer legal protections and even less law enforcement that might punish whatever illegal business are being conducted on the rented servers (Moore, Clayton, & Anderson, 2009). A now defunct cybercrime group that was infamously known for providing such bullet proof hosting was the Russian Business Network, which was in the business of offering a multitude of illegal services. Cybercrimes of varied types could be carried out on their rented server plans, such as phishing scams, child pornography, and malware distribution. While the Russian Business Network is no longer operation, other service providers have sprung up to take their place, as tends to be the case with temporary victories against cybercrime.

Those wishing to host a phishing website can expect to pay $300 a month, with an additional fee of $100 for setting up the hosting plan (Brody, Mulig, & Kimball, 2007). However, a phishing site is useless without traffic to direct to it. Spam sending software is required, which can also be installed on the server to direct victims to the site. The costs of such software can add an additional $1,200 a month to the hosting plan. A third monthly charge of $1,900 can be used to add a listserv of email addresses to spam, proxy servers, and other add-on services.

Cybercrime has become both profitable and low in risk (Kanich, Weaver, McCoy, Halvorson, Kreibich, Levchenko, Paxson, Voelker, & Savage, 2011; Paul, 2006), with the added bonus of being easy to carry out with little technical skills to bring to the table. Spam is no small part of the problems that cybercrime poses, for it is often used when the distribution of malware, advancing fraudulent scams, and soliciting others for various illegal products and purchases.
Without the sending of spam, the cybercrime world would be significantly diminished. Without spam, many of the dangers of the internet would no longer be present. The harms of spam are many, whereas some are harmless and benign, others can be outright dangerous. Spam is not given as much attention as other forms of illegal behavior, but it still poses some substantial dangers to internet users.

**Harmful Consequences of Spam**

Spam is generally considered a nuisance, as it floods inboxes with useless messages to be disposed of and wastes time sorting between the spam and the “ham” (non-spam messages). Distinguishing between the spam from the rest of one’s emails can often be laborious, where risks of deleting wanted email may be present when sifting through all the junk. Most recipients do not make the mistake of purchasing any of the advertised products spam messages seek to sell (Larkin, 2009); nor are most deceived by fraudulent spam. Therefore, the most common problem of spam is the wasted time filtering it all.

Email spam is the most prominent form of spam (Rao & Reiley, 2012). Web spam follows email spam as a close second. The wasted time that goes into removing unwanted email spam affects both businesses and individuals. Loss of human resources due to the nuisance of spam was estimated to be at $22 billion in 2004 (Lachhwani & Ghose, 2012). In additional to wasted human resource time spent sorting spam, spam also costs network traffic time as more than half of all traffic on the internet is spam. Spam is costly both in terms of time wasted and money lost for internet service providers, businesses, and individuals.

While this may sound trivial, the problems spam poses go further than this. While wasted time can be a nuisance, especially for employees attempting to sort emails which can compromise business productivity, risks can go beyond these concerns. Successful spam targets
the victim’s finances, personal information, and even the victim’s own computer; and even in some rare circumstances, victims have lost their lives (Smith, Holmes, & Kaufmann, 1999).

**Spam Scams**

Some spam genuinely attempts to sell a legitimate product, but much of it does not. There is often dishonesty in the products being sold via spam, and often there is no product as the entire message is outright fraud. A substantial proportion of spam emails contain some level of deception, with falsified headers (Sanchez, Duan, & Dong, 2010), misleading subject lines (MacFarlane, Harrington, Salsburg, & Goodman, 2003) or outright fraud (Nhan, Kinkade, & Burns, 2009). The strategy involved in each spam email tends to fall under one of four categories. The methods employed by spammers include: advertising a product or service the recipient can purchase (spamvertising), wire the spammer money as an advance fee for a business or similar offer (advance fee fraud), divulge personal information such as bank account credentials (phishing), or be recruited by the spammer to help launder money (by becoming a mule). Deception is almost always employed when perpetrating any one of these schemes.

For cases involving spamvertised products, one of four things could occur: (1) the item purchased arrives on time and in working condition, (2) the item is late to arrive or is of a quality not properly advertised when buying, (3) the product is never mailed and the spammer takes your money anyway, and (4) the product is never mailed and the spammer uses your credit card details to empty as much from your bank account as possible (Saltzman, 2009). Of these four possible outcomes, the most likely one is unclear. An experiment to determine as much would be costly, as most spamvertised purchases run for about $100 (Kanish, Kreibich, Levchenko, Enright, Voelker, Paxson & Savage, 2008), compounded by the added risk of a stolen credit card.
Yet the risks imposed by spamvertised emails are low in comparison to email solicitations that are wholly fraudulent, such as phishing and advance fee fraud. Fully fraudulent emails such as these may not be as common of regular advertisements, but they are substantially more harmful. In 2003, 8% of spam emails were considered to be in these categories of scams, while in 2004 the number dropped to 6% (Hulten, Penta, Seshadrinathan & Mishra, 2004). Today fraudulent emails are estimated to be at a low .5% (Wood, 2012).

One of the more common forms of email scams includes phishing, where the spammer attempts to masquerade as an entity the recipient knows or trusts in order to get personal information from them. The email may appear to be from the recipient’s bank or online retail account, or even an online gaming service (Brody, Mulig, & Kimball, 2007). Any website a user has registered an account with may be fair game for such targeting, as there are a multitude of ways to turn them into cash, whether by accessing your funds directly, impersonating the victim to withdraw funds, or using your account to send even more spam. Identifying these scams is sometimes easy, but other instances may be very convincing. A more deceptive form of phishing would be that of spear phishing, where the spammer targets someone or some business known very well be the perpetrator, such as what bank the recipient uses. The ploy of spear phishing is more effective, as it fabricates a more realistic story that the victim may not immediately identify as fraud (Aycock, 2007).

Businesses take on more substantial losses when falling victim to phishing attacks, which tend to be targeted by spear phishing attempts more often where employees of the business are sent phishing emails. The financial losses from phishing attacks in the United States totals almost $53 billion per year (Brody et al., 2007). Ninety percent of the victims of such attacks are businesses and financial institutions, with the remaining victims being individuals. The average
losses on a case-by-case basis for falling prey to a phishing scheme are $10,200 for businesses and $1,180 for individuals (Brody et al., 2007).

When the phishing scheme is successful, the stolen information can then be used open new credit card accounts, apply for loans, file and acquire someone else’s tax return, and sell the information to other cybercriminals (Brody et al., 2007), make online purchases, take out mortgages, and ruin the financial status of the victim’s lives as they have to sort out the damages (Stone & Levy, 2005). The most common personal information sought after are credit cards, often bought and sold in the online cybercrime markets (Giles, 2009). Yet building an entire website that can accept form submissions to defraud victims into giving up their credit card details is not the only way to orchestrate an email scam. Often it is easier to convince a target to wire the spammer money directly, saving them the trouble figuring out how to turn personal information into cash.

Called advance fee fraud, or 419 scams (named after the criminal code in Nigeria regulating fraud (Tive, 2006), fraudsters utilize more social engineering than technical skill. Advance fee fraud utilizes spam email messages that describe some sort of narrative story introducing the spammer as a potential trusted business partner or even possible romantic interest to the recipient. Some sort of deal is suggested, such as help transferring money or meeting the scammer in person; but before any such actions are to be taken, an advance fee is required of the recipient for whatever reason (filing fees, investments, plane tickets). By cooperating, the recipient expects to be rewarded somehow. Of course, the reward never comes, instead, after a first payment, a second and third payment will be requested, with new details added to the story to justify them. Someone who is willing to wire money once is also probably likely to do so a
second or third time. This will continue until as much money as possible is extracted from the victim.

The stories fabricated by the spammer are many, yet the goals are always the same. Some are contacted and told they have won the lottery (Dryud, 2005), or that a wealthy investor, widow, or orphan wants to move to and transfer all of his/her money into the recipient’s home country (Tanfa, 2009). To play along with these schemes, the victim is required to pay a multitude of up-front fees. The victim can be strung along for months or even years as they are made to pay more fees to play along with the story, eventually realizing that the deal is a scam. Once the victim realizes this and ceases contact with the fraudster, the spammer may send new emails pretending to be law enforcement officials looking to recover the victim’s stolen funds. However, there is a small fee that is required before they can begin their investigation (taxes, etc.) (Dryud, 2005). Of course, the stolen money is never returned, and the victim is repeatedly re-victimized.

The median losses to the victims of such scams is $34,000 (Durkin & Brinkman, 2009), according to an FBI report. However, in some reported instances, money is not the only thing a victim stands to lose. In the case of Nigerian scams (419 scams conducted by spammers in Nigeria), some scams require the victim to travel to Nigeria to complete a transaction. Once there, the offenders may hold them hostage and demand more money to be paid in order to be released. The victim is sometimes provided with a forged visa to encourage travel, which is leverage to use by threatening to reveal their illegal status to the authorities. Once discovering that they have been scammed, some victims travel to Nigeria of their own accord, attempting to recover their stolen funds (Ampratwum, 2009). Some have even been killed during these encounters; since 1992, 17 people have been murdered in Nigeria attempting to recover their
stolen money, with 25 murders or disappearances of Americans in total that have been linked to 419 scams. The US State Department has documented over 100 cases where US citizens have been rescued from Nigeria involving such scams (Ampratwum, 2009).

Not all of the victims of email fraud stand to lose money should they cooperate. In some cases, the victim is recruited to help launder money stolen from other victims. Called mules, these victims are offered a cut of the profits from such illegal activity, with the mule completely unaware that the nature of the business is illegal. The purpose of the mule is to accept and cash checks sent to their address, whereby they are then wired to the cybercriminals, with a small percent of the transaction going to the mule. Unknown to the victim, the checks mailed are part of other fraudulent schemes with the fraudster manipulates other victims into paying for nonexistent goods. The scammer uses the mules address for the checks to be mailed, resulting in the mule taking on most of the risk (Goodin, 2007).

As one example, the scammer might create a false Ebay posting offering to sell some commercial product (e.g. TV, laptop). The buyer who bids for the product is sent a private message by the offender that the payment can only be made by check, rather than through Ebay itself. The victim is provided with the mules address where the check is to be sent (Goodin, 2007). However, these arrangements may be short, as one of the many victims reports the mule to the police or confronts them at their home address directly. Eventually the bank overdrafts the stolen amount from the mule’s bank account, leaving him/her in debt. The scammer then cuts off all dealings with the former mule, seeking new partnerships elsewhere.

As much social engineering goes into spam, it would not be nearly as successful without the proper technology to mass produce spam email solicitations. While spam can be highly deceptive, most are keen to the fraudulent motives behind the spam filling their inbox. The
solution is to reach as large an audience as possible, as probabilistically one recipient will eventually make a purchase. The ability to mass produce high volumes of spam depends on more than just the offender’s own personal computing resources. Much of spam needs to rely on the parallel processing of hundreds, if not thousands, of compromised PCs working in tandem to mass mail millions of recipients. Botnets have leant much needed computing power to spammers, who otherwise might find it difficult to carry out their operations.

**Botnets**

A botnet is a network of infected computers managed remotely by one or more command and control (C&C) servers. The C&C servers are operated by the cybercriminal or bot herder/master, which are used to send commands to the multitude of infected machines, called zombies or botnet clients. One goal of the bot herder is to recruit as many infected machines into his/her own bot network as possible, which serve as valuable resources for perpetrating a myriad number of cybercrimes. A bot herder can own millions of bots simultaneously, which can be operated in parallel to achieve computational tasks not possible on one machine alone (Schiller, Binkly, Harley, Evron, Bradley, Willems & Cross, 2007). Some estimates suggest that botnets compose almost 27% of all malicious activity on the internet (Guerra, 2009).

Botnet can serve many purposes, such as distributed denial of service (DDoS) attacks, which can take websites offline by flooding them with requests, or the bots can serve as spyware by scanning the contents of the very computers they infect for sensitive information such as passwords and financial data. The bots themselves can also be rented out to other cybercriminals. Yet among its many uses, the botnets most common purpose is to send spam (Schiller et al., 2007). A majority of spam is actually sent from botnets, up to 88% of the spam sent and received (Stone-Gross, Holz, Stringhini, & Vigna, 2011). The sending of millions of
messages a day in order to make a profit would not be nearly as possible without the utilization of hundreds or thousands of PCs to lending computing power to the task; which is why a botnet is often necessary. Not only is the botnet a more efficient means to send bulk email, but the spam sending hosts are not even owned by the spammer him/herself, making the traceability back to the spammer much more difficult. The number of internet-connected computers that are infected with at least one botnet is estimated to be at 14% (Kindsight, 2012).

A botnet can aid the spammer in multiple ways. Botnets can automatically register multiple email accounts online from which to send spam, as bulk email often requires a bulk number of email accounts. The botnet client can also crawl internet web pages to scan for email addresses posted publicly online, of which to send more spam. The botnet can even act as spyware, and read the victim’s contacts books and spam their friends, family, and coworkers posing as the victim (Schiller et al., 2007). Hijacking a victim’s email account can make for more convincing fraud if the recipients believe they know the sender.

When a botnet has acquired a large enough list of emails to spam via either crawling the internet or from a victim’s contacts, the bot master can then use the list for one or more of three things. The spammer can send blank messages to those recipients to see if they bounce and to record whether they are legitimate emails or not. Second, if the spammer builds a large enough list that has verified addresses for real persons in it, the list can be sold to other spammers on the cybercrime market. Third, the list can be distributed to the spammer’s own botnet to send spam to those recipients. Often botnets are modularized, meaning that they can receive updates to install different features or “modules” to add functionality as the bot herder sees fit. For example, a botnet client may initially have been installed without any spam sending functionality, but the bot master can send commands for the bot to download and install such
functionality (Schiller et al., 2007). Many botnets start out with minimum functionality to facilitate faster installs, only updating themselves as needed.

Botnets are a substantial advantage to spammers. Without botnets, spam would be reduced drastically (Wood, 2012). Yet with as many as 14% of all household computers having a botnet infection present (Kindsight, 2012), significantly reducing such a threat has remained elusive. But anti-spam technology is available, and there is always an arms race between cybercriminals and cybersecurity professionals to improve their respective technologies. While spamming technologies are always improving, anti-spam solutions are also being developed and improved upon constantly.

**Defenses Against Spam**

The most common complaint from spam is the time it takes sorting it all. It is very easy for an adept spammer to send millions of messages automatically; but it is much more difficult for recipients to sort through it all to determine which is spam and which is ham (desirable email). The most common defense against spam is filtering, whereby each message is automatically classified as spam and prevented from reaching the recipient’s inbox in the first place. Anti-spam software can classify messages at a much faster rate than end users, freeing up valuable time. Technological solutions to the spam problem are by far the most prevalent.

Among some of the more predominant technical anti-spam solutions include authentication, such as authenticating the spam sending service, user, computer, or domain name from which the message originated (Bishop, 2005). Authorization services can also be employed, such as whether the computer or IPA used to transmit a message is authorized to do so using the domain name associated with it. A challenge-response can also be imposed, requiring a sender to verify they intended to send email to the intended recipient. Blacklists of
blocked servers and hosts can also be considered, as well as whitelists allowing for only a select number of hosts to send messages to a given recipient. Lastly, messages can be filtered, by either imposing keyword matches or general spam filtering that seeks to classify messages as spam (Bishop, 2005).

The most highly rated solution to the spam problem by businesses are spam filters (Siponen & Stucke, 2006). Spam filters determine the certainty a message is spam, and redirect messages highly likely to be spam to a spam folder or are otherwise prevented from reaching their destination. One of the most common filters are Bayesian filters, which statistically calculate the probability a message is spam. The calculations done are completed in multiple steps. For instance, first the probability that a message is spam is computed, regardless of the content of the email message. Following that, the probability that the message is spam is calculated based on each word used in the message, and also the probability that the same keyword appears in ham is computed, and so forth (Zdziarski, 2005). All such probabilities are based on prior machine learning training, whereby the spam filter is provided with a sample so spam already classified as such, and a sample of ham. The base rate probabilities are based on the contents of such training files. Messages with a high probability of being spam are either removed or marked as spam in some way.

Yet spam filters must be constantly built upon and improved, as spammers will and have adapted to this technology by crafting messages that can bypass the filters. Some spammers might attach an image instead of sending plaintext messages, with the spam message contained as text on the image (Chitu, 2007). A filter would need to be able to perform character recognition on the image file in order to analyze its keyword frequencies. Using images as a tactic to fool filters initially led to an increase of spam traffic by 334% in 2006 (Mosher, 2007).
Spammers will also randomly change the ordering or spelling of words in spam that the spam filters have not been trained on yet (Kestenbaum, 2006).

Another method of classifying messages as spam is by verifying the identifying information in the email. Called authentication, filters check header information in each message to determine if it has been tampered with or spoofed. For example, a third of all emails are authenticated with sender ID, in which the domain name of the sender’s address (e.g. gmail.com) is compared with the domain name servers (DNS is used to associate domain names with IP addresses) for that email domain (Lemos, 2006). If that DNS is not paired with the actual IP address of the domain name used in an email, the from field is not genuine. Another method involves compared the sender’s return address with a whitelist of trusted domain names registered with the spam service. Messages with return addresses not appearing on the trusted list are rejected (Wong & Schlitt, 2006).

Perhaps one of the most effective and latest technological defenses against spam would be collaborative spam filters. Collaborative spam filters are trained to detect spam based on what every day email users identify as spam. Google implements just this functionality in its own email service (Gmail). If a given message is marked multiple times by varying Gmail users as spam, then the probability that the Gmail servers will block such a message will correspondingly be higher (Parathaneni, 2011). A single email user cannot be expected to mark all spam in order for it to be blacklisted, but collectively millions of users classifying messages in such a way can much more easily expect to do this.

Technological solutions to the spam problem prevent messages from reaching their intended recipients. That is, the spam messages or cyberattacks are what is stopped. However, these technologies do nothing to prevent the spammers themselves from continuing to violate
spam laws in the first place. The predominant focus in the fight against spam has been the further development of these spam-fighting technologies. Much less attention has been spent improving legal and law enforcement efforts to crack down on spam offenders. While anti-spam technology can stop spam, anti-spam laws are needed to stop the spammer. Legal measures are less mature in their development than technological solutions. The effectiveness of such measures has yet to be fully tested and improved upon.

The CAN SPAM Act is one such recent legal measure created in response to cybercrime. Such laws are not as rigorously tested and improved upon as that of technological defenses against spam. Further evaluation of the CAN SPAM Act is desirable, so that stakeholders can pick apart those elements of spam and the CAN SPAM Act that warrant further development and improvement. As the law tends to lag behind technology, further development of existing laws is necessary to keep them up to date.
Chapter 3: The CAN SPAM Act: Legal Recourses to Spam and their Effectiveness

Traditionally a more sizable portion of the efforts to combat spam and cybercrime have focused on technological solutions to the cyberattack itself. Legal developments intended to reduce cybercrime victimizations have been slower, in part because of the nature of the law and additionally because of the complications of jurisdiction that cybercrime creates. Cybercriminals can conduct cyberattacks that breach international borders and are not constrained to any given geographical location. This multiplicity of crimes in cyberspace, that a cyberattack can be propagated exponentially over the internet to reach millions of victims, has made legal recourses difficult, but it also can mean that when legal actions are in fact successful, crime rates can be significantly, if temporarily, impacted.

Specifically, a disproportionate amount of spam is accounted for by a minority of spammers. This phenomenon has been observed elsewhere and is termed the Pareto principle, which suggests that for many events, approximately 80% of the effects originate from just 20% of the causes (Koch, 2011). The rule has been observed to apply to traditional street crimes, with a disproportionate number of the arrests originating from a smaller subset of offenders (Visher, 1987). However, the nature of crimes committed in cyberspace have magnified this effect several times over.

As much as 80% of the spam received by internet users in North America and Europe can be traced to almost 100 spam gangs (ROKSO, 2013). Each spam gang or spam operation is estimated to be composed of between one to five spammers totaling no more than 400 individuals. In addition to the spammers themselves, the Pareto principle applies to other elements of spam. Namely, a disproportionate amount of spam is sent from a minority of ISPs which permit spamming on their networks in exchanges for higher service fees (Spamhaus,
For those purchasing the goods advertised in spam email solicitations, just a handful of banks process about 95% of the purchases for such products (Levchenko, Pitsillidis, Chachra, Enright, Félegyházi, Grier, Halvorson, Kanich, Kreibich, Liu, McCoy, Weaver, Paxson, Voelker, & Savage, 2011). Lastly, about half of all affiliate programs which promote their products via spam have their websites hosted on just eight percent of the domain registrars (Levchenko et al., 2011). The idea this highlights is that legal recourses to the spam problem might have potential to make a significant dent in global spam rates were the law to successfully target this small minority of offenders, at least temporarily.

The laws against spam that are of interest to this research are those included under the United States Controlling the Assault of Non-Solicited Pornography and Marketing Act (CAN SPAM Act) of 2003. The legislation was passed by Congress in 2003, going into effect on January 1, 2004. The regulations inherent in the bill set requirements that electronic commercial messages must adhere to when sending advertisements to recipients electronically (including email and other electronic means of communication). The bill does not prohibit unsolicited commercial emails, but rather it regulates the way they are sent and the content that is delivered. The messages must be truthful and not fraudulent. The sender must also comply with a recipients express request to opt-out of all future emails. Violators of the CAN SPAM Act are usually fined, but can also receive prison time for additional aggravating violations perpetrated when sending spam.

During the formulation of the bill that led to the CAN SPAM Act, the United States Congress expressed an understanding that the growing expansiveness of the internet across the globe was becoming increasingly more ubiquitous in our everyday lives and requiring greater legal considerations. Deliberations in Congress led to some conclusions about the importance of
the internet and electronic commerce. The CAN SPAM Act states that the right to freedom of speech and expression is an important facet of the internet; a freedom which has allowed substantial progress and growth. Unsolicited commercial messages, or “spam,” therefore was considered an integral part of the free markets and the continuation of one’s business. Hence why spam is not criminalized outright by the Act.

However, Congress also found that in order for electronic commerce to continue being a strong market force, both ISPs and internet users must be free of the inefficient burden that spam can impose. Spam can be a burden on both the intended recipients and the ISPs that transmit internet communications, including emails. The high volume of spam can congest internet traffic and add an increased load that ISPs have to deal with when routing information over the internet. For the end user, a high bulk of unwanted mail can be costly in terms of sorting through it all to find the mail the user actually wants to read.

In addition to the costs of sorting and transmitting the messages themselves, the actual content and method in which they are sent come with their own problems. The messages contained in spam may be obscene or unsavory, and so email users may wish to block them from reaching their inbox or their children. Spam email rarely provides a means for a recipient to request the sender discontinue mailing them such messages. Many senders falsify email header information, such as the sending or return address. Much of spam is also sent from malware that has been illegally installed without authorization on someone else’s machine.

When considering these issues, Congress decided that there was a substantial government interest in regulating commercial electronic mail. Such a substantial government interest is required in order to impose restrictions on commercial advertisements; as such advertisements are considered free speech and therefore protected under the first amendment (Arora, 2005).
During the drafting of the Act, it was decided that bulk commercial mailers were not permitted to send any deceptive content in their messages and had to honor any requests by the recipients to remove them from any future mailing lists (CAN SPAM Act of 2003). Spam was argued to impose substantial costs to ISPs, consumers, and businesses by congesting network traffic and slowing internet services. Congress cited that spam costs subscribers globally $9.4 billion each year, and fighting spam adds an average additional cost of $2 per month to an individual’s internet bill (Senate report 108-102, 2003). The CAN SPAM Act was created to address these and other concerns.

Prior to the CAN SPAM Act, laws regulating electronic mail were created at the state level. Spam sent or received in one state fell under the jurisdiction of that state. The CAN SPAM Act preempts the majority of state laws that address spam. Of the laws left to the states to regulate include any fraudulent content of electronic mail. State laws that restrict falsified headers or deceptive contents of email are still remain under that state’s jurisdiction, should those laws exist locally. Lastly, the CAN SPAM Act does not supersede state laws on computer crime in general (CAN SPAM Act of 2003). While the CAN SPAM Act replaces most state laws regulating spam, the CAN SPAM Act is superseded by all Federal laws relating to obscenity and the sexual exploitation of children.

**Regulations of Spam**

There are three tiers of regulations set forth in the CAN SPAM Act, with separate gradations of punishments for their violation. First, there are general requirements each message must conform to that relate to the contents encoded in such messages. Second, there are additional aggravating factors that enhance existing punishments for email content violations. The aggravating offenses relate to how each message is sent, not the contents of the message.
Third, there are additional stacked penalties for sending pornographic content that does not meet CAN SPAM Act requirements.

The regulations only apply to commercial electronic mail messages. Commercial electronic mail messages are defined as any electronic message whose primary purpose is to promote a product or service. This definition excludes messages which are considered transactional or relationship messages. Such messages may be transmitted to finalize the completion of a commercial transactions (such as emailing a receipt confirming an online purchase of a product), or other messages sent from a commercial institution that facilitates a prior relationship or agreement with that commercial institution or sender. Messages that contain both advertisements and transactional or relationship content are considered primarily commercial (and therefore subject to CAN SPAM Act regulation) if the subject line or the majority of the message’s content appears to be commercial in nature (Ervin, Loeffler, Drye, LLP, 2010). Additionally, unsolicited bulk emails that do not promote a service or product, such as political or religious emails, are also free of any kind of CAN SPAM regulation. Only commercial messages must adhere to the provisions of the CAN SPAM Act (CAN SPAM Act of 2003).

Regulations of Commercial Email

One of the first and basic rules set forth in the Act forbids the falsification of email headers. The headers of an email message include the recipient’s address, the sender’s address, the return or bounce address, and additional routing details contained in the headers. Spammers will often fabricate false header details, such as an email address or domain name, or anything that might obscure their actual location or otherwise manipulate the recipient’s trust or perceptions. Any such tampering with a message’s headers is prohibited.
Another header field in an email subject to restriction is the subject header. Senders are not permitted to write header titles intended to mislead the recipient on what the contents of the message body are before opening the email. Subject headings must relate to the contents of the email. The email subject must also indicate that the message is an advertisement, as the subject is often the first part of the message email users read.

The sender must also include a genuine return address in the email, one that the sender checks and can receive responses from (unless they make a notice that the email is unmonitored). The recipient must be able to contact the sender through the return address, and the sender must be able to respond. Additionally, emails sent from an address that is not owned by the spammer (from a hijacked email account or botnet), or emails sent from accounts that there never was intention to check (such as hundreds of automatically registered email accounts), are prohibited.

The sender must provide a channel for the recipient to opt-out of further advertisements. Opt-out is the ability of the recipient to make a request to discontinue receiving spam, and the willingness of the spammer to honor such opt-out requests. The spammer is not required by law to get opt-in from the recipient. Specifically, the spammer does not need the recipient’s express permission beforehand to send them their first message. While opt-in is not necessary, opt-out choices must be made available. The means for providing opt-out choices are not discussed in the 2003 version of the CAN SPAM Act, just so long as they are included. A user may be able to follow a link where they submit their opt-out request in an online form. The opt-out may also be as simple as replying to the email making their request.

The sender’s valid physical postal address must also be included somewhere in the body of the email, or at least an address where the sender may be contacted. The purpose of this requirement is to reduce some of the anonymity that can incentivize misbehavior on the internet.
Lastly, the message contents or subject heading must identify itself as an advertisement. Providing this notice can help the intended recipient decide if they want to read or continue reading the email when sorting mail. The notice does not have to be made if the recipient opted in to receiving mail, however.

*Aggravating Factors of Illegal Spam*

In addition to the basic requirements spam email must comply with, aggravating offenses exist that can triple the maximum penalties under the CAN SPAM Act. The first aggravating offense include acquiring an email list by unethical means, such as address harvesting or dictionary attacks. The purpose of these techniques are to compile as large a list of intended recipients as possible of which to send spam, regardless of whether each email is real or not. Quantity is always preferred over quality in the case of spam, so spammers seek larger email lists over accurate email lists.

Address harvesting involves mining random internet web pages by automated software to match and pull any publicly posted email addresses. The harvesting is conducted by a bot that scans a web pages text and even HTML content that specifies how to render text, looking for anything that resembles an email address. While scanning the page, the bot also stores a list of all links on that page, of which it will follow and scan next after processing the current page. If those linked pages have more links, then they will likely be scanned as well, in an attempt to find as much information as possible. When the bot finds a string of text that appears to be an email address, it is saved to a list to later be spammed.

An alternative to email harvesting is the dictionary attack, which is much simpler to carry out, but also will generate a spam list with more erroneous email addresses on it. The dictionary attack involves automatically generating email addresses from a digital dictionary and affixing
them on a common email domain name (@gmail.com, @yahoo.com) such that the dictionary word becomes the username of the email (word@gmail.com). Considering the hundreds of thousands of words in the English dictionary, and the multitude of email domain names available, many of these generated email addresses are going to belong to actual people.

Another aggravating offense is the automated registration of multiple email accounts. Instead of harvesting emails to send spam to, an automated bot harvests email accounts of which to send spam from. The bot will register for multiple free online email accounts that will be used to send spam from. The purpose of this method is to send as many spammed messages as possible, as sending millions of emails from one email account will be denied by the mail servers. Instead the process is executed over a multitude of accounts and email service providers, making it less clear to any one single email provider whether a user is sending spam.

The third and final aggravating offense is the use of a protected computer or computer network to relay spam through without authorization. The spam can either be relayed through an protected computer that the spammer has illegally installed a botnet on, or an open relay service intended to reroute email message after stripping originating header information from. The term protected computer used here is from Title 18, Section 1030 of the Computer Fraud and Abuse Act. A protected computer is any computer owned by the government, a financial institution, an international or interstate commerce institution, or a computer used in telecommunications.

Sexually Explicit Material

The majority of the violations discussed so far only have financial penalties associated with them. However, illegal spam that contains sexually explicit material warrants added penalties of a prison sentence. Any spam message that contains sexually explicit material (images, text, or links to such) must include a warning label in the subject indicating the nature
of the content, so that users can decide whether to open and read the message beforehand. The warning is not required if the user had previously opted in to receiving such content. Violations of this requirement can result in a prison sentence of up to five years.

**Penalties for CAN SPAM Violations**

There are three major authorities that are authorized to pursue offenders who commit CAN SPAM Act violations. They include (1) the attorney general for violations within a state, (2) internet access providers, and (3) the Federal Trade Commission (FTC). The Federal Trade commission is the primary enforcer of the Act, however, and is authorized to further update CAN SPAM Act regulations in light of new emerging technology that warrants similar regulatory changes to keep pacing. Additionally, the FTC has the authority to bring suit for any violation detailed in the Act, whereas the remaining two authorities (ISPs and states) may only enforce a subset of the Act’s provisions. Rewards are also offered for any individual who provides information that helps lead to the conviction of a CAN SPAM Act offender to one of these authorities.

Other institutions may also enforce CAN SPAM violations if the violations fall under their jurisdiction laid out by other federal codes. The Department of Justice may bring criminal charges against a spammer who violates CAN SPAM regulations in the furtherance of a federal crime, such as fraud, obscenity, sexual exploitation of children, and racketeering. Other agencies, such as national banks, the Federal Communications Commission (FCC), and the Securities and Exchange Commission (SEC) may sue when CAN SPAM violations involve areas of their jurisdiction (such as fraudulent stock market spam in the case of the SEC, or spam sent to a wireless device in the case of the FCC). However, the provisions of the Act primarily concern the original three authorities, the FTC, states, and ISPs.
**Enforcement by States**

An attorney general can bring a civil suit against a spammer on behalf of the victims residing in that state. The state can only bring suit when the spam messages transmitted to state residents fail to label sexually explicit material contained in the message, do not include a return address for recipients to respond to, disregard opt-out requests or fail to include an opt-out mechanism, or do not include a notice that the message is an advertisement. The state attorney general is not authorized to bring suit if the message violates other CAN SPAM provisions.

The state may fine the spammer in the amount of the number of recipients victimized, or the number of spam messages sent in total, multiplied by $250. The fine this amounts to may not exceed two million dollars, unless the offender falsified any header information when committing the offense. If headers were falsified in such a way, the fine that can be imposed is unlimited. All fines determined can be tripled if the offender willfully and knowingly transmitted the emails. This is opposed to someone who contracted the spammer’s services without being fully aware of the illegal nature and methods employed by the spammer. The fines can also be tripled if the offender is guilty of any of the aggravating violations contained in the CAN SPAM Act.

**Enforcement by Internet Access Providers**

Internet *service* providers (ISP) are generally considered to fill the category of internet *access* providers. However, as it is defined in the CAN SPAM Act, ISPs are not the only private entity permitted to file suit for CAN SPAM violations. CAN SPAM indicates that an internet access service or internet service provider has the same meaning as that defined in 47 U.S.C. Sec 231 (CAN SPAM Act of 2003). Namely that an internet access service provides a service that enables users to access content, information, or other services provided over the internet.
(Restrictions of access by minors, 2013). The definition does not include telecommunications services.

The definition is important because its literal interpretation allows other private entities in addition to ISPs to file suit against CAN SPAM violators. Initial conceptualizations of internet access providers foresaw ISPs and email service providers enforcing the Act, but a number of recent court cases have found the broad language of the Act to allow any private website owner where electronic messages can be transmitted or posted to be able to file suit (Reid, 2010). Namely, social networking sites, such as MySpace and Facebook, have successfully brought CAN SPAM charges against users abusing their private messaging services. While private messages are regulated, status updates are considered public, and not intended for a specific electronic mail address, and are therefore not regulated (Thomas, 2010). The ninth circuit court has issued a final ruling that this definitional inclusiveness ends with questionable internet access service providers, such as those registering websites whose servers they don’t maintain themselves to prevent frivolous lawsuits (Reid, 2010).

When legally permitted to enforce CAN SPAM provisions, internet access providers (IAP) are to file suit in a district court that has jurisdiction over the defendant. IAPs may only sue for a subset of the violations that the FTC can. Specifically, they can only bring suit if the defendant falsified message headers, applied misleading subject headings to the message, failed to include a warning label on sexually explicit content, or disregarded opt-out requirements. IAPs can sue for the damages the defendant caused by congesting traffic on their networks via bulk email or malware distribution or otherwise misusing the services that the IAPs provide. If the defendant falsified header information during the commission of the crime, IAPs can opt not to sue for damages and instead levy a fine of $100 per email sent. If the defendant did not falsify
header information, a fine of up to $25 can still be charged for any other violation. The total amount of the fine may not exceed one million dollars, unless the false headers rule is violated. As is the same with the attorney general, the fines can be tripled if any aggravations were committed, or if the defendant knowingly and willingly committed the offense(s).

**Enforcement by the Federal Trade Commission**

The Federal Trade Commission (FTC) can enforce any penalties for violations under the Federal Trade Commission Act as if they were contained in the CAN SPAM Act. The Act directs the FTC to treat a violation of its statutory provisions as if such a violation were an unfair or deceptive business practice prohibited in the Federal Trade Commission Act (FTC Act, 2007). The FTC holds jurisdiction over any crimes interfering with or affecting interstate commerce, which includes spam. The legal actions the FTC can bring are described in Title 18, Section 1037 of the United States Code (Fraud and Related Activity, 2006). Penalties can include a maximum of up to five years in prison in addition to a fine if the defendant is a repeat offender, or the offense is committed in the furtherance of a felony. The prison term is included, but cannot exceed three years, if the defendant only used unauthorized access to a compute, automatically registered multiple email accounts to send spam from, or accumulated $5,000 or more in damages. If the defendant is not found guilty of any of the aggravating offenses above, then a prison sentence cannot exceed one year.

There is one additional law in the CAN SPAM Act that only the FTC may enforce, which is the use of affiliate spam. Affiliate spam is the hiring of another to send spam on one’s behalf, such as a business hiring a spammer to promote their products. Affiliation with the spammer is only criminalized if the defendant knew the affiliate sent spam, received a monetary benefit from the illegal spam, and took no action to prevent or sever ties to these illegal activities.
Rewards for Reporting Violations

While only ISPs, the FTC, and state attorney generals may prosecute spamming offenders, individuals may still participate by reporting spamming suspects or evidence to the Federal Trade Commission. A cash reward is offered for information leading to the successful prosecution and conviction of an offender under the CAN SPAM Act. The citizen who knows of a spammer’s illicit actions must notify the FTC, and the information provided must lead to the successful collection of a civil penalty by the FTC. Given these conditions are met, the informant is entitled to 20% of the civil penalties collected from the offender.

Updates to the CAN SPAM Act

The CAN SPAM Act became effective on January 1, 2004. However, the Act authorizes the Federal Trade Commission (FTC) to issue updates to the regulations of CAN SPAM to better implement its provisions (CAN SPAM Act of 2003). In 2008, the FTC published updates and clarifications to the rules of the Act in the Federal Register that became effective on July 7, 2008 (Definitions and implementation under the CAN-SPAM Act, 2008). Prior to the FTC’s final ruling, the FTC invited comments on the Act from email marketers and their associations, non-profits, email recipients and other interested parties.

Among the updates were three definitions contained in the Act. Specifically, the term “person,” “sender,” and “valid physical postal” address were further expanded and clarified. The language of the Act regulates how a “person” may send commercial electronic mail, but some confusion existed as to who or what would be the “person” responsible for transmitting such a message. The FTC further clarified the definition of person to broadly include an individual, group, unincorporated association, corporation, or other business entity, not limited to natural persons.
Confusion by email marketing stakeholders also existed regarding the CAN SPAM definition of “sender,” which was originally defined as a person who initiates an email whose product or service is promoted by the commercial message. Often a single electronic message will contain advertisements promoted by multiple marketers or business entities, so it was uncertain if one or all of the senders would be liable for any violations the message contains. The final ruling by the FTC requires a single marketer among the group to be specifically designated as the single sender of the message, which must be detailed in each message to determine responsibility for its contents.

Lastly, the remaining definitional issue regarding the Act related to what a valid physical mailing address includes. Lawful electronic messages are required to provide the businesses’ or marketers’ valid mailing address. It was not certain whether a PO box address was considered valid or not. The FTC ruled that such addresses would be considered valid so long as they were registered with the United States Postal Service. Some commenters objected to registering only within the United States. However, the FTC insisted that United States PO Box registration regulations make it easier for authorities to track registrants, hence why the requirement was important.

A final regulatory update was enacted concerning the provision of opt-out procedures for recipients to discontinue receiving promotional messages from a given sender. As it was initially passed, the CAN SPAM Act did not regulate the methods for which users can choose to opt-out of a specific spam list. Such a loophole led to exploitations where email recipients were compelled to pay a fee in order to successfully opt-out, or were required to navigate a complex menu or internet website links that exposed the recipient to more promotional advertisements before successfully finding the final opt-out page. FTC filled in this loophole by requiring opt-
out mechanisms to involve no more effort than visiting a single web page or responding to the spam message they received, after which the recipient must receive no further commercial mail after ten business days of opting out.

**Efficacy and Evaluation of the CAN SPAM Act**

The original formulation of the CAN SPAM Act contained plans by the Federal Trade Commission to conduct an evaluation of the Act’s effectiveness following its implementation. In addition to evaluating the Act’s impact on spam in the United States, the FTC was also required to conduct an investigation on the merits of a do-not-email registry, similar to a do-not-call registry. The do-not-mail registry was determined to be inadvisable, because the registry of those opting-out of unsolicited email would be a publicly available list of emails. The email list would likely be sought after by spammers who do not care about the legality of their methods (Muris, Thompson, Swindle, Leary & Horbour, 2004). The list would likely increase spam, in fact.

By December, 2005, the FTC had completed its report to Congress evaluating the CAN SPAM Act. The report contained both good and bad news on the FTC’s findings. The FTC found some aspects of the CAN SPAM Act to be successful, which included the result that there had been over 50 prosecutions of spammers at the time the report was written. The FTC concluded that many legitimate online marketers were now in compliance with the terms set out in the CAN SPAM Act (Majoras, Leary, Harbour & Leibowitz, 2005). While the Act likely compelled compliance among legitimate businesses, full-time spammers and other cybercriminals are not likely to be deterred by the Act’s enforcement.

The FTC also evaluated spam activity itself, and concluded that there were some changes. The FTC determined that there had been a decrease in sexually explicit content
contained in spam messages overall. The FTC also claimed that the rate of spam sent had begun to flatten out, slowing in its noticeable trend upwards over time. It was also acknowledged that the amount of spam received in inboxes had been lessened from better spam filtering technology (Majoras et al., 2005).

The FTC also acknowledged some limitations of their findings. The FTC admitted that there were more malware attachments in spam email, and that there had been no perceptible decrease in the amount of falsified information when registering for domain names (Majoras et al., 2005). Registering for a domain name with false personal registration details allows for anonymity of the registrant. It should be noted that nothing in the FTC’s report indicates the statistical methods used or statistical significance.

The FTC was not the only authority to evaluate the effectiveness of the CAN SPAM Act. Other independent researchers who also tested the impact the Act had mostly consisted of computer security firms and spam filtering technology companies. Two predominant questions of interest to researchers were whether spam rates had been impacted and whether compliance with the conditions of the Act increased.

With regards to spam rates, it appeared that the volume of spam sent had in fact gone up following the passing of the Act. According to Scott Chasin, Chief Technology Officer of the spam and malware filtering company MX Logic, spam had increased (Gross, 2004). According to MessageLabs, another anti-spam and cybersecurity firm, spam had grown by 50 to 80% one year following the passing of the Act (Zeller, 2005).

Compliance with CAN SPAM Act regulations was also considered when evaluating the Act. According to MXLogic, more than 99% of spam was not fully in compliance with one or more regulations set out in the Act (Gross, 2004), based on a random sample of 1,000 spam
messages. Commtouch Software, another spam filtering company, wrote software which analyzed millions of spam messages to measure compliance, including whether emails contain a return address or have meaningful subject titles. The results indicated that less than 1% of the messages were in compliance. A third spam filtering vendor, Audiotrieve, analyzed 1,000 spam messages and determined that only 10% were in compliance (Gross, 2004). However, none of these reports capture spam compliance levels before the passing of the CAN SPAM Act. There is no baseline from which to draw conclusions. Compliance may sound low, but it may in fact have been higher following the Act’s passing. Additionally, the follow up times of these reports do not extend beyond 2004, shortly after the Act began to be enforced.

Grimes (2007) also conducted an analysis of CAN SPAM compliance on spam emails, with a slightly longer follow up period. Repeated measures were captured, once at six months following CAN SPAM’s passing, and a second two years afterward. Five honeypot email accounts were registered, which are email accounts used to bait spammers by posting the addresses publicly online. After six months following the Act’s passing and a sample of spam messages had been collected by the honeypot, 1,100 spam emails were randomly selected. A second sample followed at the two-year mark, selecting 800 additional spam emails at random.

At the first follow up time of six months, only a little more than 14% of spam was found to be compliant. During the second follow up at two years following the Act, almost 6% of spam was found to be compliant (Grimes, 2007). Unfortunately, there is no analysis contained in the report beyond descriptive statistics. There is also no measurement before the passage of the CAN SPAM Act, only six months after and a two year follow up.

A more thorough analysis of violations of the CAN SPAM Act contained in email messages was conducted by Szde Yu (2011), who performed a qualitative assessment of 3,983
spam emails. The messages were collected from the spam folders/labels of multiple Gmail accounts which had been received between May 2010 and August 2010. Each individual spam message was analyzed to assess compliance with the CAN SPAM Act.

Only 24% of the sample messages violated the deceptive subject line requirement, indicating compliance was likely high for this type of regulation. However, compliance was likely for reasons unrelated to the CAN SPAM Act. Most messages that were sexually explicit also indicated as much by the wording of the subject, although they did not explicitly state that the content was sexual in nature.

Many of the emails contained falsified headers to some degree. Twenty seven percent of emails failed a DNS validation check, indicating that the server for the sender’s email address was not valid or had been falsified. Another 14% contained false header information for other reasons, including using a false name (e.g. “me” instead of an actual name), using the same address for both the sender and recipient, and some were marked as “unknown sender.” Almost 51% were not honest about the recipient’s own email address, as an incorrect recipient address was shown. Using a BCC method may account for this, preventing recipients from seeing everyone who received the message.

Lastly, a little over 7% of messages failed a validation check against sending email through a server designated for sending or forwarding emails. Sending email through a third party server without authorization is an aggravation under the CAN SPAM Act. Of the 3,983 spam emails analyzed, only 2.7% were found to be in complete compliance with the CAN SPAM Act. While a compliance rate of under three percent sounds low, there was no account of compliance before CAN SPAM Act was passed, but rather one measurement of spam during 2010.
A study that sufficiently accounted for both spam sent before and after the passing of the Act was conducted by Kigerl (2009), who performed a time series analysis to assess the CAN SPAM Act’s impact on a number of spam dimensions and outcomes. Among those outcomes included the volume of spam sent, compliance with the CAN SPAM Act, and locality from which spam was sent, all measured from among a sample of over two million spam messages sent between 1998 and 2008 and received in a series of bait email inboxes within the United States. The intervention measure of the CAN SPAM Act was a dichotomous predictor, false before January 1, 2004, and true on or after, the date the CAN SPAM Act went into effect.

There was no significant impact found for the CAN SPAM Act on the absolute number of spam messages received. Spam volume was unaffected following the passing of the Act. Additionally, three measures of compliance with the CAN SPAM Act were measured. They included (1) providing an unsubscribe option in the message, determined by matching keywords such as “unsubscribe” or “opt-out”; (2) providing a valid mailing address in the message, matched via software recognizing the format of addresses included in the message (number, direction, street name, and street suffix); and (3) including an accurate subject line, identified by software that matches keywords in the subject line with words in the body of the message. Only one of these three measures of compliance was significant, and the one measure found to be significant (accurate subject headings), was actually found to decrease following the CAN SPAM Act.

Lastly, a final time series model was created to address the question that spam might not have decreased following the CAN SPAM Act, but rather moved spam operations overseas where the CAN SPAM Act had little to no jurisdiction. Spam sending location was measured by capturing the first IP address of an email and geolocating it to a nation of origin. This measure
failed to produce a significant effect, suggesting that the CAN SPAM Act had no impact on spam messages received from IP addresses appearing to originate from within the United States.

There are some limitations to this study that should be mentioned. The measure of the CAN SPAM Act included in all time-series models tested was a dichotomous measure, only indicating an abrupt but permanent impact the CAN SPAM might theoretically have been suspected to have immediately following the date the Act became enforceable. A binary measure may not be sufficient to capture the true variation and influence a legal code might have on actual behavior, be it spamming behavior or otherwise. A continuous measure of CAN SPAM would be more desirable, such as the number of prosecutions under the Act, or the amount of media attention given to the Act in the news over time.

Furthermore, the time series models use were also not completely sufficient to partial out the impact of the CAN SPAM Act. The models presented were only simple time series regression models, with a single predictor per model (the CAN SPAM Act). A multiple time series regression would be a more rigorous indicator of whether the CAN SPAM Act might have had any kind of influence. Specifically, additional control variables ought to be included, such as the growth in IT technology in the United States, or the number of internet users per capita over time. Many studies suggest spam actually increased since the passing of the Act, but this increase might simply be a function of the similar parallel expansion of the internet in general. It may not be that criminality has increased, but rather there are simply more internet users today to send and receive spam. Including control variables ought to help rule out variation in spamming activity caused by the spread of technology and the internet, so that we can find the unique effects the CAN SPAM might have had. The CAN SPAM’s impact might otherwise be difficult
to discern should spamming variation be largely influenced by changing trends in technological development.
Chapter 4: Theories of Cybercrime: Understanding Spamming Behavior and the Theoretical Foundations Underpinning Anti-Spam Efforts

A proper evaluation of the CAN SPAM Act and its effectiveness ought to rely on an appropriate theory underpinning cybercriminal and spammer behavior and motivations. Being a form of crime, existing criminological theory lends support in understanding spamming offenses. However, spam, as well as all cybercrime, is novel and unlike traditional forms of crime one is familiar with as the offense takes place in cyberspace, which has dimensions and properties dissimilar from conventional space where crimes may occur.

Early discussions about predicting an offense in cyberspace focused on whether a new theory was required, considering the historical novelty of cyberspace itself. Cyberspace is an environment quite unlike physical reality, where traditional laws and understandings about the physical world do not apply. Therefore debate has focused on whether a different theory customized to cyberspace was needed to predict criminal behavior.

A contrast can be made between physical space and digital space. In cyberspace, physically defined locations lose their importance. Whether a website is hosted in the next room or across the world is less meaningful, as both websites can be opened simultaneously without costing much in terms of time. This stretches out social relations across space as well, as online communities can be composed of members all over the world. Each community is not as isolated by the constraints of physical distance from other communities. Although the regular visitors of an online community may be socially isolated from other communities by choice alone, there are no real barriers to joining many different membership communities.

Anonymity in cyberspace also means that fewer participants’ identities are certain, as they can assume a new pseudonym and claim to be anyone. The internet allows for the multiplication of exchanges of goods, ideas, and persons from across the world. Ideas and
information can propagate much faster over decentralized networks (Adams, 1998). Distance and time have less meaning over the internet. A single individual can come into contact with literally millions of other participants or resources via electronic transmissions anywhere in the world where there is internet access. Such transmissions are not as limited by physical space or time, unlike traditional behaviors. A single action in cyberspace can affect thousands of users and cross dozens of jurisdictions. This limitless ability for offender and victim to meet in cyberspace is one of the potential reasons cybercrime is so prevalent (Yar, 2005).

The new space of the internet has two implications for offenses committed within such space (Felson, 2011). Ordinary users can become exposed to the risks of becoming victims of cybercrime simply by surfing the web or reading email or other electronic messages. The cyberspace user also has access to a shadowy world of vice many such users would never consider venturing into physical space. The user also may have more incentive to become an offender his/herself, as the internet also offers financially motivated and technically sophisticated people anywhere in the world direct access to an endless supply of possible targets or victims. Legal responses to offenses in cyberspace can also be blunted, further nudging otherwise law-abiding citizens to commit illegal acts. Investigation of cybercrime requires unprecedented use of technology by law enforcement and international cooperation between nation states (Felson, 2011). The realization of such differences of cyberspace from physical space originally had prompted suggestions that a new theory was required (Capeller, 2001; Snyder, 2001).

An existing theory of criminal behavior, routine activity theory (RAT), has been put forth as appropriate, rather than developing an entirely new theory specialized to the internet (Pease, 2001). Routine activity theory can be powerfully descriptive of deviance on the internet because it focuses on three different conditions being true and meeting in space that allow a crime to
occur. Considering the lessened restrictions of time and space over a global network, these three conditions have a magnified ability to intersect in cyberspace.

Routine activity theory predicts that if three conditions are met in space (or cyberspace), then a crime will occur. These conditions include the ease in which an offense can be carried out in a particular setting, and the psychological willingness of the perpetrator to capitalize on the opportunity to offend that the setting provides. These conditions are (1) suitable targets (e.g. to steal), (2) a motivated offender (e.g. someone willing to steal), and (3) the absence of a capable guardian (e.g. to prevent theft) (Cohen & Felson, 1979).

Routine activity theory (RAT) has been used to understand cybercrime, as RAT takes into consideration both the environment and the individual. The internet itself is an environment so large yet easily traversable that countless situations conducive to crime can exist or be sought ought deliberately by a willing participant. There are about 2.4 billion users of the internet worldwide (Internet World Stats, 2012), many of which are or could be motivated offenders able to travel just about to any location available on the public internet. Cyberspace as an environment also applies to illegal spam, as anyone connected to the internet can send and receive email. Even without email, other electronic communication channels exist on the internet, such as chat, forum, comment, and social networking methods, all of which can be used to propagate spam.

**Routine Activity Theory**

The three conditions of routine activity theory have been used to describe, understand, and predict crime. The three conditions, motivated offenders, suitable targets, and the absence of capable guardians, will be discussed in turn.
Motivated Offender

A motivated offender is a person willing to commit a given crime. The other two conditions, suitable targets and capable guardians, matter little if there is no one willing or motivated to capitalize on any opportunities to offend. Offenders are not to be simply motivated; they must also perceive there to be a suitable target and the absence of a capable guardian. The motivated offender condition is the crime from the criminals perspective. Their perspective can be influenced by their personality, temperament, abilities, experiences, or anything that describes him or her that might explain their motivation to offend. (Akers & Sellers, 2004).

Suitable Targets

A suitable target is something desired that another might wish to obtain (Cohen & Felson, 1979). The target can be money or merchandise, or possibly something less tangible, such as status or recognition. The target must also be suitable, not just desired. It must be capable of being obtained (through illegal means) and be visible to the offender, such as merchandise placed conveniently near an exit door or small enough to conceal on one’s person. A suitable target is, from the offender’s perspective, one that can more easily be stolen rather than bought, or otherwise acquired via an illicit means.

Felson and Clarke (1998) have disaggregated the condition of suitable targets into four constituent parts, represented by the acronym, VIVA. VIVA stands for value, inertia, visibility, and access. Value is the worth a motivated offender believes the target to have. If the target is expensive, then the offender may place more value on it. Inertia represents the ability to move or otherwise manipulate the target. If the target is lightweight and easily movable, the motivated offender can take it into his or her own possession more easily. Visibility represents how easily the target is able to be noticed by the offender. If merchandise is placed near a window, then the
offender may notice it and decide he/she wants it. Lastly, accessibility is how easily the offender can get at and withdraw the target from another’s custody. Merchandise placed near an exit reduces the number of steps necessary to access and acquire the target. If the motivated offender would have difficulty acquiring the target or otherwise interacting with it, it is unlikely to be considered suitable (Felson & Clarke, 1998).

**Absence of Capable Guardians**

A capable guardian is something or someone ready and able to protect a target from the motivated offender. Often law enforcement is considered a capable guardian, and the absence of law enforcement can be conducive to the commission of a crime (Cohen & Felson, 1979). However, a guardian is not limited to law enforcement, and can be anything. Capable guardians can be police patrols or security guards as well as padlocks, fences, lighting, alarms, surveillance cameras, suspicious neighbors, or even a victim willing to defend against an assailant.

**Applicability to Cybercrime**

With the emergence of the internet, a vast number of participants have migrated to this new environment, allowing new possibilities for human interaction. With trillions of web pages and billions of internet users present online, any number of them can be suitable targets to any other participant in cyberspace. A website or web server connected to the internet can have any number of vulnerabilities an offender might exploit. If a web page has none, there are trillions more to choose from. And chances are, any internet ready host or server has some undiscovered vulnerabilities. Additionally, the technology itself may not be a suitable target, but the individuals using it may be. This goes equally for motivated offenders, as any web page or internet user can be malicious, and willing to exploit vulnerabilities found in cyberspace. There
are few borders on the internet, so motivated offenders and suitable targets can easily find each other in cyberspace.

*Motivated Offender*

The near irrelevance of geographical borders over the internet makes motivated offenders, no matter their location in the real world, a present danger to all users. Cybercrime used to require a certain level of technical skill to be successful. Today, with exploit kits, cybercrime services, and free or cheap crimeware, anyone motivated enough can perpetrate cybercrime, given they’re willing to pay for them. While the skill level of the motivated offender is not discussed often in writings on RAT, someone with the skill to carry out a crime would likely be more motivated to do so than someone who had to learn the skills first. Perpetrating cybercrimes now requires much less skill today, possibly motivating more offenders. For example, some versions of existing crimeware (botnets, phishing kits) is even free or open source, requiring just a little patience to repurpose for an offender’s chosen crime. Other crimes that require some level of skill to commit (e.g. securities fraud, embezzlement) often require the offender to have a certain profession (e.g. stock broker, bank teller) for whom the opportunities for such crimes exist exclusively. This is not the case with cybercrime, as the only access a motivated offender needs is to the internet, where there are sufficient resources to learn the technical skills required of the crime. Many carding forums (online forums for trading in illegal credential goods, such as credit cards) have boards where tutorials are posted, containing details on how to conduct cybercrime anywhere from launching a botnet to exploiting technical vulnerabilities in web servers. The ease and convenience in which someone can learn these skills securely at home behind a screen is likely pretty motivating.
The anonymity the internet provides can embolden those to commit deviant acts that they otherwise would be unlikely to consider. The emboldening effect of cyberspace has been termed the online disinhibition effect (Suler, 2004). The effect is the lowered inhibitions participants experience while secure at a keyboard, reducing the natural resistance to being disrespectful, offensive, or hostile. The disinhibition can also lower an aversion to cybercrime, making committing an offense in cyberspace more likely than if the victim is face to face with the offender. Online offenders are likely more motivated to offend in cyberspace than elsewhere.

Not only are the risks lower for online offending, with fewer opportunities to be caught, but there are substantial rewards in the form of profits that can be gained from illegal acts over the internet. The low risk and high yield cybercrime can offer can be a substantial enticement for those looking to make money off spam or other cybercrimes. A typical spammer can earn a six figure income while working at home (Kanish, Kreibich, Levchenko, Enright, Voelker, Paxson, & Savage, 2008), in which case there are some unambiguous motivations. Working full-time as a spammer can be a lucrative and rational career choice.

**Suitable Targets**

The number of internet users world-wide grows every day, expanding the content available on the web. More internet users means more people who use internet banking, make purchases online, and more people comfortable filling out online forms with personal information. Because there are more users willing to pay for goods online, they are also more likely to purchase goods advertised in spam. There is therefore more money, and those willing to exchange it, over the internet. Suitable targets in cyberspace vary; they can include bank accounts, credit cards, social security numbers, email logins, intellectual property, server hosting,
botnet bots, cheap stolen property to buy, etc. The continued growth of the internet will expand the list of suitable targets that a motivated offender may seek out.

The multiplicative nature of being able to send electronic messages to a nearly unlimited number of recipients makes the growth of the internet and internet users more compelling to spammers or other motivated offenders. An internet user almost anywhere in the world can be a suitable target for a spammer, as geographical and physical boundaries lose their importance over cyberspace. This means there are an unlimited number of targets, and email spam succeed via quantity, not quality, as only a tiny fraction of the messages sent require a buyer or victim for the spam operation to be successful.

Absence of Capable Guardians

Capable guardians can be grouped into three separate categories. Capable guardianship can describe the street smarts or resistance to being socially engineered that can protect against fraud or other manipulations. A capable guardian can also be the technology in place, by the victim or otherwise, which can prevent a cyberattack. Lastly, the laws and their enforcement prohibiting cybercrimes can also be a deterrent. Capable users may be wise enough to the fraudulent nature of a phishing scheme not to fall victim to it. They may also have current and up-to-date antivirus software installed or important applications regularly patched. While a user may be capable, the software itself must also be capable. Software updates that patch a technical vulnerability need to be available to install as soon as possible once the exploit is in the wild. This goes equally for updating the malware signatures for an antivirus database to identify a new malware variant. Finally, when these defenses fail, law enforcement must punish the offender in order for them to decide that their crimes do not pay. If law enforcement is to be a capable
guardian, the chances of prosecution and conviction must be high, as well as sufficiently swift and severe (Mendes, 2004).

The online disinhibition effect applies to capable guardians as well. Some may be apprehensive about walking the streets alone at night, but such fears are greatly reduced while traversing the internet. The disinhibition at the computer screen may not motivate a user to install antivirus, or to check the internet address of the website they have been directed to by a spam email. Despite this, most users ought to recognize an email scam that reaches their inbox, and many more ought to understand the need to sufficient antivirus. However, cybercriminals need just one victim among countless targets present in cyberspace who are successfully conned, or who fail to protect themselves, or buy their product.

Research on RAT and Cybercrime

The existing literature suggests that routine activity theory (RAT) applies to cybercrime. RAT has been found to explain phishing scams, DDoS attacks, malware infections, and even cyber-bullying. Holt and Bossler (2009) defined online harassment as a type of cybercrime, and conducted a survey study to test its relation to RAT. About 600 college students were surveyed on their internet use and computer skills (i.e. suitable targets). Some responses surveyed did not relate to online bullying victimization, such as owning a computer or not, internet speed available, frequency of internet use, computer skills, and the use of a firewall and antivirus. Responses that were found to increase online harassment victimization was the amount of time in chat room, direct involvement or friends involved in computer deviance (e.g. illegal file sharing, guessing other people’s passwords), and being female. Time spent in chat rooms was found to increase exposure to offenders (suitable targets), as well as associating oneself with computer deviants (motivated offenders) increases the susceptibility to a type of cybercrime.
Receiving threats online has also been compared to receiving threats face-to-face or in physical space, thus contrasting RAT as applied to cybercrime/deviance to traditional RAT experienced on the street (Wilsem, 2011). Wilsem conducted a study to compare victimization in both cyberspace and physical space. Data was acquired from a longitudinal monthly survey of households in the Netherlands. Routine activities on the internet that make a user a suitable target, such as online shopping and webcam usage, predicted an increased risk of receiving digital threats over the internet. Social networking also predicted receipt of digital threats, but it also predicted face-to-face threats in physical space as well. Routine activities in physical space also related to digital victimization, such that shopping in a physical commercial building and longer commute hours increased risk of this kind. Lastly, participants who sent digital threats themselves received more threats of both kinds, perhaps because of retaliation.

While sending threats online may not necessarily be criminal, and as such may not be a form of cybercrime, cyberstalking has been classified as criminal by many jurisdictions and has also been analyzed in terms of its applicability to RAT (Reyns, Henson, & Fisher, 2011). Cyberstalking is the repeated pursuit of an individual using electronic or internet capable devices. A random sample of undergrads in a Midwest university was surveyed using this measuring cyberstalking using this definition. Routine activities such as the number of social networks participants had active, the number of strangers they added as friends to such networks, and using a profile tracker to view who has visited their online profile all predicted increased odds of cyberstalking. However, by far the biggest predictor of cyberstalking was the number of deviant peers who engaged in such behaviors as digital piracy, sending sexually explicit pictures to someone, or harassing another person online.
Sexual harassment online, also a minor offense which can be illegal, has also been studied. Marcum, Higgins, and Ricketts (2010) looked at the routine activities of a sample of 744 college freshmen students in terms of their sexual harassment victimization via self-report. Participants were surveyed on their experiences currently as college students, but also their past experienced during high school. Socializing online and providing personal information online predicted sexual harassment victimization online in both high school and college. Using chat rooms increased online sexual harassment victimization in high school only. Lastly, using a computer in a less private area, such as a living room instead of a bedroom, predicted less online sexual harassment while in high school.

While much of the research linking RAT with computer related offenses, most of it has focused on minor crimes which require very little sophistication for a perpetrator to carry out. More serious cybercrimes have also been investigated, however. Hutchings and Hayes (2009) looked at phishing scam victimizations among a sample of 104 participants pulled randomly from a telephone directory aged 18 and up. Participants were interviewed by phone. Their first hypothesis proposed that users with less internet experience and more frequent use of online banking would be more susceptible to a phishing attack. Unfortunately, they lacked sufficient evidence to make this claim as there were no subjects in the sample had ever been a victim of a phishing attack to be able to answer this question.

There was sufficient support for the second hypothesis, however. The more time that was spent on a computer or using online banking was associated with a higher risk for receiving a phishing scam email. Internet use in general was not related to a phishing attack. A third hypothesis, that the use of spam filters reduce the risk of receiving a phishing email was not
supported. While spam filters did not act as sufficient capable guardians, routine activity on the computer in general does appear to expose one to greater risks.

Additional research on targets of online fraud looked exclusively at suitable targets (Pratt, Holtfreter & Reisig, 2010). Internet fraud was the dependent variable, which could be a phishing scam, advance fee fraud, or other forms of online fraud. A phone survey of 922 randomly sampled subjects found that those who were younger, more educated, male, and had higher incomes spent more time online and made more website purchases. Thus, they were more likely to be suitable targets. It is worth noting that those with higher education and income often perceive there to be fewer risks online, and therefore use the internet more often (Reisig, Pratt & Holtfreter, 2009). Risk takers make better offenders but they also make better targets for offenders. Both hours spent online and the number of website purchases predicted being the target of internet fraud, controlling for the personal characteristics mentioned. The impact these personal characteristics (e.g. age, income) have on being the target of cyberattacks were fully mediated by time spent online and the number of website purchases (Pratt, Holtfreter & Reisig, 2010).

Risk of malware infection has also been analyzed. Bossler and Holt (2007) collected a sample of 788 college students who were surveyed on any loss of time or data due to a malware infection. The student’s routine activities were captured by measuring internet connection speed, online shopping habits, email and chat use, any programming experience, and use of social networking websites. A majority of the routine activity measures were unrelated to malware victimization. Capable guardian measures, such as use of antivirus, keeping passwords safe, and computer skills were also recorded. Findings revealed that such measures were not sufficient capable guardians, and did not relate to malware outcomes. It may be that protections such as
antivirus installations were put in place only after experiencing a malware attack. Finally, measures of computer deviance, such as hacking, downloading online pornography, and digital piracy, were recorded. Deviant computer behavior was positively associated with malware victimization, but only weakly. Having friends who scored high on computer deviancy was also found to predict malware victimization. The findings mirror research on cyber-bullying (Holt & Bossler, 2009), suggesting that deviant behavior or deviant associates (motivated offenders) exposes one to more cybercrime victimization.

Choi (2008) also investigated malware victimizations, conducting a survey of 172 college students in an attempt to look at the predictors of the number of malware infections and the number of hours and dollars lost from such victimizations. The presence of capable guardians was measured by whether participants used antivirus, antispyware, and firewalls. The suitability of each target (student) was indicated by the amount of time spent online both at home and at work. Both the suitability of a given target, and the absence of the measured capable guardians, was predictive of subsequent malware victimization.

Reyns (2013) was the first to measure identity theft victimization among surveyed respondents and its relation to RAT. The data was taken from a British crime survey between 2008 and 2009 involving almost 6,000 respondents. Subjects were asked if they had any of their bank cards used without their permission. Respondents who engaged in routine activities over the internet such as use of internet banking, shopping online, using either email or instant messaging, or downloading files, were all more likely to report that their identity had been stolen. In addition to RAT measures, older, male, and wealthier respondents were also more likely to be victims. Finally, perceiving there to be a higher risk of identity theft also predicted being a victim, although the temporal ordering of this findings is uncertain.
Ngo and Paternoster (2011) looked at both computer virus and phishing victimization in the past 12 months via an online survey administered to about 300 college students. Phishing victimization was measured by receipt of a phishing email, not actually falling prey to it and having one’s identity stolen. The authors actually found little support for RAT and its applicability to cybercrime. Measures of capable guardians, such as having security software installed, actually increased the chances of reporting a virus victimization. Target suitability, such as following links in emails from strangers, clicking pop-ups, opening email attachments from strangers, actually lowered computer virus victimization outcomes. Remaining RAT measures included were not found to be significant in terms of virus victimizations. Measures that were significant included younger, nonwhite respondents who engaged in “computer deviance” (piracy, using pornography, etc.), all of which predicted higher incidences of malware victimization. The authors conclude that there was no support for routine activity theory under the conditions studied. However, the authors’ findings are at odds with the bulk of college sample surveys measuring RAT and cybercrime, and the study may be an anomaly.

Not all research linking RAT to cybercrime has relied on survey samples. As routine activity theory was originally formulated in terms of predicting aggregate level crime rates (Cohen & Felson, 1979), more recent research on cybercrime has conceptualized criminal behavior in this form when attempting to apply routine activity theory to it. Maimon, Kamerdze, Cukier, and Sobesto (2013) analyzed intrusion detection system recorded data on digital attacks against a large university computer network over time. Attacks against the network included port scanning the network for vulnerabilities by malware, actual exploits of the network (e.g. installing malware), and DDoS attacks. Aggregate level data about the university’s network
users over time (age, foreign student status, etc.) was included in a model to predict cyberattacks on the network.

The majority of attacks on the network were between 9:00 AM and 4:59 PM. It is likely that this period is when legitimate users of the network (suitable targets) are most active, provoking the greatest number of attacks. Percent of network users who were foreign born increased the number of attacks originating from those user’s home country. International students likely browse pages located or in the language of their home countries more often, exposing them to motivated offenders in those countries. Of countries outside the United States which attacks originated from, there was a greater propensity for the attacking nation to have higher GDP and more internet users. Naturally countries that have more internet users have more motivated offenders, or have more internet ready devices (suitable targets) that can be infected with malware (malware which can attack the university network from abroad).

Almost no literature has looked at the predictors of spam and its relation to routine activity related measures. Some consider phishing attacks, which are only a narrow type of spam, but this is only included in survey sample data, not aggregate level spam/crime rate data as RAT was originally intended. Kigerl (2012) conducted an analysis of RAT and cybercrime exclusively in the context of national predictors of cyberattacks using spam activity as an outcome measure. Data on a sample of 132 countries was collected to measure what RAT characteristics of each nation predicted higher spam and phishing attacks. Spam was measured from a sample of about 700,000 spam messages sent in 2008 and operationalized as a spam rate sent per capita within each country. Phishing attacks were measured by the number of top level domains (TLD) linked to by phishing scam emails in 2010, based on a per capita rate within each country. A TLD is a domain name portion such as .com or .net, some of which are owned by or
associated with specific countries (e.g. a TLD of .us indicates the attacks is associated with the United States).

The number of internet users per capita within each country predicted higher spam rates, suggesting either more motivated offenders within those countries that might send spam, or more internet ready machines on which to install a spambot (suitable targets). Nations with higher GDP per capita also sent more spam and phishing attacks. Whether a nation was a participant to the Convention on Cybercrime, an international agreement to collaborate with neighboring states to prosecute cybercriminals (a measure of capable guardians), was not related to either spam or phishing rates. The Convention on Cybercrime is still in its nascent stages and may not be ready to be expected to deter or incapacitate offending yet. Lastly, there was a significant interaction between internet users and unemployment on spam rates, such that high unemployment rates increased the positive effect the number of internet users had on spam attacks. It may be that technologically sophisticated internet users who face unemployment may turn to cybercrime to make ends meet, thus turning them from legal employment into motivated offenders.

Applicability to the CAN SPAM Act

The most appropriate condition of routine activity theory for understanding the impact of the CAN SPAM Act would be capable guardians. The Act is intended to serve as a capable guardian and deter or incapacitate any motivated offenders who reside in the United States from sending illegal spam. Of those offenders not deterred, it should be expected that the Act would be used to punish them and put a halt to their existing spam operations.

The capable guardianship element of routine activity theory does not have to be one of deterrence, especially when it comes to cybercrime and spam. Technological fixes to the spam problem, such as authentication protocols and message filters, are technologically preventative
and do not require any psychological experience or reaction of deterrence. Technological capable guardians do not apply to criminology or human behavior, and spam is essentially a type of human behavior. While technological defenses can prevent a cyberattack, they do not stop the cyberattacker.

That is why cybercrime laws are important. The CAN SPAM Act is intended to serve as a capable guardian where deterrence plays a role. An intended component of the Act is to deter a would-be offender from engaging in illegal spam activity as the potential profits that can be made from spam would not exceed the potential losses from legal fines and possible prison time. Regarding theories of cybercrime, the CAN SPAM Act is intended to serve as the capable guardian condition of routine activity theory, and while capable guardians do not necessarily require a psychological component of deterrence, in the case of CAN SPAM the Act is in part intended to operate through psychological deterrence specifically in its method of lowering spam crimes.

**Deterrence Theory**

Deterrence is defined as the omission of an act as a response to the perceived risk and fear of punishment for contrary behavior (Gibbs, 1975). In the context of the CAN SPAM Act, the penalties under the Act are intended to serve as such a deterrent effect. In the broader context of the elements predicting cybercrime involved in routine activity theory, the CAN SPAM Act also serves as a capable guardian. Namely, the Act’s purpose is to prevent motivated offenders from perpetrating their crimes against suitable targets. The mechanism behind such guardianship is via deterrence.

Deterrence theory was first postulated in the late 1700s by Cesare Beccaria and Jeremy Bentham and stated that the rate of a particular crime varies inversely with the celerity, certainty,
and severity of punishment of that crime (Beccaria, 1963; Bentham, 1962). In order for a given punishment to be effective and to produce a deterrent effect on a potential offender, the punishment ought to be highly probable a consequence for each crime committed (certainty), the punishment ought to follow the offense with a certain swiftness after the offending act is carried out (celerity) in order for the offender to learn to associate the crime with a paired punishment, and the punishment must be unpleasant to such an extent that an offender would be averse to receiving it and take action (or inaction) to avoid it (severity) (Gibbs, 1975).

There are multiple channels in which a punishment can deter subsequent or future offending. General deterrence is the threat of punishment which the public or potential offenders are aware of beforehand and thus resist committing a given crime because of the perception of that possible punishment (Gibbs, 1975). That is, someone is deterred from having committed a crime in the first place, an individual who might have otherwise committed the offense. The punishment is threatened, but never actually enforced. General deterrence is contrasted with specific deterrence, specific deterrence being the deterrent effect of an already enforced punishment against an offender having already committed a first or initial offense. General deterrence or the lingering unenforced threat of punishment was not sufficient to prevent the initial crime, but an actual observed punishment following that initial crime deters the offender from reoffending a second time. That is, a punishment follows a specific offense, which is a sufficiently effective punishment to convince the offender to desist.

The CAN SPAM Act is amenable to either types of deterrence. Offenders may avoid the spam business altogether or comply with the regulations of the Act when sending spam, or a subsequent conviction under the CAN SPAM Act may end an individual spammer’s career. Whether the typical career spammer can be deterred via such general deterrence strategies, or
even specific deterrence, is uncertain, although legitimate online email marketers were in full compliance with the Act after its passing (Majoras, Leary, Harbour & Leibowitz, 2005). If CAN SPAM prosecutions or media attention are associated with a reduction in overall spam rates or events, there may be a general or specific deterrent effect.

A punishment does not have to prevent an offense entirely in order to be effective (Zimring & Hawkins, 1973). A punishment can simply reduce the approximate severity of the offense itself. Termed partial deterrence, if an offender commits an offense lower in severity due to an expected threat of punishment than he or she otherwise would have, that is also considered to be a deterrent effect. An example would be a motor vehicle driver observing a speed limit of 65 MPH, and subsequently reduce his/her speed from 90 to 70 MPH; a less offensive speed but one still over the posted speed limit. Thus the offender would be partially deterred.

Such deterrence may also be relevant to the CAN SPAM Act. While it would likely not be profitable enough for the typical career spammer to comply with all the provisions of the CAN SPAM Act, he or she might avoid committing some aggravations of the Act or reduce the number of violations contained in each spammed message. The spammer might avoid criminal sanctions under the CAN SPAM Act by opting not to involve fraudulent strategies or malicious software in the spam sent, instead only accruing civil penalties by continuing to not honor opt-out requests and trying to fool spam filters.

Deterrence strategies can also be classified in terms of marginal and absolute deterrence (Zimring & Hawkins, 1973). Absolute deterrence is the impact a given threat of punishment has on behavior versus no threat of punishment at all. Thus it is the absolute effect size of the punishment’s influence relative to a baseline of nothing at all threatened. Marginal deterrence is the deterrent effect of one punishment relative to a second, alternate punishment. Some
punishments or legal deterents may be better than others. It may be up for consideration what about the CAN SPAM Act can be expected to provide a deterrent effect. Media attention to prosecutions under the CAN SPAM Act may create a general deterrence for those potential spammers exposed to such stories. The count of prosecutions over time may also be marginally better than simple public awareness of the threat the Act imposes. Or perhaps the CAN SPAM Act in its entirety has absolutely no impact on how much spam is sent or the nature of the messages contained in spam.

It should be mentioned that simply because a type of crime is lower following the introduction of a new threat of punishment imposed against the offenders of such a crime does not necessarily prove a deterrent effect. Gibbs (1975) suggests ten types of psychological and physical effects explaining the reduction of an offense in response to the behavior being criminalized, only one of those ten being that of deterrence. Among them are included incapacitation, where the offender is unable to reoffend due to detention by law enforcement, who otherwise might offend were he/she able. Enculturation is another alternative, where an individual feels obligated to obey a law for reasons of conformity, not because they feel it is rationally within their self-interest to avoid punishment (the person would obey even if they could evade punishment). There is also habituation, which may follow a deterrent effect, such as the reduction of speeding over a posted limit. However, over time, complying with the speed limit may simply be a habit, and no longer due to fear of punishment.

For any of these three alternative crime avoidance outcomes (or the entire ten Gibbs speaks about), is deterrence involved? That is, the citizen does not engage in a crime because a law forbids the act, but not due to reasons of fear of punishment. Punishment avoidance does not have to be due to fear, but also rational forethought, yet even considering this definition, for
something like enculturation, deterrence plays no role. This is certainly a concern with testing the legal impact a law may have on crime rates, as a reduction in criminal activity may be due to any of the ten reasons cited by Gibbs, not just deterrence. However, it is unlikely that the CAN SPAM Act is intended to influence cybercrime culture, or change norms without first relying on deterrence, as the statute lists a myriad number of criminal and civil sanctions to be imposed on offenders. Changing norms might itself be a type of deterrence, as individuals actors may feel violating the norm would be an embarrassment or shameful act, thus they rationally choose not to offend (Stevens, 2012). However, the anonymity of cyberspace can make it more difficult for the community to shame a cyber-offender were he/she to engage in cybercrime (Guitton, 2012). Incapacitation might be the only credible alternative to deterrence, as incapacitating individual spammers can impact global spam rates, albeit temporary (Mosher, 2007). Despite this, this research focuses on CAN SPAM Act media attention and prosecutions, and does not capture the incapacitation of spammers outside the scope of the media. The amount of media attention towards the CAN SPAM Act, if it has any effect at all, ought to be one of deterrence.

Deterrence and Cybercrime

Can a cyberattack be deterred with the threat of punishment? Deterrence has been suggested to generalize from application to street crime to application to cyber threats and information security (Kunreuther & Heal, 2005; Png, Tang, & Wang, 2006). Research exists discussing deterrence of cyberattacks, but the bulk of that research focuses on cyberwarfare, not cybercrime (Chun, 2011). Cyberwarfare is an attack on a neighboring country’s critical infrastructure (power plants, banks, government, etc.) over the internet, by a foreign nation’s government or terrorists hostile to the target country. Cybercrime is the focus of this research, involving local or international breaches of legal codes via illicit uses of technology which
warrant criminal or civil penalties, namely illegal spam in this case. However, the literature does touch on deterrence of cybercrime to some extent.

The majority of existing strategies against both cybercrime and cyberwarfare are defensive and not deterrent in their intent (Glaser, 2011). Specifically, defense is intended to reduce the effectiveness of a cyberattack when it happens, such as the detection and quarantine of malicious software, network intrusions, or spam filters. A deterrence strategy would instead convince an adversary not to attack in the first place, via a lingering and convincing threat of punishment or retaliation.

Just as cyberspace changes the nature of routine activity theory, it also has different implications for deterrence theory than can be expected from deterrence in physical space (Taipale, 2010). Digital space is governed by different physics, as computer networks are unbounded, infinitely scalable, and abstract. In physical space, both modalities for attack and the severity of outcomes scale consistently and predictably. For instance, a truck bomb does more damage than a car bomb. However, in cyberspace, an offender can scale infinitely, as a small time offender can acquire as large a botnet as his/her skillset and determination allow. Less motivation and fewer resources are required to scale an attack upward, making deterrence more difficult.

The temporality also becomes less meaningful in cyberspace (Taipale, 2010). An offender stealing uranium gives law enforcement advance warning before an attack can follow. Yet the time between a digital probe or port scan and an actual attack on the network can be almost zero. That gives law enforcement almost no time between detection of a threat and the consequences of that threat. Preventative deterrence is not as straightforward. The economics are also less costly, as a virus can be copied and multiplied an infinite number of times.
Something physical, such as a firearm, does not have such properties. There are fewer barriers to committing cybercrime that might otherwise persuade an offender to desist.

Factors such as these may make cyberspace seem less risky. There may be certain personality characteristics less receptive to deterrence strategies, which can explain why many offenders continue to offend even with a credible threat of punishment (Zimring & Hawkins, 1973). Individuals who are less forward thinking, more optimistic about their chances of success, impulsive and less anxious, may be less likely to be deterred. There is some evidence cybercriminals may be similar to street criminals in such a way. At least in the case of digital piracy, seeing oneself as a risk taker (Goles, 2007) or self-reporting as having low self-control (Higgins, 2007) is associated with higher incidences of software and movie piracy. Self-reports of participants at a hacking convention who rate themselves as likely to take risks, also have made more hacking attempts in their lifetimes (Bachmann, 2010). Compound a risk-taking personality with the reduced risk inherent in the anonymity of cyberspace, and the nature of deterrence becomes quite a bit different than what can be expected from crimes in physical space.

That is not to say that deterrence is not an appropriate recourse to crimes committed in cyberspace, it is just that individuals who would not otherwise be committing offenses may feel incentivized to do so in cyberspace. It can also be argued that cybercriminals are more rational than street offenders, as more planning and technical sophistication goes into the commission of their crimes (Guitton, 2012). It is still recommended that the law seek to deter these new types of offenders and that governments still maintain a cyber-deterrence strategy (Jensen, 2012). Criminal prosecution is still the most common response to cyberattacks. While governments can more easily enforce anti-cybercrime laws within their own borders, international cooperation
remains a necessity. There exist international laws and mutual legal assistance treaties to address such concerns, but there is not unanimous support and consent from all nations, and extradition and jurisdiction issues remain problematic.

Criminal prosecutions may not always be sufficiently frequent enough and enforcing and investigating cybercrimes is often more burdensome than the resources deployed to enforce other types of crimes. That is why civil penalties are also recommended in attempting to impose a deterrent effect on cybercriminals (Kesan & Hayes, 2011). With the ability to bring civil action against offenders, an expansion of the parties able to impose sanctions on offenders makes the certainty of punishment more likely. Victims themselves, not just governments, may hold offenders accountable. ISPs can be made liable for failing to monitor and secure the data being transmitted over their infrastructure, such as the distribution of spam or malware via their networks. The cybercriminals themselves can and should also be made to be liable to private entities. Unfortunately, under the CAN SPAM Act, only ISPs, states, and the FTC can bring civil action against a spamming offender, severely limiting the base of possible enforcers of the CAN SPAM Act.

There has been some empirical testing of deterrence theory as applied to cybercrime and cyberattacks. Png and Wang (2007) suggested that government enforcement has a deterrent effect on hackers perceived benefit from a given attack. Panel data for eight countries was analyzed, testing the impact newspaper reporting of hacking prosecutions and convictions had on hacking incidences. Hacking incidences were collected from the SANS Institute Internet Storm Center database, which tracks hacking attempts collected from intrusion detection systems and firewalls in over 50 countries. Government enforcement was measured via media articles within
each country on hacking prosecutions, and the conviction status and sentence length for the offender was recorded.

In all eight countries examined, reports of hacking crime enforcement by local governments had a significant negative effect on the number of hacking attempts and breaches within each nation. In addition, unemployment rates were also found to be positively associated with hacking crime rates. The authors suggest that unemployment may provide time and motivation for idle citizens with technical skills, but that government enforcement appears to provide a significant counterbalancing force that convinces some potential offenders that the risks of violating the law are too high.

Guittion (2012) also investigated media deterrence of cyberattacks at the national level, but conducted an additional analyses at the individual level as well. Guittion considered the degree of rational choice among cybercriminals and whether a deterrent such as attribution, which is the identification of a cyber-offender by removing their anonymity, has an influence on their decision-making and behavior. The first analysis considered deterrence and attribution at the nation state level. Computer and cybercrime arrest, investigation, and conviction rates per country were recorded. Newspaper and media attention indicating lack of attribution to cybercriminals, or revealing uncertainty among law enforcement as to the identities of a given culprit in a cybercrime investigation, were also recorded. Finally, the number of businesses undergoing cyberattacks was captured from business survey data.

The number of cybercrime cases solved by the police was positively related to the number of business experiencing cyberattacks. Also, the number of articles published in newspapers showing lack of attribution was associated with a higher number of attacks. Conversely, media reports that showed high attribution were associated with lower rates of
cyberattacks. The findings suggest that media reporting a higher certainty of being apprehended for a cybercrime acts as a deterrent effect, and additionally that cybercriminals decisions to offend or not may be influenced by perceptions of the certainty of being caught. Cybercriminals may be rational and subject to deterrence.

Guitton (2012) also conducted a separate and additional analysis assessing a set of 46 trial cases of individuals being prosecuted for breach of an information system committed in France, Germany, and the UK between 2003 and 2010. Each case was coded from reports on the trials in terms of rational decisions the offenders made, such as erasure of log files, use of privacy enhancing technology, or even claiming the actions they committed, under a handle or even under their real names. Claiming credit for the crimes suggests offenders perpetrated the offenses to achieve recognition and status. It was determined that the offenders could be broken into two separate groups: those seeking status as skilled hackers, and those seeking financial gain. Offenders who stole information tended to seek fame, those seeking to steal financial details such as credit card numbers of course sought money. The financially motivated hackers used more rational strategies to cover their tracks and remain anonymous. Guitton suggests these types of offenders can be deterred by threatening their anonymity, or attribution status. Those seeking fame and recognition, however, would not be considered as rational in the traditional sense, and thus would be more difficult to deter.

Anti-cybercrime laws have also been tested in terms of their deterrent effect. The Convention on Cybercrime, which is an international bill participating nations can sign that facilitates the cooperation between neighboring states to prosecute cybercrimes that cross national jurisdictions. Hui, Kim, & Wang (2013) investigated possible marginal deterrent effects the Convention on Cybercrime might have on DDoS attacks at the national level. DDoS is a
distributed denial of service attack where a victim server is flooded with more internet packets/requests to connect than it can possibly handle, thus making it unable to respond to actual legitimate requests to connect (bringing it offline). Marginal deterrence suggests that a punishment may decrease one type of crime by shifting criminals to another type of crime the initial punishment does not target. For instance, a law forbidding drunk driving may result in motorists turning to drugs rather than alcohol. Hui and colleagues recorded daily DDoS attacks on IP addresses associated with computers from 240 different countries.

Nations with higher GDP were found to be more intensely attacked via DDoS. Intense attacks were measured by more packets sent per individual attack. Nations that were rated as being higher on education indexes were attacked with greater frequency. Frequency of attack was measured by more individual attacks, or more servers/IP addresses attacked within that nation. Most importantly, if a given nation signed into law the Convention on Cybercrime, the frequency of attack that nation would experience would drop afterward. However, this drop was associated with an increase in the severity of the existing attacks that remained after signing the Convention. The authors conclude that this is evidence of a marginal deterrent effect, with many offenders desisting from committing cyberattacks, but perhaps shifting those that remain to more serious offenses.

It should also be mentioned that Kigerl (2012) looked at the possible deterrent effects of the Convention on Cybercrime at the national level, only in terms of spam and phishing attacks from those countries, rather than on those countries. Without including control variables, nations who signed the Convention on Cybercrime were the source of more spam and phishing attacks. After control measures were included the significance of these findings were suppressed. However, Kigerl did not account for the temporal nature of the passing of the Convention on
Cybercrime, as Hui and colleagues did, and may not have sufficiently captured the deterrent effect of anti-cybercrime law. The present study, however, seeks to capture the temporal ordering of the CAN SPAM Act and subsequent cyberattacks (spam related activity).

**Conclusion**

There are four research questions that this paper is aimed to answer. They are (1) whether CAN SPAM has affected spam frequency (spam rates or the amount of spam sent), (2) whether CAN SPAM has affected spam compliance with CAN SPAM regulations, (3) whether the CAN SPAM Act has affected spam severity (law violations not included in the CAN SPAM Act, such as malware distribution and fraud), and (4) whether CAN SPAM has affected the appearance of spam location (the amount of spam sent from IPAs associated with the United States, spam linking to domain names associated with the United States, etc.). All four potential effects on spamming behavior that relate to CAN SPAM Act enforcement or public awareness have implications for deterrence, as the CAN SPAM Act is intended to serve as a capable guardian. If the first question is answered such that spam rates are inversely related to CAN SPAM enforcement or media attention, it may be that some spammers choose to desist in response to such a legal threat. If the second and third questions are answered in such a way that the severity of spam violations of the law is reduced, this may have similar implications for deterrence. However, an increase of spam severity, while a corresponding drop in spam frequency, may suggest a marginal deterrent effect (Hui, Kim, & Wang, 2013). It has been suggested that the CAN SPAM Act has reigned in minor spammers who are more receptive to not violating the law, while the more serious career spammers have become more frequent and more severe (Jonsson, 2009). Remaining offenders may switch to more serious types of spam crimes. Finally, the fourth question concerns spam’s nation of origin, which may relate to
cybercrime law (Kigerl, 2012). If any shift is detected in the amount of spam sent under the CAN SPAM Act’s jurisdiction (the United States), it may be indicative of the Act’s impact on spamming behavior.

There is evidence cybercrimes can be subject to deterrence effects (Guitton, 2012; Hui, Kim, & Wang, 2013). However, to date, reports on the efficacy of the CAN SPAM Act in deterring spam have been mixed or negative (Majoras, Leary, Harbour, & Leibowitz, 2005; Kigerl, 2009). Yet no evaluation of the CAN SPAM Act has sufficiently captured the variation of the CAN SPAM Act that the law is intended to impose, such as media attention on spam and the CAN SPAM Act or prosecutions under the Act. Similar measures have been found to serve as a deterrent to certain forms of cybercrime, but has never been applied to spam within the United States.

Additionally, no research on the CAN SPAM Act has considered other elements of routine activity theory, only the capable guardian element of the statute. This research seeks to expand the explanatory power of spam related activity by including additional theoretical measures in attempting to capture suitable targets (internet ready PCs) and motivated offenders (internet users, unemployment rates). These additional measures to be included in subsequent models would also help to alleviate concerns of spuriousness inherent in all CAN SPAM Act evaluations conducted to date. Specifically, many reports mention spam increasing after the passing of the Act. However, spam was not the only facet of technology that is and continues to experience a growth in prevalence and frequency. Higher spam rates might be due to more internet ready computers and internet users, all of which can be implicated in spam’s fluctuations. Accounting for this variation would allow for the capture of any possible unique variation the CAN SPAM Act might have on spamming activity.
Chapter 5: Methodology

Four research questions involving spam volume, compliance, severity, and locality will be investigated by analyzing a sample of spam messages. The spam sample has been collected from publicly available internet spam archives in preparation for processing into a final tabular dataset for analysis. Eighteen time series measures of spamming behavior, each of which addresses one of the four research questions, will be created from the spam sample by software written to build the dataset. The different types of spamming behaviors will serve as the dependent variables, which will be regressed on a dozen measures of CAN SPAM Act enforcement, attention, public attitudes, and awareness to assess a possible deterrent effect of the CAN SPAM Act on spam.

Research Questions

There are four research questions this paper intends to address. They include whether the CAN SPAM Act has influenced (1) spam volume or frequency (the amount of spam sent), (2) spam compliance with regulations set forth under the CAN SPAM Act, (3) spam illegality and severity by the violation of anti-cybercrime laws beyond the scope of the CAN SPAM Act, and (4) nation of origin associated with elements of spam emails.

Regarding question (1), if the CAN SPAM Act can be shown to have an effect on spam volume, this might suggest a deterrent effect. There is some evidence cybercrimes of other types can be deterred in such a way (Guitton, 2012; Hui, Kim, & Wang, 2013). The same is to be tested against spam crime.

Question (2) is intended to determine to what extent spam emails are in compliance with the CAN SPAM Act. The CAN SPAM Act did not illegalize the sending of spam; instead it simply regulated the means to send and the content of spam emails. Just because spam volume is
unaffected by CAN SPAM Act activity, does not mean the Act had no impact. This question is intended to capture whether the CAN SPAM Act increased compliance (and perhaps the legality) of spam emails, regardless of how much spam is received on average.

Question (3) accounts for illegal spam that is in violation of laws extraneous to the CAN SPAM Act (such as fraudulent and malicious spam), but that the CAN SPAM Act might still relate to (e.g. phishing, malware distribution). It is questioned whether the CAN SPAM Act might have a marginal deterrent effect, with CAN SPAM regulations effecting law abiding email marketers who are risk averse, yet increasing the severity and illegality of that spam which remains (Hui, Kim, & Wang, 2013). If spam increases with severity in relation to CAN SPAM Act activity, it may suggest this marginal deterrent effect. Additionally the measures capturing the severity of spam might also be similar to research question 2 on CAN SPAM Act compliance, as both can capture the severity of spam emails.

Question (4) relates to the locality associated with spam emails. The CAN SPAM Act only has jurisdiction in the United States, the question remains whether the Act’s introduction in the US and its enforcement relates to whether spam is associated with internet addresses and domain registrations associated with the United States. There may be a displacement effect, with emails shifting to other countries in response to CAN SPAM activity. Spam locality has been found to be unrelated to the CAN SPAM Act in prior research (Kigerl, 2009). However, the current research seeks to extend these findings with more sophisticated measures of both the CAN SPAM Act and measures of spam locality as well.

Sample and Data

The sample data, that of spam emails and spam behavior, was taken from publicly available spam archives from which spam emails are collected and stored for subsequent
download by researchers. The data was retrieved from the Untroubled Software website (http://untroubled.org/spam) on December 18, 2013. The Untroubled Software website is maintained by Bruce Guenter, and the available spam archives are collected by posting multiple “honey net” email addresses publicly online for spam crawlers to harvest. The honey net approach is intended to bait spammers to add a given email address to a spam listserv, with the goal of intentionally receiving spam emails. When an email address is posted on internet websites such as forums, message boards, and on personally hosted web pages, web bots may scan and identify them as email addresses, extracting and storing them in a spamlist or database for subsequent spam targeting. The data used for this research was collected in such a manner and made available for download at the Untroubled Software website.

The data collected includes a purposive sample of all individual spam emails hosted for download which were received in bait honeypot email accounts between March, 1998 and November, 2013, totaling 5,490,905 email messages in total. No emails are excluded that fall between these dates, and the sample includes the entire population of available emails starting in March, 1998 up until November, 2013 when the data was pulled. Each email is encoded in an individual text file containing the contents of the spam email, which includes header information, body of the message including any scripts or HTML, and any file attachments the email contained, converted to a plaintext format stored at the end of the file with an encoding scheme called BASE64. Each of the approximately five and a half million emails were scanned by software written specifically for this study that extracted 17 variables or features representing each individual spam email.
Procedures

In order to be analyzed in a time series design, the spam email sample was coded on a number of dimensions relevant to this study’s research questions. However, the five and a half million emails gathered were too numerous for a human rater to code manually. Software was written in Java to parse each and every message in the sample which served as an automated rater. The software read each email individually, file-by-file, evaluating its textual content, headers, attachments, and any other relevant features of the email that must be logically identified. The data points were then extracted by the software and saved in a tabular dataset for cleaning and analysis. Each row of the dataset represents an individual spam email message, with 5,490,905 rows total.

Each email is in a plaintext format, including any file attachments included in them. File attachments are able to be stored as plaintext in a text file because the binary files are encoded using BASE64. Base64 can store binary data in a plaintext ASCII string represented as a number with 64 unique digits total, including alphanumeric and some special ASCII characters (Josefsson, 2006).

For some of the emails in the sample, the entire contents of the emails are stored as BASE64 text, including the body message of the email, but not including the headers. Library files available to the Java programming language that can decode BASE64 text have been incorporated into the software. The software was written to handle such emails, recognize any BASE64 text encoding, and decode it back into human readable text so that the software can code the variables based on the email contents.

In some cases text or HTML files are attached in the emails. These are also stored in BASE64 strings in each email. For the purposes of coding some of the variables, the software
decodes these files and parses the text contained in them for scoring some of the measures. For instance, for determining whether an email links to an executable file (likely malware), the software needs to scan the textual content of the link. However, some spam includes the links in attachment files (text files, etc.), not in the body of the email itself. Therefore the software can decode any plaintext files and read them to search for malicious links as well as the email body itself. Non-plaintext binary files will be ignored, as they cannot contain malicious links, etc., in plaintext that is readable by the software.

The resulting output of the software was a comma separated values (CSV) file with a row for each of the emails in the sample. Because some of the variables which are identified by the software are complex and contain elements of subjectivity (fraudulent vs. non-fraudulent spam messages based on the message’s textual content, or spam emails with informative subject field lines vs. those with deceptive subject lines, etc.), the accuracy of the software is not entirely certain. That is, before the data can be analyzed, some of these more complex measures will have to be tested for interrater reliability, by comparison of the software’s ratings to that of an actual human rater. Some of the measures can be more easily detected by a human being. For instance, whether the subject field is meaningful is something humans can easily identify, but software may make more errors. A random selection of 150 cases from the software’s output dataset will be coded by a human rater for scoring of some of the more complex variables. Cohen’s kappa will be used to test for agreement between the human and machine raters.

The contents of the spam CSV data was imported into SPSS in order to be aggregated into a monthly time series dataset. Each of the spam variables were aggregated by month (count of spam emails per month, percent of emails with malicious links per month, etc.), subsequently
shrinking the size of the dataset to just 189 observations, as there are 189 months in the sample. The smaller dataset was then exported to R for analysis\(^1\).

**Measures and Variables**

There are three sets of measures that are included in subsequent models. The first set includes the dependent variables, or measures of spam activity intended to answer each of the four research questions (regarding spam rates, spam compliance, spam severity, and spam locality). The second set includes independent variables representing CAN SPAM Act activity, including enforcement, CAN SPAM Act attention and public awareness, attitudes toward the CAN SPAM Act, and lack of spammer anonymity due to attribution to the spammer’s identity due to the CAN SPAM Act’s enforcement. Finally, a number of economic and technological time series predictors are available for inclusion in each model, to serve as control variables. Each of these is thought to possibly relate to spam or illegal behavior in some way, and are discussed in more detail below.

*Dependent Variables: Spam Activity*

The dependent variables, spamming activity, were all extracted from the spam mining software discussed in the procedures section above. There are 17 different measures capturing one of the four aspects of spam this research hopes to address (spam volume, compliance, severity, and locality). Some of the measures are similar enough that they were tested for reliability such that they can be collapsed into a single scale measuring spam behavior. Each of the spam measures is broken down into one of the four research questions the predictor is intended to capture. All measures of spamming outcomes are lagged by one month in all subsequent regression analyses.

\(^1\) The data will not be aggregated in R as R is unable to handle large datasets.
Research Question 1: Spam Volume

The spam emails were aggregated by month and incorporated into a measure capturing the absolute number of spam emails received per month. The spam mining software records a column representing the date each email was received in the bait email account inbox used to collect the spam data. The date was extracted from the header information of each email.

Emails have a number of dates or timestamps representing a message’s transit over the internet towards its destined recipient. Each timestamp is associated with an email’s “hop” or a transfer between routers or servers on a network. Each hop appends a new header record to the top of an email message, including information such as the server facilitating the hop, the date the message was received by the server while in transit, any authentication details about the message itself, among other things. Because each hop results in a header appended to the top of the email message headers, the top-most timestamp found in the message’s headers can be assumed to be the most recent hop. Therefore, it can be assumed to be the date the message was received. It is this date that the software will record.

The date received will be used and not the date sent, because the date the message was sent would be recorded lower in the message’s headers, which are more likely to be falsified. Email sender’s can forge the initial headers before an email is sent, as the email sending protocol (SMTP) does not always authenticate messages prior to forwarding them on (Haskins, 2004). Once sent, the remaining headers appended to the message are more likely to be accurate, as the spammer has less control over the routing servers.

The date is identified by the software via regular expressions written in Java (a language for matching search patterns in textual data). The software will pull the first date matched from the top of the headers. The pattern to match is any numerical digit one or two characters in
length (the day), followed by a three character string representing the month (‘jan’, ‘feb’, ‘mar’, etc.), followed by a four digit number (the year). A rare few emails contain no dates in them, an obvious sign of header forgery and tampering. Other emails contain impossible dates (12/95/2005). Emails in the sample with invalid or missing date information are eliminated from the dataset. The data was then aggregated by month based on the date received measure. A time series plot of this measure can be seen in Figure 1.

Figure 1
Plot of Spam Volume Per Month Time Series

Note the large spike in spam volume during late 2006. Such a large and sudden increase in spamming activity was noted from the spam archives source website to not be due to a genuine increase in the spam population, but rather due to a technical change to how the honeynet bait email client used to collect the spam sample was set up. During the three months of August through October of 2006, wild card addresses were accepted by the mail server, allowing misspelled recipient user names in the recipient address field to be successfully
delivered anyway. To correct for these three outliers in the data, the three months were deleted and replaced with interpolated values based on the trend and contiguous values of the time series data.

*Research Question 2: Spam Compliance with the CAN SPAM Act*

There are five regulations stipulated in the CAN SPAM Act that the software is intended to rate each spam email on in terms of compliance. There are 10 separate dichotomous measurements of whether a message complies with any of the five rules in the CAN SPAM Act. The regulations in the Act include the prohibition of falsified headers, requirement of non-deceptive subject lines, provision of an opt-out mechanism in the email, inclusion of a valid physical postal mailing address, and the notice that the message is an advertisement somewhere in the contents of the message.

The 10 binary indicators of CAN SPAM Act compliance will also be tested for scale reliability, to determine if compliance with one regulation predicts compliance with another. The measures may be able to be collapsed into a single dimension or scale to be used as a dependent variable during time series analysis. However, lack of significant reliability amongst this construct does not necessarily suggest the measures of compliance are poor. Rather it might indicate that compliance with one CAN SPAM regulation is due to chance, and not due to regulations in the CAN SPAM Act at all. That is, spammers comply for reasons other than CAN SPAM Act regulation. Separate tests for reliability will have to be performed both before and during the CAN SPAM Act’s enforcement, as well as another reliability test combining the two time periods.

*Falsified Headers.* The prohibition against falsifying email headers is the first regulation set forth in the CAN SPAM Act (CAN SPAM Act, 2005). Unlike most of the rules contained in
the Act, the prohibition against falsified headers applies to email messages whether they are considered traditional commercial electronic mail messages are not, which is the central type of message the Act is intended to target. The rule also applies to relational or transactional messages, not just commercial solicitations or advertisements contained in email. These include confirmation receipt emails delivered after a purchase, or emails the recipient has opted in to receive.

The Act forbids any of these messages from containing email header information that is materially false or misleading. Email headers are equivalent to traditional physical mail envelope headers in that they include routing information, such as names, sender and recipient addresses, and specific to email, router and mail server details (timestamp, IP address, etc.). Any falsification of a message’s from, to, or other header address fields are considered a violation under the Act.

The software is written to identify seven different true or false measures of evidence for email header falsification. The first measure, duplicate from address, is coded ‘1’ if the sender’s email address is the same as the recipient’s email address. If they are different, the measure is coded ‘0’. The software also records a second variable containing the actual email address the software identified as being identical between the sender and recipient fields. This will not be used in analysis, but instead might be needed as a source of evidence and documentation for the software’s reasoning. A time series plot of this measure can be seen in Figure 2. The measure is operationalized as the percent of email message falsifying headers in such a way per month. An Augmented Dickey Fuller test revealed this measure to not be significantly trend-stationary (Dickey-Fuller = -3.2, p = .09). Single differencing of the measure resulted in significant
stationarity ($\text{Dickey-Fuller} = -6.68, \ p < .01$). The differenced version will be used in the final regression model.

Figure 2  
Plot of Percent Spam with Duplicate From and To Fields Per Month Time Series

Email senders can also optionally set a reply-to field in an email, that may be different from the sender’s “from” email address. Most email clients will use the reply-to field, and not the from field, so long as it is available, when a user attempts to reply to an email. This is usually a legitimate feature for email senders, as commercial marketers may have separate departments for sending and receiving email for a given marketed product. That is, one department may send the email advertisement, while another is intended to handle recipient responses or questions to that email advertisement (Definitions and implementation under the CAN-SPAM Act, 2008).
However, similar to having a sender address that is the same as the recipient’s address, if the reply-to field is identical to the recipient’s address, then the message is clearly falsified. When email users attempt to reply to this spam message, the reply ends up being routed back to their own inbox. Messages such as this are coded ‘1’ under duplicate reply-to address, for any messages with a reply-to email address that matches the recipient’s address. A second column is also created to store the address the software identified as a duplicate for reference purposes.

The measure is operationalized as the percent of duplicate “from” and reply-to fields per month, seen in Figure 3.

Figure 3
Plot of Percent of Spam with Duplicate ‘To’ and Reply-To Fields Per Month Time Series

The same methodology for identifying falsified headers above was also applied to a message’s return-path field. That is, if a return-path specified matches the recipients address, a duplicate return-path column is coded ‘1’, and the email address in question is recorded in an additional column. The return-path of an email is typically created by the recipient’s mail server,
and appended to the top of the headers. The return-path is a reference for where to send email replies when replying to a given email message, assuming a reply-to field is not specified (Wood, 2009). However, spammers can inject falsified return-path fields into the headers when the message is composed. The software will identify the first return-path field in a message (the lowest one matched in the headers), as that would be the field subject to possible falsification.

The time series measure can be seen in Figure 4.

Figure 4
Plot of Percent of Spam with Duplicate ‘To’ and Return-Path Fields Per Month Time Series

Another indicator of sender falsification of the “from” address would be if the from field or address is missing entirely. Some spam messages do not have any clear from address contained in the message headers (Spykerman, 2012). These messages can be forwarded on to the recipient anyway as the originating mail servers did not authenticate the sender’s message.
Messages with a sender address found are coded ‘1’. The resulting time series is represented in Figure 5.

Figure 5
Plot of Percent of Spam with ‘From’ Field Found Per Month Time Series

Senders of email can also falsify their names, not just email addresses, when sending spam. Senders can also fail to include their name entirely. The software codes all emails that have a sender name included ‘1’, with any emails with missing names coded ‘0’. A time series plot of the resulting data is represented by Figure 6. The measure was found to be serially dependent ($Dickey-Fuller = -0.81, p = .96$), and regular differencing significantly removed such dependency ($Dickey-Fuller = -5.056, p < .01$).

Figure 6
Plot of Percent of Spam with Sender Name Found Per Month Time Series
Finally, if the recipient’s email address is missing entirely, this is also unlawful under the CAN SPAM Act. Missing “to” fields in email may indicate the sender included the recipient’s email in a BCC field, obscuring the recipient’s email address as well as the likely numerous other recipients the spammer contacted. If the “to” field is identified as present by the software, a column will be coded ‘1’, otherwise it is coded ‘0’ to indicate an absent ‘to’ field. A plot of the time series can be seen in Figure 7. The data was not found to be significantly trend stationary \( (Dickey-Fuller = -1.83, p = .647) \). Regular differencing resulted in significant stationarity \( (Dickey-Fuller = -7.61, p < .01) \).

Figure 7
Plot of Percent Spam with ‘To’ Field Found Per Month Time Series
The six falsified header indicators will be tested for scale reliability, to attempt to predict if falsification of one header field relates to falsification of other header fields. Some of the binary measures are mutually exclusive, however. For instance, if the message has missing “from” or “to” fields, then the message cannot possibly also have used the same address in the “to” field as that of the sender’s “from” field. Some of these measures will have to be collapsed into a single dichotomous indicator before inclusion in any scale reliability tests.

**Misleading Subject Heading.** The CAN SPAM Act requires meaningful subject headings contained in emails. Recipients of a commercial email message must be able to determine the contents of the message are a commercial advertisement or promotion of a product or service from reading the email subject line alone. The intent is for recipients to be able to decide if they
want to open and read the message or not, or whether to just delete the email as spam. Deceptive subject lines (e.g. including “hey there!” in a subject field for an email promoting knock off Rolex watches) do not indicate what the contents of the email message is.

Figure 8
Plot of Percent Spam Subject Field Compliance Per Month Time Series

If the software identifies a message as complying with the CAN SPAM Act’s requirement of a meaningful subject heading, the measure is coded ‘1’; otherwise the measure is coded ‘0’ for failure to comply with the regulation. In order to accomplish this, the software breaks down every word, separated by spaces or the beginning or end of a line, into a list. Highly common words, such as function words (e.g. the, and, or, not, with) will be eliminated from the list. The words contained in the email body will also be composed into a list, each list item being a single word. If any word contained in the subject field list matches any other word contained in the email body list, the measure will be coded ‘1’ as it seems likely to have a meaningful subject heading. The software will also record the subject being matched in a
different column for later reference if needed. A figure representing this measure over time can be seen in Figure 8.

It is not certain the accuracy of such an algorithm, so the measure will have to be tested for interrater agreement before use in any statistical models. A random subset of spam emails will be selected. The same random subset will also be provided to an independent human rater who will rate each message in terms of its compliance and how meaningful the subject line is. Depending on the significance of agreement between human and software, the dependent variable will either be used or discarded for consideration as a time series outcome.

Dishonoring Opt-Out requests. Commercial emailers are required to both provide a means for recipients to request emailers discontinue sending them solicitations and honor those requests within 10 business days. Emailers are required to provide a link to a web page where recipients can opt-out, or instruct the recipient to respond to the email message requesting they be taken off the mail list. The opt-out mechanism may not require more than visiting a single web page or responding to a single email message.

A message is coded ‘1’ to indicate it complies with the requirement of providing an opt-out mechanism if the body contains one or more of three possible keyword-based matches in the body of the email. If the terms “opt out”, “opt-out”, or “unsubscribe” are not contained anywhere in the email body, regardless of whether the matches are surrounded by spaces or other letters, the message will be considered to not be in compliance with the opt-out regulation and coded ‘0’. The matching is not case sensitive, as all text will be converted to lower case. The matched pattern (“unsubscribe”, “opt out”, etc.) is also recorded in a separate column. The monthly time series plot of this measure is displayed in Figure 9. The measure was not found to
be trend stationary ($D\text{ickey-Fuller} = -1.85, p = .64$). Differencing of the data resulted in sufficient stationarity ($D\text{ickey-Fuller} = -6.7505, p < .01$).

Figure 9
Plot of Percent Spam Opt-Out Compliance Per Month Time Series

The opt-out measure will also be included in the list of variables a human rater will be required to code in order to test for interrater agreement. Other keywords not included in the software might be used by spammers to inform recipients of their unsubscribe options. Terms such as “opt out” might also be used in a context other than offering an actual unsubscribe mechanism.

Invalid Physical Address. Spammers are required to provide their valid physical mailing address, so as to reduce their anonymity. Initially, after the formulation of the CAN SPAM Act,

---

2 Additionally, emails might also simply not genuinely honor opt-out requests, but that is beyond the scope of both the software and the human rater.
there was no clear indication that PO Box addresses were considered acceptable physical addresses to be included in commercial email messages. In 2008 the FTC clarified that such PO Box addresses were in compliance with the Act (Definitions and Implementation Under the CAN-SPAM Act, 2008).

Figure 10
Plot of Percent Spam Address Compliance Per Month Time Series

The software creates a dichotomous variables indicating if a regular street address is matched or a PO Box address is matched. The street address compliance measure will be coded ‘1’ if the software matches one or more numbers followed by one or more spaces, followed by a non-cases-sensitive street direction (NE, SW, N), followed by one or more alphanumeric characters (street name), and ending with a street suffix of some sort (ave, st, ct, rd, blvd, etc.) or if it matched a PO Box patter such as “po box” or “post office box” (non-case sensitive) followed by a number. A time series measure of this variable can be seen in Figure 10. The
measure will be tested for interrater agreement to determine if a human rater agrees with the software’s judgment on email compliance with the regulation.

**Fail to Include Notice of Advertisement.** Messages must contain clear and conspicuous identification that the message is an advertisement or solicitation. The CAN SPAM Act does not specify an exact formal means for identifying emails in such a way, but mostly leaves it up to the bulk emailer to include such “inconspicuous” identification. Typically such emails include a description at the bottom of the message that states the message is an advertisement.

The software will mark a message as in compliance if one of three of the following patterns are matched in either the subject field or email body: the message includes the string “advertisement”, “advertisment” (misspelled), or “adv”, so long as “adv” is not surrounded by letters, digits, or underscores. The matching will not be case sensitive. Because of the broad number of ways text can be written to explicitly communicate that a message is an advertisement, the measure will need to be tested for interrater agreement. The software also records the matched expression in an additional column to facilitate understanding why some matches fail to find agreement with a human rater. A time series of this measure is represented in Figure 11.

---

**Figure 11**  
Plot of Percent Spam Advertisement Compliance Per Month Time Series
Research Question 3: Severity and Illegality of Spammer Behavior

The software also captures spam behavior that is not explicitly forbidden under the CAN SPAM Act, such as outright fraud and malware distribution. These types of violations are left to other laws and jurisdictions (both Federal and state) to enforce. These offenses are not regulated under the CAN SPAM Act in part because they are more severe than the type of spam the Act was intended to deter, and are thus already regulated by other and existing statutes. While not under the jurisdiction of the CAN SPAM Act, they are still behaviors attributed to email spam, and subsequently may be influenced by the CAN SPAM Act’s enforcement. Specifically, it is questioned if there is a marginal deterrence effect that the CAN SPAM Act might impose, such that the frequency of spam violations may decrease, but that the severity of spam’s illegality
might actually experience an increase. Such evidence has been supported while studying DDoS attacks as a form of cybercrime (Hui, Kim, & Wang, 2013) and so the same is intended to be investigated for spammer behavior.

**Probability a Message is Fraudulent.** The software uses a re-purposed spam filter to calculate the probability a message is fraudulent, as opposed to just regular and non-fraudulent email spam. Spam filters are used to calculate the probability a message received is spam, as opposed to legitimate email that the users wants to receive/read (ham) (Zdziarski, 2005). Spam filters typically are trained on a sample of two data sources, a collection of spam emails that we already know are spam, as well as a collection of ham emails (legitimate messages) a trainer has already identified as legitimate. Spam filters scan the two data sources and calculate base probabilities that a given keyword found in an email is spam based on the keyword frequencies found in these two training data samples. The spam filter can then predict the probability a new set of test or actual email is spam based on these probabilities.

The spam filter employed for this research that the software uses is a naive Bayes classifier, which attempts to classify an instance of text as either being in one of two dichotomous categories based on trained probabilities associated with the text’s keyword frequencies (Conway & White, 2012). Naive Bayes is a Bayesian statistical method, an alternative to a frequentist statistical approach, in that it calculates the probability of the hypothesis (that the message is spam) given the data (the keyword frequencies). This is contrasted with the frequentist methods which calculates the probability of the data (that demonstrates an effect size) given the hypothesis (that there is no effect size in the population).

However, to calculate the probability of the hypothesis (that the message fraudulent), the probabilities that a keyword is found in fraudulent emails must first be calculated. An existing
open source naive Bayes spam filter was acquired from the Code Project website (http://www.codeproject.com/KB/recipes/BayesianCS.aspx, retrieved September 12, 2010). The software is written in C# and was repurposed to create a dataset of trained probabilities that a word is found in a fraudulent message.

The filter is intended to build probabilities to identify spam, but can be re-purposed to classify any text on any dichotomous category given the training algorithm is provided with examples of true positive (fraudulent spam text files) and true negatives (non-fraudulent spam text files). In the context of spam, fraudulent spam emails are those that tend to fall under the categories of either advance fee fraud or phishing scams. That is, the message is intended to deceive the recipient into wiring the fraudster money or entering personal information the scammer wants to misuse, typically for financial gain. Non-fraudulent emails tend to be selling a legitimate or semi-legitimate product, such as pharmaceuticals. There may be some level of deception in these emails, such as promoting the sale of genuine Rolex watches but really the product delivered is a knock-off Rolex watch. However, these emails are not outright fraud, and it is expected that the fraudulent spam classifier would label these as simply regular spam.

There’s a two-step process in order to employ fraud classification on the spam archives data sample: training the base keyword probabilities and then incorporating those probabilities into the software that can calculate whether an email instance is fraudulent. The training of the classifier will be performed using two spam email samples, one taken from the spam archives itself, and another pulled from a separate fraud spam database available online. Approximately 1,000 spam emails will be taken from the existing spam archives used in this sample that are confirmed via a human rater to be non-fraudulent in nature. This will be used as the non-
fraudulent training sample, and will be excluded from the final dataset building process by the software to avoid classifying messages the software was trained on.

The second sample, consisting of the fraudulent training emails, has already been pulled from the Scamdex website (http://www.scamdex.com), which hosts a publicly available database of user-submitted reports of email scam contents. Because the accuracy of a spam filter can degrade if the sample used to train the classifier is older than the new messages it is intended to scan (Cormack & Cruz, 2009), fraudulent spam messages were sampled from the years 1998 to 2013 in order to capture a temporally wide range of messages.

Fraudulent spam messages available on the website cannot be downloaded in large bundles, but are instead available as single web pages per message. In order to download a large sample of fraudulent spam messages, a web crawler written in C# was created to automatically download just the body of each scam message as a text file. The Scamdex website was crawled on March 13, 2013, downloading the first two index pages of scam email samples per year from 1998 to 2013, of about 155 messages per year. A total of 2,339 scam messages were downloaded to be used as a scam training sample.

The reason the scam training sample was sought from outside the actual spam archives the software will be classifying is due to the greater proportion of non-fraudulent spam emails relative to fraudulent ones in the current sample. In order to identify a single fraudulent email, many emails would have to be read manually, either consuming time or possibly shrinking the sample size of the fraud training set. To expand sample size, a separate source of exclusively scam-related emails was sought.

After both the scam and non-scam sample are prepared and ready, a dataset of keyword probabilities will be computed based on these two data sources using the aforementioned
software acquired from the Code Project website. The software was derived from spam filtering
code written by Paul Graham (2002), and originally formulated in the programming language
Lisp. An example of the algorithm the software employs can be seen in formula (1) below,
written in Lisp.

\[
\begin{align*}
&\text{let } ((g (* 2 \text{ (or (gethash word good) 0)})))
\text{ (b (or (gethash word bad) 0)))) \\
&\text{unless } (< (+ g b) 5)
\text{ (max .01}
\text{ (min .99 (float (/ (min 1 (/ b nbad))}
\text{ (+ (min 1 (/ g ngood))}
\text{ (min 1 (/ b nbad)))))}))
\end{align*}
\]

The algorithm of the software will break down the samples into two lists of keywords,
one list for non-scam emails (“good”, in the formula above) and another for scam emails (“bad”
in the formula above). Accompanying each keyword (“word”) is the frequency the word appears
throughout the sample (“g” for non-scam and “b” for scam emails). All keywords that sum
together between both lists that exceed a frequency of four will be analyzed, as is done in the
above algorithm. Those that do not exceed four will be dropped.

Referring to the algorithm in formula (1), a ratio of the frequency of a given keyword
found in scam emails divided by the count of all scam messages (“nbad”) in the sample over the
frequency of the word among both scam and non-scam emails divided by the total number of
emails (“nbad” and “ngood”) between both samples will calculate the probability that the
keyword is found in scam emails. If there is a higher probability of the message being
fraudulent, the probability will be greater than .5. If the keyword is found only in scam emails, it
is automatically assigned a probability of .99. If only found in non-scam emails, it is assigned a
probability of .01.

Following training of the classifier, a dataset of keyword probabilities will be created.
The probabilities dataset will be uploaded to Amazon Web Services along with the spam mining
software. The software will compute a measure (fraud) that is the continuous predicted probability a given message is fraudulent based on the trained keyword file. The software employs a combined probability algorithm, an example of which can be found in formula (2) below. The software will select the 15 highest scored (highest chance of being fraudulent) keywords from each email message and combine their probabilities to compute the fraud measure.

\[
\frac{Pr_1 \times Pr_2 \times ... \times Pr_n}{Pr_1 \times Pr_2 \times ... \times Pr_n + (1-Pr_1)(1-Pr_2)...(1-Pr_n)}
\]

(2)

To date there is little literature testing the efficacy of this algorithm in successfully classifying messages based on a training sample. The true positive rate, and therefore accuracy, of the algorithm is not entirely known. For this reason, the measure will need to be tested for interrater agreement on a smaller subset of spam before inclusion in any statistical tests. If a human rater agrees with the classifier significantly and substantially, the measure will be used.

The fraud probability measure was aggregated into two time series variables, one representing the average probability a message is fraudulent per month (see Figure 12) and the percent of emails with 85% or higher predicted probabilities of being fraudulent per month (see Figure 13). Both average fraud probability per month (Dickey-Fuller = -1.24, \( p = .9 \)) and percent of messages classified as fraudulent (Dickey-Fuller = -1.41, \( p = .82 \)) were not found to be trend stationary on their own. Regular differencing resulted in significant stationarity for fraud probability (Dickey-Fuller = -7.41, \( p < .01 \)) and fraud classification (Dickey-Fuller = -8.26, \( p < .01 \)) per month.

Figure 12
Plot of Average Spam Fraud Probability Per Month Time Series
Executable Download Link. Likely a more common attack vector involving the distribution of malware in spam is the inclusion of malicious web links in spam emails. Because the email client is not able to scan or detect the scripts or executable nature of the website the user is being linked to, the local software or mail server is less able to warn and protect the recipient from following a link that looks suspicious. A common method is the drive-by-download, where simply visiting a web page with the recipient’s browser launches scripts on that page that exploit a vulnerability in the user’s browser that automatically downloads and installs malicious software hosted on the web server.

While this is a popular and effective method among cybercriminals for malware distribution, it is beyond the scope of the spam mining software as the spam sample dates back to 1998, in which case most of the malicious links or websites are certainly no longer online today.
and cannot be verified to be malicious. Instead, the software detects if an email links directly to a file download, and will determine if the file extension is executable if it matches a list of known executable file extensions. The list of executable file extensions used was found on the About.com website (http://pcsupport.about.com/od/tipstricks/a/execfileext.htm, retrieved March 21, 2013).

Linking directly to a malicious file download usually doesn’t exploit any vulnerability that triggers an automatic install after download. Instead the file is downloaded and it is up to the spam recipient whether they choose to open the file or not. A limitation is that a link may not directly include the filename in the URL, but rather dynamically route the victim to a downloadable file after following the link. However, for the purposes of this research, the software will only detect direct download links, although that may bias the data in favor of more novice malware distributors.

The software scans both the body of the email as well as any plaintext attachments for these executable download links. Some of the emails attach HTML files, and so may contain the malicious links in there. All plaintext attachments will be decoded from the BASE64 encoding for scanning described above. The software will match any text that begins with “http://” and ends with a dot (.) followed by an executable file extension that matches the available list of executable file extensions discussed above. The first executable link found in the email will be recorded and coded ‘1’ for executable download link found, otherwise ‘0’. If multiple executable links are present in the email, only the first one will be matched and coded as being found. A time series representation of the measure can be found in Figure 14.

Figure 14
Plot of Percent of Spam with Executable Links Per Month Time Series
**HTML Scripts.** A second means in which emails can be used to distribute malware is by running executable scripts that are embedded in the email itself which the email client, not the operating system or web browser per se, executes. Emails can be formatted with HTML, the same language used to design web pages. HTML can also include script tags, which an HTML interpreter or script engine, such as that in a web browser or email client, can execute. Most email clients disable running scripts in email because of its possibility for abuse, so it is rare for an email sender to include script tags in email. When an email does include such tags, it is likely for malicious purposes, such as installing malware.

The software matches any opening script tag embedded in the email body, or if applicable, any attached HTML files in the email. If the software matches “<script”, followed by zero or more of any character of any length so long as there is no line break, ending with a closing bracket (“>”), the software codes an executable script tag variable as ‘1’ for true,
otherwise ‘0’. The measure was then aggregated by month as the percent of emails with executable script tags over time. The aggregate measure is represented in Figure 15.

Figure 15
Plot of Percent of Spam with Embedded Scripts Per Month Time Series

Research Question 4: Location Associated with Spam

A possible effect of a law against cybercrime, such as spam, is that cybercrime is rarely hindered by international borders. Law enforcement, however, is hindered by jurisdiction and international borders. An increase in enforcement may have an impact in the given location in which it has jurisdiction, which may influence cybercriminals or attacks situated there. However, cyberattacks from outside the jurisdiction from other countries in the world can be just as problematic among that locality’s residents. In light of this, it is questioned whether CAN SPAM Act enforcement or attention has had any influence on the country associated with spam

---

3 Scripts in emails tend to be of the form “<script language='JavaScript'>”, and so the pattern matcher would identify such text.
email received in the United States. Perhaps it displaced spam activity to jurisdictions beyond the CAN SPAM Act’s reach.

Any information contained in a spam email message that can be geolocated to a given country is not an accurate estimate of where the spammers themselves live, as spam is almost never sent from the spammer’s own home. However, it may be an estimate of where spammers choose to host their spam sending bots, or where they choose to register their web domain names. It is suspected that the nation-state associated with elements contained in a spam email are meaningful and nonrandom in their causes, and some evidence suggests these variables are in fact predictable in certain ways (Kigerl, 2012).

**Country Associated with Email IP Address.** While an email message is in transit, routing servers will append address information into the email headers, including IP addresses (IPA). IPAs are finite and allocated to different countries. That is, if an IP address falls within a range that belongs to the United States, it can be assumed that that IPA is an internet address for a server located in the US. As an email is routed to its destination over the internet, it may cross national borders, and contain a record of this transit in the headers as new routing locations are appended to the top of the headers. The lowest-most IPA that can be found in the headers is assumed to be the first server the message was forwarded from, and so ought to be closest to the country of origin from which the email was sent, assuming the headers are not forged in any way.

The Software matches the lowest-most IPA found in the headers that is in the form of four numbers (octets) separated by a dot (e.g. 255.255.255.255), so long as each octet is not less than zero or more than 255. The software will then look up the matched IPA in a geographical IP address database (downloaded from http://software77.net/geo-ip/ on March 24, 2013) to
determine the country associated with the IPA. If the software returns with an error indicating the IPA is private (not an internet IPA, but rather one associated with a private internal network not associated with a country), a private IPA count variable will be incremented to count the number of private IPAs found. If private, the software will move on to the next header in the email and extract the IPA, repeating the process. If there is a match, the software will record the country name in a country IPA column.

Tracking the number of private IPAs found will be important, as the more hops up we move in an email headers means the IPAs are less likely to be associated with the spam bot’s or sender’s country of origin, assuming no tampering with the headers. Private IPAs associated with the sender of an email do not necessarily indicate fraud, as some email clients do this purposely to respect user privacy. It can represent fraud, however, if an email is sent through a relay or VPN. The software is not able to distinguish this, though.

The categorical measure of IPA country is dichotomized as ‘1’ if the country of origin is from the United States, and ‘0’ for those associated with any other country. Emails that could not be geolocated were eliminated. The measure was further aggregated into a time series as the percent of emails associated with the United States per month. Preliminary analyses will be conducted using this time series metric to determine if it relates to any of the falsified header time series measures.

Figure 16
Plot of Percent US IPA Spam Per Month Time Series
Specifically, it is questioned whether emails with falsified header information are more or less likely to appear to originate from within the United States. For instance, if the sender email address is falsified, it is likely that the IPA is falsified as well. Whether these types of header tampering evident in the sample predict the country of the IPA will be investigated. Of course, just because an email is not identified to have falsified headers by the software does not mean that the IPA was not falsified.

The time series measure can be seen in Figure 16. The measure was not found to be sufficiently trend stationary \((\text{Dickey-Fuller} = -2.62, \ p = .315)\). Differencing of the data resulted in significant stationarity \((\text{Dickey-Fuller} = -6.28, \ p < .01)\). All analyses of this measure will rely on the differenced version.
**Link Country Top-Level-Domain (TLD).** Spam emails may provide a web link for recipients to follow. The hosted web pages found by following such links have a top-level-domain (TLD) (e.g., .com, .net, .org). Some TLDs are internationalized, meaning they are associated with a specific country (e.g., .us for the United States, .ca for Canada). Internationalized TLDs are typically reserved only for the country they are associated with, meaning they can only be registered within the nation (Kigerl, 2012). Most do not place citizenship restrictions on who can register the domain, but some do. In this way, TLDs may be more accurate in some cases than the IPA.

The software matches any link that has a TLD which belongs to a specific country. The software scans the body of the email and any plaintext attachments the email may have. Text will be matched that begins with “http://”, is followed by any number or combination of alphanumeric symbols, hyphens (-), or dots (.), that end with a single dot and a string of characters that match a list of internationalized TLDs. The first internationalized TLD is matched, if any are found, ignoring any that might follow.

The country code is recorded under a column for link country TLDs. Like the country IPA measure above, country TLD is dichotomized (‘1’ for the United States, ‘0’ for any other country or no US TLD found), and aggregated as the percent of TLDs associated with the United States per month. The aggregate measure can be seen plotted in Figure 17.

---

Figure 17
Plot of Percent US TLD Spam Per Month Time Series
The measure will be tested against the IPA variable as well, to determine if the country associated with an email’s IPA is related to that associated with a linked TLD. Unlike some nationalized TLDs, the .us TLD requires proof of US citizenship to register (NeuStar Inc., 2013). However, private citizens were not permitted to register the US TLD domain until April of 2002, so time series analyses will begin during this month in addition to the usual analyses that begin in March of 1998, to determine if such differences change outcomes.

Independent Variables: CAN SPAM Act Activity

The impact to test against the spam outcome measures (of spam volume, compliance, severity, and locality) is the activity, enforcement, and attention of and towards the CAN SPAM Act. Most of the measures are taken from news and media attention about the CAN SPAM Act, but an additional time series metric has been created from Google search history data. Theoretical underpinnings of deterrence regarding the CAN SPAM Act are intended to be captured and tested with the independent measures, including attention towards CAN SPAM
convictions and monetary damages won, attribution and lack of anonymity of spammers in the media, news that is critical of the CAN SPAM Act vs. news that is not, as well as internet search activity of the CAN SPAM Act.

All measures that derive from news reports on the CAN SPAM Act are acquired from a series of LexisNexus searches. The search results are limited to only those news sources that are located within the United States. All of 347 unique news articles that were returned from a combination of CAN SPAM search terms (can spam act, can-spam, etc.) were downloaded and coded on the four sets of measures that derive from news sources.

**CAN SPAM Act Enforcement**

The quantity and severity of CAN SPAM Act enforcement and prosecutions highlighted by the media were captured from the news results, such as the number of prosecutions, convictions, the amount of damages awarded during lawsuits, etc. Prior research has used this methodology to test law enforcement efforts against hacking and cybercrime in respect to its effect on hacking attempts against the countries in the sample (Png & Wang, 2007). Cybercrime enforcement was associated with a reduction in malicious intrusions collected by intrusion detection systems at the national level. The same methodology will be tested using CAN SPAM Act prosecutions on spam related activity.

There are eight time series measures of CAN SPAM Act enforcement and deterrence from the news article sample coding. The damages awarded in CAN SPAM Act lawsuits is operationalized as two different time series variables, one representing the sum of damages awarded in USD among all articles each month, the other simply representing a count of articles that award any damages per month. Whether a suspected spammer was detained in jail or prison is also operationalized into two separate measures in such a way. The sum of days
detained among all articles per month and the count of articles mentioning spammer detention per month were created. There is also a variable representing the count of articles mention any arrests of suspected spammers per month amongst the news articles.

There are also three count variables describing trials involving spammers under the CAN SPAM Act. The first is a count of all articles that describe an ongoing, so far unresolved trial under the CAN SPAM Act. The second indicates the count of articles mentioning a spammer was convicted under the CAN SPAM Act. Finally, a count of spammers acquitted of CAN SPAM charges per month is included. There is also a measure representing the percentage of articles coded that related to the CAN SPAM Act, out of all articles total. Not all articles were about the CAN SPAM Act, but still pertained to spam in some way.

News Critical of the CAN SPAM Act

Much of the initial attention the CAN SPAM Act received when it was introduced was not positive (Arora, 2006; Grimes, 2007; Lee, 2005; Zeller, 2005). Users, experts, and bloggers criticized the Act for not requiring opt-in, for preempting stricter state laws forbidding spam, and for being weakly enforced as spammers were convicted but never were forced to pay the damages they owed. Naturally this kind of reporting could have the opposite effect of deterrence, emboldening spammers located in the United States.

For each news article, three dummy variables were coded indicating the opinion expressed by the author. The measures captured whether the article is negative towards the CAN SPAM Act, neutral towards it, or positive in the CAN SPAM Act’s worth and effectiveness. All articles related to spam, but not all articles described the CAN SPAM Act. Articles were also coded on their attitudes towards spam in general in addition to their attitudes towards the CAN SPAM Act, whether they were optimistic about efforts to combat spam or believed that the
growing spam problem would only worsen. Four time series variables were created from the coded articles. Two related to attitudes towards the CAN SPAM Act, representing both the percent of all articles negative about the Act per month and also the percent of all articles positive about the Act per month. The remaining two measured just attitudes towards spam in general, operationalized as the percent of articles negative about the state of spam in the United States per month and the percent of articles positive about our odds of fighting spam per month.

**Attribution of Spammers**

The impersonal and anonymous nature of crimes perpetrated in cyberspace, such as that of spam, can attenuate some of the deterrent effect a legal punishment might impose. Individuals who would otherwise not commit risky illegal acts might be tempted to do so if they feel safe behind a computer screen while committing a crime. This problem involving cybercrimes is considered “lack of attribution”, where a potential cyberattack cannot be deterred if the defenders/enforcers are unable to attribute the source of the attack (such as revealing the identity of the individual behind the attack) (Glaser, 2011).

Attribution has been operationalized in prior literature from news articles and tested for its possible impact on cybercrime arrest and conviction rates at the national level (Guitton, 2012). Attribution was measured in the media by the number of news articles that discussed cybercrime where the attacker was known, vs. discussing cybercrime when it was unknown the identities of the cybercriminals. Attribution of this sort was associated with fewer cybercrime attacks within each country studied.

This research seeks to perform the same analysis for spam related cybercrimes. News articles discussing spam without identifying specific spammers or spam gangs and their locality have been coded ‘0’ indicating lack of attribution. News reporting that reveals the identity of
any given spammer (such as during a prosecution) has been coded ‘1’ for true attribution. The coded articles have been aggregated into two monthly time series: the percent of all articles with attribution or where the spammer is known, and also the percent of all articles without attribution where the spammer is not known.

*Dichotomous Impact of the CAN SPAM Act*

A simpler measure of the CAN SPAM Act will also be used, representing a before and after intervention variable representing the months in which the CAN SPAM Act was being enforced and in effect as a law. The CAN SPAM Act impact measure is coded ‘0’ for any time before January 1, 2004, before the CAN SPAM Act went into effect. Any time on or after this date the measure is coded ‘1’, true, for CAN SPAM enforcement. This measure has been used in prior assessments of the CAN SPAM Act (Kigerl, 2009), but was only used for a smaller dataset with fewer outcome measures of spam and no control variables.

*Google CAN SPAM Act Search History*

Google offers reporting of time series plots for popular terms searched for using the Google search engine, called Google Trends (http://www.google.com/trends). Users can plot the search history of a given set of keywords and download the data in a spreadsheet. Some research has used Google Trends search histories to predict actual outcomes, such as flu outbreaks (Carneiro & Mylonakis, 2009). Searches using terms associated with crime and criminal justice have been associated with economic time series, such as consumer confidence (Choi & Varian, 2012). However, research utilizing Google Searches is in its nascent stages and should be considered new ground.

It is suspected that Google searches for the CAN SPAM Act ought to reflect public awareness of and attention towards the law. If awareness of the CAN SPAM Act by spammers
has an influence their behavior (e.g. a deterrent influence), then general searches of the CAN SPAM Act ought to be correlated with spammer awareness and perhaps even deterrence. Multiple time series of different CAN SPAM search queries (e.g. “can spam”, “can spam act”, “can-spam act”) were downloaded from the Google Trends service and merged into a single time series representing the count of all searches related to the CAN SPAM Act per month. Searches were limited to only those within the United States.

Control Variables: Technological, Economic, and Demographic Predictors

Possible influences on spam from sources other than enforcement and awareness of the law are also to be taken into account. Little research evaluating the effectiveness of spam has investigated the temporal ordering of the CAN SPAM Act followed spam related activity (either volume or compliance). Additionally, no known studies have considered possible confounding measure that theoretically predict spamming behavior. Some have pointed out that spam had risen immediately following the passing of the CAN SPAM Act (Zeller, 2005). However, spam had been on the rise since its inception, not likely due to any increased criminality among internet users capable of sending spam, but simply from the mere growth in technology and the number of those internet users over time. Evaluations of spam law should take into account these possible confounds before making any conclusions.

Accounting for control variables also allows this research to better approximate causality in drawing conclusions about the data. A simple time series regression can estimate a predictor time series (the CAN SPAM Act) relation to an outcome (spam), lagged by one or two months. Such an estimate would establish correlation and temporal ordering, but would not be able to rule out spurious factors. That is why control measures would be appropriate, so as to make it more
difficult to argue that spam increased simply because of a parallel increase in information and communication technology and the number of internet users per capita.

The control measures are also important to create a model that better captures the elements of routine activity theory (RAT), which is important for understanding cybercrime. CAN SPAM Act enforcement is specified here as the capable guardianship of RAT, whereas control measures such as the number of internet ready PCs or unemployment rates might serve as suitable targets and motivated offenders, respectively.

**Technological Predictors**

**Internet Users Per Capita.** Time series estimates on the number of internet users per capita in the United States were acquired from the Pew Internet Research website (http://pewinternet.org/Trend-Data-(Adults)/Internet-Adoption.aspx, retrieved March 25, 2013). The data is the result of regular Pew Internet and American Life project surveys on respondents self-reported internet usage. The data is an irregularly spaced time series, meaning the intervals of time between each observation are not constant (some observations are two days apart, some two weeks apart, some one month apart). Any observations less than one month apart have been averaged for that month, and any observations with more than one month between them have their missing months interpolated. The data begins in June, 1995 and ends in March, 2013. Prior research using cross sectional data of internet users per capita at the national level suggests the number of internet users is positively correlated with spam volume (Kigerl, 2012). Internet users per capita are perceived to be a measure of both suitable targets and motivated offenders (as victims and spammers alike must be internet users). The measure was not found to be sufficiently trend-stationary ($Dickey-Fuller = -2.13, p = .52$). Regular differencing resulted in stationarity ($Dickey-Fuller = -5.66, p < .01$).
Tech Jobs: Computer Systems Design and Related Services. A measure of the availability of technology jobs was downloaded from the Federal Reserve Economic Data website (http://research.stlouisfed.org/fred2/series/CES6054150001, retrieved December 17, 2013) which was initially compiled from the Bureau of Labor Statistics. The series measures the number of employees over time in occupations such as software developers, computer systems analysts, and database administrators. However, the definition the Department of Labor employs of “computer systems design and related services” is very broad, and so also includes employed in occupations such as financial specialists and loan officers, which is a limitation of the data. The data is monthly and begins in January, 1990, and ends in November, 2013. The measure was not found to be sufficiently trend-stationary (Dickey-Fuller = -2.66, p = .3). First and second differencing resulted in significant stationarity (Dickey-Fuller = -6.76, p < .01).

Wilshire Internet Market Index. The Wilshire Internet Market Index was created in 1997 to capture the performance of the entire internet market sector. The index measures the price and the total returns on investments (the performance) of publicly traded internet stocks. The data measuring the index was acquired from the Federal Reserve Economic Data website (http://research.stlouisfed.org/fred2/series/WILLWWW, retrieved December 17, 2013). The observational time period begins in January, 1997 and extends into March, 2013. Missing values are interpolated. The metric is intended to serve as a proxy measure of the market demand and success of those pursuing careers involving internet tech sector companies. It may relate to spam volume, as well as serve as a measure for motivated offenders, should the market be weak for legal internet and related career opportunities. The measure was not found to be sufficiently trend-stationary (Dickey-Fuller = -2.3, p = .45). Regular differencing resulted in stationarity (Dickey-Fuller = -5, p < .01).
Economic Predictors

Real Disposable Personal Income Per Capita. Disposable income is an individual’s personal income minus taxes owed on that income. The data used was downloaded from the Federal Reserve Economic Data website (http://research.stlouisfed.org/fred2/series/A229RX0Q048SBEA, retrieved December 18, 2013). The metric captures real disposable income (adjusted for inflation for 2009 US dollars), per capita. The data begins in January of 1947 and ends May of 2013. Missing values for the time series measure have been interpolated so that the data ends in November. The measure was not found to be sufficiently trend-stationary (Dickey-Fuller = -0.71, p = .97). Regular differencing resulted in stationarity (Dickey-Fuller = -5.78, p < .01).

Gross Domestic Product (GDP) Real Growth Rate. GDP real growth rate was acquired from the Organization for Economic cooperation and Development’s StatExtracts database (http://stats.oecd.org/Index.aspx?QueryId=350, retrieved December 17 26, 2013). The measure is adjusted for inflation, and captures the growth in GDP over a quarterly basis, instead of the level of GDP itself. Because of this, the measure might not have to be differenced as there is likely to be less trend or serial correlation in the series. Differencing data de-trends a time series by replacing each observation in time with the difference between itself and one or more lagged observations. Because the data is quarterly, intermediate monthly values have been interpolated. While it is uncertain how growth in GDP might relate to spam, tests using this same archival spam data has linked actual GDP per capita to higher spam volume at the national level (Kigerl, 2012). The data begins in October, 1976 and ends in the quarter of September, 2013. The measure was not found to be sufficiently trend-stationary (Dickey-Fuller = -2.99, p = .16). Regular differencing resulted in stationarity (Dickey-Fuller = -4.23, p < .01).
Unemployment Rate. Monthly unemployment rate data was acquired from the Federal Reserve Economic Data website (http://research.stlouisfed.org/fred2/series/UNRATE, retrieved December 17, 2013). The data originally comes from the US Department of Labor, Bureau of Labor Statistics. The series starts January, 1948 and ends November, 2013. Unemployment may also serve as a metric for motivated offenders, as high rates of unemployment coupled with a high number of internet users within a country has been associated with more spam and more phishing TLDs being sent from that nation (Kigerl, 2012). Unemployment has also been linked to hacking attempts at the national level (Png & Wang, 2007). The measure was not found to be sufficiently trend-stationary ($Dickey-Fuller = -2.13, p = .52$). Regular differencing resulted in stationarity ($Dickey-Fuller = -8.21, p < .01$).

College Graduate Labor Force Participation Rate. College graduate labor force participation is the percent of the civilian labor force in the United States with a Bachelor’s degree or higher presently employed. The data includes only those 25 years of age or older. The data originates from the U.S. Department of Labor, Bureau of Labor Statistics. The time series variable was downloaded from the Federal Reserve Economic Data website (http://research.stlouisfed.org/fred2/series/LNU01327662, retrieved December 25, 2013). The data is monthly and begins in January, 1992 and ends in November, 2013.

Consumer Price Index. The consumer price index measures the inflation level and spending power of the average United States household to purchase from a fixed list of consumer goods. The data originates from the U.S. Department of Labor, Bureau of Labor Statistics. The time series variable was downloaded from the Federal Reserve Economic Data website (http://research.stlouisfed.org/fred2/series/CPIAUCSL/, retrieved December 18, 2013). The
measure was not found to be sufficiently trend-stationary \((Dickey-Fuller = -2.49, p = .37)\).

Regular differencing resulted in stationarity \((Dickey-Fuller = -6.4, p < .01)\).

**Financial Stress Index.** The financial stress index measures the amount of financial stress in the markets and is build from 18 time series datasets capturing interest rates, yield spreads, and other indicators. The metric is scaled so that higher scores represent more stress on the United States markets. The data originates from the Federal Reserve Bank of St. Louis. The time series variable was downloaded from the Federal Reserve Economic Data website (http://research.stlouisfed.org/fred2/series/STLFSI/, retrieved December 18, 2013). The data begins in January, 1994 and ends in November, 2013. The measure was not found to be sufficiently trend-stationary \((Dickey-Fuller = -2.96, p = .17)\). Regular differencing resulted in stationarity \((Dickey-Fuller = -5.81, p < .01)\).

**Demographic and Other Predictors**

**United States Population Size.** Population size has been found to predict phishing TLDs (Kigerl, 2012) and so will be tested in terms of predicting overall spam activity, including spam link TLDs. The data was downloaded from the Federal Reserve Economic Data website (http://research.stlouisfed.org/fred2/series/POP, retrieved December 17, 2013). The series is monthly, and the first observation begins in January of 1952 and ends November, 2013. The measure was not found to be sufficiently trend-stationary \((Dickey-Fuller = -2.08, p = .54)\). Differencing resulted in stationarity \((Dickey-Fuller = -20.85, p < .01)\).

**Percent of Population 15-24 Years of Age.** Youth and younger adults in their late teens or early twenties tend to account for a disproportionate amount of crimes committed (Blonigen, 2010). It is also true that younger generations tend to be more involved with emerging technology (Bauerlein, 2011), so there might be several reasons why age might relate to spam
crimes. Data on percent of the population ages 15 to 24 years was acquired from the Federal Reserve Economic Data website (http://research.stlouisfed.org/fred2/series/LFWA24TTUSM647S, retrieved December 17, 2013). The data originates from the Organization for Economic Cooperation and Development (OECD). The data begins January, 1977 and ends October of 2013. November of 2013 was the only missing data point in the series and so was interpolated. The measure was not found to be sufficiently trend-stationary (Dickey-Fuller = -.76, p = .97). Regular differencing resulted in stationarity (Dickey-Fuller = -6.54, p < .01).

Arrest Rates. Monthly Uniform Crime Reports on the number of arrests per capita was downloaded from ICPSR (http://www.icpsr.umich.edu/icpsrweb/content/NACJD/guides/ucr.html, retrieved December 25, 2013). The measure represents the number of arrests per month per 100,000 individuals in the population. The series begins in January of 1990 and ends December of 2010. All missing values have been interpolated.

Analytic Plan

Interrater Agreement

Five measures will need to be tested for interrater agreement. They include the CAN SPAM compliance measures of an accurate subject heading, notice of advertisement, unsubscribe option, valid physical mailing address, and whether the message appears fraudulent. A human coder would be better able to identify these five features of a spam message. However, a human coder would not be able to rate five million emails. The study proposes to test the software’s ratings against a human’s with a smaller sample of emails a human coder could read. If statistically significant agreement is observed between the human and software, the software
will be rated as reliable and thus able to be used in subsequent analyses. A sample of 150 spam emails will be randomly selected from the entire spam archive for a human to read, about 10 emails per each of the 16 years (from 1998 to 2013). The 150 emails will be coded by the software during a test run, and the emails will also be provided to an independent coder. Cohen’s Kappa will be used to test the agreement among the five dichotomous variables (Cohen, 1960). The measure of fraud will be converted into a dichotomous measure for the purposes of these tests (any value greater than .85 will be coded ‘1’, otherwise fraud will be coded ‘0’).

**Scale Reliability**

Unidimensionality will be tested for amongst some of the spam measures, to determine if they can be combined into a scale. The measure groupings to be tested will include the falsified header measures (seven items), the CAN SPAM Act noncompliance measures (with and without falsified headers included, four to eleven items), malware distribution (two items), and country or origin (a two item scale). Cronbach’s alpha will be used to determine if a scale represents a single construct (Cronbach, 1951), with a cut-off of around .7 as an indicator of the scale’s unidimensionality (Santos, 1999). Measures that can be combined will be used as a dependent variable, aggregated by month to represent the average scale value (by adding each dichotomous item together per email) of spam emails during that time period.

**Missing Monthly Values**

As mentioned, months represent the unit of analysis. However, some of the measures are not measures in months and contain missing values. Because time series measures tend to be serially correlated, with each observation dependent on the next and preceding contiguous observations closest to it in time, interpolation is appropriate to estimate missing values. Interpolation estimates missing time series values based on these adjacent and serially correlated
observations (Chow & Lin, 1971). The method of interpolation used for the data is the Kalman filter, which breaks a univariate time series into three components: trend, seasonal, and level/noise time series (Brookner, 1998; Pizzinga, 2012). Trend represents the slope or basic direction of the time series, seasonal movements are the cycles that repeat themselves at regular intervals in the data, and level/noise represent unaccounted for variation or the precise value of the time series not including the trend or seasonal cycles. The Kalman filter smoothes across these time series to create estimates of missing values, and finally combining them back into the original series, only without missing cases.

*Serial Correlation Tests*

Prior to correlation analysis, each time series measure has been checked and corrected for serial dependency. Due to series trends, time series are often correlated with other time series, increasing the risk of false positives. For instance, if two series both have upwards trends, comparing them observation for observation will appear as though they covary. As observations are continually higher in time for both series due to trend, a higher score for one series predicts a higher score for the other series at the same time point. The Augmented Dickey Fuller has been used to check for significant serial correlation (or trends) amongst the time series measures. For series with a significant trend, differencing has been used to ensure the series are trend stationary (Shumway & Stoffer, 2011). A time series that has no trend is said to be “trend stationary.”

*Multivariate Analyses*

Once the data has been sufficiently differenced, multiple generalized least squares (GLS) regression models will be conducted to test the impact the CAN SPAM measures have on the separate spam activity time series outcomes lagged by one month, net other economic and technological controls. GLS allows for a non-constant variance among the residuals to be
controlled for, as may be the case in time series models due to serial correlation of the residuals (Judge, Griffiths, Hill, Lutkepohl, & Lee, 1986). The gls function available in the nlme R package will be used (Pinheiro, Bates, DeRoy, & Sarkar, 2013).

Predictor variables included in each of the 17 regression models were selected via backward stepwise regression based on model AIC. This functionality is provided by the “step” function included in the “stats” package in R. Stepwise regression utilizing AIC is conducted via backward elimination of predictor variables that ends once elimination results in a higher AIC score. Predictors that increase the AIC score when eliminated are retained. The AIC method was chosen over other stepwise selection procedures because non-criterion based selection tends to bias $R^2$ upwards, p-values downwards, and is more susceptible to overfitting (Whittingham, Stephens, Bradbury, & Freckleton, 2006). While other methods might yield more highly significant models, the accuracy of the significance level is less trustworthy. Because the purpose of significance is to determine the generalizability of the findings to the population being examined, AIC methods would produce less biased p-values. All 17 regression models used started with 33 independent and control variables, following backwards elimination to select the strongest models per measure of spamming outcomes.

With each of these models specified, the residuals will be tested to investigate signs of serial autocorrelation. Autocorrelation and partial autocorrelation will be tested in the residuals for each model, and for any substantial serial correlation, autoregressive or moving average arguments will need to be specified in final GLS models to account for heteroscedasticity (McLeod & Mahdi, 2011). This process will be repeated for each regression model, one for each of the 17 spam outcome time series, with varying combinations of CAN SPAM Activity accounted for.
Chapter 6: Results

Descriptives

The initial dataset was built with Java software deployed on the spam sample. The software generated a dataset consisting of 5,490,905 cases, each case representing an individual spam email sent and received during the study period. The software coded each message according to 16 different variables and recorded their results in the dataset. Table 1 depicts descriptives for each of the 16 variables broken down into two periods: that of individual spam emails prior to the CAN SPAM Act’s enforcement, and individual spam emails after the CAN SPAM Act went into enforcement.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Before the CAN SPAM Act (n = 97,528)</th>
<th>After the CAN SPAM Act (n = 5,393,377)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M/SD (%)</td>
<td>n</td>
</tr>
<tr>
<td>US IPA</td>
<td>41.89 (40.664)</td>
<td>52.16 (2,813,017)</td>
</tr>
<tr>
<td>US TLD</td>
<td>.61 (592)</td>
<td>.15 (8,037)</td>
</tr>
<tr>
<td>Opt Out Compliance</td>
<td>23 (22,282)</td>
<td>7 (375,734)</td>
</tr>
<tr>
<td>Address Compliance</td>
<td>&lt; .0001 (104)</td>
<td>&lt; .0001 (2,797)</td>
</tr>
<tr>
<td>Subject Compliance</td>
<td>79 (77,344)</td>
<td>41 (2,228,338)</td>
</tr>
<tr>
<td>Advertisement Compliance</td>
<td>4 (3,975)</td>
<td>&lt; .0001 (21,975)</td>
</tr>
<tr>
<td>Fraud Probability</td>
<td>4.19 (18.5)</td>
<td>6.54 (2.25)</td>
</tr>
<tr>
<td>Fraudulent, true/false</td>
<td>3.06 (2,988)</td>
<td>4.75 (256,054)</td>
</tr>
<tr>
<td>Malicious link</td>
<td>&lt; .0001 (64)</td>
<td>&lt; .0001 (600)</td>
</tr>
<tr>
<td>Malicious Script</td>
<td>&lt; .0001 (282)</td>
<td>&lt; .0001 (5,000)</td>
</tr>
<tr>
<td>Header ‘From’ Found</td>
<td>99 (96,121)</td>
<td>97 (5,241,778)</td>
</tr>
<tr>
<td>Header ‘To’ Found</td>
<td>86 (83,983)</td>
<td>98 (5,263,488)</td>
</tr>
<tr>
<td>Duplicate ‘From’ ‘To’ Fields</td>
<td>5.13 (5.005)</td>
<td>5.53 (298,453)</td>
</tr>
<tr>
<td>Duplicate ‘To’ Return-Path</td>
<td>4.57 (4.457)</td>
<td>1.47 (79,081)</td>
</tr>
<tr>
<td>Duplicate ‘To’ Reply-To</td>
<td>1.39 (1.356)</td>
<td>.44 (23,902)</td>
</tr>
<tr>
<td>Sender Name Found</td>
<td>60 (58,919)</td>
<td>82 (4,446,030)</td>
</tr>
</tbody>
</table>

Between the beginning of the study period in March, 1998 and before the enforcement of the CAN SPAM Act prior to January, 2004, only 97,528 spam emails were sent and received in the spam sample. After this period and ending in November, 2013, a total of 5,393,377 spam
emails were sent and collected in the spam sample. The first two measures in Table 1, US IPA (IP address) and US TLD (top-level domain), indicate the locality of spam, or the percent of emails associated with the United States. The percent of emails with US IPAs appear to be higher following the passing of the CAN SPAM Act, but only slightly by 10 percentage points. The percent of emails associated with US TLDs, however, appear to have gone down following the passing of the CAN SPAM Act. Also note that US TLDs tend to be fairly rare in the sample, with only 8,629 emails total that link to them out of millions.

The next four measures include the CAN SPAM Act compliance variables: opt out provision, physical mailing address inclusion, meaningful subject field, and notice of advertisement compliance. For each of these four measures, the percent of emails in compliance is lower following the passing of the CAN SPAM Act.

The next four measures represent spam severity, whether the message is fraudulent or attempts to distribute malicious software. The first measure, Fraud Probability, represents the average predicted probability a message is fraudulent. Prior to the CAN SPAM Act, the average probability of fraud was 4.19%. Following the Act, the average probability is not much higher at 6.54%. The second measure, Fraudulent, true/false, represents the percent of emails with a predicted probability of fraud 85% or greater. Prior to the Act, only 3.06% messages were fraudulent. Following the Act, the percent of fraudulent emails are more than twice as large at 6.54%.

The remaining two measures out of the group represent malware distribution, Malicious Link and Malicious Script. Both are lower following the CAN SPAM Act, suggesting less malware distribution by these methods. However, the number of emails marked as distributing malware in such ways is very small (664 and 5,282, respectively). It is likely that these are
simply novice methods of distributing malware that more experienced spammers would no
longer use and that would slowly be phased out as more effective methods are developed. More
sophisticated malicious links would not have the executable file extension contained in the URL.
Rather, the executable portion of the malware would be embedded in the website being directed
to itself. Additionally, most email clients today would likely block scripts contained in emails.
The reductions in these measures likely represents a shift to other malware distribution

The remaining six measures represent header falsification. Two header tampering
methods change little over time (Duplicate ‘From’ ‘To’ Fields and Header ‘From’ Found
matches). However, the remaining methods suggest increased compliance with header
tampering regulations following the CAN SPAM Act, with fewer duplicate ‘to’ and return-path
and reply-to instances, and a higher percentage of emails including a sender name and recipient

The data was then aggregated into a monthly time series, with 17 time series spam
variables in total. Table 2 depicts descriptives of the 17 time series variables. There are 17
instead of the 16 variables in Table 1 because of the additional Spam Count measure
incorporated into the time series data. The average number of spam emails sent per month prior
to the CAN SPAM Act was 1,393. Following the CAN SPAM Act, the average was 45,322.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Before the CAN SPAM Act (n = 70)</th>
<th>After the CAN SPAM Act (n = 119)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spam Count</td>
<td>M (SD)</td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td>1,393 (1,925)</td>
<td>33 – 8595</td>
</tr>
<tr>
<td>Percent US IPA</td>
<td>52.89 (15)</td>
<td>23 – 88</td>
</tr>
<tr>
<td>Percent US IPA</td>
<td>.14 (0)</td>
<td>0 – 3</td>
</tr>
<tr>
<td>Percent Opt-Out Compliance</td>
<td>29.62 (9)</td>
<td>1 – 55</td>
</tr>
<tr>
<td>Percent Address Compliance</td>
<td>.17 (0)</td>
<td>0 – 1</td>
</tr>
<tr>
<td>Percent Subject Compliance</td>
<td>85.78 (8)</td>
<td>66 – 96</td>
</tr>
</tbody>
</table>
The remaining measures represent the percent of emails per month that were
dichotomously classified by the spam mining software according to each variables criteria, with
the exception of Average Fraud Probability, representing the average fraud probability per
month. The results of the descriptives tend to mimic that of the individual spam descriptives in
Table 1.

**Interrater Reliability Testing**

Because of the difficulty in writing an artificial intelligence that can make assessments
just as accurate as the human brain, the sample coded by the software needed to be compared
with that of a human rater. Some of the codings were of narrative text, a type of data that are
much more easily processed and understood by humans than by software. Initially a random
sample of 150 emails was drawn from the software coded spam dataset, with cases selected at
least once per each of the 16 years of the study period. However, this was not a large enough
sample to capture a sufficient amount of variation among all the variables selected for interrater
agreement testing. All variables were dichotomous, and some measures, such as notice of
advertisement compliance and physical mailing address inclusion, were rarely marked as true
coded ‘1’ by the software. Therefore after initial random sampling of 150 emails, an additional
purposive sampling with replacement was conducted until all variables marked for testing were true at a minimum of 10 times out of the 150 emails. This allowed for sufficient variation for significance testing.

The measures tested for interrater agreement were the probability a message is fraudulent at 85% certainty, as well as four of the CAN SPAM Act compliance measures. Eighty five percent certainty was selected as it was the most significantly related to the 33 predictor variables tested for out of a series of certainty levels (60, 75, 80, 85, 90, 95, 99). The compliance measures included opt-out provision, address inclusion, meaningful subject field, and advertisement notice compliance. A dataset containing the file names for each of the 150 emails was created with all machine coded variable results stripped out. A new set of five blank variables was created, and a human coder read and rated each of the emails in the sample for each of the five variables.

Following coding, the machine-rated and the human-rated datasets were matched and compared for interrater agreement using Cohen’s Kappa. All measures were determined to have significant interrater agreement between machine and human, although not all effect sizes were large. The agreement on message fraud classification was significant and moderate ($Kappa = .48, p < .001$). Agreement on opt-out compliance ($Kappa = .6, p < .001$), address inclusion ($Kappa = .71, p < .001$), and meaningful subject headings ($Kappa = .69, p < .001$) was higher. Agreement on notice of advertisement compliance was low but significant ($Kappa = .32, p < .001$), still suggesting better than chance agreement between raters. Given that all results were at least statistically significant at a sample size of only 150, all measures will be included in subsequent regression analyses.
Scale Reliability Testing

Investigations were completed to determine if any measures coded by the software could be combined into scales measuring higher dimensionality of spam behavior. All tests were conducted using Cronbach’s Alpha. The four major CAN SPAM Act compliance measures were tested together for scale reliability. These included opt-out provision, address inclusion, meaningful subject field, and notice of advertisement compliance. The items were not found to be reliable ($\alpha = .17$), suggesting compliance with one CAN SPAM Act requirement does not predict compliance with other requirements. The finding also suggests that the compliance found is not actually compliance with the CAN SPAM Act, but rather something the spammer would have done anyway with their spam message construction. For instance, some might include an address as part of their spam business to solicit mail orders, rather than doing so because of having read the CAN SPAM Act. Any compliance found is likely a coincidence.

United States IPA and United States TLD were tested for reliability. Only cases with a nationalized TLD link included in the email were used. Not only were the two unreliable, but the alpha was negative due to negative average covariance among the two items ($\alpha = -.002$). A Pearson correlation between the two yielded an extremely small negative correlation coefficient ($r = -.003$, $p = .003$). There appears to be no relationship between a US TLD and a US IPA of the email sent. While US TLDs are highly likely to be registered by United States citizens, there is probably no relationship between the nation of an email IPA and the spammer’s own country of residence.

Lastly, the falsified header measures were tested for scale reliability. A scale composed of missing recipient address, sender address, and sender name was found to not exceed reliability alpha criteria ($\alpha = .38$). A scale composed of duplicate ‘from’ and ‘to’ addresses, duplicate
‘to’ and reply-to addresses, and duplicate ‘to’ and return-path addresses were also found to be unreliable ($\alpha = .25$). Falsification of one type of email header is not a consistent indicator for falsification of another, at least to a sufficient degree to warrant combining items into a unidimensional scale. Due to the unreliability of items, no scales were created nor used in subsequent multivariate analyses.

**Multivariate Analyses**

*Time Series Regression of Spam Count per Month*

A backward stepwise regression using the model AIC was conducted with spam count per month regressed on an initial 33 predictor variables total. Backward elimination yielded seven predictor variables to be included in the final time series model. The selected predictors were included in a linear regression model and the residuals were computed to test for autocorrelation. A Durbin Watson test of the residuals indicated significant autocorrelation of the residuals ($dw = .36, p < .001$). The Durbin Watson test measures the presence of autocorrelation of the residuals, testing the null hypothesis that there is no autocorrelation present. All variables in the model were differenced and included in a second OLS regression model. Differencing data is a technique used to attempt to remove the trend from a time series. The process replaces each observation in time with the difference between that observation and the contiguous observation in the series. Instead of representing the level of a time series observation in time, the resulting data represents the change in movement of one observation to the next in time. The residuals of the differenced model were not found to have significant autocorrelation ($dw = 1.77, p = .088$), although they were marginally significant. The differenced version of the model was therefore used in the final analysis.
After inspection of the autocorrelation function (ACF) and partial autocorrelation function (PACF) correlograms of the regression (see figures A.1 and A.2), an ARMA(1,2) model was identified. A correlogram is a visual plot of the correlation between a time series and itself at a given lag, for successive increments of lags one and up (McDowall, McCleary, Meidinger, Hay, 1980). The ACF is a simple correlation of a time series with itself at a given lag, while a PACF represents the correlation of a time series with itself at lag k, controlling for all lags in between itself and lag k. While differencing of the data is sufficient to create a trend stationary time series process, there may still be some degree of serial dependency in the data. ACF and PACF functions can reveal such serial dependency and indicate that said processes need to be controlled for in any subsequent regression models. If there are q spikes in an ACF correlogram, a moving average model of order q should be controlled for (ARMA(0,q)). If there is a decay pattern in the ACF function, an autoregressive parameter at order p should be controlled for (ARMA(p,0)), p being inversely proportionate to the speed of decay. The reverse interpretation is required of the PACF correlogram, with a spike indicating an autoregressive process and decay patterns indicating a moving average process.

A final generalized least squares time series regression model was run using the seven selected and differenced predictors and included in Table 3. Percent of the United States population who are internet users was found to be significant and negative ($B = -0.023$, $p = .035$), suggesting that more internet users per capita were associated with less spam sent the following month. This finding is contrary to what routine activity theory would suggest, as internet users ought to reflect the number of motivation offenders or suitable targets. Spammers have to be internet users, and recipients of spam also have to be internet users.

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Generalized Least Squares Time Series Regression of Spam Count per Month, ARMA(1,2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 189$</td>
<td></td>
</tr>
</tbody>
</table>
The population size aged 15-25 is positively associated with spam at lag 1 \((B = .028, p = .031)\). This finding might indicate that the population concentration of youth is related to cybercrime, or at least spam crime, in similar ways as it is related to traditional street crime. However, the population size in general is also predictive of spam \((B = .055, p = .032)\), and has a stronger effect size. This may simply mean that more people also means more spammers, and may have little to do with youth. The outcome is absolute, not a rate, and population size is also absolute, therefore one can anticipate this association.

The remaining predictors in the model are the CAN SPAM Act and deterrence independent variables. The count of articles that mention an unresolved ongoing CAN SPAM trial per month is associated with less spam sent the following month \((B = -.027, p = .025)\). The count of articles that mention a spammer being convicted is also associated with less spam, but only approached significance \((B = -.024, p = .052)\). Both findings suggest a possible deterrent effect of CAN SPAM trials on spam volume sent per month. However, the count of articles mentioning a spammer being detained per month is positively associated with spam the following month \((B = .043, p = .004)\).

Finally, the percent of articles that are negative about our ability to combat spam as a proportion of all articles on spam is associated with higher spam volume \((B = .013, p = .037)\). A
dejected attitude of authors in the news regarding spam may embolden spammers to continue engaging in their spam sending operations.

*Time Series Regression of Percent US IPAs Per Month*

A backward stepwise regression using the model AIC was conducted with percent of spam with United States IPAs per month regressed on an initial 33 predictor variables total. Backward elimination yielded five predictor variables to be included in the final time series model. The selected predictors were included in a linear regression model and the residuals were computed to test for autocorrelation. A Durbin Watson test of the residuals found no evidence of significant autocorrelation ($dw = 2.03, p = .58$).

After inspection of the ACF and PACF correlograms of the regression (see figures A.3 and A.4), an ARMA(0,1) model was identified. A final generalized least squares time series regression model was run using the five selected predictors and included in Table 4. The monthly financial stress index is the only control variable that made it into the model. However, it only approaches significance ($B = -.115, p = .07$). The measure’s direction does suggest that financial stress in the United States shifts spam IPAs outside of the country the following month.

**Table 4**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Beta</th>
<th>SE</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>.04</td>
<td>.071</td>
<td>.565</td>
<td>.5725</td>
</tr>
<tr>
<td>Financial Stress Index</td>
<td>-.115</td>
<td>.063</td>
<td>-1.821</td>
<td>.0702†</td>
</tr>
<tr>
<td>Count of Trial Ongoing Articles</td>
<td>-.175</td>
<td>.081</td>
<td>-2.162</td>
<td>.032*</td>
</tr>
<tr>
<td>Count of Spammer Detained Articles</td>
<td>.28</td>
<td>.082</td>
<td>3.438</td>
<td>.0007***</td>
</tr>
<tr>
<td>Percent of Articles Negative About CAN SPAM</td>
<td>.068</td>
<td>.036</td>
<td>1.859</td>
<td>.0647†</td>
</tr>
<tr>
<td>Percent of Articles Negative About Spam</td>
<td>-.122</td>
<td>.039</td>
<td>-3.089</td>
<td>.0023**</td>
</tr>
</tbody>
</table>

$R^2 = 11.43\%$

*< .05, **< .01, ***< .001, †< .1
The count of articles mentioning ongoing, as yet unresolved trials per month was significant and negative ($B = -0.175, p = 0.032$). That is, more CAN SPAM trials were associated with fewer spam emails with IPAs associated with the United States the following month. The finding could represent a possible deterrent or displacement effect, as spammers choose not to host their spambots in the United States where the CAN SPAM Act has jurisdiction. However, the count of articles mentioning a spammer being detained was associated with more United States spam IPAs ($B = 0.28, p = 0.0007$), which is contrary to a deterrent explanation. Note that these two variables’ theoretical directions are identical here as they were for Table 3 regarding spam counts. Ongoing trials suggest deterrence, while spammer detentions indicate the opposite.

The remaining two predictors relate to author attitudes towards the CAN SPAM Act and spam in general. While the proportion of articles negative about the CAN SPAM Act is associated with more US IPAs the following month, the predictor is not quite significant ($B = 0.068, p = 0.065$). Despite this, the effect size direction might suggest an emboldening effect. Contrary to this, however, the percent of articles negative about the state of spam indicates the opposite ($B = -0.122, p = 0.002$). Less optimistic attitudes about spam in the news within the United States appears to shift spam IPAs to other countries.

*Time Series Regression of Percent US TLDs Per Month*

Prior to analyses, all observations older than April, 2002 were deleted from the US TLD dataset. This was done as registration of United States TLDs did not become available to the American citizenry until April, 2002. This resulted in an $n$ of 140 cases. A backward stepwise regression using the model AIC was conducted with percent of spam linking to URLs with United States TLDs per month regressed on an initial 33 predictor variables total. Backward elimination yielded 18 predictor variables to be included in the final time series model. The
selected predictors were included in a linear regression and the residuals were computed to test for autocorrelation. A Durbin Watson test of the residuals found significant autocorrelation ($dw = 1.29, p < .001$). Regular differencing of the entire dataset yielded a second regression with no evidence of serial correlation among the residuals ($dw = 2.16, p = .942$).

After inspection of the ACF and PACF correlograms of the regression (see figures A.5 and A.6), an ARMA(2,1) model was identified. A final generalized least squares time series regression model was run using the 18 selected predictors and included in Table 5. Very few of the control variables are significant. Of the seven included, only one is significant at the .05 level. GDP growth per month predicts fewer United States TLDs linked to in spam emails the following month ($B = -.059, p = .044$). Wealthier countries may be more likely to be targets by spammers looking to make a profit. GDP growth in the United States may make the nation a bigger suitable target by foreign spammers.

### Table 5
Generalized Least Squares Time Series Regression of Percent US TLD Spam Link Per Month, ARMA(2,1)

<table>
<thead>
<tr>
<th>Measure</th>
<th>$Beta$</th>
<th>$SE$</th>
<th>$t$</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>.016</td>
<td>.023</td>
<td>.68</td>
<td>.4972</td>
</tr>
<tr>
<td>Percent Internet Users</td>
<td>.015</td>
<td>.032</td>
<td>.468</td>
<td>.6402</td>
</tr>
<tr>
<td>Unemployment Rate</td>
<td>-.073</td>
<td>.046</td>
<td>-1.595</td>
<td>.1126</td>
</tr>
<tr>
<td>Internet Market Index</td>
<td>-.048</td>
<td>.036</td>
<td>-1.353</td>
<td>.1778</td>
</tr>
<tr>
<td>Population Aged 15-25</td>
<td>.314</td>
<td>.311</td>
<td>1.009</td>
<td>.3144</td>
</tr>
<tr>
<td>Percent Population Aged 15-25</td>
<td>-.344</td>
<td>.398</td>
<td>-.865</td>
<td>.3881</td>
</tr>
<tr>
<td>GDP Growth</td>
<td>-.059</td>
<td>.029</td>
<td>-2.034</td>
<td>.0435*</td>
</tr>
<tr>
<td>Financial Stress Index</td>
<td>-.049</td>
<td>.038</td>
<td>-1.276</td>
<td>.2038</td>
</tr>
<tr>
<td>Count of CAN SPAM Articles</td>
<td>-.32</td>
<td>.068</td>
<td>-4.706</td>
<td>&lt;.0001***</td>
</tr>
<tr>
<td>Sum of CAN SPAM Damages Awarded</td>
<td>-.036</td>
<td>.027</td>
<td>-1.336</td>
<td>.1832</td>
</tr>
<tr>
<td>Count of Spammers Arrested</td>
<td>.075</td>
<td>.03</td>
<td>2.483</td>
<td>.014*</td>
</tr>
<tr>
<td>Sum in Days of Spammers Detained</td>
<td>.21</td>
<td>.029</td>
<td>7.196</td>
<td>&lt;.0001***</td>
</tr>
<tr>
<td>Count of Trial Spammer Acquitted Articles</td>
<td>.084</td>
<td>.035</td>
<td>2.436</td>
<td>.0159*</td>
</tr>
<tr>
<td>Count of all Spam Articles</td>
<td>.343</td>
<td>.07</td>
<td>4.874</td>
<td>&lt;.0001***</td>
</tr>
<tr>
<td>Percent of Articles on CAN SPAM</td>
<td>.047</td>
<td>.027</td>
<td>1.717</td>
<td>.0878†</td>
</tr>
<tr>
<td>Percent of Articles with Spammer Attribution</td>
<td>-.055</td>
<td>.021</td>
<td>-2.556</td>
<td>.0115*</td>
</tr>
</tbody>
</table>
The remaining 11 predictors relate to the CAN SPAM Act or spam deterrence measures. The number of articles on the CAN SPAM Act published per month is associated with fewer US TLDs the following month ($B = -0.32, p < 0.0001$). The finding might indicate a displacement or deterrent effect. However, the count of articles mentioning spammers being arrested ($B = 0.075, p = 0.014$) and the sum of days per month of spammers being detained ($B = 0.21, p < 0.0001$) both indicated more US TLD links contained in spam emails. While not consistent with either routine activities or deterrence theory, the spammer detained variable is consistent with the prior two models finding more detentions of spammers is associated with higher spam volume, more US IPAs, and now more US TLDs the following month.

The count of spammers acquitted in the news was associated with an increase in US TLDs ($B = 0.084, p = 0.016$). Potentially acquittals could be emboldening spammers who reside in the United States. The count of all articles on spam per month, regardless of whether they were about the CAN SPAM Act, predicted more US TLDs among spam emails ($B = 0.343, p < 0.0001$). This can be contrasted with the count of all CAN SPAM articles mentioned previously, which predicted a decrease in US TLDs. CAN SPAM articles may serve as a deterrent or displacement motivator, as such articles may discuss prosecutions and other aspects of the Act which might appear more dangerous to spammers. Articles on spam in general, however, may not draw attention to prosecutions or other deterrents, but instead may focus on the technical details of spam in America without any mention of spammers being punished.
The percentage of all articles published that contain attributions to individual spammers and their identities is associated with a decrease in US TLDs the following month ($B = -0.055, p = 0.012$). Removing the impression of the anonymity from sending spam may prevent some spammers from including domains they registered in the United States in emails. The percentage of articles that are optimistic about the CAN SPAM Act are also associated with a decrease in UD TLDs the following month ($B = -0.065, p < 0.0001$). If attitudes towards spam law are positive, then spammers might assume the laws are more effective and choose not to advertise their US residency by including links in such emails.

Google searches within the United States for keywords relating to the CAN SPAM Act predicts more US TLD links the following month. Google searches may simply reflect US citizens’ interest in spam, including existing spammers residing in the United States with US domain names. Interest may lead to sending more spam amongst US residents. Finally, the CAN SPAM Act enforcement dummy, coded ‘1’ after the passing of the CAN SPAM Act, is associated with a reduction in US TLDs. Spammers might have been more confident about registering US domain names prior to the CAN SPAM Act.

*Time Series Regression of Percent Spam Opt-Out Compliance Per Month*

A backward stepwise regression using the model AIC was conducted with percent of spam providing recipients with opt-out mechanisms per month regressed on an initial 33 predictor variables total. Backward elimination yielded 4 predictor variables to be included in the final time series model. The selected predictors were included in a linear regression and the residuals were computed to test for autocorrelation. A Durbin Watson test of the residuals found no evidence of autocorrelation ($dw = 2.69, p = 1$).
After inspection of the ACF and PACF correlograms of the regression (see figures A.7 and A.8), an ARMA(1,2) model was identified. A final generalized least squares time series regression model was run using the 4 selected predictors and included in Table 6.

Table 6
Generalized Least Squares Time Series Regression of Percent Opt-Out Compliance Per Month, ARMA(1,2)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Beta</th>
<th>SE</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-.024</td>
<td>.034</td>
<td>-.696</td>
<td>.4872</td>
</tr>
<tr>
<td>Unemployment Rate</td>
<td>.02</td>
<td>.056</td>
<td>.35</td>
<td>.727</td>
</tr>
<tr>
<td>GDP Growth</td>
<td>-.036</td>
<td>.051</td>
<td>-.706</td>
<td>.4812</td>
</tr>
<tr>
<td>CPI</td>
<td>-.063</td>
<td>.029</td>
<td>-2.197</td>
<td>.029*</td>
</tr>
<tr>
<td>Percent of Articles with Spammer Attribution</td>
<td>.05</td>
<td>.028</td>
<td>1.785</td>
<td>.0759†</td>
</tr>
</tbody>
</table>

$R^2 = 1.68\%$

* < .05, † < .1

The only significant predictor in the model is the Consumer Price Index ($B = -.063, p = .03$). Higher CPI values predict lower opt-out compliance among spam emails the following month. Higher consumer spending power might draw the attention of those using illegitimate or unlawful email marketing techniques.

Only one independent variable made it into the model, and it only approaches significance. The percentage of articles per month where the spammer’s identity is known is associated with higher opt-out compliance the following month ($B = .05, p = .076$). If this effect size exists in the population of emails sent, it might suggest that the impression of reduced anonymity from sending spam incentivizes compliance with spam regulations. Attribution may serve as a deterrent to sending illicit spam emails.

Time Series Regression of Percent Spam Address Compliance Per Month

A backward stepwise regression using the model AIC was conducted with percent of spam including a physical mailing address per month regressed on an initial 33 predictor
variables total. Backward elimination yielded 13 predictor variables to be included in the final time series model. The selected predictors were included in a linear regression and the residuals were computed to test for autocorrelation. A Durbin Watson test of the residuals found significant autocorrelation ($dw = 1.19, p < .001$). Differencing of the dataset resulted in sufficient stationarity of the residuals ($dw = 2.3, p = .99$).

After inspection of the ACF and PACF correlograms of the regression (see figures A.9 and A.10), an ARMA(1,1) model was identified. A final generalized least squares time series regression model was run using the 13 selected predictors and included in Table 7.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Beta</th>
<th>SE</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-.028</td>
<td>.021</td>
<td>-1.316</td>
<td>.1899</td>
</tr>
<tr>
<td>Internet Market Index</td>
<td>-.116</td>
<td>.058</td>
<td>-2.013</td>
<td>.0457*</td>
</tr>
<tr>
<td>Population Aged 15-25</td>
<td>-.103</td>
<td>.06</td>
<td>-1.725</td>
<td>.0863†</td>
</tr>
<tr>
<td>Financial Stress Index</td>
<td>-.178</td>
<td>.061</td>
<td>-2.905</td>
<td>.0042**</td>
</tr>
<tr>
<td>Count of CAN SPAM Articles</td>
<td>-.177</td>
<td>.069</td>
<td>-2.564</td>
<td>.0112*</td>
</tr>
<tr>
<td>Count of Trial Ongoing Articles</td>
<td>.278</td>
<td>.086</td>
<td>3.219</td>
<td>.0015**</td>
</tr>
<tr>
<td>Count of Trial Spammer Conviction Articles</td>
<td>.207</td>
<td>.072</td>
<td>2.854</td>
<td>.0048**</td>
</tr>
<tr>
<td>Count of Trial Spammer Acquitted Articles</td>
<td>.196</td>
<td>.057</td>
<td>3.468</td>
<td>.0007***</td>
</tr>
<tr>
<td>Count of Spammer Detained Articles</td>
<td>-.17</td>
<td>.071</td>
<td>-2.392</td>
<td>.0178*</td>
</tr>
<tr>
<td>Count of Damages Awarded Articles</td>
<td>-.188</td>
<td>.084</td>
<td>-2.245</td>
<td>.026*</td>
</tr>
<tr>
<td>Percent of Articles on CAN SPAM</td>
<td>-.049</td>
<td>.043</td>
<td>-1.14</td>
<td>.2557</td>
</tr>
<tr>
<td>Percent of Articles Negative About Spam</td>
<td>.05</td>
<td>.029</td>
<td>1.709</td>
<td>.0893†</td>
</tr>
<tr>
<td>CAN SPAM Act Google Search Volume</td>
<td>-.395</td>
<td>.113</td>
<td>-3.486</td>
<td>.0006***</td>
</tr>
<tr>
<td>CAN SPAM Act in Enforcement</td>
<td>1.899</td>
<td>.443</td>
<td>4.285</td>
<td>&lt;.0001***</td>
</tr>
</tbody>
</table>

$R^2 = 28.99\%$

* < .05, ** < .01, *** < .001, † < .1

Two of the three control variables are significant. A higher Wishire Internet Market Index is associated with lower address inclusion compliance ($B = -.116, p = .046$). A more profitable internet market sector might be a better suitable target for less ethical spammers.
Higher market financial stress is also associated with a decrease in address compliance ($B = -0.178, p = .004$). While only approaching significance, youth population size also predicts reduced spam compliance ($B = -0.103, p = .086$). Because the entire population size of the US predictor did not make it into the model, the youth population measure may represent motivated offenders.

The number of articles published per month pertaining to the CAN SPAM Act is associated with a reduction in address compliance ($B = -0.177, p = .011$). However, the next three predictors on CAN SPAM Act trials, the count of unresolved trials, trials where the spammer is convicted, and trials where the spammer is acquitted are all associated with an increase in address compliance ($B = 0.278, p = .002; B = 0.207, p = .005; B = 0.196, p < .001$, respectively). That means that acquittals of spammers have the same impact on compliance as convictions of spammers. Note that excluding all three CAN SPAM trial predictors results in the count of CAN SPAM articles measure becoming nonsignificant (not shown in tables). So possibly a higher proportion of CAN SPAM articles that are trial related results in higher compliance or deterrence. Additionally that would mean that holding trials constant, as CAN SPAM articles increase, compliance goes down. CAN SPAM articles unrelated to prosecution of spammers might not serve as a deterrent.

Both the number of articles where the spammer was detained and where the spammer had damages judged against him/her during a lawsuit were predictive of less compliance ($B = -0.17, p = .018; B = -0.188, p = .026$, respectively). This finding is contrary to theory but the spammer detention effect size is consistent with all prior models in terms of worsening the spam situation. That is, more detentions result in higher spam volume, and more spam US IPAs and US TLDs.
Detentions might simply foreshadow more serious spam offending the following month that would have occurred anyway.

While only approaching significance, the percent of articles negative about spam is associated with a decrease in compliance \( (B = .05, p = .089) \). The direction of this relationship is consistent with prior findings that negativity about spam predicts more spam and more spam being sent into the United States from foreign IPAs. It may simply be that news portrayals of spam as a worsening problem are actually accurate, and the spam problem continues to worsen the following month.

Google search interest in the CAN SPAM Act is associated with a reduction in CAN SPAM compliance \( (B = -.395, p < .001) \), even though compliance was higher following the passing of the CAN SPAM Act \( (B = 1.9, p < .0001) \). The latter might suggest the Act increased compliance, although not during times when the public was interested in or as aware of the Act.

**Time Series Regression of Percent Spam Meaningful Subject Compliance Per Month**

A backward stepwise regression using the model AIC was conducted with percent of spam with a descriptive subject heading per month regressed on an initial 33 predictor variables total. Backward elimination yielded 5 predictor variables to be included in the final time series model. The selected predictors were included in a linear regression and the residuals were computed to test for autocorrelation. A Durbin Watson test of the residuals found significant autocorrelation \( (dw = .15, p < .001) \). Differencing of the dataset resulted in sufficient stationarity of the residuals \( (dw = 2.21, p = .943) \).

After inspection of the ACF and PACF correlograms of the regression (see figures A.11 and A.12), an ARMA(2,0) model was identified. A final generalized least squares time series regression model was run using the 5 selected predictors and included in Table 8.
Table 8
Generalized Least Squares Time Series Regression of Percent Subject Compliance Per Month, ARMA(2,0)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Beta</th>
<th>SE</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-.019</td>
<td>.011</td>
<td>-1.733</td>
<td>.0847†</td>
</tr>
<tr>
<td>Percent Internet Users</td>
<td>.048</td>
<td>.012</td>
<td>3.828</td>
<td>.0002***</td>
</tr>
<tr>
<td>CPI</td>
<td>-.016</td>
<td>.01</td>
<td>-1.689</td>
<td>.0928†</td>
</tr>
<tr>
<td>Count of Trial Spammer Acquitted Articles</td>
<td>-.033</td>
<td>.011</td>
<td>-3.124</td>
<td>.0021**</td>
</tr>
<tr>
<td>Percent of Articles on CAN SPAM</td>
<td>.012</td>
<td>.009</td>
<td>1.351</td>
<td>.1785</td>
</tr>
<tr>
<td>Percent of Articles with Spammer Attribution</td>
<td>-.013</td>
<td>.008</td>
<td>-1.725</td>
<td>.0863†</td>
</tr>
</tbody>
</table>

R² = 14.96%
** < .01, *** < .001, † < .1

The percentage of the US population with internet connectivity is associated with an increase in meaningful subject compliance ($B = .048, p < .001$). The result might indicate that more internet users means more legitimate email marketers who send CAN SPAM Act compliant spam emails. However, theory might propose that more internet users means more suitable targets for illegitimate spammers or more motivated offenders to engage in illegal spam, but that is not what is found. It should be noted that internet users was associated with a reduction in spam the following month in the prior spam count model above. It is possible that internet connectivity reduces the severity and volume of spam, but it is not certain why.

More acquittals of spammers in the news during CAN SPAM Act trials is associated with a decrease in compliance with the meaningful subject requirement ($B = -.033, p = .002$). Acquittals might embolden spammers to violent CAN SPAM regulations. While not achieving significance, a higher percentage of articles with spammer identity attribution predicts a drop in meaningful subject compliance ($B = -.013, p = .086$). Lacking anonymity may simply increase deception used in emails to conceal as much as possible.
Time Series Regression of Percent Spam Advertisement Notice Compliance Per Month

A backward stepwise regression using the model AIC was conducted with percent of spam including a notice of advertisement per month regressed on an initial 33 predictor variables total. Backward elimination yielded 11 predictor variables to be included in the final time series model. The selected predictors were included in a linear regression and the residuals were computed to test for autocorrelation. A Durbin Watson test of the residuals found significant autocorrelation ($dw = .97, p < .001$). Differencing of the dataset resulted in sufficient stationarity of the residuals ($dw = 2.68, p = 1$).

After inspection of the ACF and PACF correlograms of the regression (see figures A.13 and A.14), an ARMA(0,3) model was identified. A final generalized least squares time series regression model was run using the 11 selected predictors and included in Table 9.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Beta</th>
<th>SE</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.008</td>
<td>0.008</td>
<td>-1.035</td>
<td>0.302</td>
</tr>
<tr>
<td>Unemployment Rate</td>
<td>-0.054</td>
<td>0.047</td>
<td>-1.147</td>
<td>0.2529</td>
</tr>
<tr>
<td>Internet Market Index</td>
<td>0.069</td>
<td>0.032</td>
<td>2.151</td>
<td>0.0328*</td>
</tr>
<tr>
<td>GDP Growth</td>
<td>-0.059</td>
<td>0.03</td>
<td>-1.985</td>
<td>0.0487*</td>
</tr>
<tr>
<td>Disposable Income Per Capita</td>
<td>-0.047</td>
<td>0.026</td>
<td>-1.793</td>
<td>0.0746†</td>
</tr>
<tr>
<td>Financial Stress Index</td>
<td>-0.085</td>
<td>0.036</td>
<td>-2.362</td>
<td>0.0193*</td>
</tr>
<tr>
<td>Count of Trial Spammer Acquitted Articles</td>
<td>0.1</td>
<td>0.033</td>
<td>3.049</td>
<td>0.0027**</td>
</tr>
<tr>
<td>Count of all Spam Articles</td>
<td>-0.032</td>
<td>0.04</td>
<td>-0.807</td>
<td>0.4209</td>
</tr>
<tr>
<td>Percent of Articles Positive About CAN SPAM</td>
<td>0.022</td>
<td>0.016</td>
<td>1.38</td>
<td>0.1693</td>
</tr>
<tr>
<td>Percent of Articles Negative About Spam</td>
<td>0.019</td>
<td>0.02</td>
<td>0.964</td>
<td>0.3364</td>
</tr>
<tr>
<td>Percent of Articles without Spammer Attribution</td>
<td>-0.038</td>
<td>0.022</td>
<td>-1.731</td>
<td>0.0851†</td>
</tr>
<tr>
<td>CAN SPAM Act in Enforcement</td>
<td>-0.114</td>
<td>0.13</td>
<td>-0.883</td>
<td>0.3787</td>
</tr>
</tbody>
</table>

$R^2 = 20.63\%$

* $< .05$, ** $< .01$, † $< .1$
A higher internet market index predicts an increase in address compliance \((B = .069, p = .033)\). A better performing internet market sector may be associated with an increase in legitimate email marketing efforts, as technology companies seek to promote their products via spam emails. However, GDP growth is associated with lower advertisement notice compliance \((B = -.059, p = .049)\), controlling for internet market level and other factors. Financial stress was the only other remaining significant control variable, predicting reduced advertisement notice compliance \((B = -.085, p = .019)\).

Only one of the six independent variables of CAN SPAM and other deterrence measures is significant. As the count of articles mentioning spammers being acquitted under CAN SPAM charges increases, advertisement compliance also increases the following month \((B = .1, p = .003)\). Also note that the effect size is the strongest of all significant predictors in the model. It should also be mentioned that none of the other CAN SPAM trial variables (conviction, ongoing) made it into the model. The finding is contrary to the expected emboldening effect of acquittals.

The percent of articles without spammer attribution (that do not reveal the identities of any individual spammers) predicts a decrease in compliance, but does not quite achieve significance \((B = -.038, p = .085)\). If the association exists, the finding might suggest feelings of anonymity result in less ethical or legally compliant spamming efforts.

**Time Series Regression of Percent Spam Header “To” Found Per Month**

A backward stepwise regression using the model AIC was conducted with percent of spam without a missing “to” field or recipient address per month regressed on an initial 33 predictor variables total. Backward elimination yielded 3 predictor variables to be included in the final time series model. The selected predictors were included in a linear regression and the
residuals were computed to test for autocorrelation. A Durbin Watson test of the residuals found no evidence of autocorrelation ($dw = 2.64, p = 1$).

After inspection of the ACF and PACF correlograms of the regression (see figures A.15 and A.16), an ARMA(0,3) model was identified. A final generalized least squares time series regression model was run using the 3 selected predictors and included in Table 10.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Beta</th>
<th>SE</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-.087</td>
<td>.051</td>
<td>-1.699</td>
<td>.0911†</td>
</tr>
<tr>
<td>UCR Arrest Rate</td>
<td>-.181</td>
<td>.055</td>
<td>-3.278</td>
<td>.0012**</td>
</tr>
<tr>
<td>Internet Market Index</td>
<td>-.128</td>
<td>.05</td>
<td>-2.546</td>
<td>.0117*</td>
</tr>
<tr>
<td>Percent of Articles without Spammer Attribution</td>
<td>.097</td>
<td>.031</td>
<td>3.134</td>
<td>.002**</td>
</tr>
</tbody>
</table>

$R^2 = 5.39 \%$

* < .05, ** < .01, † < .1

While only three predictors were selected for model inclusion, all three achieve significance. Higher UCR arrest rates predict a reduction in spam emails including the recipient’s address in the ‘to’ field ($B = -.181, p = .001$). Decreases in header ‘to’ field inclusion represents an increase in header falsification prohibited under the CAN SPAM Act. So in this case higher street crime predicts an increase in this type of header forgery, a type of cybercrime.

A more profitable internet market sector also predicts header forgery ($B = -.128, p = .012$). It could be that a higher internet market index represents an increase in more sophisticated motivated offenders who are able to forge email headers. The finding could also represent more suitable targets that draw the attention of more sophisticated offenders.

Finally, the higher the percentage of spam articles published without revealing the identities of individual spammers predicts an increase in recipient address inclusion among spam emails ($B = .097, p = .002$). Lack of attribution in the media may increase the impression of
anonymity among spammers, reducing their need to use deception or tamper with header information when sending spam emails.

**Time Series Regression of Percent Duplicate Email From and To Fields Per Month**

A backward stepwise regression using the model AIC was conducted with percent of spam with a duplicate ‘from’ and ‘to’ field per month regressed on an initial 33 predictor variables total. Backward elimination yielded 7 predictor variables to be included in the final time series model. The selected predictors were included in a linear regression and the residuals were computed to test for autocorrelation. A Durbin Watson test of the residuals found no evidence of autocorrelation ($dw = 2.52, p = .99$).

After inspection of the ACF and PACF correlograms of the regression (see figures A.17 and A.18), an ARMA(2,2) model was identified. A final generalized least squares time series regression model was run using the 7 selected predictors and included in Table 11.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Beta</th>
<th>SE</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>.076</td>
<td>.037</td>
<td>2.038</td>
<td>.043*</td>
</tr>
<tr>
<td>UCR Arrest Rate</td>
<td>.081</td>
<td>.069</td>
<td>1.188</td>
<td>.2363</td>
</tr>
<tr>
<td>Percent Internet Users</td>
<td>-.118</td>
<td>.064</td>
<td>-1.825</td>
<td>.0696†</td>
</tr>
<tr>
<td>Population Size</td>
<td>-.077</td>
<td>.065</td>
<td>-1.173</td>
<td>.2422</td>
</tr>
<tr>
<td>Percent Population Aged 15-25</td>
<td>-.253</td>
<td>.098</td>
<td>-2.586</td>
<td>.0105*</td>
</tr>
<tr>
<td>CPI</td>
<td>-.098</td>
<td>.031</td>
<td>-3.138</td>
<td>.002**</td>
</tr>
<tr>
<td>Count of Trial Spammer Conviction Articles</td>
<td>.121</td>
<td>.031</td>
<td>3.885</td>
<td>.0001***</td>
</tr>
<tr>
<td>Percent of Articles without Spammer Attribution</td>
<td>.015</td>
<td>.023</td>
<td>.655</td>
<td>.5131</td>
</tr>
</tbody>
</table>

$R^2 = 6.03\%$

* < .05, ** < .01, *** < .001, † < .1

A higher percentage of the US population aged 15-25 predicts a decrease in duplicate header forgery in spam emails ($B = -.253, p = .011$). The relationship is contrary to what was proposed a priori, with younger generations thought to be more criminally inclined and perhaps
more technologically sophisticated in such ways that would increase email header forgery. The consumer price index also predicts a decrease in duplicate email header forgery \((B = -.098, p = .0001)\). While not quite significant, the number of internet users per capita predicts decreased header forgery as well \((B = -.118, p = .07)\).

The only significant independent variable is the number of spammer convictions mentioned in the news per month \((B = .121, p = .0001)\), which predicts an increase in email header forgery. Spammer convictions ought to have a deterrent effect, but an increase in header forgery as a result could suggest spammers are simply more cautious with their identities when sending spam and therefore falsify their return addresses. The finding is consistent with that in Table 10, in which case feelings of security and anonymity decrease header tampering.

*Time Series Regression of Percent of Emails with Sender Name Found Per Month*

A backward stepwise regression using the model AIC was conducted with percent of spam including a sender name per month regressed on an initial 33 predictor variables total. Backward elimination yielded 5 predictor variables to be included in the final time series model. The selected predictors were included in a linear regression and the residuals were computed to test for autocorrelation. A Durbin Watson test of the residuals found no evidence of autocorrelation \((dw = 1.99, p < .451)\).

After inspection of the ACF and PACF correlograms of the regression (see figures A.19 and A.20), an ARMA(1,1) model was identified. A final generalized least squares time series regression model was run using the 5 selected predictors and included in Table 12.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Beta</th>
<th>SE</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>.053</td>
<td>.055</td>
<td>.95</td>
<td>.3433</td>
</tr>
</tbody>
</table>

Table 12

**Generalized Least Squares Time Series Regression of Percent Sender Name Found Per Month, ARMA(1,1)**

\(n = 189\)
The internet market index is significant and predicts an increase in compliance via inclusion of a sender name in emails ($B = .164, p = .011$). This is the opposite effect for recipient address inclusion in Table 10, as the internet market sector predicts reduced compliance. More technology jobs predicts a decrease in compliance ($B = -.123, p = .009$). That is, the more technology sector jobs there are available, the fewer spam emails are sent that include a sender name the following month.

The percentage of articles published out of all spam articles that relate to the CAN SPAM Act predicts a decrease in sender name inclusion the following month ($B = -.074, p = .023$). The pattern found here is consistent with the prior two header tampering models, suggesting methods of deterrence actually increase the severity of illicit spam emails when it comes to header forgery. Header forgery may be an extra precaution adopted by spammers in response to awareness of possible legal threats to their spam business.

**Time Series Regression of Percent of Emails with Header “From” Found Per Month**

A backward stepwise regression using the model AIC was conducted with percent of spam including a ‘from’ address per month regressed on an initial 33 predictor variables total. Backward elimination yielded 4 predictor variables to be included in the final time series model. The selected predictors were included in a linear regression and the residuals were computed to test for autocorrelation. A Durbin Watson test of the residuals found significant autocorrelation.
(\(dw = .89, p < .001\)). Regular differencing of the dataset yielded sufficient stationarity (\(dw = 2.52, p = .99\)).

After inspection of the ACF and PACF correlograms of the regression (see figures A.21 and A.22), an ARMA(0,2) model was identified. A final generalized least squares time series regression model was run using the 4 selected predictors and included in Table 13.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Beta</th>
<th>SE</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-.018</td>
<td>.022</td>
<td>-.786</td>
<td>.4329</td>
</tr>
<tr>
<td>Unemployment Rate</td>
<td>-.125</td>
<td>.078</td>
<td>-1.606</td>
<td>.1101</td>
</tr>
<tr>
<td>Technology Jobs</td>
<td>.115</td>
<td>.084</td>
<td>1.368</td>
<td>.1729</td>
</tr>
<tr>
<td>Percent of Articles Negative About Spam</td>
<td>-.048</td>
<td>.032</td>
<td>-1.521</td>
<td>.13</td>
</tr>
<tr>
<td>Percent of Articles without Spammer Attribution</td>
<td>.048</td>
<td>.037</td>
<td>1.298</td>
<td>.196</td>
</tr>
</tbody>
</table>

\(R^2 = 6.5\%\)

Not a single predictor included in the model achieves significance, nor does any predictor approach significance. None of the economic, technological, demographic, or deterrent variables appear to relate to whether a spammer conceals his or her email address from the recipients begin spammed. The possible influences of this type of header forgery are not captured by any of the data gathered for this research.

**Time Series Regression of Percent of Emails with Duplicate “To” and Return-Path Fields Per Month**

A backward stepwise regression using the model AIC was conducted with percent of spam with a duplicate ‘to’ and return-path addresses per month regressed on an initial 33
predictor variables total. Backward elimination yielded 12 predictor variables to be included in
the final time series model. The selected predictors were included in a linear regression and the
residuals were computed to test for autocorrelation. A Durbin Watson test of the residuals found
significant autocorrelation ($dw = .81, p < .001$). Regular differencing of the dataset yielded
sufficient stationarity ($dw = 3.04, p = 1$).

After inspection of the ACF and PACF correlograms of the regression (see figures A.23
and A.24), an ARMA(1,1) model was identified. A final generalized least squares time series
regression model was run using the 12 selected predictors and included in Table 14.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Beta</th>
<th>SE</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-.018</td>
<td>.023</td>
<td>-.781</td>
<td>.4358</td>
</tr>
<tr>
<td>UCR Arrest Rate</td>
<td>-.009</td>
<td>.027</td>
<td>-.319</td>
<td>.7502</td>
</tr>
<tr>
<td>Unemployment Rate</td>
<td>-.067</td>
<td>.031</td>
<td>-2.166</td>
<td>.0317*</td>
</tr>
<tr>
<td>Internet Market Index</td>
<td>.01</td>
<td>.022</td>
<td>.444</td>
<td>.6575</td>
</tr>
<tr>
<td>GDP Growth</td>
<td>-.003</td>
<td>.023</td>
<td>-.152</td>
<td>.8795</td>
</tr>
<tr>
<td>Financial Stress Index</td>
<td>.011</td>
<td>.022</td>
<td>.496</td>
<td>.6206</td>
</tr>
<tr>
<td>Sum of CAN SPAM Damages Awarded</td>
<td>-.02</td>
<td>.019</td>
<td>-1.067</td>
<td>.2874</td>
</tr>
<tr>
<td>Count of Spammers Arrested</td>
<td>-.026</td>
<td>.022</td>
<td>-1.179</td>
<td>.24</td>
</tr>
<tr>
<td>Count of Trial Ongoing Articles</td>
<td>.049</td>
<td>.035</td>
<td>1.407</td>
<td>.1612</td>
</tr>
<tr>
<td>Count of Spammer Detained Articles</td>
<td>-.033</td>
<td>.023</td>
<td>-1.411</td>
<td>.1602</td>
</tr>
<tr>
<td>Count of Damages Awarded Articles</td>
<td>-.026</td>
<td>.032</td>
<td>-.813</td>
<td>.4176</td>
</tr>
<tr>
<td>Percent of Articles on CAN SPAM</td>
<td>.018</td>
<td>.018</td>
<td>.955</td>
<td>.3407</td>
</tr>
<tr>
<td>Percent of Articles without Spammer Attribution</td>
<td>-.045</td>
<td>.015</td>
<td>-2.982</td>
<td>.0033**</td>
</tr>
</tbody>
</table>

$R^2 = 11.51\%$

* < .05, ** < .01

Out of the 12 predictors included in the model, only two are significant. A higher
unemployment rate is associated with a reduction in duplicate ‘to’ and return-path fields in the
spam sample ($B = -.067, p = .032$). Previously it was hypothesized that unemployment might
serve as an indicator of the number of motivated offenders, but it might not relate to the severity or legality of existing offenses, which is in part what this header forgery measures captures.

The percentage of news articles published without spammer attribution predicts a reduction in email header forgery \((B = -.045, p = .003)\). Again the finding has the same theoretical direction as all prior significant deterrence measures tested against email header forgery outcomes. Feelings of anonymity remove the need to forge email headers when sending spam emails. The finding may suggest that a primary motivation for header forgery is to conceal the sender’s identity.

_Time Series Regression of Percent of Emails with Duplicate “To” and Reply-To Fields Per Month_

A backward stepwise regression using the model AIC was conducted with percent of spam with a duplicate ‘to’ and reply-to addresses per month regressed on an initial 33 predictor variables total. Backward elimination yielded 2 predictor variables to be included in the final time series model. The selected predictors were included in a linear regression and the residuals were computed to test for autocorrelation. A Durbin Watson test of the residuals found significant autocorrelation \((dw = .73, p < .001)\). Regular differencing of the dataset yielded sufficient stationarity \((dw = 3.07, p = 1)\).

After inspection of the ACF and PACF correlograms of the regression (see figures A.25 and A.26), an ARMA(0,2) model was identified. A final generalized least squares time series regression model was run using the 2 selected predictors and included in Table 15.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Beta</th>
<th>SE</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>.001</td>
<td>&lt; .0001</td>
<td>1.224</td>
<td>.2224</td>
</tr>
</tbody>
</table>

Table 15
Generalized Least Squares Time Series Regression of Percent Duplicate To and Reply-To Fields Per Month, ARMA(0,2)
Only two predictors made it into the model, and neither of them is significant. Variations in this method of header forgery are not captured or explained by the predictors tested for. The lack of significance might seem odd considering that the duplicate ‘to’ and return-path field model and the duplicate ‘to’ and ‘from’ field model were found to be predictable based on the data. However, emails found to falsify headers in this way were seen to be more rare than the other two methods just mentioned, so there might not be enough variation to explain.

*Time Series Regression of Average Fraud Probability Per Month*

A backward stepwise regression using the model AIC was conducted with the average predicted fraud probability per month regressed on an initial 33 predictor variables total. Backward elimination yielded 11 predictor variables to be included in the final time series model. The selected predictors were included in a linear regression and the residuals were computed to test for autocorrelation. A Durbin Watson test of the residuals found no evidence of significant autocorrelation ($dw = 2.53, p = .99$).

After inspection of the ACF and PACF correlograms of the regression (see figures A.27 and A.28), an ARMA(1,1) model was identified. A final generalized least squares time series regression model was run using the 11 selected predictors and included in Table 16.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Beta</th>
<th>SE</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-.019</td>
<td>.027</td>
<td>-.704</td>
<td>.4823</td>
</tr>
<tr>
<td>UCR Arrest Rate</td>
<td>.089</td>
<td>.029</td>
<td>3.056</td>
<td>.0026**</td>
</tr>
<tr>
<td>Percent Internet Users</td>
<td>-.127</td>
<td>.058</td>
<td>-2.202</td>
<td>.029*</td>
</tr>
</tbody>
</table>
An increase in arrest rates in the United States predict an increase in the average fraud probability in the spam email sample ($B = .089, p = .003$). Again, street crime appears to relate to cybercrime. However, the percentage of the population with internet connectivity is associated with a decrease in fraud prediction per month ($B = -.127, p = .029$). Internet connectivity appears to decrease fraud, a conclusion which matches that proposed for the deceptive subject heading model, as internet connectivity decreased that form of noncompliance as well. It should also be mentioned that internet user rate was also associated with lower spam volume. Internet connectivity may have a positive impact on spam, although that is contrary to what was expected.

The number of technology jobs predicts a decrease in fraud probability the following month ($B = -.068, p < .0001$). Technology jobs might offer an alternative and legitimate means of earning an income that might prevent some from using their technology skills to profit illegally. That is, lack of technology jobs may be a measure of motivated offenders.

The sum of days spammers are detained in the news per month predicts a decrease in fraud probability the following month ($B = -.152, p = .002$). Note that this effect size is the largest in the model. The result is consistent with deterrence theory, suggesting that knowledge
of other spammers being detained might deter other spammers from perpetrating fraud when sending emails.

**Time Series Regression of Percent Fraudulent Emails Per Month**

A backward stepwise regression using the model AIC was conducted with the percentage of emails classified as fraudulent at 85% certainty per month regressed on an initial 33 predictor variables total. Backward elimination yielded 9 predictor variables to be included in the final time series model. The selected predictors were included in a linear regression and the residuals were computed to test for autocorrelation. A Durbin Watson test of the residuals found no evidence of significant autocorrelation ($dw = 2.65$, $p = 1$).

After inspection of the ACF and PACF correlograms of the regression (see figures A.29 and A.30), an ARMA(1,1) model was identified. A final generalized least squares time series regression model was run using the 9 selected predictors and included in Table 17.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Beta</th>
<th>SE</th>
<th>$t$</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>.042</td>
<td>.026</td>
<td>1.58</td>
<td>.1159</td>
</tr>
<tr>
<td>UCR Arrest Rate</td>
<td>.107</td>
<td>.025</td>
<td>4.213</td>
<td>&lt; .0001***</td>
</tr>
<tr>
<td>Percent Internet Users</td>
<td>-.137</td>
<td>.047</td>
<td>-2.908</td>
<td>.0041**</td>
</tr>
<tr>
<td>Unemployment Rate</td>
<td>-.116</td>
<td>.034</td>
<td>-3.433</td>
<td>.0007***</td>
</tr>
<tr>
<td>Technology Jobs</td>
<td>-.072</td>
<td>.011</td>
<td>-6.446</td>
<td>&lt; .0001***</td>
</tr>
<tr>
<td>Population Aged 15-25</td>
<td>-.171</td>
<td>.074</td>
<td>-2.304</td>
<td>.0224*</td>
</tr>
<tr>
<td>CPI</td>
<td>-.06</td>
<td>.026</td>
<td>-2.277</td>
<td>.024*</td>
</tr>
<tr>
<td>Count of Trial Spammer Acquitted Articles</td>
<td>.059</td>
<td>.05</td>
<td>1.172</td>
<td>.2428</td>
</tr>
<tr>
<td>Percent of Articles Negative About CAN SPAM</td>
<td>-.007</td>
<td>.013</td>
<td>-.532</td>
<td>.5955</td>
</tr>
<tr>
<td>Percent of Articles without Spammer Attribution</td>
<td>.006</td>
<td>.018</td>
<td>.339</td>
<td>.735</td>
</tr>
</tbody>
</table>

$R^2 = 5.81\%$

* < .05, ** < .01, *** < .001

Some of the findings are similar to the prior fraud probability model. Arrest rates continue to predict a higher amount of fraud ($B = .107$, $p < .0001$). More internet users continues
to predict a reduction in fraud \((B = -.137, p = .004)\). Technology jobs also continue to predict a reduction in fraud \((B = -.072, p < .0001)\).

However, unemployment is significant in the new model, with higher unemployment predicting less fraud \((B = -.116, p = .0007)\). CPI is also significant, and predicts less fraud as well \((B = -.06, p = .024)\). There is little consistency with the economic predictors of unemployment, technology jobs, and CPI, other than they all predict a reduction in fraud. However, unemployment represents a struggling economy, while the remaining two predictors indicate stronger economies. Finally, youth population size is associated with less fraud \((B = -.171, p = .022)\). The direction of this relationship is not consistent with the prior assumptions proposed during variable selection.

*Time Series Regression of Percent of Spam with Malicious Links Per Month*

A backward stepwise regression using the model AIC was conducted with the percentage of emails containing links to executable files per month regressed on an initial 33 predictor variables total. Backward elimination yielded 5 predictor variables to be included in the final time series model. The selected predictors were included in a linear regression and the residuals were computed to test for autocorrelation. A Durbin Watson test of the residuals found significant autocorrelation \((dw = 1.13, p < .001)\). Regular differencing of the variables in the model resulted in sufficient trend stationarity \((dw = 2.86, p = 1)\).

After inspection of the ACF and PACF correlograms of the regression (see figures A.31 and A.32), an ARMA(0,2) model was identified. A final generalized least squares time series regression model was run using the 5 selected predictors and included in Table 18.
Table 18
Generalized Least Squares Time Series Regression of Percent of Spam with Malicious Links Per Month, ARMA(0,2)

\[ n = 189 \]

<table>
<thead>
<tr>
<th>Measure</th>
<th>Beta</th>
<th>SE</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>.001</td>
<td>.015</td>
<td>.064</td>
<td>.9487</td>
</tr>
<tr>
<td>Unemployment Rate</td>
<td>.072</td>
<td>.091</td>
<td>.79</td>
<td>.4304</td>
</tr>
<tr>
<td>Count of Spammers Arrested</td>
<td>-.128</td>
<td>.067</td>
<td>-1.913</td>
<td>.0574†</td>
</tr>
<tr>
<td>Count of Trial Ongoing Articles</td>
<td>.165</td>
<td>.086</td>
<td>1.92</td>
<td>.0564†</td>
</tr>
<tr>
<td>Count of Damages Awarded Articles</td>
<td>-.205</td>
<td>.09</td>
<td>-2.278</td>
<td>.0239*</td>
</tr>
<tr>
<td>Percent of Articles with Spammer Attribution</td>
<td>.13</td>
<td>.044</td>
<td>2.981</td>
<td>.0033**</td>
</tr>
</tbody>
</table>

\[ R^2 = 8.16\% \]

* < .05, ** < .01, *** < .001, † < .1

The count of articles mentioning damages being awarded and judged against a spammer per month is associated with a decrease in malicious links (\( B = -.205, p = .024 \)). The relationship is in the theoretically expected direction and would suggest a deterrent effect. However, the percentage of articles with spammer attribution predicts an increase in malicious links (\( B = .13, p = .003 \)). The relationship is opposite the theoretically expected direction.

It should be mentioned two other predicts in the model which are not quite significant. The count of articles mentioning arrests of spammers per month predicts a reduction in malicious links (\( B = -.128, p = .057 \)). The count of ongoing trials mentioned in the news per month predicts an increase in malicious links in the spam sample (\( B = .165, p = .056 \)). The two measures which predict a decrease in malicious links both represent punishments of spammers (arrests and fines). The two measures which predict an increase in malicious links do not necessarily pertain to punishments of spammers (ongoing trials and attribution). The attribution articles mention spammer identity, but do not necessarily involve punishments, and may even describe spammers being acquitted. The attribution measure may be similar to the ongoing trial measure in this way, and so they may not have a deterrent effect. Punishment in the news, however, may be a deterrent to this type of cybercrime.
**Time Series Regression of Percent of Spam with Embedded Scripts Per Month**

A backward stepwise regression using the model AIC was conducted with the percentage of emails containing embedded script tags per month regressed on an initial 33 predictor variables total. Backward elimination yielded 5 predictor variables to be included in the final time series model. The selected predictors were included in a linear regression and the residuals were computed to test for autocorrelation. A Durbin Watson test of the residuals found significant autocorrelation \( (dw = .89, \ p < .001). \) Regular differencing of the variables in the model resulted in sufficient trend stationarity \( (dw = 2.36, \ p = .99). \)

After inspection of the ACF and PACF correlograms of the regression (see figures A.33 and A.34), an ARMA(0,2) model was identified. A final generalized least squares time series regression model was run using the 5 selected predictors and included in Table 19.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Beta</th>
<th>SE</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>&lt; .0001</td>
<td>.032</td>
<td>- .004</td>
<td>.997</td>
</tr>
<tr>
<td>Percent Internet Users</td>
<td>-.088</td>
<td>.045</td>
<td>-1.966</td>
<td>.0508†</td>
</tr>
<tr>
<td>Percent Population Aged 15-25</td>
<td>-.236</td>
<td>.071</td>
<td>-3.335</td>
<td>.001**</td>
</tr>
<tr>
<td>Count of CAN SPAM Articles</td>
<td>-.144</td>
<td>.057</td>
<td>-2.506</td>
<td>.0131*</td>
</tr>
<tr>
<td>Count of Spammer Detained Articles</td>
<td>.067</td>
<td>.044</td>
<td>1.509</td>
<td>.133</td>
</tr>
<tr>
<td>Percent of Articles without Spammer Attribution</td>
<td>.071</td>
<td>.031</td>
<td>2.294</td>
<td>.0229*</td>
</tr>
</tbody>
</table>

\( R^2 = 17.15\% \)

* < .05, ** < .01, † < .1

The percent of the population aged 15-25 is associated with a decrease in embedded script tags per month \( (B = -.236, \ p = .001). \) The direction is contrary to theory but consistent with the only other time the measure was significant, predicting a reduction in falsified headers in Table 11. The number of articles on the CAN SPAM Act published per month predicts a decrease in embedded scripts \( (B = -.144, \ p = .013), \) consistent with what might be expected from
a deterrent effect. The percentage of articles without spammer attribution predicts an increase in malicious scripts ($B = .071, p = .023$), consistent with an emboldening effect of feelings of anonymity. However, attribution was found to increase malicious links in the prior model. Attribution appears to have the opposite effect on malicious scripts.

Finally, the percent of the population who are internet users does not quite achieve significance, but is associated with a decrease in embedded scripts ($B = -.088, p = .051$). The direction of the relationship is consistent with all significant findings involving this variable in prior models. More internet users predicts less spam (Table 3), more compliance (Table 8), less header forgery (Table 11), less fraud (Table 17), and now a reduction in malicious scripts. Internet connectivity appears to reduce the negative effects of illicit spam.
Chapter 7: Discussion

Data on multiple measures of spamming behavior have been collected for this research and tested against a previously untested number of time series capturing CAN SPAM Act and other possible deterrents against spam crimes. A spam sample totaling 5,490,905 email messages transmitted in the United States has been collected and analyzed. Seventeen time series variables representing four dimensions of spam to address each of four research questions were constructed for inclusion in multiple time series regression analyses. The four categories, and by extension, the four research questions this paper intends to address, relate to email spam’s volume, spam law compliance, seriousness, and locality over time within the United States. This research has found that while CAN SPAM Act activity relates to spam in different ways, it is not an obvious deterrent effect. The CAN SPAM Act may worsen certain aspects of spam while deterring others. Additionally, the type of CAN SPAM Act activity reported in the news varies, and these different variations can have opposing influences on spam and spammer behavior.

Implications of Findings

Spam Volume

Prior literature has mostly concluded that spam has increased following the CAN SPAM Act (Arora, 2006; Lee, 2005; Zeller, 2005). The FTC concluded that spam levelled off following the passing of the CAN SPAM Act (Majoras, Leary, Harbour, & Leibowtiz, 2005). Neither of these conclusions are necessarily incorrect, as spam is definitely higher today than prior to 2004, during the nascent stages of email spam. However, after inspecting a time series plot of spam over time, it may appear that spam has levelled off, albeit at a higher level, following the CAN SPAM Act. Despite this, causality cannot be drawn from these conclusions, as very few, if any, statistical measures were employed to identify any kind of deterrent effect.
The only evaluation of the CAN SPAM Act utilizing adequate statistical methods concluded that a dichotomous measure of the Act had no significant impact on spam volume (Kigerl, 2009). The current study mirrored these findings, as a dichotomous measure of spam was also included in the analysis. A binary measure of CAN SPAM was included, but not selected for, during the backward stepwise selection process, suggesting that the measure did not decrease the AIC of the spam volume model. The conclusion would be the same as that in Kigerl, 2009, suggesting no difference in spam volume following the passing of the act.

Yet the present research went beyond a simple dichotomization of the CAN SPAM Act, including multiple other measures of the Act as well as additional deterrence measures. Regarding these measures, the impact of the CAN SPAM Act on spam volume at first look appears to be mixed. The number of ongoing CAN SPAM trials and convictions reduces the amount of spam sent. However, the number of spammers being detainted increases spam volume. Spammer detentions, and even spammer arrests, may capture something not theoretically foreseen prior to coding CAN SPAM news articles. These two measures of the CAN SPAM Act (detentions and arrests) appear to consistently worsen spam outcomes in many of the models tested, such that there is more spam or more severe spam sent.

Looking over the coded news articles mentioning arrest and detentions tend to be paired with ongoing trials without a resolution. Articles mentioning detentions of spammers reveals that some are not positive about the prospect of bringing the suspect to justice, or there are some hurdles along the way that suggests the consequences the suspect will face are not severe. Additionally, spammers are not detained long, and articles often mention them going free after their time is spent being detained. Detentions and arrests tend to be earlier in the investigation and prosecutorial process and likely not a particularly strong deterrent. Therefore if the number
of ongoing trials is held constant, an increase in spam detentions may be associated with more early trial stages, because detentions are made earlier during the prosecutorial process, with fewer detentions being associated with ongoing trials that are further along and closer to conviction or resolution.

Non-CAN SPAM related deterrents also appear to relate to spam volume. Authors that are negative about the current state of spam and our ability to combat it increases spam volume, suggesting an emboldening effect. It is possible that the articles published of this nature were in response to a growing spam problem that continued to grow after negative articles were published. It might have been that an increase in spam volume would have occurred anyway.

*Spam Locality*

It has been argued that the CAN SPAM Act simply displaced spammer operations overseas where the CAN SPAM Act would have no impact (Lee, 2005). Spam email IPA was used for this research to attempt to capture the nation of origin from which an email was sent. Similar to a prior evaluation of the CAN SPAM Act (Kigerl, 2009), a dichotomous measure of the Act failed to predict US IPAs contained in spam emails, even after introducing a multitude of control variables for potential inclusion in the model. Yet in addition to spam IPA, nationalized TLDs were also used as an approximation for where spam is sent from or which nation it is associated with. Contrary to US IPA, a dichotomous measure of the CAN SPAM Act does predict US TLD, such that there are fewer US domain name registrations following the Act. The finding suggests a deterrent effect.

However, a dichotomous measure of the CAN SPAM Act was not the only variable intended to capture deterrence included in the locality models. With regards to all of them, and similar to the spam volume model, the possibilities of the CAN SPAM Act to deter spammers
appears to be mixed at first glance. Many CAN SPAM measures relate to United States TLD links in theoretically expected ways and is suggestive of deterrence, or maybe even displacement. The more articles published on the CAN SPAM Act and the fewer acquittals of spammers under the Act predict fewer US TLDs the following month. As mentioned, the dichotomous measure of the CAN SPAM Act suggests a deterrent effect. Considering that the CAN SPAM Act only has jurisdiction in the United States, this may make spammers think twice about registering US domain names and including them in their spam emails. However, similar to that of spam volume, more detentions of spammers predict an increase in both US TLDs and US IPAs, as well as more arrests of spammers predicting more US TLDs. Arrests and detentions of spammers in the news may have a different psychological influence on readers than originally anticipated.

Also, while the CAN SPAM Act may deter spammers from registering US TLDs, the finding is not necessarily evidence of displacement. Deterrence does not necessarily shift spam operations overseas, but instead persuades United States spammers to simply take more precautions when registering website domains for their spam business. However, articles that were more negative about the CAN SPAM Act were associated with an increase in spam with a US IPA. This finding might represent a displacement effect, but only one relating to where spambots are being hosted over the internet. Spammers may be more willing to host spam sending servers in the US if they feel the CAN SPAM Act is ineffective.

**Spam Compliance with the CAN SPAM Act**

The FTC report to congress suggested that compliance with the CAN SPAM Act had increased following the passing of the act (Majoras, Leary, Harbour, & Leibowitz, 2005). However, the report only referred to legitimate email marketers, who would be expected to
comply with regulations anyway. Other reports suggested compliance had decreased following the CAN SPAM Act among genuinely unlawful spam (Grimes, 2007). However, Kigerl (2009) found no impact of a dichotomized version of the CAN SPAM Act on compliance, with the exception of meaningful subject heading compliance, which decreased following the Act. Different results were found with the present study, finding no significance of the CAN SPAM Act on meaningful subject heading, but instead a binary measure of the act was found to predict higher address inclusion compliance, net all other variables. Yet this research expanded on prior literature by including other measures of CAN SPAM Act compliance, namely falsified header compliance measures.

The implications of deterrent strategies and their impact on spam compliance with the CAN SPAM Act vary depending on which of these two types of compliance outcomes are studied. CAN SPAM Act measures were not consistent across the four compliance measures of opt-out provision, address inclusion, meaningful subject field compliance, and advertisement notice. Some CAN SPAM Act measures were in a theoretically expected direction towards these measures of compliance, while others were not. However, CAN SPAM Act measures were decidedly consistent when predicting header falsification, whenever they were significant across any of the six measures of header forgery.

In fact, all significant IV measures of deterrence were predictive in the same direction: increased deterrence is associated with decreased compliance with header requirements. This was not in a theoretically anticipated direction, although after consideration might be explainable after the fact. While it is expected that deterrent strategies ought to increase compliance with spam regulations, they appear to decrease compliance when violating those requirements and enhances the secrecy and anonymity of the offender when sending spam emails. Header forgery
may be an effective way to prevent investigators from revealing the offender’s identity. If an offender feels the risks are higher for sending illicit spam, or that their anonymity is not secure or for certain, they may take some extra precautions to conceal their origins. Header falsification would be an appropriate means to do this.

Header forgery was highlighted as one of the primary concerns of illegal spam by Congress during construction of the Act (CAN SPAM Act of 2003). Yet the CAN SPAM Act may worsen what it initially sought to reduce, not because the Act is not a deterrent, but because it in fact is a deterrent. It might be argued that spam is worse because the CAN SPAM Act is effective in one of its primary goals.

However, for the remaining four CAN SPAM Act compliance measures, the deterrent prospects of the Act are not clear. The CAN SPAM Act trials of all types, including acquittals, appear to increase compliance with the address inclusion requirement. Acquittals also increase compliance for the advertisement notice requirement. However, acquittals decrease compliance for the meaningful subject field regulation. The deterrent prospects of attribution is also not clear. Attribution increases compliance for opt-out provisions and advertisement notice requirements, but decreases compliance for meaningful subject header inclusion. The meaningful subject field model appears to be the most different of the four. However, the meaningful subject field measure captures efforts on the part of the spammer to deceive, with a deceptive subject heading intended to trick the recipient into opening the spammer's email message. The deceptive nature of this measure may indicate a better fit with the header forgery models, and that might be suggested due to the finding that attribution (a deterrent phenomenon) actually decreases subject heading compliance, just as it decreases header forgery compliance. Prior literature on attribution suggests that higher attribution predicts reductions in cybercrime at
the national level (Guitton, 2012). The opposite effect was found here, with attribution exacerbating noncompliance with header forgery regulations.

*Spam Severity and Illegality*

The CAN SPAM Act and other deterrent influences appear to have little impact on email fraud, yet are predictive of malware distribution among the spam email sample. There were two models for email fraud, and only one of them contained a significant IV. The number of days spammers were detained was associated with a reduction in fraud probability. However, nothing was significant in the percent fraudulent model. It is unclear what a reduction in fraud probability means, as spam emails are either fraudulent or they are not. There is no in-between, and a reduction in keywords associated with fraud is not certain to be a reduction in actual fraud. The percent of emails classified as fraudulent at 85 percent certainty ought to be more valid, yet no IV was predictive of this outcome.

It may be that email spammers who reside in the United States are not as likely to rely on fraud, but rather rely more heavily on spamvertised goods and products. Fraudulent emails received in the United States may be more likely to originate from other countries (Nigeria, Russia, etc.), and therefore the fraudsters would likely not even read US news on the CAN SPAM Act, let alone be deterred by it. However, utilizing the individual spam dataset, there is no correlation between a dichotomous measure of US IPA and a dichotomous measure of fraud at 85 percent certainty ($r = -.037, p < .001$). Of course, the relationship between US IPA and the spammer’s actual home country of origin is not certain and cannot be determined from the present dataset. Therefore fraudulent emails may still be more likely to originate from other nations.
However, this research also proposed an alternative effect of the CAN SPAM Act, that of a marginal deterrent effect. Namely, deterrence decreases the prevalence of less severe offenders, with the more severe perpetrators (such as those perpetrating fraud) being the ones to remain. More serious offenders may be less likely to be deterred, as they have already taken on greater risk than less serious offenders by the nature of their crimes. Specifically it was questioned whether deterrence increases severity or fraud. That was also not found to be the case with fraud models. Assuming the measure of fraud is valid, the CAN SPAM Act seems to have little effect.

For the malware distribution models, however, deterrence efforts may have an impact. Yet the direction of effect sizes is not consistent with the malicious links model. Arrests and fines under CAN SPAM Act prosecutions reduce the number of malicious links, while ongoing trials and attribution increases malware. It may be that articles mentioning punishments of spammers deter malicious links, whereas articles only mentioning spammer identities (attribution) and ongoing trials without a punishment increase malicious links. Yet attribution decreases malware in the malicious scripts model, even though it increases malware in the malicious links model. These two types of attack vectors for distributing malware may be different from each other such that deterrent efforts may have different impacts on them.

The two influences of malicious scripts were consistent with prior discussions of deterrence theory, though, with CAN SPAM articles and attribution predicting reduced malicious scripts. Neither of these measures necessarily mention punishments of spammers. Regardless of effect size direction, the CAN SPAM Act appears to influence malware, whereas it has no effect on fraud. There also appears to be no evidence of a marginal deterrent effect, with an increase in severity associated with a deterrent stimulus. Prior research has suggested such a marginal
deterrent influence when the measure of cybercrime is a DDoS attack (Hui, Kim, & Wang, 2013), yet the same does not appear to be the case when looking at fraud and certain types of malware distribution.

**Limitations**

The spam sample and the procedures for extracting the time series metrics from the sample have some limitations that ought to be mentioned. The sample itself was acquired from only a single web archive, collected by an uploader ([http://untroubled.org/spam](http://untroubled.org/spam)) which may not have been completely consistent in the process for baiting spam messages during the entire 15 years of data collection. That is, changes might have been made or slowly introduced, such as the number or frequency of bait email addresses or address posts on the internet made over time. The spike in the spam volume in 2006 due to the temporary use of wildcard addresses is one such example. While the data ought to reflect genuine spamming activity, there may be fluctuations or systematic changes in time for different observations that would not be accounted for by the existing predictors in the model.

Additionally, the means by which the sample was gathered likely means the spam emails used in this analysis are not representative of the population of unsolicited bulk commercial emails received in all inboxes in the United States. Many legitimate email marketers or senders would not acquire their recipient mail lists via web crawlers scanning public web pages. That is, they would not email the honeynet email addresses. Instead, the spammer may be a company with an internet presence that sends unsolicited commercial email to existing customers, such as Amazon or Microsoft. These types of emails would not appear in the current spam sample.

The possible bias this would introduce into the findings would underestimate compliance with the CAN SPAM Act, as legitimate businesses would undoubtedly be aware of and comply
with most federal regulations. If such emails were able to be included in the spam sample, then the CAN SPAM Act would likely be found to be more effective. However, that option was beyond the scope of this research, and instead the sample more accurately represents illegal spammers and cybercriminals.

The spam sample used might also be skewed towards certain spammers who send more spam than usual to the same recipient’s inbox. For the individual email unit of analysis dataset, each observation is not independent of other observations or spam that was received. That is, multiple emails can easily be sent by the same spammer, and many of those emails are likely identical to each other in the types of variables coded by the spam software. So the sample is biased towards spammers who send more spam to the same recipient. This is likely another example where the data is more representative of more serious, unlawful spammers.

The locality measures of a spam message’s national origins are also uncertain in terms of accuracy. The IPA of a message, for instance, may represent where a spambot is located, where the spammer was located at the time (home address or wifi hotspot), where a rented SMTP server is located, where a proxy server is located, or the IPA itself may be made up by the spammer entirely. In all of these scenarios, the spammer has some say in what the IPA is, but the differences in the causes of each IPA make the interpretation of findings ascertained from this measure uncertain.

Even though significant results were identified utilizing these two measures of spam locality, interpretations come with many of these caveats. It is beyond the scope of this research to differentiate between the different possible sources that caused an IP address to become associated with a given nation. The internationalized TLD may be a better measure than header IPA, however, as forgeries or spoofing of an domain name would be less likely, and it is more
difficult to register a US TLD without being a US citizen, whereas hosting a botnet in the United States is probably easier than hosting in most other countries considering the high internet connectivity of the US.

The database used to geolocate the header IPA also may have some limitations. IP addresses are allocated to different nations, companies, and networks, but this allocation changes over time. Databases to track IPA space allocation changes over time are necessary to determine which IPA ranges belong to which country. The IPA database that the software uses is as recent as 2013, so it may be less accurate for geolocating IPAs included in spam messages in the late 1990s.

Another limitation of the software might exist in the deceptive subject header computation. The software matches keywords used in the subject field with keywords found in the message body. However, the software is not able to account for differences in spelling of said keyword matches. One word may appear in the subject line of an email message and in the body as well, but be spelled differently, such as one being plural, and the software would declare there to be no match. The software would also fail to match words deliberately spelled differently to fool spam filters.

There may be some limitations of the CAN SPAM Act measures as well as limitations associated with the spam sample. All measures of the CAN SPAM Act come from news and media reporting on the Act. Nothing comes directly from CAN SPAM Act prosecutions. This research does not use a comprehensive list of all CAN SPAM lawsuits as well as the outcomes of each lawsuit. Instead measures of law enforcement are taken from news and blog reporting. However, news reporting may be more effective as a deterrent or other psychological influence
on spammer behavior, as this is information they are likely to be made aware of, rather than actual trial case instances.

However, despite these limitations, this study is the first to utilize many of the measures discussed, be they somewhat limited or not. It is also the first to apply the degree of statistical and scientific rigor in evaluating the CAN SPAM Act. Prior literature was also beset with many of these limitations, among many more, without this study’s strengths. Despite the limitations just mentioned, the findings are the first of their kind to reveal many multifaceted and significant influences that the CAN SPAM Act has on spamming behavior.

**Policy Recommendations**

The CAN SPAM Act is not consistent in its impact on spammer behavior, although it appears to be associated. While the CAN SPAM measures chosen were not actually direct CAN SPAM Act prosecutions but rather the responses to those prosecutions filtered through the news, the CAN SPAM Act undoubtedly caused these news reports, and such news potentially reaches the notice of actual spammers. Improvements to the CAN SPAM Act have a similar potential to affect actual spammers, but the ways the Act can influence them is not entirely clear.

Even news reporting that mentions punishment of spammers, such as detentions and arrests, may not necessarily be described in articles positive about the CAN SPAM Act. Spammers who are arrested or detained may quickly be let free, and spammers with damages judged against them may not pay the fines as the courts make little effort on collection. The CAN SPAM Act might need to be expanded or be given more teeth, or perhaps be enforced more consistently.

Much of the CAN SPAM Act goes unenforced, as it is limited to a narrow number of possible authorities that would enforce it, as well as being conservative as to what is classified as
spam. The Act might do well to have its definition of illegal spam broadened or simplified to facilitate easier enforcement. The definition of illegal spam could be simplified to only include unsolicited commercial mail, requiring opt-in. Such a revision might cut down on the annoyances legitimate marketers impose (at the very least), as they would likely comply with the law. As of now, unsolicited spam is not illegal, so long as the spammer abides by specified regulations.

The law might apply to more than just commercial electronic message marketers as well, such as requiring internet service providers to authenticate all email sent from their networks. Also, the entities authorized to file lawsuits against spammers for damages might be expanded, such as permitting consumers to file suit in addition to the FTC, states, and internet access providers. Enforcement under the CAN SPAM Act is not frequent, as the FTC is underfunded and ISPs have little profit motive to bother pursuing spammers on their networks which may not actually cost them that much in the way of direct damages (Rutenberg, 2011); where consumers and recipients themselves absorb most of the negative influences of spam. Opening the option of lawsuits to everyone who might be affected would likely increase enforcement.

Regulation of spammers themselves may not be the most effective route. Spammers are numerous and incapacitating one might simply lead to others emerging to take his/her place in acquiring profits in the spam business. However, many spammers profit from spamvertised products, such as the sale of pharmaceuticals or other products via email solicitations. These products require banks to facilitate the financial transaction between recipient and spammer. Evidence suggests that just a small handful of banks account for a majority of these transactions (Levchenko, Pitsillidis, Chachra, Enright, Félegyházi, Grier, Halvorson, Kanich, Kreibich, Liu, McCoy, Weaver, Paxson, Voelker, & Savage, 2011). The CAN SPAM Act is not consistent in
either reducing spam or increasing compliance, so perhaps newer regulations that target banks might be more effective. However, many of the banks are beyond US jurisdiction, therefore the legislation would, like most problems of cybercrime, require international cooperation. However, this solution still might be more easily implemented than pursuing the numerous spammers themselves.

**Conclusion**

This research is the first of its kind on a number of dimensions. It is the first to use a series of continuous measures capturing the enforcement and awareness of the CAN SPAM Act in its possible influence on the spam as the law is intended to deter. It is also the first to incorporate control variables while conducting impact assessments of the legislation on spammer activity, taking into consideration the fact that there are likely many causes of spam. Finally, it includes a diverse number of measure of spam itself, some of which have not been tested in prior literature addressing the effectiveness of spam legislation. Some of these novel measures include national TLD registrations, fraud classification, and malware distribution.

Many sources have attempted to comment on the efficacy of the CAN SPAM Act. There are some substantial limitations with all of them. This research has improved on these limitations to make a final evaluation of the CAN SPAM Act that is more conclusive. Most research on the CAN SPAM Act to date has concluded that the Act has little impact on spam. Contrary to prior literature, the CAN SPAM Act does appear to influence spam both in ways tested for in existing studies and in ways previously untested or even questioned. The findings indicate the Act has some potential to influence illegal spam outcomes in the United States, and hence, abandonment of the Act is not warranted. Rather, further revisions to the Act that seek to improve enforcement are strongly recommended.
References


Criminology, Atlanta Marriott Marquis, Atlanta, Georgia.


Cyber Security Synergy, 47.


Restrictions of access by minors to materials commercially distributed by means of world wide web that are harmful to minors, 47 U.S.C. Sec 231 (2013).


study. *Hawaii International Conference on Systems Science.*


Appendix A

Figure A.1

ACF - Spam Count

<table>
<thead>
<tr>
<th>Lag</th>
<th>ACF</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-0.2</td>
</tr>
<tr>
<td>5</td>
<td>0.2</td>
</tr>
<tr>
<td>10</td>
<td>0.4</td>
</tr>
<tr>
<td>15</td>
<td>0.6</td>
</tr>
<tr>
<td>20</td>
<td>1.0</td>
</tr>
</tbody>
</table>

The ACF plot shows the correlation between lagged observations of spam count. The dashed lines indicate the confidence intervals.
Figure A.2

PACF - Spam Count

Lag
Figure A.4

PACF - Percent US IPA
Figure A.5

ACF - Percent US TLD

![ACF Plot]

Lag

ACF

1.0
0.8
0.6
0.4
0.2
0.0
-0.2

0
5
10
15
20
Figure A.6

PACF - Percent US TLD

Lag

PACF

-0.15 -0.10 -0.05 0.00 0.05 0.10 0.15
Figure A.7

ACF - Percent Opt-Out Compliance

ACF

-0.2 0.0 0.2 0.4 0.6 0.8 1.0

Lag

0 5 10 15 20
Figure A.9

ACF - Percent Address Compliance

ACF

Lag
Figure A.10

PACF - Percent Address Compliance

Lag

PACF

-0.15 -0.10 -0.05 0.00 0.05 0.10 0.15
Figure A.11

ACF - Percent Subject Compliance

ACF

Lag
Figure A.12

PACF - Percent Subject Compliance

PACF

Lag

206
Figure A.14

PACF - Percent Advertisement Compliance

Lag
Figure A.15

ACF - Percent Header 'To' Found

ACF

Lag
Figure A.16

PACF - Percent Header 'To' Found

Lag

PACF
Figure A.17

ACF - Percent Duplicate From-To Field

ACF

0.0

-0.2

Lag

0 5 10 15 20
Figure A.18

PACF - Percent Duplicate From-To Field

Lag

PACF
Figure A.19

ACF - Percent Sender Name Found

ACF

Lag
Figure A.20

PACF - Percent Sender Name Found

Lag

PACF

Values range from -0.20 to 0.15.
Figure A.21

ACF - Percent 'From' Field Found

![ACF Plot]

ACF

Lag

-0.2

0

0.2

0.4

0.6

0.8

1.0
Figure A.22

PACF - Percent 'From' Field Found

PACF

Lag
Figure A.23

ACF - Percent Duplicate To-Return-Path Found

ACF

0.0

0.5

1.0

Lag

0 5 10 15 20

-0.5

Dashed lines indicate the 95% confidence interval for the ACF.
Figure A.25

ACF - Percent Duplicate To-Reply-To Found

ACF

-0.5

0.0

0.5

1.0

Lag
Figure A.26

PACF - Percent Duplicate To-Reply-To Found
Figure A.28

PACF - Average Fraud Probability

Lag
Figure A.29
Figure A.30

PACF - Percent Fraudulent

PACF

Lag
Figure A.31

ACF - Percent Malicious Links

Lag

ACF

0 5 10 15 20

0.2

-0.2

-0.4

0.0

0.4

0.6

0.8

1.0
Figure A.32

PACF - Percent Malicious Links

Lag
Figure A.33

ACF - Percent Malicious Script Tags

ACF

Lag
Figure A.34

PACF - Percent Malicious Script Tags

PACF

Lag