AFTER THE BREACH: THE MONETIZATION AND ILLICIT USE OF STOLEN DATA

HEARING

BEFORE THE

SUBCOMMITTEE ON TERRORISM
AND ILLICIT FINANCE

OF THE

COMMITTEE ON FINANCIAL SERVICES
U.S. HOUSE OF REPRESENTATIVES

ONE HUNDRED FIFTEENTH CONGRESS

SECOND SESSION

MARCH 15, 2018

Printed for the use of the Committee on Financial Services

Serial No. 115–81

U.S. GOVERNMENT PUBLISHING OFFICE
31–386 PDF WASHINGTON : 2018

HOUSE COMMITTEE ON FINANCIAL SERVICES

JEB HENSARLING, Texas, *Chairman*

PATRICK T. MCHENRY, North Carolina, *Vice Chairman*
PETER T. KING, New York
EDWARD R. ROYCE, California
FRANK D. LUCAS, Oklahoma
STEVAN PEARCE, New Mexico
BILL POSEY, Florida
BLAINE LUETKEMEYER, Missouri
BILL HUIZENGA, Michigan
SEAN P. DUFFY, Wisconsin
STEVE STIVERS, Ohio
RANDY HULTGREN, Illinois
DENNIS A. ROSS, Florida
ROBERT PITTENGER, North Carolina
ANN WAGNER, Missouri
ANDY BARR, Kentucky
KEITH J. ROTHFUS, Pennsylvania
LUKE MESSER, Indiana
SCOTT TIPTON, Colorado
ROGER WILLIAMS, Texas
BRUCE POLIQUIN, Maine
MIA LOVE, Utah
FRENCH HILL, Arkansas
TOM EMMER, Minnesota
LEE M. ZELDIN, New York
DAVID A. TROTT, Michigan
BARRY LOUDERMILK, Georgia
ALEXANDER X. MOONEY, West Virginia
THOMAS MACARTHUR, New Jersey
WARREN DAVIDSON, Ohio
TED BUDD, North Carolina
DAVID KUSTOFF, Tennessee
CLAUDIA TENNEY, New York
TREY HOLLINGSWORTH, Indiana

MAXINE WATERS, California, *Ranking Member*
CAROLYN B. MALONEY, New York
NYDIA M. VELÁZQUEZ, New York
BRAD SHERMAN, California
GREGORY W. MEEKS, New York
MICHAEL E. CAPUANO, Massachusetts
WM. LACY CLAY, Missouri
STEPHEN F. LYNCH, Massachusetts
DAVID SCOTT, Georgia
AL GREEN, Texas
EMANUEL CLEAVER, Missouri
GWEN MOORE, Wisconsin
KEITH ELLISON, Minnesota
ED PERLMUTTER, Colorado
JAMES A. HIMES, Connecticut
BILL FOSTER, Illinois
DANIEL T. KILDEE, Michigan
JOHN K. DELANEY, Maryland
KYRSTEN SINEMA, Arizona
JOYCE BEATTY, Ohio
DENNY HECK, Washington
JUAN VARGAS, California
JOSH GOTTHEIMER, New Jersey
VICENTE GONZALEZ, Texas
CHARLIE CRIST, Florida
RUBEN KIHUEN, Nevada

SHANNON MCGAHN, *Staff Director*

SUBCOMMITTEE ON TERRORISM AND ILLICIT FINANCE

STEVAN PEARCE, New Mexico *Chairman*

ROBERT PITTENGER, North Carolina, *Vice Chairman*
KEITH J. ROTHFUS, Pennsylvania
LUKE MESSER, Indiana
SCOTT TIPTON, Colorado
ROGER WILLIAMS, Texas
BRUCE POLIQUIN, Maine
MIA LOVE, Utah
FRENCH HILL, Arkansas
TOM EMMER, Minnesota
LEE M. ZELDIN, New York
WARREN DAVIDSON, Ohio
TED BUDD, North Carolina
DAVID KUSTOFF, Tennessee

ED PERLMUTTER, Colorado, *Ranking Member*
CAROLYN B. MALONEY, New York
JAMES A. HIMES, Connecticut
BILL FOSTER, Illinois
DANIEL T. KILDEE, Michigan
JOHN K. DELANEY, Maryland
KYRSTEN SINEMA, Arizona
JUAN VARGAS, California
JOSH GOTTHEIMER, New Jersey
RUBEN KIHUEN, Nevada
STEPHEN F. LYNCH, Massachusetts

CONTENTS

AFTER THE BREACH: THE MONETIZATION AND ILLICIT USE OF STOLEN DATA

Thursday, March 15, 2018

U.S. HOUSE OF REPRESENTATIVES,
SUBCOMMITTEE ON TERRORISM
AND ILLICIT FINANCE
COMMITTEE ON FINANCIAL SERVICES,
Washington, D.C.

The subcommittee met, pursuant to notice, at 2:03 p.m., in room 2128, Rayburn House Office Building, Hon. Stevan Pearce [chairman of the subcommittee] presiding.

Present: Representatives Pearce, Pittenger, Rothfus, Williams, Poliquin, Hill, Emmer, Zeldin, Davidson, Budd, Kustoff, Perlmutter, Maloney, Himes, Foster, Kildee, Sinema, Vargas, Gottheimer, Kihuen, and Lynch.

Chairman PEARCE. The subcommittee will come to order.

Without objection, the Chair is authorized to declare a recess of the subcommittee at any time.

Members of the full committee, who are not members of the Subcommittee on Terrorism and Illicit Finance, may participate in today's hearings.

All members will have 5 legislative days within which to submit extraneous materials to the Chair for inclusion in the record.

This hearing is entitled, "After the Breach: The Monetization and Illicit Use of Stolen Data."

I now recognize myself for 5 minutes to give an opening statement—for 2 minutes to give an opening statement.

I want to thank everyone for joining us today.

In today's hearing, we will examine the economics of cyber crime, the monetization of stolen data from cyber attacks, the role the dark Web marketplaces play in helping criminals profit from their theft, and how illicit proceeds are laundered into our financial system.

Last month, the Council of Economic Advisors released a report estimating that malicious cyber activity cost the U.S. economy between $57 and $109 billion in 2016. And this cost is expected to climb as more devices become Internet connected.

Most commonly, these cyber attacks against private and public entities include ransomware attacks, requesting payments in cryptocurrencies, denial of service attacks, and a business e-mail of compromise scenarios. These attacks lead to property destruction; business disruption; and the theft of proprietary data, intellectual property, and sensitive financial information.

Unfortunately, this activity is only becoming more widespread as criminal organizations realize the low cost of entry, the ease of using hacking tools, and the difficulty law enforcement faces trying to apprehend the hackers.

It is estimated that in 2017, there were 610 public breaches in the United States, triggering the exposure of 1.9 billion records.

This sensitive information, including stolen credit card numbers and personally identifiable information, is monetized and sold on the dark Web, often for a few dollars or less, making cyber theft a lucrative endeavor and providing anonymity for the criminals.

Cyber theft is particularly damaging because of the sensitive information being stolen, including Social Security numbers, and is difficult or sometimes impossible to change.

The victim of a breach can become a victim repeatedly as their identity can be used to apply for credit cards, mortgages, and other financial products over and over again.

In today's hearing, I hope to discuss how we are currently combating cyber attacks that lead to electronic identify theft, credit card and other types of fraud, including what tools and partnerships are working well in the effort to detect and disrupt criminal actors.

I would also appreciate any comments about deficiencies in our system that may impede our ability to predict or stop future breaches.

I would like to thank our witnesses for being here today. I look forward to their expert testimony on these very important issues.

Now, the Chair recognizes the gentleman from Colorado for 2 minutes for an opening statement.

Mr. PERLMUTTER. Thank you, Mr. Chair. And thanks to the witnesses for joining us today, and we look forward to your testimony.

I doubt there is a person in this room who hasn't been effected, whether they know it or not, by a data breach. In the Equifax breach alone, 147 million Americans were effected and impacted.

Every day, hackers steal an additional 780,000 records. And according to the Identity Theft Resource Center, there were a total of 1,579 U.S. data breach incidents in 2017.

Criminals have grown more sophisticated, more organized, and so have the markets for purchasing the stolen data. In many cases, the cyber criminals are encouraged and supported by governments.

In terms of state-sponsored cyber criminals, the most pervasive actors are Russia and North Korea, both of which heavily target financial institutions.

And, as we all know, as a fact, Russia used its cyber capabilities to interfere in the 2016 election. I was glad to hear today's news from the Department of Treasury announcing sanctions on 19 Russian operatives and 5 organizations. Many of whom were identified by Special Counsel Robert Mueller.

I am glad to see the Department of Treasury is beginning to take this Russian cyber threat seriously. I hope President Trump will understand the importance of this issue soon as well.

With that, I thank you, Mr. Chairman, for holding this hearing, and I look forward to today's discussion.

And I yield back.

Chairman PEARCE. The gentleman yields back.

The Chair now recognizes the gentleman from North Carolina, Mr. Pittenger, for 2 minutes.

Mr. PITTENGER. Thank you, Mr. Chairman and Ranking Member Perlmutter, for holding this hearing today. Thank you to each of our distinguished panelists for giving their expertise to our subcommittee this afternoon.

Cyber crimes, whether they are sponsored by states or not, are one of our Nation's biggest and most pressing national security threats.

In recent years, we have seen the frequency and size of cyber crimes increase exponentially. The dark net, for online activities and transactions, are largely untraceable. And proliferation of cryptocurrencies has made it easier for criminals to monetize illicit activities.

Of particular concern are easily accessible dark net marketplaces where criminals can, with startling ease, sell or buy stolen data and wide—and a wide variety of other illicit cyber services.

Cyber crimes have wreaked havoc on our businesses and upended the lives of countless Americans. Yet, we must recognize the complex and multi-layered landscape of this threat. We know loan actors and criminal syndicates are behind many of these crimes but so are hostile states.

Notably, for years now, China has used strategic foreign investment through joint ventures to acquire American companies and access their data, intellectual property, proprietary technologies.

Many of China's targeted transactions evaded the purview of the outdated Committee on Foreign Investment in the United States, commonly known as CFIUS. This is the chief body tasked with screening foreign investments for national security risk.

To remedy this problem and safeguard our intellectual property, data and proprietary technology, I have introduced, with Senator Cornyn, legislation to modernize CFIUS and strengthen its ability to identify and stop malicious foreign investments.

The scope and landscape of illicit cyber activities is rapidly evolving. Cyber crimes are becoming more damaging, more frequent, more creative, and are impacting more Americans.

In many ways, we find ourselves alarmingly vulnerable in large uncharted waters. It is imperative we address these threats with the utmost seriousness and remain vigilant and proactive in our efforts to combat all forms of the furious cyber activities.

Thank you, Mr. Chairman.

I yield back the balance of my time.

Chairman PEARCE. The gentleman yields back.

Today, we welcome the testimony of our panelists.

First, we have Ms. Lillian Ablon. Ms. Lillian Ablon is an Information Scientist at the RAND Corporation. She conducts technical and policy research on topics spanning cyber security, emerging technologies, privacy and security in the digital age, computer network operations, among many others.

Ms. Ablon's recent research topics include the intersection of commercial technology companies and public policy; black-markets for cyber crime tools and stolen data; as well as the white-, gray-, and black-markets for zero-day exploits, social engineering

and open source intelligence, tools, and technology for greater cyber situational awareness and many others.

Prior to joining RAND, Ms. Ablon worked for some of the most cutting-edge technologies and encryptos—cryptography network exploitation and vulnerability analysis and mathematics. She has won an Uber Black Badge at the DEF CON 21 Computer Industry Conference. And holds a bachelor's degree in mathematics from the University of California, Berkley and a master's degree in mathematics from John Hopkins University.

Mr. Joe Bernik has over 2 decades of experience creating and implementing cybersecurity management programs at global financial institutions, while serving as Chief Information Security Officer and Head of Information Risk and Security at ABN AMRO Bank, Fifth Third Bank, and BNY Mellon.

Mr. Bernik led global teams dedicated to protecting customer data and complying with data-related laws, regulations, and managing incident response programs.

Mr. Bernik started his career with the U.S. Department of Defense. He is an avid speaker and writer and has held posts on several industry groups, including the Federal Reserve Council on Fraud and the Financial Services Information Sharing and Analysis Center, and the open Web Application Security Project.

Mr. Bernik holds a bachelor's degree in information systems from the University of Mary Washington and has completed graduate studies in business administration at the City University of New York.

Dr. Nicolas Christin is an Associate Research Professor at Carnegie Mellon University, jointly appointed in the School of Computer Science and in Engineering and Public Policy. He is affiliated with the Institute for Software Research and a core faculty member of CyLab of the university-wide information security institute.

He also has courtesy appointments in the Information Networking Institute and the Department of Electrical and Computer Engineering. He was a researcher in the School of Information at the University of California, Berkeley, prior to joining Carnegie Mellon in 2005.

His research interests are in computer and information systems security. Most of his work is at the boundary of systems, networks, and policy research. He has most recently focused on security analytics, online crime modeling, economic and human aspects of computer security.

He holds a degree in engineering from a prestigious French University and both a Master's Degree and PhD in computer science from the University of Virginia.

Dr. James Lewis is a Senior Vice President at the Center for Strategic and International Studies (CSIS). Before joining CSIS, he worked at the Departments of State and Commerce as a Foreign Service Officer and as a member of the Senior Executive Service.

He served on several Federal Advisory Committees, including a Chair of the Committee on Commercial Remote Sensing, as well as a member of the Committees on Spectrum Management and International Communications Policy, and as an adviser on the Security Implications of Foreign Investment in the United States.

Dr. Lewis has authored numerous publications since coming to CSIS on a broad array of topics, including innovation space, information technology, globalization deterrence and surveillance. He was director for CSIS as commissioned on cyber security and is an internationally recognized expert on cybersecurity.

Dr. Lewis received his PhD from the University of Chicago.

Each of you will be recognized now for 5 minutes to give an oral presentation of your testimony. Without objection, each of your written statements will be made part of the record.

Now, Ms. Ablon, you are recognized for 5 minutes.

STATEMENT OF LILLIAN ABLON

Ms. ABLON. Good afternoon, Chairman Pearce, Ranking Member Perlmutter, and distinguished members of the subcommittee. Thank you for inviting me to testify.

As you mentioned, in 2017, there were more than a thousand data breaches, exposing over a billion records of sensitive data. To gain an understanding of what the attackers are doing with the stolen data and how they are monetizing it, we first need to understand who they are and what motivates them.

First, attackers, or cyber threat actors, can be grouped by their sets of goals, motivations, and capabilities. Four groups of note are: Cyber criminals, state-sponsored actors, cyber terrorists, and hacktivists.

I discuss each actor in my written testimony, but the two I would most note for this hearing are cyber criminals and state-sponsored actors. I emphasize the distinction between these groups as they tend to seek different types of data and use or monetize that data in different ways.

Cyber criminals are motivated by financial gain. They care about making money as quickly and efficiently as possible. Often, the data that they steal ends up for sale on underground black-markets.

State-sponsored actors advance the interests of their particular nation's state. They tend to keep the data that they steal for their own purposes, rather than trying to monetize it on underground black-markets.

State-sponsored actors are believed to be responsible for the cyber attack on Sony, the theft of millions of dollars to the Swiss Banking software, and the data breach of millions of records from the Office of Personnel Management (OPM).

Turning to the cyber crime black-markets. They are quite advanced. Full of increasingly sophisticated people, products, and places to conduct business transactions. They are resilient in the face of takedowns and are constantly adapting to the new tactics and techniques of law enforcement and computer security vendors.

They are easy to enter and very easy to get involved in, at least at the most basic level. Essentially, all you need is an Internet connection and a device to become part of the cyber crime ecosystem.

Participants in these markets range across all skill levels. There are often hierarchies and specialized roles. Administrators at the top; followed by brokers, venders, and middlemen; and, finally, mules, the moneychangers who use multiple methods to turn the stolen data into money.

Cyber crime markets offer a diverse slate of products for all phases of the full cyber crime lifecycle. From initial hack all the way through to monetizing the stolen data.

In recent years, as a service offerings, ransomware, malware, and point-of-sale credit card schemes have become popular.

Prices in these markets can range widely depending on hardness of attack, sophistication of the malware, whether something is do-it-yourself or as a service, and the freshness of the data.

For example, credit cards stolen from Target in 2013 appeared on the black-markets within days. Those cards initially fetched anywhere from $20 to $135, depending on the type of card, expiration and limit.

But, eventually, they went on clearance for just a few dollars a card. Although prices, in general, range widely, similar products tend to go for similar amounts.

And anonymous cryptocurrencies like Bitcoin, among others, are preferred for making transactions.

So, how did stolen data get monetized on these markets? Cyber criminals use financial information, things like credit card data and bank account numbers, to withdraw cash, purchase gift cards for resale, or harness a money mule to make fraudulent orders to purchase goods, like expensive electronics, which can, then, be shipped overseas to be sold on other black-markets.

They might use stolen credentials, things like usernames, passwords and e-mail addresses, to get access to a victim's contact list for further spam or phishing campaigns.

Both cyber criminals and state-sponsored actors might use credit report information. Things like addresses, States of birth, and other personally identifiable information, like that taken in the 2017 data breach of Equifax, to create a comprehensive profile of a victim.

Cyber criminals could use that kind of data to create a custom dictionary of possible passwords that can be used to attempt to crack a victim's bank or financial account or for identify theft purposes.

State-sponsored actors, on the other hand, might use this information to build profiles of who to target for exploitation or espionage campaigns or as leverage to gain other types of information.

Unfortunately, there is no easy policy prescription to completely stop data breaches or monetization of stolen data. But a combination of information sharing between the public and private sectors strengthened international cooperation between law enforcement and increased efforts to tarnish the reputation of these black-markets can all help.

Thank you for the opportunity to testify. I look forward to the discussion.

[The prepared statement of Ms. Ablon can be found on page 32 of the Appendix.]

Chairman PEARCE. Thank you.

Mr. Bernik, you are now recognized for 5 minutes.

STATEMENT OF JOE BERNIK

Mr. BERNIK. Good afternoon, Chairman Pearce and Ranking Member Perlmutter. Thank you for the opportunity to testify. My

name is Joe Bernik and I am the Chief Technical Strategist for McAfee, representing the financial services sector.

We are happy that you have addressed this important issue. The financial services sector represents a very sensitive part of our Nation's infrastructure, and I am pleased to see the committee addressing these issues.

According to the SCIS report, recently produced by McAfee, banks continue to be the favorite target of criminals, as we know, probably because the money is held in these institutions. The banks—the attacks, however, are not always directed directly at the banks.

We are now seeing attacks directed at the seams within the institutions themselves. This can be seen in the instance of the swift attack in which alleged North Korea stole or attempted to steal over a billion dollars.

And smaller, less sophisticated organizations are more vulnerable to this type of attack. The practice of not directly attacking institutions, such as was the case—excuse me, such as the case within the Equifax attack, represents a vulnerability within the banks. They all depend on Social Security numbers and, therefore, that type of attack has a lasting and devastating impact on the banks themselves.

All the financial institutions rely heavily on Social Security numbers as a form of identifier. This reliance, as you stated, is a vulnerability as the numbers have all ultimately been lost, resulting in the numbers being somewhat useless as a means of identification.

The methods used are exceedingly commoditized. Malware and phishing attacks are used across the sector. And although new attacks, such as artificial intelligence and machine learning, are available, we have not seen them used because, thus far, the commoditized nature of the attacks doesn't require their usage. So, therefore, the simple attacks continue to be the main methods being exploited and used.

One method of attack that is of extreme importance and urgency right now is the use of social media attacks. Social media—the anonymous-nature of social media, allows for criminals and nation states to use it, to manipulate markets.

I believe and we believe that this type of attack, using social media, will continue to be prevalent and will continue to be devastating against financial markets, given that you can set up an identification without any kind of verification or authentication requirements.

As far as the stolen data goes, and the question that the community had raised, as the previous speaker said, the information is sold on the dark Web for profit by criminals. This information can be easily accessed. The information can be bought for varying prices.

We have seen everything from credit card details sold for $50, Amazon accounts sold for $9.00, passports sold for $62. And the prices vary, depending on the markets and the—and the freshness of the data.

However, the concern that we have, really more so than the data that is being sold today, is the data that we have not seen sold as

of yet. Meaning the Equifax data, which I know everyone is interested in, has not been widely made available in any markets.

It is, therefore, assumed that this data is being collected for other purposes. Potentially for nation state-level attacks. So, that, obviously, the unknown-unknown nature of that type of attack makes it all the more concerning. And we wait and—we are waiting to see what sort of attacks will come from that sort of data that was stolen.

Large institutions have been preparing for cyber-war or cyber attacks for a long time. So, we have seen the sharing of information amongst the banks with the Department of Homeland Security, and scenarios being played out simulating cyber attack. This has been happening for a number of years.

However, since we haven't had one of these events, these large events, occur yet, it is—we are not sure whether we are actually prepared for such an event when it does occur.

As far as policy recommendations go, we offer the recommendations, obviously, to address the Social Security issue and replace a Social Security number with a better identifier, promote cyber security inoperability, pass national breach legislation, and enhance information sharing, such that all organizations can benefit from the intelligence and information that is made available to, currently, some of the largest organizations.

Thank you.

[The prepared statement of Mr. Bernik can be found on page 50 of the Appendix.]

Chairman PEARCE. Mr. Christin, you are recognized for 5 minutes.

STATEMENT OF NICOLAS CHRISTIN

Dr. CHRISTIN. Thank you, Mr. Chairman.

Chairman Pearce, Ranking Member Perlmutter, members of the subcommittee, thank you for hosting this important hearing today and for giving me the opportunity to testify.

My name is Nicolas Christin. I am an Associate Research Professor at Carnegie Mellon University, jointly appointed in the School of Computer Science and in the Department of Engineering and Public Policy.

My research focuses on computer security. For the past decade, I have been studying online crime. In particular, my research group and I have conducted a series of measurement studies on dark Web marketplaces.

We attempt to better understand the potential economic impact of these markets, including the role as retail channels for stolen data. This is the topic at hand today.

In the past 25 years, online retail channels for stolen data have evolved from dial up forums, to online chat rooms, to specialized Web forums, to online anonymous marketplaces, also known as dark Web marketplaces.

Business models have also become increasingly complex to facilitate the sale and purchase of stolen data on a large scale by less sophisticated actors.

Similar to industrial supply chains, the modern market for stolen data shows specialization. The number of technically savvy actors

responsible for data breaches is rather small. After the stolen data is broken down in what is suitable for individual resale, retail-level vendors will offer the data to the general public.

Although criminals provide services surrounding stolen data, such as mule services, or money laundering tutorials without directly interacting with stolen data.

Using measurements we collected between 2011 and 2017, from most of the major online anonymous marketplaces are heightened during the timetable, we can make four observations.

First, revenue generated by criminals engaged in monetizing data breaches continues to pale in comparison to the potential costs of the remedies.

In the early- to mid-2000's, although researchers estimated that criminals made in the orders of tens of millions of dollars per year from the sale of the required data. Meanwhile, the societal costs of those breaches were thought to be in the billion-dollar range.

Our measurements indicate that this asymmetry still exists today. The overall revenue from the entire trade of illicitly acquired data remains rather low, compared, for instance, to the online trade of narcotics.

Stolen credit card numbers are often sold for a few dollars each. More expensive offerings, including Social Security numbers or date of birth, may reach in the order of a hundred dollars apiece.

However, recovering from the damage incurred by each individual theft is far more expensive, due to second-order effects, such as impact on credit ratings.

Second, the dark Web marketplace ecosystem, as a whole, has shown strong resiliency to law enforcement. Shutting down a marketplace has, so far, mostly seemed to result in criminals moving to a different one.

Long-term impacts on the overall illicit trade are uncertain. Takedowns also may potentially lead some of the actors to move the activity to less pubicly observable forums.

Third, 80 percent of the revenue is generated by 10 percent of the vendors. A few successful individuals attract relatively large numbers of amateurs that do not profit much, if at all, from their activities.

These unsuccessful actors, nevertheless, contribute to the overall problem by making the market for stolen data larger and more complex.

Fourth, these marketplaces are international in nature. And even when certain actors are identified, jurisdiction issues may complicate prosecution or arrest.

These findings indicate that focusing on preventing breaches from happening in the first place is probably more economically efficient than attempting to disrupt retail and distribution channels.

Prevention is also likely to be more effective than recovering from a data breach, once it has happened.

Finally, measurements of dark Web marketplaces solely focus on the retail end of the stolen data ecosystem. They, thus, are an imperfect signal, particularly when it comes to tracing stolen data back to a specific breach.

Nevertheless, these measurements give us important information on the health and evolution of the market for illicitly acquired data and on the monetization techniques in use.

Thus, it is important to continue supporting these documentation efforts, to understand the criminals' business models, determine the most specific strategies to disrupt them and improve overall security.

Thank you very much.

[The prepared statement of Mr. Christin can be found on page 57 of the Appendix.]

Chairman PEARCE. Thank you, sir.

Dr. Lewis, you are recognized for 5 minutes.

STATEMENT OF JAMES LEWIS

Dr. LEWIS. Thank you, Mr. Chairman and Ranking Member Perlmutter. I appreciate the opportunity to testify.

Cyber crime is big business. I think you have heard that from all my colleagues. We have conducted three studies with the support of McAfee to estimate the cost.

In interviews for our studies, one senior official called this the greatest transfer of wealth in human history, while another said it was a rounding error in a $14 trillion economy. So, we hope to bring a little more precision to this range.

Estimating the cost of cyber crime is difficult because data collection is willfully inadequate. Most countries don't collect statistics on cyber crime. And many victims prefer not to report their losses.

Our reports looked at a broad range of costs, including recovery costs, I.P. theft and damage to brand.

Our most recent study estimated that cyber crime cost the world between $450 and $600 billion a year, a 20 percent increase in 2 years.

This increase can be explained by the growing sophistication of cyber criminals, by the increase in the number of Internet users and by improvements in the ability of cyber criminals to monetize stolen data.

This has always been a problem for cyber crime. You can take personally identifiable information or intellectual property, but then turning it into actual cash can be a challenge.

One of the reasons cyber crime continues to grow is that criminals have become better at monetization, in part because of the availability of cryptocurrencies. Cryptocurrencies make cyber crime easier by increasing anonymity and by simplifying money transfers.

Cyber crime activity on what is called the dark Web, the hidden Web, also contributes to the growth in cyber crime. This hidden Internet is a safe space for cyber crime.

And I was—in preparation for the testimony, I was looking at some of these sites this morning, and I found one that offered a money-back guarantee if you bought data from them, stolen data. And it didn't work. They would—they would refund your—so, it is a very sophisticated market.

Another reason for the growth as you—of the cyber crime, as you heard, is state-sponsored cyber crime. Russia is a haven for the

most cyber—advanced cyber-criminal groups in the world. The Kremlin sees Russian cyber criminals as a strategic asset.

The other state that extensively supports cyber crime is North Korea. It uses hacking by its principle intelligence agency, the Recognizance Genera Bureau, to generate hard currency for their regime.

So, this is a daunting set of problems. You have protected spaces on the dark Web, innovative and dynamic cyber criminals, cryptocurrencies in countries that provide safe havens and support for cyber crime.

But there are actions we can take to reduce risk. As you heard earlier, we won't be able to eliminate cyber crime, but we can make better efforts to manage it.

This would include the U.S. and its allies, developing an effective strategy for punishing states that support cyber crime, greater regulation of cryptocurrencies, and expanded efforts to disrupt criminal networks, in partnership with our allies in other countries.

Finally, all nations would benefit from a serious effort to collect data on cyber crimes' cost. I think that would be helpful.

I thank the committee for the opportunity to testify and for its work on illicit finance and for our CFIUS modernization and look forward to any questions.

Thank you.

[The prepared statement of Mr. Lewis can be found on page 66 of the Appendix.]

Chairman PEARCE. Thank you, sir.

The Chair now recognizes himself for 5 minutes for questions.

So, I think, Dr. Lewis, I would ask you, first, that estimating losses is hard, according to what you are saying. I think we understand that.

Is there—what, sort of, effort is there, internationally, to, maybe, join together countries? First of all, which country is probably the best at intercepting and stopping the cyber crime? And then, are there international efforts where countries are joining together?

Dr. LEWIS. Thank you, Mr. Chairman.

There is a good correlation between countries that have strong law enforcement systems and punishment for cyber crime.

So, if you are a cyber-criminal and you live in the U.S. or the U.K. or France or Germany, your life expectancy is probably only about 3 years before you are caught and go to jail.

In places that have weak cyber-security laws, like Brazil or countries—other developing countries, you see a growth in criminal activity.

So, the effort here is to have strong cyber-security laws. The U.S. leads in that with the Budapest Convention and to develop new ways to cooperate on the exchange of evidence and on the efforts to take down networks.

So, currently, there is no central place that does this. The U.N. has a committee on crime that is trying to develop a more common approach. But the differences among nations make it hard to get—nations make it hard to get cooperation.

Thank you.

Chairman PEARCE. Thank you.

Mr. Bernik, you had mentioned North Korea as being one of the state actors. And then, the testimony of others indicated Russia.

Who are the other major players, as far as state-sanctioned, state actors?

Mr. BERNIK. Another major player would be China. They have invested a lot of resources in building capabilities and also have been attributed to some of the most significant hacks at recent times.

So, the Anthem hack, as you will recall, which was the big one that occurred a few years back, where a lot of medical and Social Security numbers were stolen. As well as the Yahoo hack has been attributed to them.

This—these hacks and this information is being amassed for a purpose. We just—we just don't know what that purpose is.

So, that threat and that capability that they are massing, raises a significant amount of risk to our—to us, as a—as a—as a country.

And, I think, that is the to-be-determined risk. What that will look like and whether those attacks, if they occur, will be targeted against infrastructure, banks, individuals. We don't have the answer right now and that is—that is one of the most concerning things, I think, for all—for all of us.

Thank you.

Chairman PEARCE. Yes. And Ms. Ablon, the lack of consequences, obviously, plays a big role in encouraging.

Are there any nations that appear to be dealing with the lack of consequences? I don't think that we are.

So, what is your comment on that?

Ms. ABLON. Specifically for the cyber crime markets, they are highly reliable. And so, products are what they say they are. People do what they say they do. Trying to tarnish the reputation is quite difficult.

In terms of specific countries that are going after it, it is, really, on a country-by-country basis. Law enforcement, here in the U.S., is getting better in going after cyber criminals. Certainly, more resources would help.

But more digital natives are entering into our law enforcement and that helps to understand the nature of cyber crime and the technical capabilities.

And also, suspects, in the last few years, are going more after big companies, rather than specific individuals. And that allows cyber crime to bubble up and be more seen and giving more opportunities for U.S. law enforcement to go after them.

Chairman PEARCE. Dr. Christin, if you were contemplating the hack into the Office of Personnel Management, what advantage does that—how is that information viable to the nation states? What do they use it for?

Dr. CHRISTIN. I tend to focus on economically motivated cyber crime. And, as such, I will not, really, be able to answer that question because it is not clear that there are actual economic incentives to use the OPM breach.

Chairman PEARCE. OK. My time is expired.

The Chair now recognizing the gentleman from Colorado, Mr. Perlmutter, for 5 minutes.

Mr. PERLMUTTER. Thanks, Mr. Chair. And this is all very interesting. And, for me, just some very basic questions.

Ms. Ablon, if you were a bad guy out there, and Dr. Lewis talked about going to the Internet today and just skimming some stuff.

So, how does somebody find out about the dark Web, and, if they want to go purchase some information? Just give us a little primer on that.

Ms. ABLON. It is pretty incredible how easy it is to get involved in these markets.

As I mentioned in my original testimony, all you need is an Internet connection and a device to get involved.

I have seen—certainly much of the markets are in the dark Web. Things where you need special tools or special services to access things like Tor, the Onion Router.

But there is plenty that can be found on the surface Web. Things that you can Google for.

For example, I have seen Google guides on how to use a particular exploit kit. I have watched YouTube videos on where to find and buy stolen credit card data.

So, this kind of stuff is easily accessible and within a few finger taps.

Mr. PERLMUTTER. Could I—could I get on there and query, where does Steve Pearce live? Or give me credit card information about Steve Pearce. Just me. Ed Perlmutter, I go on. I want to know something. I want to pick up something on him.

Ms. ABLON. So, in terms of just getting general fungible data, so things that are reusable, you can certainly find that in mass quantities. Random Social Security numbers. To find a particular targeted person, that would require a little more work.

Now, as I mentioned, as service offerings are increasing, you can hire someone to try and find that particular data. Or with enough information, try to go after a particular e-mail account and guess the password of whoever you are trying to target for in order to get their information.

Mr. PERLMUTTER. OK, thank you.

Mr. Bernik, you ticked through some major hacks. I seem to be—so, yes, I have Anthem. You forgot J.P. Morgan, Equifax, Target, Department of Personnel and you forgot the DNC. OK?

So, you didn't want to speculate as to—who wants this information? What do you think they can do with it? They could get credit card information and maybe steal something?

Mr. BERNIK. Right.

Mr. PERLMUTTER. Let us go bigger. Let us go, one, who are the big purchasers? Is Russia? Is North Korea? And what are—what are—what would they do with this stuff?

Mr. BERNIK. We have done a lot of studies with Dr. Lewis on this and trying to—trying to analyze that very question.

The reality is that we are at a cyber—some say a cyber war with these nations now. It is a cold war, if you will. It is not—we are not full-fledged.

We are gathering the constant—they are gathering the constant information. We are gathering this information to use it, potentially to understand how corporations operate individuals of interest.

They may be able to use this as leverage, by having information about an individual, their medical conditions. There is a lot of power in having information as—

Mr. PERLMUTTER. So, these states could be both the hackers and the buyers of information?

Mr. BERNIK. In some senses, they are the—right. They could be the buyers, the aggregators of the information. They are the perpetrators, in some—in some cases of the attacks, themselves.

So, although we are not certain, in many cases, because attribution—the anonymous nature of the Internet makes attribution very difficult, as has been stated.

So, we cannot, 100 percent, guarantee that these are the attackers. But all indicators point to them, to China, North Korea. and, in some cases, Russia.

They are gathering this information for a—to launch attacks against our populists, potentially, to influence, to direct individuals to do things on their behalf, we know that.

So, I think that is what we are going to see more of in the future. We haven't seen it yet.

Mr. PERLMUTTER. Dr. Lewis, you mentioned the cryptocurrencies and the camouflage or the obscurity of these things. Can you—can you expand on that just a little bit?

Dr. LEWIS. Sure. The way that you can acquire these currencies can make it difficult to trace back who is actually buying them.

And so, good trick would be to steal your credit card, buy the cryptocurrency, while using your credit card, and then, it is—it can be anonymous as to who is actually acquiring it after that.

And you can—just as you have done with money laundering, you can go through a number of steps to help obscure the trail.

One of the interesting things, as we all know about Bitcoin—and Bitcoin isn't anonymous enough for cyber criminals, so they are developing a range of new cryptocurrencies that are even harder to track. So, this is a gift to money laundering.

Mr. PERLMUTTER. OK, thank you all for your testimony.

Chairman PEARCE. And if the gentleman is going to really search my data, you probably ought to do it quick because it is—it is about to be emptied anyway. So, move fast.

The Chair will now recognize Mr. Pittenger for 5 minutes.

Mr. PITTENGER. Thank you.

Dr. Lewis, I do appreciate you mentioning CSIS in your written testimony. As I have previously noted, Senator Cornyn and I have introduced legislation to reform and modernize the CSIS process.

Could you please elaborate on how the Chinese are using joint ventures to steal our critical technologies and know how?

Dr. LEWIS. Yes, thank you, Mr. Chairman. And I should congratulate you. Didn't you have a journal op-ed?

Mr. PITTENGER. Yes, sir.

Dr. LEWIS. Good op-ed.

Mr. PITTENGER. Thank you.

Dr. LEWIS. Let me touch on two cases that are recent that we know about that illustrate this and answer some of the questions that came up earlier.

Just last week, or just this week, we saw the President block Broadcom from acquiring Qualcomm. And a few months ago, we

saw CFIUS block the Ant, Chinese, Financial company's efforts to acquire an American company.

And we can think about Chinese behavior as, really, an intelligence activity. It is an effort to acquire data.

If you look at what the Chinese are doing, they are investing in artificial intelligence and big data analytics in quantum computing and quantum communications. And they may, actually, be ahead of us there.

And they are building a global communications network, using their telecom companies which have close links to the states.

So, if China is building an intelligence capability, one of the things they need to do is populate that with data. And so, acquiring U.S. companies that would ease that acquisition of data.

The thing that is interesting to me is we are all fairly familiar with what CFIUS used to do. So, the first bill blocks acquiring military technology. First of all, FINSA blocked terrorist and Homeland Security concerns.

And now, I think it is time for modernization, as the bill you have put forward does to think about how China uses this, not just for military advantage but for intelligence advantage.

Mr. PITTENGER. Could you, Dr. Lewis, give us some greater detail of the types of critical technology and intellectual property that China and other countries are trying to steal?

Dr. LEWIS. Sure. And one easy way to track that is to just look at Chinese activity in Silicon Valley.

So, a lot of attention to artificial intelligence, a lot of attention to big data. They are also looking at sensor technology which can be useful, both on the Internet and for your military application.

They are looking at space technologies. So, they are looking at autonomous vehicle technology. And when I say looking, I should probably say looking to acquire.

So, the Chinese have identified the crucial technologies for modern military and are seeking to use joint ventures, greenfield efforts in the valley, acquisitions of U.S. companies or other western companies.

And you all probably remember KUKA, the German robotics firm, that the Chinese were able to acquire. They have a good strategy for acquiring the technologies for a 21st century military.

Mr. PITTENGER. Thank you.

If you could just elaborate some more on how this threatens our U.S. businesses and international security.

Dr. LEWIS. Sure.

So, one of the problems is that Chinese state-supported investment in high-tech companies crowds the market. So, if the market can support 10 companies and the Chinese subsidize 3 more, everyone's revenue share falls down. Every company is made weaker. Every company invests less in R&D. And that will hurt us.

Our dependence on some Chinese technologies creates intelligence vulnerabilities that we have seen China exploit in other countries.

Chinese efforts to modernize its military have gone into high gear. And when you look at anti-satellite efforts, precision guide and admissions, economic strike, cyber attack, they have found

that China, itself, has become very strong, as an innovator, but they still gain advantage from borrowing other people's technology.

And I think those are the areas I would look at.

Mr. PITTENGER. Yes, sir. They sought to acquire semi-conductor companies. I think they have acquired 20 over the last few years.

What impact, do you believe, that this has already had and how critical and what kind of crisis are we in now to try to do something about reforming CFIUS?

Dr. LEWIS. So, China has—had creating a domestic semi-conductor industry as a goal since it opened to the west in the early 1980's. And they have failed, each time, despite spending billions of dollars because it is hard to make semi-conductors.

And so, their most recent strategy is, let us just buy the whole company. And I think CFIUS has done a good job at blocking that.

But the Chinese are persistent. They are well resourced. And they have not given up on this goal in more than 30 years.

The effect on the U.S. is that we could become dependent on sensitive technologies from China that the Chinese could take advantage of. That is a real concern. That is a supply chain concern.

The second one is that U.S. companies could find themselves hard pressed to continue to invest, hard pressed to innovate. And the market could tilt away from the U.S. and toward China.

Mr. PITTENGER. Thank you.

My time has expired.

Chairman PEARCE. The gentleman's time has expired.

The Chair now recognizes the gentleman from Connecticut, Mr. Himes, for 5 minutes.

Mr. HIMES. Thank you, Mr. Chairman. And thank you, all, for your testimony.

I have heard a theme reiterated today that I first heard from Gartner which happens to be in my district in Stamford, Connecticut.

And the point made is that there aren't a lot of new attacks, new technology, new software, zero-day software. There is just not a lot out there.

That most of the successful attacks are using techniques and malware that are readily identifiable. And that the problem is that people simply aren't using good cyber hygiene. That they don't update their security software. That sort of thing.

Setting aside, for a moment, the question of policy, which we have discussed here a little bit. We, as Members of Congress, interact a lot with the—with the public and our constituents.

I would love to just take my time to cycle through the witnesses. And apart from the obvious, and by the obvious, use of two-factor authentication, not using your birthdate as a password, changing your password periodically.

Apart from the obvious, what would you suggest to us are other measures that our constituents, that the American public should take to try to increase the overall level of network security and the—and the safety of their data?

Ms. ABLON. As you mentioned, there is no new attacks just new attack surfaces. So, as things, like the Internet, if things come up, and there are a lot more digital devices people are not necessarily

thinking about securing their thermostat, like they are their computer.

So, there is certainly the normal cyber hygienes that can be applied to those new attack surfaces.

I would also say that it is not possible to be 100 percent secure. A determined attacker will get through no matter what. So, if we can make it more expensive, in terms of time, resources, and research for an attacker to get through, then that can—that can be helpful.

Something—humans are the weak element. So, if we can educate people to be aware of the kind of attacks that might be facing them, that is something that is an obvious cyber hygiene thing. But the more that we can do it, the better.

Mr. HIMES. Mr. Bernik?

Mr. BERNIK. So, we, at McAfee, would suggest that you invest in software to protect your computer. I think it is pretty basic, at this point. There are a lot of different options. I had to say that, didn't I? It is the correct answer.

But beyond that, I will—

Mr. HIMES. Let us take a commercial break.

Mr. BERNIK. Beyond that, I would say that—don't use the same password for everything. People just do that because it is easy.

And, I think, people—it used to be said, don't write your password down. People would say that. But I think they are going to change it. Write it down. It doesn't matter.

Just don't use the same password for everything. Because once you get attacked once, you are hacked on everything if you use the same password which most people do.

Lock your Social Security report. If you—if you are not applying for credit, then lock your report. Everybody should do that. Because if they have your Social Security number, they can probably—maybe hackers can probably do something against your—against your credit.

So, by default, you should lock your report at all times, if you are not applying for credit. That is basic. And that is free.

What else? Those—I think those two—those three things, using protection on your computer, keeping it patched and up to date, not using the same password, and locking your Social Security report would be—it is supposed to be your credit report. Pardon me. Your credit report would be, in my advice, for individuals.

Thank you.

Mr. HIMES. Before we get to Mr. Christin, Mr. Bernik, since you brought it up, what, in 20 seconds or so, is your take on some of these password protection apps, like Dashlane and others? Are they secure?

Mr. BERNIK. Well, so, what they do is they control—use an app you install that controls all your credentials in one place and it is, basically, in the cloud, effectively. It is stored in a database on the Internet. It is one key that unlocks all keys.

I, personally, think they are useful because they let you change and create random credentials which is more effective than what most people do which is use one series, and they just change the last couple of numbers. Or where they don't have to change it, use the same password for everything.

So, I would say that they are useful tools if used correctly. If you use a weak password or weak credential, and you use that credential as the key, then you are, basically, creating a disaster for yourself.

So, used the wrong way, that could be very disastrous.

Mr. HIMES. Great, thank you.

And, very briefly, Dr. Christin, Dr. Lewis, anything to add?

Dr. CHRISTIN. Yes, I would echo the previous witness, his comments on the password materials. They are very useful and they should be used to generate passwords, as opposed to simply recalling them. Because computers are really good at generating long, random unguessable strings.

That would be my main recommendation.

Mr. HIMES. Thank you.

Dr. LEWIS. Think about where you go online. You probably saw in the indictments today—pardon me, in the sanctions today that one of the tactics that cyber criminals use is what they call waterhole attacks.

Think about where you go. Think about what you put online. Think about what you click on. Be cautious with social media.

Do the basic hygiene. People still don't do that.

And, finally, back up your data. If you would use iCloud or one of the other cloud services, it makes you a little more difficult to suffer from a ransomware attack.

Mr. HIMES. Thank you.

I yield back, Mr. Chairman. Thank you.

Chairman PEARCE. The gentleman's time has expired.

The Chair now recognizing the gentleman from Pennsylvania, Mr. Rothfus, for 5 minutes.

Mr. ROTHFUS. Thank you, Mr. Chairman.

I want to go to Dr. Lewis.

In your testimony, you said that Russia is a haven for the most advanced cyber-criminal groups. And that they use cyber criminals as a strategic asset.

Is the Russian government directly profiting monetarily from cyber crime?

Dr. LEWIS. It would be safe to say that members of the Russian government profit directly from cyber crime.

Mr. ROTHFUS. Do we have any estimate of the revenue that they would generate?

Dr. LEWIS. We could probably come up with one. I did not for this hearing, so it may be a question.

I don't know what the other panelists think. But we know that this is a very profitable line of activities. So, at a minimum, it is probably in the hundreds of millions of dollars.

Mr. ROTHFUS. How do state-sponsors of cyber crime recruit or obtain the services of cyber criminals that carry out illicit activity?

Dr. LEWIS. In countries like North Korea, it is very easy because they are members of either the military or the intelligence services.

In places like Iran or China, and to some extent Russia, they are hackers who come to the attention of the security services. And it is suggested that they cooperate with the state.

In Russia, there are both state programs to identify potential hackers and a linkage between the security services and cyber criminals.

So, each one is a little bit different. But if you monitor the Internet, you can always see when somebody is doing something bad. And then, you go to their house and say, jail or play ball.

Mr. ROTHFUS. Mr. Bernik, in your testimony, you discussed how ransomware is the fastest growing form of cyber crime. Can you discuss the various reasons why ransomware is becoming a more popular tool used by cyber criminals?

Mr. BERNIK. Certainly. It is a very commoditized tool. The ransomware can be purchased on the dark Web through exchanges. It is a commercial-grade software so it is very effective.

As Dr. Lewis mentioned, there are situations where you can get a money-back guarantee on that ransomware. So, you can pay for it. You can pay with cryptocurrencies. So, it leverages all the best and worst parts of the technology available to the criminals and that is why it is effective.

And the punishment for not paying is you don't get your data back. So, the damage is you may be out of business and you may have lost all your personal information, depending on whether you are a company or an individual.

This is the reason why ransomware is so fast growing and so effective.

Mr. ROTHFUS. Which type of cyber attack methods are companies and governments currently most and least equipped to prevent?

Mr. BERNIK. That is a good question.

So, as was previously mentioned, malware and ransomware has become commoditized. The difference between them is just the update with the latest vulnerabilities.

So, if you take a new vulnerability that just came out, say yesterday, and you add it to an existing kit, it will be very effective because that vulnerability will have no protection. It is often referred to as a zero-day because there is no protection for it, the first day.

So, that is the most dangerous scenario for any organization where they have a missing configuration or patch issue, as was the case in Equifax.

So, as we move the window from availability of a vulnerability to its inclusion in a kit, the danger is greater. Because no one—fewer companies will have the protection, if at all.

And that is the biggest fear of these organizations. That a destructive type zero-day attack will occur. Where they are racing machines at fast clip—at a fast pace. And there is no, necessarily, protection that you can have for that type of attack.

And that would be the worse-case scenario and the one we are least prepared for, as a country and as organizations.

Mr. ROTHFUS. I was intrigued when Mr. Perlmutter was talking about looking for some data on Mr. Pearce.

This is a question I am going to ask Ms. Ablon. What—how would—if you went out looking for the data and wanted to, then, buy the data, what payment methods are being used to buy this illicit data?

Are they using Bitcoin? Are they using—do they send cash through Western Union? How do—how does one pay for data like that?

Ms. ABLON. You can pay with it with any method. Cyber criminals will accept money in any way that they can get it.

So, absolutely you can pay with PayPal. You can pay with digital currencies that aren't crypto, that aren't hidden. So, things like Web money, Western Union. You can also pay a crypto card.

Mr. ROTHFUS. Are they do—you can but are they? Do we have—do we know what they are doing?

Ms. ABLON. Yes. Yes. So, there are people that pay with that more and more. There is crypto card—

Mr. ROTHFUS. With what?

Ms. ABLON. With—pay with non-cryptocurrencies.

But more and more, the trend is to go toward cryptocurrencies because of their anonymity—anonymous properties.

The thing about cryptocurrencies is that they are anonymous until you get to the exchange. The crypto—the bitcoins' exchange—the cryptocurrency exchanges is when you actually turn the digital money into actual cash, Euros or dollars. And that is the point where you can tie a human being to the wallet, to the digital currencies.

That is, really, the weak point to go after.

Mr. ROTHFUS. My time is expired.

Chairman PEARCE. Anybody that would pay a hacker for—with a credit card is just asking for trouble, it looks like.

The Chair would recognize Mrs. Maloney for 5 minutes. Oh, I am sorry. Ms. Sinema for 5 minutes.

Ms. SINEMA. Thank you, Mr. Chairman. And thank you to our witnesses for being here today.

Mr. Chairman, more than most, Arizonans value their privacy and that is why we have been outraged by data breaches, like the one in Equifax. And we are frustrated there has been so little action by Congress, the CIPB and others to hold Equifax accountable and prevent future breaches.

We all know this is a growing problem that requires action. Just in the last year, there were over 1,000 breaches that exposed over 1 billion records of sensitive data, according to the Identify Theft Resource Center.

And that makes fraud significantly more likely which is why we are working across the aisle to protect Arizonans from its identity theft and financial fraud.

Arizona's 1.1 million seniors are especially at risk, which is why I am working to pass the Senior Safe Act.

Our bill with Congressman Poliquin, of Maine trains employees at banks, credit unions, and other financial institutions to spot financial fraud against seniors and report to law enforcement. Our bill was recently endorsed by AARP and it passed the House with the support of both parties.

But seniors aren't the only ones with significantly greater risk of financial fraud. We are also working to protect Arizona's children from synthetic identify theft which occurs when a criminal takes a Social Security card—or Social Security number.

And uses it to open bank accounts and lines of credit under a fraudulent name. This type of I.D. theft is often targeted at children because they have no prior credit history.

In Arizona, a 17-year-old girl discovered, to her horror, that a scammer had accumulated over $725,000 of debt in her name. Her information was linked to 8 suspects who opened 42 accounts, including mortgages, auto loans, and credit cards.

So, targeting our kids and running up massive debts in their name is both shameful and cowardly. We have to fight back to ensure they have the change to build their futures.

So, we have introduced the Protecting Children From Identify Theft Act which is a common-sense fix that modernizes Federal fraud detection to stop criminals and protect Arizona's kids.

Every—Arizonan deserves financial peace of mind and we are going to get these bills signed into law.

Mr. Chairman, last month, I requested more hearings on Equifax and these data breaches, and I am glad we are now getting the opportunity to dig deeper into these important issues.

So, with that, I have a question for Ms. Ablon from the RAND Corporation. So, thank you for being here today.

The two bills that I mentioned today focus on enhancing cooperation between Government, law enforcement, and the private sector to catch cyber criminals and protect law-abiding Americans.

Your testimony has noted the importance of these efforts, and there are highlighted steps that we could be taking to disrupt cyber crime markets, it was the clearing houses for criminals, sell our personal and stolen information.

Identity theft operations vary in both scope and sophistication. So, I have two questions for you. What percentage of these illicit operations would you say directly rely on the use of reliable cyber crime markets to be profitable? And which Federal agency is best equipped to infiltrate and thwart these markets?

The second question is, what additional authorities and resources should Congress provide to crack down on these cyber crime markets?

Ms. ABLON. I can't give a specific number of the percentage of identity theft victims or identity theft directly related to the cyber crime markets. However, I would posit that it is quite high, given the accessibility, the availability, and the reliability of the markets.

In terms of what authorities can do to crack down. I mentioned three things in my testimony: International cooperation, information sharing, and then tarnishing of the reputation of the markets.

With international cooperation, this is an effective strategy, especially as I mentioned before, these bitcoin exchanges are the weak point in identifying who the attackers are, who the cyber criminals are.

More and more, these bitcoin exchanges are hosted overseas, so having good international relations with other countries can help law enforcement in the U.S. work with law enforcement overseas and try to get to the actual people to attribute—to detect, attribute, and then interdict the cyber criminals.

In terms of information sharing, information sharing is something that gets talked about a lot. As one of my RAND colleagues

has mentioned, information sharing is not a cyber-security panacea. It won't solve all problems, however it can be very helpful.

Information sharing between law enforcement and banks can be useful as well as small businesses, to let them know what they should be doing. What they should be looking for. What bad or odd behavior looks like in order to, then, notify law enforcement.

Also, sharing information with consumers about who are the victims of data breaches of what they should be looking for as well, can be useful for them to call their credit cards—credit card companies or call places like Equifax or other places that might have their identify information to shut those down so that the cyber criminals can't monetize those or can't take advantage of those.

Ms. SINEMA. Thank you, Mr. Chairman. My time has expired.

Chairman PEARCE. The gentlelady's time has expired.

The Chair now recognizes the gentleman who has been selected as the preseason all-star from Texas, Mr. Williams.

Mr. WILLIAMS. Thank you for that introduction, Mr. Chairman.

In 2017, more than 1.9 billion records were exposed to public cyber breaches. As of this year, we only have—half way into March, cyber breaches have already exposed nearly 20 million records across the Nation. Important cyber information, including intellectual property and personal information continues to be the target.

What is alarming to me is that terrorist and state-sponsored regimes, like North Korea or China, are often behind these attacks, as we talked about. They will continue to take advantage of America's cyber-security weakness. We cannot let that happen.

And I hope the testimony today begins to let us come up with solutions on this pressing matter. And I want to thank the—all of you for being here.

The first question real quick, Ms. Ablon, is what advice would you have for everyday citizens to do if they become victims of stolen data, ransomware, or other crimes?

Ms. ABLON. The one piece of advice I would give consumers, who are more and more becoming victims, is to be alert. Be aware of what is going on. Be—as Dr. Lewis mentioned, look where you are going online.

And then, also, be a little paranoid. I think it is safe for everyone to be a little paranoid about what—where their data is going and their activities online.

Mr. WILLIAMS. OK, thank you.

Mr. Bernik, what lessons, in dealing with the aftermath of mass hacking attacks, like we have seen in the last few years in the breaches, as we have spoken, again, Equifax, Home Depot, Target, and J.P. Morgan, has the industry learned as the result of those attacks?

Mr. BERNIK. The industry has learned to prepare more effectively through scenarios. So—and, obviously, the sharing of intelligence.

So, when an organization becomes aware of a threat, they will run a scenario where they will, basically, self-assess themselves against that threat and understand what the implications might be should they become impacted.

Another thing they have done is prepare for the outcomes. These are corporations now—to prepare for the outcomes of those attacks, meaning preparing for destructive-type malware that erases sys-

tems, creating backups, offline backups that are separated from their online backups.

So, they are really gearing up for what they feel will be, essentially, inevitable scenarios that will play out for them. And that is something they learned.

Mr. WILLIAMS. Good.

Dr. Lewis, you mentioned in your testimony that monetization is easiest for criminals when they can transfer funds directly from the victim to the bank account.

Are there particular jurisdictions that we—that are especially vulnerable to hosting criminal accounts like these?

Dr. LEWIS. Yes, thank you. The interesting part for me here too is, this will fall certainly within the interest of the committee, is that it very closely parallels money laundering.

So, when you think about Malta, Cypress, some of the other countries where you would want to do money laundering, Eastern European banks have, in the past, been a good target.

Usually, there are multiple hops. So, it goes from your bank account to another one and then to a third one and then, maybe, to one of these money laundering centers.

Now, it may just disappear in the void because, at some point, as Ms. Ablon has said—oh, I am sorry. It looks like money laundering. It tracks very closely with how money laundering is carrying out.

And its cryptocurrencies are changing that a little bit by making it easier to hide the tracks of where it goes.

But if you know how money laundering works, and, of course, the members of this committee do, that is a very similar pattern.

Mr. WILLIAMS. OK, thank you.

Dr. Christin, you mentioned the sale of services surrounding data breaches, like data verification and money laundering. Could you discuss these services or steps we might be able to take to prevent those services?

Dr. CHRISTIN. Yes, thank you.

So, for instance, an example of services, what is called money mules, and at a high-level, very simply the way they work is that somebody is being recruited online for a work-from-home type of opportunity.

And the way it works is that this person is instructed to transfer moneys from a stolen account. They don't know it is stolen, they are just being given a number into an overseas account or into their own account before transferring it to an overseas account.

So, that is one of the avenues that is being used for money laundering. Very similar to what drug dealers are using for the transport of drugs.

To address this kind of problem, I think that, what Dr. Lewis was mentioning earlier, in terms of putting some pressure on certain financial institutions, is probably the best—the best avenue.

Thank you.

Mr. WILLIAMS. Thank you. And I yield the remainder of my time back.

Thank you, Mr. Chairman.

Chairman PEARCE. The gentleman's time has expired.

The Chair would now recognize the gentlelady from New York, Mrs. Maloney, for 5 minutes.

Mrs. MALONEY. Thank you, Mr. Chairman and Mr. Ranking Member and all the panelists. It has really been very insightful and, actually, very disturbing.

Unfortunately, we have seen that hacking has become more—much more lucrative because of cyber criminals and the cryptocurrencies, like Bitcoin.

And I have this report that I want to put in the record and share with my colleagues on "Sex, Drugs, Bitcoin: How Much Illegal Activity Is Financed Through Cryptocurrencies."

And this report points out they believe 72 billion of illegal activity is taking place on Bitcoin. And—

Chairman PEARCE. Without objection.

Mrs. MALONEY. —my question for all the panelists is, would cracking down on these cryptocurrencies reduce the incentive for cyber criminals to steal data from companies and governments?

And this report also says that roughly 25 percent of Bitcoin users were using and half their activity was illegal activity. It is disturbing to see ads to buy women on the Internet through Bitcoin and drugs and other illegal activities.

So, I would like to—I would like to ask Mr. Nicolas Christin your response to that question.

Dr. CHRISTIN. Thank you. I think that cryptocurrencies are just a means of payment. And let us assume that tomorrow, cryptocurrencies become completely illegal. I doubt that it would actually stop the criminals in their tracks.

Because cryptocurrencies are a relatively recent phenomenon. Bitcoin, for instance, started appearing in 2008–2009.

And before that, we already had cyber crime. People were just using different tools. Liberty, Reserve, WebMoney, and so forth.

So, I don't necessarily think that clamping down on the payment system itself, or even interdicting it, would necessarily improve the situation very much. People would just find other ways of getting paid.

Mrs. MALONEY. Well, I want to ask you and also Mr. Lewis. Mr. Lewis this question about nation states.

And when a nation state is behind a hack, sometimes it is hard to figure out what it is, what they want the money for.

We know, as you have testified earlier, the—North Korea was behind the hacks in Bangladesh for $81 million. That was clear, they needed money. They got money.

But, in other cases, when a nation state steals data from a company like Equifax, and then they don't sell the data on the black-market, and it doesn't seem to appear some other place, it really isn't clear what their motivations are.

So, when a nation state hacks into U.S. companies and steals data but doesn't sell the data on the black-market, why do you think—what is the explanation of why they did it? Are they collecting data for espionage purposes?

What is the—I would like to thank Mr.—ask Mr. Christin and Mr. Lewis and then all the panelists to answer. What is the motive? Are they phishing?

Are they just—what are they doing when they steal? And they don't seem to use it, or we can't track what they are doing with it.

Dr. CHRISTIN. So, I will start to answer that by saying that sometimes we don't even know who is the perpetrator of the breach, so we have no idea who is behind the actual breach.

When it is not being sold, it can be for a variety of reasons. Maybe it doesn't have an economic value but has other types of value, leverage, espionage as you mentioned, and others.

Very simply put, we just don't necessarily know who is behind every single breach, and what they are using the breach for.

Dr. LEWIS. Thank you. The nature of the intelligence business has changed dramatically in the last few years, and data is at the center of those changes.

So, you can use digital technologies to identify persons of interest, either for recruitment or, more importantly, for counterintelligence purposes.

So, we are seeing a world where it is going to be much harder to operate covertly, simply because of things like the Equifax breach. And when I see a big breach like that and the data doesn't appear on the market, I usually assume that it is an espionage-related case.

Mrs. MALONEY. It is a—pardon me, a what?

Dr. LEWIS. An espionage-related case.

Mrs. MALONEY. An espionage-related case.

Any other comments?

Ms. ABLON. I would add to that aggregating this data can be very valuable for state-sponsored actors. For example, some people believe that the state's same country carried out the attacks on OPM, Anthem, and United Airlines.

And so, combining all that information would get some of the most sensitive personal and health information, as well as information about where people travel, to build a comprehensive profile of who to target, who to leverage, how to leverage for future information, or for exploitation of espionage purposes.

Mrs. MALONEY. Well, when you—when you see all these—this theft taking place, Mr. Lewis or Dr. Lewis and Mr. Christin and others, of all the cyber crime affecting the U.S., which percentage tends to be committed by state actors, versus criminal actors, versus terrorist organization or other activities? Who do you see doing this?

Starting with you, I guess, Dr. Lewis and just going down the line.

Dr. LEWIS. There have been some classified studies on this question. In the past, China was the leader, by far, of espionage, largely in its dealing with intellectual property. Russia was number two, focused on financial crime.

That has changed a bit in the last few years. The Russians are, for some reason, much—

Chairman PEARCE. If I could get the panelists to—tighten the answers up.

Dr. LEWIS. The Russians have changed and focused now as much—they still focus on financial crime but they also look at coer-

cion, as we know. And the Chinese have become much quieter. Iran and North Korea are also actors. But—

Mrs. MALONEY. When you say the Russians want coercion, what does that mean?

Chairman PEARCE. The gentlelady's time has expired.

Mrs. MALONEY. What are they trying—who are they trying to coerce? I have been hacked twice by the Russians. That is why I am curious.

Dr. LEWIS. You have probably all been hacked by the Russians.

But Russian military doctrine changed in 2010 to emphasize a psychological warfare and online political activities. And so, we have seen them implement that doctrine across all NATO countries.

Chairman PEARCE. The gentlelady's time has expired.

The Chair would now recognize the gentleman from Ohio, Mr. Davidson, for 5 minutes.

Mr. DAVIDSON. Thank you, Chairman.

I really appreciate these witnesses and I thank the committee for doing the work to have a hearing on this topic. I think it is vital that we get after this.

It is critical, really, first, for the American people. The American people are sick of the vulnerability and the helplessness that comes with knowing something like, the Russians have probably already hacked all of you. What a shocking statement to go public with that.

But it is not something that truly will be shocking because not only have the Russians probably hacked us, the Chinese have probably hacked us. And, frankly, many of the companies that we buy from or share our data with are actively hacking, in the sense that they know far more than the average consumer knows.

Frankly, your car has probably hacked a lot of things about you, including your weight if you have a newer car. And it will tell where you have been, how long you have been there. And you aggregate the data and they might be able to speculate about what you bought when you went in the convenience store.

So, all this is really changing the landscape in the economy. But because of that, there are some real national security concerns.

And, frankly, when we talk about all the ways that the data can be used, I am curious about all the data that is collected.

And I think it is vital that, in law, that this Congress establishes that in every case, it is your data. The individual has a property right in their own data. In every platform, in every way.

And they should be choosing how their data is used. Certainly, they can give consent. Perhaps they can give consent for compensation. But they should always be given the opt-in, in my opinion.

But in the case of the data that is collected and it is swept up. I am just curious, Mr. Lewis, your assessment of what is more valuable or easier to obtain or maybe bigger, is personally identifiable information or intellectual property?

Dr. LEWIS. They are—thank you. They are both easy to acquire but probably the bulk of the data we have seen taken, at least in numbers, if not in value, is personally identifiable information.

Mr. DAVIDSON. Thank you for that.

And, Mr. Bernik, your company has built its reputation on protecting some of this data. Lots of folks use your service or one similar to it.

And I am just curious what sort of risk controls are effective at protecting personally identifiable information?

Mr. BERNIK. The types of controls that organizations can implement to protect information are things like encryption, encrypting the data, both in transit and at rest. So, when it is being transmitted as well as when it is being stored.

And making sure that high levels of authentication are used when information is accessed so that it is not so simple to get access to the information at rest. Meaning you should use more than just a user name and password.

And, I think, historically, that is all the security we really had, in a lot of cases. Thus, we have a lot of compromised information.

Mr. DAVIDSON. Thank you.

And I would add that if the data is not online, then it is harder to be accessed.

Mr. BERNIK. Absolutely.

Mr. DAVIDSON. So, it is not collected in the first place. It is not there to be hacked.

And so, I guess, is there anything specific about that that differentiates the risk, whether the database is a government database or a commercial database?

Mr. BERNIK. In terms of the—so, my view that I would take on that is that organizations should only be permitted to save this information where they have implemented certain controls. And so, what they can't determine or demonstrate.

And that is an interesting way of looking at it. That they have the controls, or they don't need the data, then they shouldn't collect it.

When you go to any office of any chiropractor or anything, they will ask you for your personal information. And you will write it down. They will put it into a database.

The question is, do they have the ability to protect it? Do they need it?

Those are questions that should be answered and should be positioned by the consumer before they provide that data. But that information didn't exist, historically.

Mr. DAVIDSON. Thank you for that. And I would add that we have offered the Market Data Protection Act. It passed the House by a unanimous consent.

We are still waiting on the Senate to take action. And this would simply require the Securities and Exchange Commission to provide an assurance to us that they do, in fact, have the controls in place to oversee that.

And so, the same governance that a board would expect of, say, Equifax, I am confident the I.T. department has a little more interaction with the board than they used to.

And I would think that would serve as good notice for governance practices around the country, whether they are in the Government or not. And since we don't have a chief technology officer for each secretary.

My time has expired. Mr. Chairman, I yield.

Chairman PEARCE. The gentleman's time has expired.

And the Chair will now recognize the gentleman from Memphis, Tennessee, Mr. Kustoff, for 5 minutes.

Mr. KUSTOFF. Thank you, Mr. Chairman. And I do want to thank the witnesses for being here.

Today's hearing has been both very interesting and very concerning. I think we would agree with that.

Ms. Ablon, if I could. Today, we have certainly had several hearings where we have talked about the use of cyber—cryptocurrency. We have talked about that being—becoming more predominantly preferred method of use on the—on the Web. I think you may have testified to that, at least becoming—turning that way.

We also know that the dark Web hosts a forum to sell and trade illicit goods and services, fire arms, drugs, et cetera. And we have talked about the personal information being bought and sold in bulk.

I know a few years ago, 3 or 4 years ago, maybe 5 years ago, there was a dark Website called Silk Road. It was shut down. Law enforcement worked very hard to shut that down but we have other dark Websites that have emerged in its place.

Given your work in studying how cyber criminals operate, can you talk a little bit more—you have talked and there has been discussion about the dark Web and online black-market sellers. But the shutdown of Silk Road, of AlphaBay, and how some of those other Websites actually interact with people and how they interact with those dark Websites.

Ms. ABLON. Sure. You mentioned some great examples of law enforcement taking down black-market Websites.

These markets, you can think of them like an Amazon or an eBay, where you point and click and you put a thing that you want to buy in your shopping cart. And then, you pay with money that you might have in your wallet.

So, it is easy to do. We are all really familiar with doing eCommerce on the surface Web, similarly as how you can do eCommerce or by purchasing things on the dark Web.

I would offer that you noted some notable takedowns. But taking down some of these big sites, like Silk Road, AlphaBay, Hansa, are good but that just leaves market share for other Websites, for other market places to come in.

So, law enforcement's efforts are like trying to drain the ocean with a cup. Every time they take out a market place, there is market share available and plenty of cyber criminals and nefarious actors to jump in to take that.

Mr. KUSTOFF. Can you also—you went through the different categories of bad actors. You talked about—one of the categories was cyber-terrorist. Obviously, I am talking about those foreign actors. Those who aren't here.

Where do they train? And do any of them train and get their education here in the United States?

Ms. ABLON. Cyber-terrorism is an interesting category of cyber-threat actor. It is—in general, they are—they combine traditional terrorism and attacks via cyber-space. For an act to be cyber-terrorism, it needs to occur through digital domain.

At this point in time, people who are cyber-terrorists or acts of cyber-terrorism are more akin to hacktivism. People in the groups like Anonymous.

Now, that is not to say a question that you might think is, well, are terrorists involved with the Internet? Are they involved with cyber in some way?

They are. They use the Internet for a number of reasons. To—information gathering, like learning how to build bombs. Recruiting, meeting, and conducting—connecting with like-minded individuals. Spreading propaganda or collecting money or other efforts in the sense that they might be cyber criminals online but terrorist in the—in the physical world.

Mr. KUSTOFF. Thank you very much.

Mr. Bernik, you testified, in relation to somebody's question, about ways to protect yourself, in terms of preventing stolen identity. Like you talked about locking the credit report.

Is that analogous to freezing the credit report?

Mr. BERNIK. Correct. It is the same thing.

Mr. KUSTOFF. Obviously, I would assume that the three credit agencies don't want that, although they do offer that service.

That could be onerous on people who are trying to, obviously, take out loans, mortgage refinance, et cetera.

Is there any other middle ground? Or is that, in fact, the most secure way to protect one's identity?

Mr. BERNIK. So, in my experience, that is the easiest way. Today, you can unlock it immediately on the Websites by pushing a button. They have all made that—all the agencies have made that feature available.

And in the event that you do need to take a loan out or you do—you are going to, you just unlock it and it is instantaneously available again.

So, it is merely a question of not allowing those kinds of hookups to be done or requests to be made of you without you first unlocking that button online and unlocking the report.

Chairman PEARCE. The gentleman's time is expired.

Mr. BERNIK. Thank you.

Chairman PEARCE. The members are advised that there is a vote in progress. A little over 6 minutes left in the vote.

For me, I would like to thank our witnesses for your testimony today.

The Chair notes that some Members may have additional questions for this panel, which they may wish to submit in writing. Without objection, the hearing record will remain open for 5 legislative days for Members to submit written questions to these witnesses and to place their responses in the record. Also, without objection, Members will have 5 legislative days to submit extraneous materials to the Chair for inclusion in the record.

I ask our witnesses to please respond as promptly as they are able.

This hearing is adjourned.

[Whereupon, at 3:23 p.m., the subcommittee was adjourned.]

A P P E N D I X

March 15, 2018

Data Thieves

The Motivations of Cyber Threat Actors and Their Use and Monetization of Stolen Data

Lillian Ablon

CT-490

Testimony presented before the House Financial Services Committee, Subcommittee on Terrorism and Illicit Finance, on March 15, 2018.

For more information on this publication, visit www.rand.org/pubs/testimonies/CT490.html

Testimonies

RAND testimonies record testimony presented or submitted by RAND associates to federal, state, or local legislative committees; government-appointed commissions and panels; and private review and oversight bodies.

Published by the RAND Corporation, Santa Monica, Calif.

Data Thieves:

The Motivations of Cyber Threat Actors and Their Use and Monetization of Stolen Data

Testimony of Lillian Ablon[1]
The RAND Corporation[2]

Before the Committee on Financial Services
Subcommittee on Terrorism and Illicit Finance
United States House of Representatives

March 15, 2018

Good afternoon, Chairman Pearce, Ranking Member Perlmutter, and distinguished members of the subcommittee. Thank you for the invitation to testify at this important hearing, "After the Breach: the Monetization and Illicit Use of Stolen Data."

Cybersecurity is a constant and growing challenge. Although software is gradually becoming more secure and developers are creating novel approaches to cybersecurity, attackers are becoming more adept and better equipped.[3] And as the world embraces more digital and hyperconnected components, the paths become more numerous for attackers to gain access to our most sensitive information.

Data breaches have become commonplace in the United States. In 2017, more than 1,000 data breaches exposed over a billion records of sensitive data.[4] From banking to retail, health care to entertainment, and even government, no sector is immune. Some of that information has been monetized by threat actors in flourishing underground markets. These cyber black markets offer the computer-hacking tools and services to carry out cybercrime attacks and sell the by-products stolen in those attacks: credit cards, personal data, and intellectual property. In other

---

[1] The opinions and conclusions expressed in this testimony are the author's alone and should not be interpreted as representing those of the RAND Corporation or any of the sponsors of its research.

[2] The RAND Corporation is a research organization that develops solutions to public policy challenges to help make communities throughout the world safer and more secure, healthier and more prosperous. RAND is nonprofit, nonpartisan, and committed to the public interest.

[3] Martin C. Libicki, Lillian Ablon, and Tim Webb, *Defender's Dilemma: Charting a Course Toward Cybersecurity*, Santa Monica, Calif.: RAND Corporation, RR-1024-JNI, 2015 (http://www.rand.org/pubs/research_reports/RR1024.html).

[4] Identity Theft Resource Center, *2017 Annual Data Breach Year-End Review*, idtheftcenter.org, no date (https://www.idtheftcenter.org/2017-data-breaches).

cases, the attackers keep the data for their own espionage purposes or use stolen funds to facilitate future operations.

To get an understanding of what the attackers are doing with the stolen data, and how they are monetizing the data, we need to understand who they are and what motivates them.

Attackers, or *cyber threat actors*, can be grouped by their set of goals, motivation, and capabilities. Four groups of note are *cyberterrorists*, *hacktivists*, *state-sponsored actors*, and *cybercriminals*.[5] It is important to understand the full environment of threat actors, so I will briefly review these four, but I will focus my testimony largely on state-sponsored actors and cybercriminals, as they are of greatest concern to businesses and the government and merit the most note for this hearing.

Today I will give a brief overview of these four types of cyber threat actors, followed by a discussion of the landscape of the black markets for cybercriminal tools and stolen data, and then finish with some of the ways that state-sponsored actors and cybercriminals use and monetize the stolen data.

Different Cyber Threat Actors Have Different Motivations

Cyberterrorists

Cyberterrorism unites two significant modern concerns: attacks via technology in cyberspace and traditional terrorism. While there is no single or globally accepted definition of *cyberterrorism*, in theory, it consists of a politically motivated extremist group or nonstate actor using cyber techniques to intimidate, coerce, or influence an audience; force a political change; or cause fear or physical harm.[6]

To date, there have been no publicly reported cases of terrorists using the internet to carry out cyberattacks; what has been done that has been attributed to cyberterrorism is more akin to hacktivism. Many terrorists, or nonstate actors, do employ cyber to further their goals. They use the internet in many ways: for information gathering, e.g., learn how to build a bomb; to recruit, meet, and connect with like-minded individuals; and to spread propaganda.[7] But just "being" in cyberspace does not make a terrorist a cyberterrorist. Cyberspace must be used somehow to

---

[5] Robinson and colleagues have documented other categories, such as script kiddies, cyber researchers, and internal actors. Neil Robinson, Luke Gribbon, Veronika Horvath, and Kate Cox, *Cyber-Security Threat Characterization: A Rapid Comparative Analysis*, Santa Monica, Calif.: RAND Corporation, 2013 (https://www.rand.org/pubs/research_reports/RR235.html).

[6] B. Hoffman, *Inside Terrorism*, New York: Columbia University Press, 2006; R. Ahmad and Z. Yunos, "A Dynamic Cyber Terrorism Framework," *International Journal of Computer Science and Information Security*, Vol. 10, No. 2012, pp. 149–158.

[7] Robert S. Mueller III, "Combatting Threats in the Cyber World: Outsmarting Terrorists, Hackers, and Spies," prepared remarks at RSA Cyber Security Conference, San Francisco, Calif., March 1, 2012 (http://www.fbi.gov/news/speeches/combating-threats-in-the-cyber-world-outsmarting-terrorists-hackers-and-spies); and Ines, von Behr, Anais Reding, Charlie Edwards, and Luke Gribbon, *Radicalisation in the Digital Era: The Use of the Internet in 15 Cases of Terrorism and Extremism*, Santa Monica, Calif.: RAND Corporation, 2013 (https://www.rand.org/pubs/research_reports/RR453.html).

commit a terrorist act.[8] The movies and media portray what cyberterrorism could be: terrorists crafting digital attacks to take down traffic lights, make trains stop on a dime, and water pipes burst. But to date, no such dramatic events have occurred.

Hacktivists

Hacktivists are typically motivated by a cause—political, economic, or social: from embarrassing celebrities, to highlighting human rights, to waking up a corporation to its vulnerabilities, to going after groups whose ideologies they do not agree with.[9]

Hacktivists may steal and disseminate sensitive, proprietary, or, sometimes, classified data in the name of free speech. Other times, they aim to deny access to a particular service or website by conducting a distributed denial-of-service (DDoS) attack, essentially denying legitimate access by flooding a website with more traffic than it can handle, causing the site to crash.[10]

State-Sponsored Actors

State-sponsored actors receive direction, funding, or technical assistance from a nation-state to advance that nation's particular interests. State-sponsored actors have stolen and exfiltrated intellectual property, sensitive personally identifying information (PII), and money to fund or further espionage and exploitation causes. In rare cases, these data appear for sale on underground black markets. Instead, these data are usually kept by the actors for their own purposes. Although the data taken from data breaches might not always appear on underground markets, what *can* appear are the tools and guides for how to take advantage of the vulnerabilities that allowed access to the vulnerable systems in the first place. As an example, a researcher published the flaw that was used to penetrate Equifax, and within 24 hours the information was published to hacking websites and included in hacking toolkits.[11] Note, however, that there has not been official attribution of who conducted the intrusion into Equifax.

---

[8] Maura Conway, "Reality Bytes: Cyberterrorism and Terrorist 'Use' of the Internet," *First Monday*, November 4, 2002 (http://firstmonday.org/ojs/index.php/fm/article/view/1001/922); and Maura Conway, "Cyberterrorism: Hype and Reality," in Leigh Armistead, ed., *Information Warfare: Separating Hype from Reality*, Washington, D.C.: Potomac Books, 2007, pp. 73–93.

[9] As a few examples, see Paolo Passeri, "List of Hacked Celebrities Who Had (Nude) Photos Leaked," Hackmageddon, August 7, 2012 (http://www.hackmageddon.com/2012/08/07/list-of-hacked-celebrities-who-had-nude-photos-leaked/); Lillian Ablon, "Hackerazzi: How Naked Celebrities Might Make the Cloud Safer," *The RAND Blog*, September 8, 2014 (https://www.rand.org/blog/2014/09/hackerazzi-how-naked-celebrities-might-make-the-cloud.html); "HBGary Federal Hacked by Anonymous," *Krebs on Security*, February 7, 2011 (https://krebsonsecurity.com/2011/02/hbgary-federal-hacked-by-anonymous/); Paula Cohen, "Anonymous Hackers' Group Declares War on ISIS," *CBS News*, November 16, 2015 (https://www.cbsnews.com/news/anonymous-hackers-declare-war-on-isis/); "Anonymous Hacks Pro-ISIS Twitter Accounts, Fills Them with Gay Pride," *CBS News*, June 15, 2016 (https://www.cbsnews.com/news/anonymous-hacks-pro-isis-twitter-accounts-fills-them-with-gay-pride/).

[10] An example is Operation Payback, which targeted websites of large corporations to shut them down temporarily.

[11] Michael Riley, Jordan Roberston, and Anita Sharpe, "The Equifax Hack Has the Hallmarks of State-Sponsored Pros," *Bloomberg*, September 29, 2017 (https://www.bloomberg.com/news/features/2017-09-29/the-equifax-hack-has-all-the-hallmarks-of-state-sponsored-pros).

In a few cases, state-sponsored actors have conducted cyberattacks—which deny, degrade, disrupt, or destroy computing systems—to send a political message. An example of this is the 2014 attack on Sony Pictures Entertainment, where North Korea wanted to advance its political agenda and, in part, to stop the release of the movie *The Interview*.[12]

Rather than viewing what they do as breaking laws, state-sponsored actors maintain that they are acting in accordance with their own laws, and most have accepted that cyberespionage is a legitimate state activity. Deterrence—diplomatic, financial, and economic consequences—is thought by some to play a role in preventing these types of attacks from happening or escalating.[13]

Cybercriminals

Cybercriminals are motivated by financial gain—they care about making money.[14] They want access to our personal, financial, or health data—in order to monetize them on underground black markets. For the retail sector in particular, the stolen data from these hacks appeared within days on black market sites.

These markets are dispersed, diverse, and segmented—rapidly growing, constantly changing, and innovating to keep pace with consumer trends and prevent law enforcement and security vendors from understanding them. They come in many forms. Some are dedicated to one product or a specialized service. Others offer a range of goods and services for a full life cycle of an attack—from the tools needed to exploit a system, all the way through to the cyberlaundering of the stolen goods. These markets are easy for almost anyone to get involved in—at least at the most basic levels.

Cybercriminals operate behind anonymous and peer-to-peer networks (such as Tor and OpenBazaar, respectively) and use encryption technologies and digital currencies (such as Bitcoin) to hide their communications and transactions.

Table 1 gives a summary of the various cyber threat actors, their main motivations, and use of stolen data.

---

[12] FBI, "Update on Sony Investigation," press release, December 19, 2014 (https://www.fbi.gov/news/pressrel/press-releases/update-on-sony-investigation).

[13] Martin C. Libicki, *It Takes More Than Offensive Capability to Have an Effective Cyberdeterrence Posture*, Santa Monica, Calif.: RAND Corporation, CT-465, 2017 (https://www.rand.org/pubs/testimonies/CT465.html).

[14] To be sure, there may be some sense of "nobility" of, for example, Russian actors victimizing Americans.

Table 1. Characteristics, Techniques, and Targets of Cyber Threat Actors

scription	Motivator	Technique	Types of Targets	Use of Stolen Data
groups or actors using hniques to , coerce, or an audience; olitical change; ır or physical	Gain support for and deter opposition to a cause; carry out dictates of an ideology	Cause kinetic damage: destroy or disrupt critical infrastructure or systems; loss of life	Determined by actors' ideology	Disrupt critical infrastructure via cyberattack; Change prescription or allergy information, switch or delete medical record; further a campaign on a particular target
ıreness to a olitical, , social); ree speech :")	Ideological activism; disruption of services or access	Steal and leak sensitive, proprietary, or classified information; conduct DDoS on websites or services	No one type of target	Gather personal information of a specific target; publicize a breach to highlight how vulnerable a particular organization is
irection, ır technical e from nation- ghly ıted and often ıost- ıted methods o-day ities); targeted stent	Advance interests of their nation-state; further political agenda	Conduct intelligence, surveillance, reconnaissance, espionage; employ watering-hole attacks; exfiltrate data (e.g., intellectual property); degrade or destroy technical components; conduct targeted attacks	Other nation-states, defense contractors, technology sector, and critical infrastructure; (rare) banks or cryptocurrency wallets	Build profiles of possible targets for follow-on targeting, exploitation, or espionage campaigns; use personal, financial, or medical information as leverage to gain other types of intelligence
ersonal, or health data ze it	Financial gain; power	Use crimeware (e.g., exploit kits, "script-kiddy" tools); rely on already known vulnerabilities, phishing, and spearphishing; smash-and-grab	Data repositories (e.g., banks, retail companies, health care) that can be monetized; cryptocurrency wallets	Use credentials (username/password combinations) and harvest contact lists for phishing attacks; exploit password reuse; conduct identity theft, tax, or medical fraud

The Challenge of Attribution

Attribution after a data breach is difficult. Often there is not enough digital evidence left behind to identify the attackers or their country of origin. That said, there are cases where various commercial security firms and threat analysis groups involved in the aftermath of a breach have found similarities in the malware used in various attacks. In particular, similar malware was used in the 2014 cyberattack on Sony Pictures Entertainment and the 2015–2016 SWIFT data breach (i.e., from North Korea). And many note the strong possibility that the malware for the data breaches of the Office of Personnel Management (OPM), Anthem, and United Airlines originate from the same place (China).[15]

There Is Overlap Between These Various Cyber Threat Actors

Although there are distinctions and differences in motivation between each of the cyber threat actors, there is some degree of fluidity between the groups. In many cases, the same tools and techniques are used by different groups, sometimes because those are the only tools available, and other times because that helps with plausible deniability and shifting the blame to a different group. In some countries, state-sponsored actors may work with "citizen hackers" or their country's cybercriminal elements to carry out an attack.[16]

Table 2 provides some more detail about the overlap between the various cyber threat actors.

The Hackers' Bazaar: How Stolen Information Gets Monetized

Turning to the cybercrime black markets, I will outline four aspects: *people*, *products*, *places* for communicating and conducting business transactions, and *prices*.[17]

---

[15] SANS Institute InfoSec Reading Room, *United Airlines May 2015 Data Breach: Suggested Near, Mid and Long-Term Mitigating Actions Using the 20 Critical Security Controls*, November 2015 (https://www.sans.org/reading-room/whitepapers/breaches/united-airlines-2015-data-breach-suggested-near-mid-long-term-mitigating-actions-th-36452); Threatconnect Intelligence Research Team, (2015, February 27). The Anthem Hack: All Roads Lead to China," February 27, 2015 (http://www.threatconnect.com/the-anthem-hack-all-roads-lead-to-china/).

[16] The attack on Yahoo! is an example of this.

[17] Most of the language and analysis in this section are drawn from Lillian Ablon, Martin C. Libicki, and Andrea A. Golay. *Markets for Cybercrime Tools and Stolen Data: Hackers' Bazaar*, Santa Monica, Calif.: RAND Corporation, 2014 (https://www.rand.org/pubs/research_reports/RR610.html).

Table 2. Overlap Between the Various Cyber Threat Actors

	Hacktivists	State-Sponsored	Cyberterrorists
Cybercriminals	May use some of the same "script-kiddy" tools for things like DDoS May have some of the same targets Can be difficult to tell them apart	Cybercriminals often have close ties (perhaps funding) to nation-states For sophisticated cybercriminals, often the groups are confused with each other, especially by the media Most of the time, they do not want attribution, though some criminal groups have taken credit for state-sponsored actions	Cybercriminals do not want recognition or attribution Cyberterrorists do typically want attribution
Hacktivists		"Citizen hackers" may have hacktivists connected with state-sponsored actors Some hacktivist groups take credit for state-sponsored actions	May in fact be the same thing: "one man's terrorist is another man's freedom fighter" Both typically want attribution Often use similar low-level attacks (e.g., website defacements, taking over Twitter accounts)
State-Sponsored			May have same goals (e.g., to intimidate, coerce, or influence an audience or force a political change)—but one is in political power and the other is not

People: Who Participates In Cybercrime Markets?

Participants in cybercrime black markets come from all over the world, range across all skill levels, and occupy different roles depending on their technical abilities and reputation.

Within these markets, there are often hierarchies and specialized roles: *administrators* sit at the top, followed by *subject-matter experts*, who have sophisticated knowledge of particular areas (e.g., exploit-kit creators, data traffickers, cryptanalysts, those who vet). Next are *intermediaries*, *brokers*, and *vendors* and then the *general membership*.

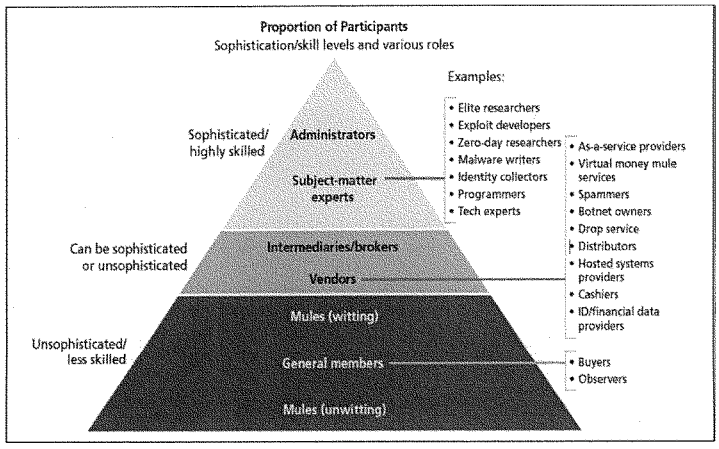
At the bottom of the hierarchy is where the market meets the consumer. This is where *mules* and virtual money mule services come into play—the final level of participation in the market. Mules use multiple methods to turn the stolen credit card or ecommerce accounts into usable money—for example, by completing wire transfers or shipping goods overseas bought with

stolen funds. Mules can be witting participants (well-informed and organized operations) or unwitting (naïve individuals duped into involvement).[18]

The number of participants in these black markets has increased as barriers to entry diminish. Barriers today to enter and participate in these markets are negligible—essentially all that is needed is an internet connection and a device. This is due to the greater proliferation of websites, forums, and chat channels where goods can be bought and sold. The greater availability of as-a-service models, point-and-click tools, and easy-to-find online tutorials makes it easier for technical novices to use what these markets offer or to just hire someone to carry out an attack for them. An increase in the number of tutorials, manuals, YouTube videos, and Google guides for "how to use exploit kit *X*" or "where to buy credit cards" also facilitates entry into the lower tiers of these markets, especially for those who wish to be buyers.

Figure 1 depicts the different participant levels in the underground market, proportionally. It also shows the sophistication and technical skill levels and gives examples of various roles.

Figure 1. Different Levels of Participants in the Cyber Black Markets



SOURCE: Ablon, Libicki, and Golay, 2014.

Surprisingly, these markets are highly reliable—reputation matters a great deal, and, for the most part, products and participants are what they say they are and do what they say they do.

[18] Examples include some "work from home" scams where unwitting people are recruited to purchase goods (using the, unbeknownst to them, stolen credit cards) and then ship them overseas.

Reputation entails either having credentials and a good reputation with others or proving oneself (for example, getting good reviews on sales).

Because contracts in black markets cannot be legally enforced, the markets are constantly plagued by *rippers*, who do not provide the goods or services they advertise and are an exception to the high reliability of the markets. Rippers tend to get reported and then quickly removed by administrators (by, for example, suspending their accounts). Although they can easily access new channels under new names, it takes time to reestablish a reputation, which inhibits cheating.

Products: What Is For Sale?

Cybercrime markets offer a diverse slate of products, including both goods (hacking tools, digital assets) and services (hacking-as-a-service, digital asset handling, fake-identity creation) for all phases of the full cybercrime life cycle—from initial hack all the way through to monetizing the stolen data.

Examples include tools to help gain initial access onto a target (exploit kits), along with the payloads (malware) and the parts and features of those payloads, services to help scale or deliver a payload, support products to ensure that infrastructure is set up correctly or to provide cryptanalytic services, and considerations for how to manage the stolen goods. As-a-service offerings (setting up botnets or conducting ransomware attacks) are on the rise.

The product slate keeps evolving with the technology. Whatever is new or novel for the traditional consumer—mobile devices, cloud computing, social media platforms—offers new entries for attack and will thus elicit a counterpart exploit on the black market.

Table 3 describes the main categories of products available.

Table 3. Goods and Services on the Black Market

Category	Definition	Examples
Initial access tools	Enable a user to perform arbitrary operations on a machine to then deliver payloads; can automate the exploitation of client-side vulnerabilities	• Exploit kit (hosted or as-a-service) • Zero-day vulnerabilities (and weaponized exploits) • Point of sale (POS) skimmers
Payload parts and features	Goods or services that create, package, or enhance payloads to gain a foothold into a system	• Packers[a] • Crypters[b] • Binders[c] • Obfuscation or evasion
Payloads	Impart malicious behavior, including destruction, denial, degradation, deception, disruption, or data exfiltration	• Bots or botnets for sale • Malware (including that focused on targeting cryptocurrency wallets)
Enabling services	Assist in finding targets or driving targets to a desired destination to use an initial access tool or payload; attack vectors and scaling methods	• Spam services • Pay-per-install and affiliates • Phishing and spearphishing services • Services to drive or find traffic • Fake website design and development
Full services (as-a-service)	Package together initial access tools, payloads, and payload parts and features to conduct attacks on a customer's behalf; can provide full attack life cycle	• Hackers for hire • Botnets for rent • Doxing • DDoS as a service (including DDoS against cryptocurrencies) • Ransomware as a service
Enabling and operations support products	Ensure that initial access tools and hacking services (enabling or full service) work as needed, are set up correctly, and can overcome "speed bumps" or obstacles	• Infrastructure (e.g., leasing services, VPN services, bulletproof hosting, compromised sites and hosts) • Cryptanalytic services (e.g., password cracking, password hash cracking) • CAPTCHA breaking
Digital assets	Items obtained from the target or victim (i.e., the hacked or stolen information)	• Credit card information • Account information (e.g., ecommerce, social media, banking) • Email login and passwords • Online payment service accounts • Credentials • PII/protected health information (PHI)
Digital asset commerce and cyberlaundering	Facilitate turning digital assets into cash	• Mule services • Counterfeit goods and services (e.g., fake documents, IDs, currency) • Card cloners, fake ATMs • Credit card processor services

SOURCE: Ablon, Libicki, and Golay, 2014.
[a] The "outer shell" of some malware—e.g., Trojan horses—hides the malware and makes detection and analysis by antivirus software more difficult. Packers can also employ antidebugging, antiemulation, anti-VM techniques, and code obfuscation.
[b] Crypters are software that can encrypt executable (.exe) files. Crypters can be used to encrypt viruses, remote access Trojans, keyloggers, spyware, etc. to make them undetectable from antivirus systems.
[c] Binders are software used to bind or combine two or more files into one file under one name and extension. They are used for the evasion of anti-virus systems.

Places: How Do These Markets Communicate and Conduct Transactions?

Communications and transactions take place through a variety of channels that span multiple tiers. Channels can include online stores, bulletin board–style web forums, email, or instant messaging platforms (allowing for private one-on-one communications or open chat rooms). Some are easy to find, and others are by invitation only and only accessible after having gone through an extensive vetting process. The highest-access tiers are usually hidden in websites on the dark web, which offer anonymizing and secure features.[19]

Transactions can cut across multiple channels and access tiers. An advertisement might be posted on a forum openly available and easily accessed, with actual transactions taking place behind encrypted VPNs, private messaging, locked-down social media accounts (e.g., private Twitter accounts), a shared email (to exchange content through draft messages), or a private server spun up on Tor for just one transaction. As digital goods, payment systems, techniques, tools, and malware continue to evolve, forums with "how to" sections remain popular.

In recent years, there has been more gravitation toward methods of communicating and conducting business transactions that offer anonymity or make it harder for law enforcement to find. Tor—the onion router—is a service one can use to access websites on the dark web and is a way to browse the internet semi-anonymously, essentially masking the location where one is coming from. As law enforcement succeeds in finding and taking down cybercriminal sites, such as AlphaBay and Hansa (two popular dark web marketplaces), there has been more of a move toward decentralized peer-to-peer networks (such as OpenBazaar), where individual users connect with each other, making it difficult for outsiders, such as law enforcement, to track.

Prices: How Much Does Everything Cost?

Prices can range widely, depending on hardness of attack, freshness of data, sophistication of malware, or whether something is as-a-service or do-it-yourself.

Easily exchanged goods—such as PII, account credentials, and financial data—are prey to the normal microeconomic fluctuations of supply and demand. Often, there is too much of a product available to sell at normal prices. By contrast, stolen-to-order, nonfungible goods—such as new technology designs, details on R&D activities, mergers and acquisitions, and other sorts of intellectual property—can command a very high price, provided that the right buyer exists.

The yield of a product influences its price. For example, a stolen Twitter account costs more to purchase than a stolen credit card because the former's account credentials potentially have a greater yield: the username/password combination could unlock access to other accounts, as well as give access to the victim's contact list in order to carry out follow-on spam or phishing attacks. Immediately after a large breach, batches of credit cards get released in the markets: freshly acquired credit cards command a higher price—as there is greater possibility for the credit cards to still be active. Over time, prices fall because the market becomes flooded with more and more batches—leveling off as the data become stale or if there has been significant time since the last breach. High or no-limit cards (e.g., the American Express Black card) or

[19] A darknet or dark web is an anonymous or semianonymous private network that uses encryption and proxies to obfuscate who is communicating with whom. An example is the onion router (Tor).

cards with chip and PIN are more valuable and can command a higher price: While a U.S. "chip and signature" card might start off for $15 a number after a breach, a European "chip and pin" card will go for closer to $120—the higher price point stemming from these types of cards being more secure, often resulting in a higher credit limit. Over time, as banks and users discover the theft and shut down the cards, a card may be discounted to $12 a number, and then drop further to $10 a number. Eventually, the credit cards may go on clearance, and one can purchase a "grab bag" of 100 credit card numbers for $700. The thought here is that, even if only two of those credit card numbers are still open for fraud, one can purchase and then illegally sell electronics, yielding a profit of more than the $700.

Access to botnets and DDoS capabilities are cheaper because there are so many more options (same for exploit kits).

Although prices, in general, range widely (e.g., hacking into accounts can be anywhere from $16 to $325-plus, depending on the account type),[20] similar products tend to go for similar amounts. Medical records can be worth up to $50 per record. Brand-name recognition also plays a role. Services can involve leasing servers, finding traffic, creating a personalized payload (or "cleaning" or obfuscating an already existing payload to avoid antivirus signatures), and setting up infrastructure.

Table 4 summarizes stolen credit card prices and markets.

Table 4. Credit Card Prices Based on Market Circumstance

Card Price (per card)	Market Circumstance
$120	Freshly acquired (EU card)
$15	Freshly acquired (U.S. card)
$10–$12	Flooded
$0.75–$7	Clearance ("stale" data)
SOURCE: Updated from Ablon, Libicki, and Golay, **2014.**	

Although transactions can be completed with nondigital currency (e.g., Western Union, cash, PayPal), black market sites have moved toward accepting digital cryptocurrencies, given the appeal of anonymity and other security characteristics. Bitcoin is currently a popular choice, although with the volatility of the cryptocurrency markets, an increase in theft of digital currency wallets, and some notable takedowns by law enforcement of Bitcoin exchanges, it is by no means the only choice. Other, more reasonably priced digital currencies, such as like Ethereum, have become more popular. Further, the possibility that such countries as Russia and China will limit uses of Bitcoin has caused those doing transactions in that virtual currency to find ways to

[20] Max Goncharov, *Russian Underground 101*, Cupertino, Calif.: Trend Micro Incorporated, 2012 (https://www.trendmicro.de/cloud-content/us/pdfs/security-intelligence/white-papers/wp-russian-underground-101.pdf); Dell Technologies, *Securworks 2017 State of Cybercrime*, February 13, 2018 (https://www.delltechnologies.com/en-us/perspectives/secureworks-2017-state-of-cybercrime-executive-summary/); Dell Secureworks, *Underground Hacker Markets: Annual Report—April 2016*, April 2016 (http://online.wsj.com/public/resources/documents/secureworks_hacker_annualreport.pdf).

move those funds into other financial system or digital currencies to get their money out and usable. The uses of WebMoney and Perfect Money, online payment settlement systems, remain popular, although because they are not anonymous, users may send their digital currency through a "tumbler" or "washer" to virtually launder the wallet through several different accounts, making the funds difficult to trace.

Making The Stolen Goods Useful: How Data Get Used or Monetized by Cyber Threat Actors

The cybercrime black markets can be more profitable than traditional markets for illicit goods: The links to end users are more direct, and worldwide distribution, being electronic, is a relatively trivial task. This is because a majority of players, goods, and services are online-based and can be accessed, harnessed, or controlled remotely, instantaneously. "Shipping" digital goods may only require an email or download at worst—or a username and password to a locked site at best. This enables lower costs and possibly greater profitability.

Cybercriminals are always looking for exotic ways to use or monetize stolen data in ways that law enforcement and security vendors are not looking for or have not yet figured out. That said, because the main motivation of cybercriminals is to make money as quickly as possible, data, which require several steps before they can be monetized, are not as valuable as that which can be monetized quickly. As such, medical records and PHI, credit report information, and extraordinarily sensitive PII are likely more valuable to state-sponsored actors, who might use this information to build profiles to target for exploitation or espionage campaigns, as leverage to gain other types of information, or to help incubate innovative new strategies for using big data against an adversarial country.

In what follows, I describe various categories of stolen data and how each can be monetized or used by cybercriminals or state-sponsored actors. While I note examples of data breaches, unless explicitly stated, I do not intend to imply attribution of a particular cyber threat actor or imply that the data from those breaches have appeared on cybercrime markets for monetization. Instead, my intent is to provide real-world examples of data breaches involving the particular types of data and illustrate what might be done with those data.

Financial information, such as credit card data and bank account numbers, can get monetized by withdrawing cash, purchasing gift cards for resale, or harnessing a "money mule" (witting or unwitting) to make fraudulent orders to purchase goods, such as expensive electronics, which can then be sold on other black markets.

Credentials, such as usernames, passwords, account login information, and email addresses, enable an attacker to get access to the victim's contact list for further spam or phishing campaigns. The actor might also take advantage of password reuse and try to access a variety of banking and ecommerce sites. With access to business email addresses, an attacker can pose as a legitimate employee in the business and request a seemingly legitimate wire transfer, whose funds end up in the account of a money mule, who may then forward on the money or withdraw it and send it to the attacker through different means.

The breaches of Target (2013) and Home Depot (2014) are examples of data breaches where attackers collected credit and debit card numbers, as well as account information and email addresses (110 million from Target and 56 million from Home Depot).[21] The 2013 breach of Yahoo! compromised the accounts of between 500 million and 3 billion users.[22] And the 2014 data breach of JP Morgan Chase, where names, addresses, phone numbers, and email addresses of the holders of 76 million households and 7 million small business accounts were taken, was the largest theft of customer data from an American financial institution at the time.[23]

Medical records and PHI, such as the information accessed from Community Health Systems in 2014[24] or Anthem in 2015,[25] may be used for medical fraud—for example, filling out prescriptions in the victim's name—or combined with other PII to create a more detailed, comprehensive profile.

Financial reports on publicly traded companies, such as those accessed during the 2016 data breach of the Securities and Exchange Commission's EDGAR database, could provide the attackers information to make illegal trades.[26]

Credit report information, such as addresses, dates of birth, social security numbers, driver's license information, and other PII can be combined, creating a comprehensive profile of a victim to create a custom dictionary of possible passwords that can be used to attempt to crack a victim's bank or financial account or for identity theft—posing as the victim to open up new lines of credit or to add new authorized users to existing credit lines. In the 2017 data breach of Equifax, approximately 149 million users had this kind of information taken.[27]

Extraordinarily sensitive PII, akin to the kind of information taken from OPM in 2014,[28] can facilitate those building comprehensive profiles of victims.[29] Cybercriminals might use this information in the same way they use credit report information.

---

[21] Target, "Target Confirms Unauthorized Access to Payment Card Data in U.S. Stores," December 19, 2013 (https://corporate.target.com/press/releases/2013/12/target-confirms-unauthorized-access-to-payment-car); Home Depot, "The Home Depot Report Findings in Payment Data Breach Investigation," November 6, 2014 (http://ir.homedepot.com/news-releases/2014/11-06-2014-014517315)

[22] Yahoo! Help, "Yahoo 2013 Account Security Update FAQs," webpage, no date (https://help.yahoo.com/kb/account/SLN28451.html).

[23] United States Securities and Exchange Commission, "Form 8-K: JPMorgan Chase & Co.," Shareholder.com, October 2, 2014 (http://investor.shareholder.com/jpmorganchase/secfiling.cfm?filingID=1193125-14-362173).

[24] Anthem, "2015 Cyber Attack Settlement Agreement Reached," (http://www.chs.net/media-notice/).

[25] "Anthem Cyber Attack Class Action Settlement Agreement," Anthemfacts.com, no date (https://www.anthemfacts.com/cyber-attack); "Anthem Data Breach: Frequently Asked Questions," Databreach-settlement.com, no date (http://www.databreach-settlement.com/Home/FAQ).

[26] Jay Clayton, "Statement on Cybersecurity," Sec.gov, September 20, 2017 (https://www.sec.gov/news/public-statement/statement-clayton-2017-09-20).

[27] Equifax, "2017 Cybersecurity Incident and Important Consumer Information," website, Equifaxsecurity2017.com, no date (https://www.equifaxsecurity2017.com).

[28] OPM, "Cybersecurity Resource Center," webpage, no date (https://www.opm.gov/cybersecurity/cybersecurity-incidents/).

[29] Data taken from OPM include up to 21.5 million records of Social Security numbers, passport numbers, birthdates, birthplaces, and multiple modes of contact information; details about victims' residential, employment,

Aggregation from Different Data Breaches Can Build a Strong Targeting Profile

Aggregating data from multiple breaches makes the potential for harm much greater. For example, combining credit report information, health care information, and travel information of a person helps to build a comprehensive profile and ideas of how to exploit this person. In this way, if the same actor carried out the hacks on OPM, Anthem, and United Airlines, they would have information on some of the most sensitive personal and health information, which, combined with travel information, can give a fairly comprehensive picture of a potential target, along with locations of frequently traveled places in order to "bump" into them.

And, in some cases, competing groups are all looking for the same data: State-sponsored actors might use credit report information to help build a dossier on a potential counterintelligence target, understanding their weaknesses and vulnerabilities, whereas cybercriminals would use the same information to create the custom dictionary for password guessing.

What Can Be Done?

Each type of cyber threat actor brings its own concerns. With cybercriminals, systems do not need to be completely secure: As long as there are easier and cheaper targets nearby, your safety is much higher. That said, the most sophisticated and determined threat actor will get through no matter what—in this case, security is more about making it expensive for an attacker (in terms of resources, money, time, research, and so on). By adopting certain core technologies, organizations can prevent cybercriminals from targeting the "low-hanging fruit" and instead turn to others with more-lax controls. Cybersecurity solutions that make it harder for threat actors to successfully breach an organization include regular patching and updating systems, employing multifactor authentication, encrypting data (in transit and at rest), and implementing regular security awareness training for people to be more aware of the individual threat and the need to protect themselves.

Recent congressional hearings on Equifax and other data breaches have been useful, highlighting several issues, including potential measures for how to improve response in the wake of a breach. One policy measure would require the extension of identity theft and credit monitoring services for victims to ten years, keeping the requirement for private companies the same as what Congress required of the U.S. government (extended from three years) in the wake of the OPM data breach. From a consumer protection perspective, this makes a lot of sense, given the permanent, unchangeable personal information that was compromised. Other solutions raised included requiring notification following a breach of security of a system containing personal information and amending the Fair Credit Reporting Act to provide access to free credit freezes for all consumers. Although the data from these particular breaches (Equifax and OPM) may not appear on black markets, breach notifications could apply to retailers and other organizations where the data stolen often do appear on black markets. The faster consumers are

---

travel, educational, criminal, financial, addiction, and mental health history; detailed information on spouses, cohabitants, other family members, and foreign contacts; and as many as 5.6 million fingerprint records.

aware of compromises, the lower the prices fetched for this information on the black market, which could disincentive the thefts. Free credit freezes could make it harder for criminals to monetize the information.

While not a cybersecurity panacea, information sharing and international cooperation between the public and private sectors can help.[30] State and local law enforcement would benefit from knowing about these markets. Communications between law enforcement, banks, and commercial sectors (health care, retail, entertainment) could help to track down nefarious actors. Although digital currency exchanges are a weak link in attributing a cybercriminal to stolen money, they are often housed overseas. Getting international cooperation is key for law enforcement to successfully pursue the attackers.

Finally, cybercrime markets are highly reliable. Finding ways of tarnishing the reputations of the markets, by wasting a criminal's time or making an exploit tool purchased on the black market ineffective, can help to prevent the loss of information and cut the value chain early in the attack cycle. Solutions might include spreading misinformation or injecting false products into the markets to breed distrust among the actors and increase the number and quality of arrests.

Thank you for the opportunity to testify, and I look forward to your questions.

[30] Martin C. Libicki, *Sharing Information About Threats Is Not a Cybersecurity Panacea*, Santa Monica, Calif.: RAND Corporation, CT-425, 2015 (https://www.rand.org/pubs/testimonies/CT425.html).

STATEMENT FOR THE RECORD OF
JOE BERNIK, CHIEF TECHNICAL STRATEGIST – FINANCIAL SECTOR, MCAFEE, LLC.
BEFORE THE U.S. HOUSE OF REPRESENTATIVES COMMITTEE ON FINANCIAL SERVICES, SUBCOMMITTEE ON TERRORISM AND ILLICIT FINANCE
ON "MONETIZATION OF STOLEN FINANCIAL AND OTHER DATA"
March 15, 2018, 2:00 PM | RAYBURN HOUSE OFFICE BUILDING ROOM 2128

Good morning, Chairman Pearce, Ranking Member Perlmutter, and distinguished members of the subcommittee. Thank you for the opportunity to testify today. I am Joe Bernik, McAfee's Chief Technical Strategist for the Financial Sector. I have two decades of experience creating and implementing cyber security management programs at global financial Institutions. While serving as CISO and head of information risk and security at ABN AMRO, Fifth Third Bank and BNY Mellon, I led teams dedicated to protecting customer data, complying with data-related laws and regulations and managing incident response programs. Since May 2016, I have developed cybersecurity practices, products and standards for McAfee. I work with the broader banking industry to align McAfee's products and roadmaps to the ever-evolving threat landscape faced by the financial services industry.

I am pleased to address the subcommittee on this important matter. My testimony will address the cybersecurity challenges financial institutions face, the threats posed by state and non-state actors, what happens to stolen data and how it is used, and general recommendations that can enhance the cybersecurity capabilities of financial institutions.

MCAFEE'S COMMITMENT TO CYBERSECURITY

McAfee is one of the world's largest cybersecurity companies, creating business and consumer solutions that help secure our digital lives. McAfee prides itself on building solutions that work with other industry peers, and we help businesses orchestrate cyber environments that are truly integrated, where protection, detection and correction of threats happen simultaneously and collaboratively. By protecting consumers across all their devices, we secure their digital lifestyle at home and while on the go. By working with other security players, we are leading the effort to unite against state-sponsored actors, cybercriminals, hacktivists and other disruptors for the benefit of all. McAfee is focused on accelerating ubiquitous protection against security risks for people, businesses and governments worldwide.

Before beginning my comments, I want to express how extremely pleased McAfee is to see the focus on improving the cyber threat landscape for financial institutions. These institutions comprise one of the most critical of all critical infrastructures. Along with governments, energy, water and telecommunications, the financial services sector is vital to the daily functioning of our economy and our overall security. Thank you for investigating ways to better protect this vital segment of our digital economy.

THE THREAT LANDSCAPE

According to a global report the Center for Strategic and International Studies (CSIS) and McAfee recently produced on Economic Impact of Cybercrime-No Slowing Down, banks are the favorite target of skilled cybercriminals – a fact that CSIS finds has been true for a decade. Yet it is also true that financial institutions, especially in the United States, invest more in cybersecurity – in both technology and information sharing efforts – than most other sectors. Finance was also the nation's first vertical to set up an ISAC (information sharing and analysis center), and it is generally seen as one of, if not *the*, leading sector in cybersecurity preparedness. There is no doubt, however, that financial institutions are still a prime target for cybercrime, for obvious reasons.

CSIS finds that North Korea and Russia are the top two nation states perpetrating financial cybercrime. Cyberattacks provide a lucrative way to supplement the North Korean government's access to foreign currency, and Russia provides a sanctuary for cybercriminals, housing some of the world's most talented cyber felons, whose attention focuses on the financial sector.

The attacks are not always directly against the largest banks, however. In fact, targeting the "seams" between well-defended networks is becoming increasingly common. The North Korean attack against the SWIFT network is a good example of exploiting weak points in the global financial network to make off with huge sums. In that situation, the North Korean Reconnaissance General Bureau (RGB) was aware of the difficulty of executing a large-scale heist from a single major western bank. Therefore, they targeted smaller, less sophisticated banks in developing countries like Bangladesh, Vietnam and Ecuador. Once they had compromised these banks' systems, they used the banks' credentials to send SWIFT fund transfer requests to larger banks in other countries. As the requests at first appeared legitimate, tens of millions of dollars were transferred fraudulently.

This practice of not targeting the largest financial institutions directly has proven to be quite effective in other situations as well. For instance, in the United States, cybercriminals zeroed in on Equifax, stealing the personal information of more than 145 million customers. The first repercussion most people thought of concerned the danger of identity theft. But that's just the beginning. Because of all the sensitive information Equifax housed, the hack was a blow to all organizations, including financial services companies. The Equifax breach and the 2015 Anthem breach before it rendered the Social Security number virtually useless as a trusted identifier throughout our economy. Social Security numbers used to be the gold standard, on which banks, credit card companies and others in the financial sector relied heavily for proof of identity. Without a secure identifier, assets become vulnerable, making the need for a digital ID much more urgent.

ATTACK METHODS AND INNOVATION

The attack methods cybercriminals use against the financial sector are similar to those used for other enterprises and organizations. Ransomware is the fastest growing cybercrime tool, with more than 6,000 online criminal marketplaces selling products and services, and ransomware-as-a-service is growing in popularity. Phishing attacks remain popular, and cybercrime-as-a-service is a big business. The dark web (the part of the World Wide Web that is accessible only by means of special software, allowing users and

website operators to remain anonymous or untraceable) offers buyers web injections, exploit kits and botnet rentals, among other tools. In some cases of ransomware, there really is no ransom that can be paid; data is simply encrypted with no way to retrieve it. This is a frightening type of cyber-attack, particularly for banks that remember the great depression of the 1930s.

There's been a lot of talk about cybercriminals innovating, and that is certainly accurate. With any new technology, there are uses for good and for bad. For instance, banks are using artificial intelligence and machine learning to enable advanced analytics to better serve and protect customers, but the sharpest cybercriminals also understand how to use it. Likewise, Amazon, Microsoft, Google and others are offering cloud technologies that can cut costs and make it easier to implement solutions, but customers don't always have the security part figured out. If companies put all their data in the cloud but don't protect it, they actually might be worsening their security posture. Our customers report that cybercriminals are also taking advantage of big data and advanced analytics for amassing and analyzing stolen information.

Yet while we know that cybercriminals have innovative tools in their arsenal, we haven't seen much of it yet because the old methods are still working. The exploits we see are more commoditized, and that fits with what we know about cybercriminals: They'll do only what they have to do to get a result, no more. If tried and true tactics like phishing are still working, why spend money to purchase a more sophisticated technique? We're seeing some new uses of exploits of WannaCry and Not Petya, and they continue to produce results. We also haven't seen a deliberate, sustained, destructive attack – the type that would render the organization inoperable – against a major American bank. This doesn't mean the financial sector has not been greatly impacted, however, as the theft of so much personal information in other hacks has weakened the fabric of the nation's financial infrastructure.

I do want to mention something that qualifies as an attack method, and that's the abuse of social media. Nation states have shown they're capable of interfering with, misdirecting and altering social media content – all of which can cause panic in financial markets. North Korea and Russia could certainly use social media to spread false information about financial conditions and markets, causing volatile markets to respond to fraudulent news. This is another example of evolving platforms such as social media, cloud and mobile offering both new opportunities as well as new means for abuse. We should be mindful that we're fighting a war, and it's very hard to keep up. We're going to need a lot of work by vendors and the government to protect citizens' information and keep major infrastructures such as power, water and finance stable.

WHAT HAPPENS TO THE DATA?

In 2015, McAfee produced a report on the market for stolen data (The Hidden Data Economy: the marketplace for stolen digital information), finding that stolen financial data is available not only on the dark web but also to anyone with a browser and the means to pay. Payment card information varies in price, depending on the amount of it out there. In 2015, the cost in the U.S. for a payment card number with a card verification value with a bank ID number and date of birth was $15. Today, it's worth approximately $12. PayPal

logins today go for $247 and online banking details for $160, according to the recently published Dark Web Market Price Index. The Index places the cost of purchasing a passport at $62, with the lowest cost stolen credentials being for Spotify – 12 cents to get a login.

Buyers of stolen information have many options, including the geographic source of a credit card and the card's available balance – both of which affect its price. Here's copy from a marketplace advertising its wares, as noted in the McAfee study referenced above:

> We are offering top quality cards. All our cards come with PPis [personal private information] and instructions. You can use them at any ATM worldwide. Our cards are equipped with magnetic strip and chip. Once you purchase, we will email you a full guide on how to safely cash out.

Everything is available on these online marketplaces, including credentials for bank-to-bank transfers and banking logins. As in other marketplaces, there are scams. One seller pitches: "Are you fed up of being scammed, and ripped? Are you tired of scammers wasting your time, only to steal your hard-earned money?" Just as in the legitimate digital marketplace, online forums are full of advice from buyers. Some sellers employ YouTube to advertise their wares. Still other types of data for sale includes access to systems within organizations' trusted networks. The types of entry vary, from very simple direct access (such as login credentials) to those that require a degree of technical competence (such as vulnerabilities).

Today, the monetization of stolen data has become easier due to the increased widespread use of digital currencies. According, again, to the McAfee-CSIS study on the Economic Impact of Cybercrime, anonymizing services like the "TOR" network and digital currencies have created an environment that gives cybercriminals both an arsenal and a sanctuary. TOR – free software that enables anonymous communication – has greatly enabled the expansion of cybercrime by allowing cybercriminals to hide their identities through a digital medium, further complicating law enforcement tracking efforts. Ransomware payments, for example, have been made more convenient and less traceable by crypto (digital) currency.

I want to emphasize, however, that these black markets represent uses of stolen data that are known – and most of these apply to criminals who hack for financial gain. Even more concerning are the evolving nation state hackers, who often are not putting the stolen data up for sale. The uses they have in mind go beyond simple financial gain, and that's the more worrisome part.

PREPAREDNESS

Our customers tell us that the three biggest problems they have are 1) dealing with conflicting regulations, 2) a constantly changing and evolving technology landscape and 3) the growing sophistication of cyber attackers. They also have to deal with cybersecurity tools that often don't work well with each other. The lack of interoperability among cybersecurity solutions limits their ability to exchange threat data on a rapid basis and creates seams of access for hackers. For our customers and also for McAfee and many

other organizations, including the federal government, a fourth problem is the insufficient supply of cybersecurity talent: There are simply too few qualified operators and professionals to enable organizations to stay on the top of their cybersecurity game.

The NIST Cybersecurity Framework provides a valuable roadmap for organizations of all sizes to evaluate their risk and see where their vulnerabilities are. We commend the U.S. government for enabling this partnership that has improved the security posture of many critical infrastructure industries. Likewise, compliance with Europe's General Data Protection Regulation (GDPR) will have a significant impact on improving both the security and privacy practices of those U.S. companies that collect data from European Economic Area residents. GDPR protects personal data in both administrative and technical manners, requiring anyone handling the data to record their uses and make sure that they are securing it.

Most major financial institutions are prepared for major cybersecurity attacks with the potential to produce system-wide failure. They have plans in place and are engaging actively in cyber sharing groups, in collaboration with the Department of Homeland Security (DHS), the Office of the Comptroller of the Currency (OCC) and the Federal Reserve. They know what they'll do first to identify and respond to a nation-state attack against economic critical infrastructure.

Our customers know that major attacks are possible. They don't know if they're imminent. So far, we haven't seen much use of the personal information stolen from Equifax and others, but the financial sector is ready and waiting. As good a job as institutions like Bank of America and US Bank are doing, they can't be expected to deter a nation state on their own. These companies value the partnership they have with the federal government in fighting cybercrime, and they look forward to having agencies such as DHS improve their cybersecurity capabilities.

Following are recommendations for ways the U.S. government can help in the constant battle against cybercrime.

POLICY RECOMMENDATIONS

Make the Social Security Number (SSN) Secure: The venerable nine-digit number first appeared as an identifier in 1936. It has become the de facto national identifier, a federal credential that people use for a range of both governmental and commercial purposes – uses for which it was never designed. Not surprisingly, the SSN is easy to guess, falsify or duplicate, and has become a premier target for cybercriminals. SSNs are sold in bulk in the cybercrime black market for as little as one dollar. Once stolen, the SSN cannot be reissued or replaced, making it a weak foundation upon with to build identity.

The steady stream of major breaches where consumers' SSNs were stolen, the most recent being 145 million stolen from Equifax, creates a compelling opportunity for change. Policymakers need to modernize the Social Security Number system. A good start is to determine what digital technologies offer strong security to create renewed confidence in the Social Security credential. A private sector eco-system of trusted identity management could then be built upon the new foundation of a modern, digitally secure SSN.

Enhance Cyber Threat Information Sharing: Although the financial services sector is a leader in information sharing, particularly in comparison to other sectors that are just now developing their own information sharing capabilities, the government can do more to help the financial services industry, and indeed, all industries, get the full benefit of cyber threat information sharing. McAfee believes that U.S. government efforts such as the DHS Automated Indicator Sharing (AIS) capability are useful but do not go far enough. There is a need to move beyond simple indicators supplied via AIS and provide a means to allow enrichment of the shared information.

Organizations need the ability to proactively collect, analyze and disseminate actionable intelligence. They need to consistently and proactively collect information and investigate it in an effort to show attribution. The government should work with the private sector to further evolve the way cyber threat information is represented, enriched and distributed in a timely fashion. Doing so will help create a high-functioning ecosystem of information sharing that enables the public and private sectors to compete with global networks of sophisticated hackers.

Implement Security and Privacy by Design: Adding or 'bolting on' security features to a system, network or device after it's already up and running has inherent weaknesses and inefficiencies, particularly in the financial sector, where companies are constantly being attacked. Policymakers should champion security and privacy by design to help incent broad adoption by the information and communication technology (ITC) ecosystem that supports all critical infrastructure sectors, including financial services. Proper protection of personal data in products needs to become an expectation throughout all data-centric industries.

Promote Cybersecurity Interoperability: Policymakers can help encourage the cybersecurity industry to continue to evolve by offering customers more solutions that benefit from an open platform model. An open platform is an architecture that makes it easier to deploy and manage a broad set of capabilities, not a business model dictating who and how others can participate. The broad set of capabilities, for instance, on Salesforce would not be possible on a closed platform.

Open cybersecurity platforms increase the rate and breadth of innovation by lowering development costs across the ecosystem. This helps leverage the power of the entire cybersecurity community to help stop the majority of unknown malware, correlate events across the broadest set of threat intelligence and produce compliance solutions appropriate for the largest population of customers. We support driving broad-based industry collaboration and adoption, partnering with standards groups to drive change toward open interfaces and allowing security products to more seamlessly integrate out of the box. We urge policymakers to reform federal procurement rules to enable faster uptake of cybersecurity solutions, particularly those based on open platforms. If procurement rules encouraged open platforms, moving the market to more standardized, open, and interoperable solutions, this alone would improve the security posture of our entire ITC ecosystem.

Pass National Breach and Security Legislation: Financial institutions are currently required to comply with the Gramm-Leach-Bliley data protection rules, and through interagency guidance, address data breach preparedness to protect their customers' data. GDPR requirements will further increase security and privacy protections for international organizations. McAfee supports the ongoing efforts to provide a U.S. federal data breach standard to enhance consumer protection to all Americans and across all sectors.

A federal data breach notification law, if implemented effectively, would provide public benefits by enhancing security and privacy, particularly in less-regulated sectors of the economy. It would support the flexibility of the NIST Cybersecurity Framework and allow organizations to focus on a single set of expectations. A U.S. federal breach law should provide a safe harbor for encrypted data and other cybersecurity protections. The ideal law would encourage good corporate behavior and incent companies to implement strong end-to-end cybersecurity programs. The U.S. federal law should be based on a technology-neutral framework that encourages effective, risk-based security strategies to protect personal information. The law should set simple procedural standards for breach notification (e.g., timing, what must be stated, who must be notified, etc.).

Requirements for federal preemption of state law, however, must be carefully considered. Pre-emption must not stifle innovation or weaken protections to those in states with strong data breach and data protection laws. While eliminating the patchwork of existing laws would provide the benefit of uniformity and additional legal certainty, this goal should not be accomplished by lessening existing strong and effective state laws.

CONCLUSION

The largest and most sophisticated companies in the financial services sector are at the top of their game in cybersecurity, particularly in comparison to smaller financial services companies and other industry sectors that have lagged in investing in the strategies, processes, people and technology needed to keep up with new threats and attackers. Thank you for giving me the opportunity to suggest ways the government can step up its cybersecurity game for the benefit of all financial services companies and consumers. I look forward to answering your questions.

After the Breach: The Monetization and Illicit Use of Stolen Data

Testimony by
Nicolas Christin, Ph.D.
Associate Research Professor
School of Computer Science, Institute for Software Research
College of Engineering, Department of Engineering and Public Policy
Carnegie Mellon University

Before the
Subcommittee on Terrorism & Illicit Finance
Committee on Financial Services
U.S. House of Representatives

The Honorable Stevan Pearce, Chairman
The Honorable Ed Perlmutter, Ranking Member

March 15, 2018

Chairman Pearce, Ranking Member Perlmutter, Members of the Subcommittee, thank you for hosting this important hearing today, and for giving me the opportunity to submit this testimony.

My name is Nicolas Christin. I am an associate research professor at Carnegie Mellon University, jointly appointed in the School of Computer Science and in the Department of Engineering and Public Policy. I am a computer scientist by training. My research focuses on computer security, and, for the better part of the last decade, I have been studying online crime. I have been focusing on the interface between technical and socio-economic aspects of computer abuse. In particular, I have conducted, with my research group, a series of measurement studies on online anonymous "dark web" marketplaces,[1,2] in an attempt to better understand the potential economic impact of these markets, including their role as retail channels for stolen data. This is the topic at hand today.

Monetizing stolen credentials and the asymmetry between societal costs and criminal revenue

The existence of an online market for stolen credentials – e.g., financial or personal data – can be traced back to the early 1990s. At the time, crooks were using dial-up forums (BBSes) to sell illicitly acquired credentials (e.g., credit card information directly stolen from postal mailboxes). As the Internet and the World Wide Web gradually rose to prominence in the mid-1990s, thieves started stealing credentials online (e.g., through "phishing" scams, tricking victims into revealing their financial information to miscreants), and dial-up forums moved to online chatrooms, primarily using the Internet Relay Chat protocol (IRC).

In an article based off seven months of data collected in early 2006,[3] Franklin et al. provided what is widely regarded as the first quantitative academic description of the online market for stolen credit card numbers. At the time, IRC chatrooms were still a very popular way for sellers and prospective buyers to transact. A vast majority of these chatrooms were open to the general public: All that was needed was freely available software (IRC clients) and the name and online location of the chatrooms used as marketplaces for purloined credentials.[4] This information could be easily found through simple web searches.

---

[1] Christin, Nicolas. "Traveling the Silk Road: A measurement analysis of a large anonymous online marketplace." In *Proceedings of the 22nd International World Wide Web Conference (WWW'13)*, pages 213-224. Rio de Janeiro, Brazil. May 2013.
[2] Soska, Kyle and Nicolas Christin. "Measuring the longitudinal evolution of the online anonymous marketplace ecosystem." In *Proceedings of the 24th USENIX Security Symposium (USENIX Security'15)*, pages 33-48. Washington, DC. August 2015.
[3] Franklin, Jason, Adrian Perrig, Vern Paxson, and Stefan Savage. "An inquiry into the nature and causes of the wealth of Internet miscreants." In *ACM conference on Computer and Communications Security*, pp. 375-388. 2007.
[4] Opening an IRC chatroom is an easy task, that requires little technical background. As a result, there exist IRC chatrooms dedicated to pretty much any topic one can imagine.

Franklin et al. described that the goods for sale included "bank logins and passwords, PayPal accounts, credit cards, and social security numbers (SSNs)."[5] Franklin et al. estimated that approximately 87,000 potentially valid credit card numbers were advertised in the chatrooms they monitored over their seven months of study, and concluded that the overall revenue for the market for stolen financial credentials sold on the chatrooms they monitored was somewhere between $37M and $95M.

In a subsequent 2013 study, Anderson et al. attempted to dimension the overall cost of cybercrime.[6] Extrapolating from estimates they produced for the United Kingdom, they projected that online card fraud (attempts to use stolen card numbers online, for instance, on electronic commerce websites) probably cost around $4.2B per year.

Anderson et al. caution that their projections "should be interpreted with utmost caution,"[7] and Franklin et al.'s study focuses on one specific distribution channel (IRC chatrooms). Furthermore, six years separate both studies. Nevertheless, the very strong disparity between both estimates, which differ by two orders of magnitude, suggests that the revenue criminals generate from the sale of stolen data is dwarfed by the costs data breaches impose on society. Later in this testimony, by considering more recent data, we will see that this asymmetry still exists today.

Evolution of the distribution channels and service professionalization

Distribution channels for stolen credentials have evolved in the past two decades. IRC chatrooms, while accessible to anybody with the proper software, only offered a relatively rudimentary text-based medium, primarily attractive to people with at least a modest amount of technological expertise. Easier to use web forums and websites with search functionality and better customer service thus became increasingly prominent since the early 2000s. These websites are generally referred to as "carding forums." Notorious examples include "carder.su," "shadowcrew.cc," among others. Original carding forums featured, in particular, forum moderators and embryonic reputation systems whose purpose was to provide better guarantees to prospective buyers and sellers. Each of these forums was run by a group of loosely affiliated criminals.

Business models also became increasingly complex. While a large number of vendors were simply selling credentials, some miscreants also started advertising "money mule recruitment services," destined to help with the transfer of funds acquired from stolen credentials to overseas accounts, or "confirmation services," which acted as external verification services to test the quality of the credentials being offered.

In short, these web-based distribution channels were designed to facilitate the sale and purchase of stolen data on a larger scale, by less sophisticated actors. Similar to industrial supply chains, the market for stolen data started to show increased specialization among its actors. Technically-savvy actors were in charge of procuring the "raw" data, i.e., causing the data breaches;

---

[5] Franklin et al., "An inquiry into the nature and causes of the wealth of Internet miscreants."
[6] Anderson, Ross, Chris Barton, Rainer Böhme, Richard Clayton, Michel J.G. van Eeten, Michael Levi, Tyler Moore, and Stefan Savage. "Measuring the cost of cybercrime." In *The economics of information security and privacy*, pp. 265-300. Springer, Berlin, Heidelberg, 2013.
[7] Ibid.

others were in charge of "commoditizing" these data, by breaking then down in lots suitable for individual resale; yet others were providing services surrounding stolen data (mule services, money laundering tutorials) without directly interacting with the data themselves.

This increased professionalization of the service, and the evolution of the distribution channels for stolen data continued with the next generation of retail channels for stolen data: "dark web" marketplaces.

The "dark web," online anonymous marketplaces, and cryptocurrencies

The World Wide Web can be fundamentally split between pages and websites that are indexed by search engines ("surface web"), and pages and websites that are not ("deep web"). While the deep web is thought to far exceed in size the surface web, most of the contents of the deep web is far from sinister. It includes, for instance, internal company pages, configuration pages (e.g., of properly configured home routers), and a large number of social network pages.

A very small portion of the deep web constitutes the "anonymous web," or "dark web."[8] All computers connected to the Internet are assigned an "IP address."[9] For instance, it is public knowledge that the computer with address 128.237.152.41 sits at Carnegie Mellon University. By looking up IP addresses, website operators and Internet Service Providers can thus easily determine who is browsing their websites, and likewise, one can usually learn on which server a given website is running. Websites in the anonymous web, however, are only accessible using freely available special-purpose software (e.g., Tor[10] or i2p[11]). This special purpose software allows the individual browsing the web to conceal their IP address, which is useful to bypass local censorship, maintain anonymity (often helpful for intelligence sources, or online investigators), or even, more mundanely, circumvent web tracking by online advertisers. The same software also allows webservers to conceal the IP address of the physical server on which they run if they so choose, yielding "end-to-end" anonymity: no one knows where the server or its visitors are located.

Starting in the mid-2000s, a number of forums promoting illicit contents started to appear in the anonymous web. These forums remained mostly confidential at the time, but the invention of the Bitcoin payment system in late 2008 drastically changed the picture. Until then, paying for illicit goods or services online was rather cumbersome: credit card payments and regular ACH transfers are very traceable and were thus a very poor fit for engaging in illicit transactions. Most miscreants instead relied on online payment systems like WebMoney or Liberty Reserve. However, those were also potentially problematic, as they were being run by centralized entities that could be pressured to

[8] We will refrain from using the "dark web" moniker, which is actually very confusing, given that most of these web sites are publicly accessible as long as the right software is used.
[9] The IP address may be public, in which case the computer is globally reachable, or private, in case a special purpose device with a public IP address, sitting on the same network as the computer, is required to allow the computer to connect to the Internet.
[10] Dingledine, Roger, Nick Mathewson, and Paul Syverson. Tor: The second-generation onion router. In Proceedings of the 13th USENIX Security Symposium, August 2004.
[11] I2P: The internet invisible project. http://www.geti2p.net.

intervene in cases of illicit activity or face legal consequences.[12] Bitcoin, on the other hand, is fully decentralized and pseudonymous, thereby offering superior privacy guarantees to its users, compared to these other, previously established payment systems.

Building on these technological developments, in February of 2011, a website called "Silk Road" opened on the anonymous web. Silk Road was the first to combine the network-level anonymity properties provided by the Tor network, with the superior privacy guarantees offered by the Bitcoin payment system, to offer a fully functional "online anonymous marketplace." Silk Road itself did not sell any product *per se*, but provided a venue where buyers and vendors could interact with each other anonymously, and with a certain level of trust. Similar to traditional electronic commerce marketplaces like eBay or the Amazon Marketplace, Silk Road offered a feedback-based review system, through which buyers could rate sellers (and, internally, sellers could also rate buyers, although this information was not public). Because Silk Road had very lax rules on what could and could not be listed, it very quickly become a haven for illicit activity.[13]

Buyers had to leave a public review for each purchase they made on the site. As a result, one could relatively precisely estimate the total number of sales taking place, and the associated revenue of the entire marketplace. Our original study[14] estimated that in the first months of 2012, Silk Road was on track to generate approximately $15M of yearly revenue and derived an overwhelming fraction of this revenue from drug sales – narcotics and prescription drugs.

Silk Road grew significantly in late 2012 and early 2013, reaching a revenue of more than $350,000 per day in the Summer of 2013.[15] The location of the server and identity of its operator were eventually discovered by authorities, which closed the site and arrested its owner in November 2013. Numerous "copycat" marketplaces immediately appeared in its stead (including a popular site named "Silk Road 2.0"), using a similar combination of network-level anonymization technologies and pseudonymous cryptocurrencies.

Most of these online anonymous marketplaces use English as the primary language. However, operators, vendors and buyers can be located anywhere in the world. Among notorious cases, Silk Road and Silk Road 2.0 were reportedly operated by U.S. nationals; AlphaBay was allegedly run by a Canadian citizen residing in Thailand; Hansa Market, by two German nationals; and Sheep Marketplace, by a Czech individual. Marketplaces in other languages do exist but tend to be generally smaller. A notable exception was the Russian Anonymous Marketplace (RAMP), which was active for a few years and was a fairly sizeable site.[16]

---

[12] Liberty Reserve was eventually shut down by U.S. authorities under the Patriot Act, and its founder charged with and convicted of money laundering.
[13] Christin. "Traveling the Silk Road."
[14] Ibid.
[15] Soska and Christin. "Measuring the longitudinal evolution of the online anonymous marketplace ecosystem."
[16] Greenberg, Andy. "How a Russian dark web drug market outlived the Silk Road (and Silk Road 2)." *Wired.* November 2014. https://www.wired.com/2014/11/oldest-drug-market-is-russian/

By 2015, through research done at Carnegie Mellon, and partially supported by the Department of Homeland Security Science and Technology Directorate, we estimated that the ecosystem of online anonymous marketplaces was generating a revenue in excess of half a million US dollars per day, that is, approximately $200M a year.[17]

As this criminal ecosystem was growing, law enforcement orchestrated a number of takedowns. In particular, authorities managed to shut down a number of online anonymous marketplaces in 2014 (including Silk Road 2.0, and other less prominent bazaars) during a joint effort between U.S. and European law enforcement agencies dubbed "Operation Onymous." We nevertheless observed that the online anonymous marketplace ecosystem, as a whole, appeared highly resilient to such disruptions. When a leading marketplace is taken down (or absconds with its customers' money), consumers appear to move relatively quickly to another marketplace, and the long-term impact of online anonymous marketplace takedowns has yet to be convincingly observed.[18]

For instance, immediately after Operation Onymous, sales on the Evolution and Agora marketplaces, which had not been affected by law enforcement intervention, ramped up significantly. When these marketplaces disappeared in 2015, the AlphaBay marketplace, which had started operating in December 2014, rose to prominence, and went on collect a revenue exceeding $800,000 per day in early 2017.[19]

Remarkably, both Evolution and AlphaBay (and other lesser known online anonymous marketplaces) initially started as carding forums, before expanding their businesses to other areas. For instance, close to 50% of all revenue on AlphaBay in its first couple of months of existence (at a time when overall revenue was still low) came from "digital goods,"[20] a category that encompasses fraudulently obtained financial credentials, forged documents, hacking kits, and so forth.

At their peak, both Evolution and AlphaBay derived a majority of their revenue from commissions on drug sales. Nevertheless, sellers of illicitly acquired data had certainly taken notice that online anonymous marketplaces were a valuable distribution channel for their products.

Digital goods and online anonymous marketplaces

An overwhelming majority (more than 80%) of the revenue of the online anonymous marketplaces we monitored comes from the sale of narcotics or prescription drugs. However, a non-negligible

---

[17] Soska and Christin. "Measuring the longitudinal evolution of the online anonymous marketplace ecosystem."
[18] Ibid.
[19] European Monitoring Centre for Drugs and Drug Addiction and Europol. "Drugs and the darknet: Perspectives for enforcement, research and policy." EMCDDA–Europol Joint publications, Publications Office of the European Union, Luxembourg. November 2017.
[20] Ibid.

portion of sales (~5-10%) concerns "digital goods," and this proportion appears to be growing (albeit slightly) over time.

In collaboration with researchers based at Delft Technical University in the Netherlands, we recently performed a thorough investigation of the market for digital goods in online anonymous marketplaces. Using data we had collected between 2011 and 2017, from most of the major online anonymous marketplaces operating during that time interval, we were able to observe the following.

First, the largest type of digital goods listings we observed (approximately 12,000 out of roughly 44,000 total offerings in the digital goods category) were "cash-out" schemes. These cash-out schemes primarily include 1) synthetic credit card numbers not associated with any real account, but that would pass rudimentary automated validity checks—those are usually not harmful to any specific individual, 2) "fullz," that denote comprehensive records, pairing for instance stolen credit card numbers, with the associated CVV codes, and in some cases the social security number or date of birth of the legitimate owner, and 3) various types of guides, including money laundering tutorials (e.g., how to recruit money mules). A smaller number of listings were for 4) bank and financial account credentials (e.g., PayPal logins) and 5) money laundering services (e.g., "Bitcoin deals," or cash payouts, such as vendors offering cash in the mail in exchange for Bitcoin).

In other words, similar to offerings on dedicated carding forums, we see a fairly large range of monetization techniques, ranging from the sale or purloined data, to integration of illicit profits in the legitimate financial system. We do not see, however, very high-value data leaks such as the OPM breach data. Highly valuable goods such as government personnel data are indeed more likely to be of value to nation-states (e.g., for intelligence purposes) than they are for financial purposes.

Second, the estimated revenue generated by cash-out schemes on anonymous online marketplaces remained low – altogether, digital goods, of which cash-out schemes are a subset, represented approximately a total revenue of $29M. We caution this is a conservative estimate, as 1) we measured only the most prominent generalist online anonymous marketplaces, and 2) due to technical difficulties inherent to large-scale data collection, the data we have may not be fully complete. Despite these caveats, these revenue numbers are strikingly low compared to the societal costs of breaches (e.g., costs to bank to reissue cards,[21] costs to individuals to restore their credit in case of a breach). This confirms the aforementioned trends observed in the late 2000s-early 2010s, which predated the use of online anonymous marketplaces as a retail channel for stolen data.

This overall low revenue can be partially attributed to the low value of most of the goods sold on online anonymous marketplaces. Overall, the median price for a cash-out listing is only around $60. Credit card numbers (without additional information) typically only sell for a few dollars

[21] Graves, James, Alessandro Acquisti and Nicolas Christin. Should Credit Card Issuers Reissue Cards in Response to a Data Breach?: Uncertainty and Transparency in Metrics for Data Security Policymaking. To appear in *ACM Transactions on Internet Technology*. 2018.

apiece, and are often sold in lots (e.g., 100 Visa cards from UK banks); comprehensive records, including social security numbers, usually go for $100 or less.

The low retail value of stolen data is partially due to improvements in fraud detection, and proactive blacklisting of financial credentials. These defenses make a significant fraction of credit card numbers and other credentials being sold online at a retail level likely to be worthless. This also explains why sellers often package stolen data in lots. Interested buyers purchase lots with the hope that some of the information purchased is not yet blacklisted, and/or would give them access to sizeable revenue. But, by and large, for a buyer, engaging in such transactions is akin to purchasing a lottery ticket from a less-than-reputable outlet.

Third, we found that, since 2014, on any given day, approximately 3,000 to 4,000 vendors were simultaneously actively offering cash-out listings. While different vendors may have been active at different times, the overall population of active vendors increased slightly between 2014 and 2017. More interestingly, the top 10% of vendors were responsible for 80% of the revenue. In other words, the business of selling purloined data (or services surrounding it) appears to be truly profitable only for the most successful criminals on these forums.

These most successful vendors manage to establish a solid reputation, and either have a steady influx of goods to sell, or a large number of goods to sell over relatively short periods (a couple of months). We see few "bulk" sales, which hints that large breaches are first broken down into smaller lots, and those smaller lots are subsequently sold independently. This commoditization process, unfortunately, generally does not play out in public marketplaces. This makes traceability challenging: Retail-level vendors on online anonymous marketplaces rarely advertise the provenance of the credentials they are offering, and in fact, may not even know how these credentials were acquired in the first place.

Summary and moving forward

With the increased digitization of records, online financial fraud and data breaches are becoming a critical problem. Our recent measurement studies of online anonymous marketplaces, or "dark web markets," allow us to get a relatively precise idea of some of the business models in use, and of the economics of stolen data.

We find that revenue generated by criminals engaged in monetizing these breaches pales in comparison to the potential costs of the remedies. For instance, stolen credit card and identity details are often sold in lots, at low retail prices. However, the owner of a stolen banking credential has to invest potentially considerable time, effort and money to attempt to repair the damage incurred by the theft.

There is also a noticeable level of activity in the sale of "services" surrounding data breaches, such as verification of the data, or money laundering and integration of the profits generated. Further, these marketplaces are international in nature, and, even if certain actors might be identified (e.g., through undercover operations), jurisdiction issues may complicate prosecution and/or arrest of individual vendors.

In addition, the online anonymous marketplace ecosystem as a whole has shown strong resiliency to law enforcement takedowns. Shutting down one or more marketplaces has so far mostly seemed to result in criminals moving to different marketplaces, and long-term impacts on the ecosystem are uncertain. Takedowns also may potentially lead some of the members of that ecosystem to move their activity to less publicly observable forums (e.g., private vendor forums).

We also observe that most of the revenue is generated by a small fraction of all criminals. This is a constant in cyber-crime, beyond stolen data markets.[22,23] A few highly successful criminals seem to attract relatively large numbers of amateurs that do not profit much, if at all, from their activities. Unsuccessful retail-level vendors nevertheless contribute to the overall problem, by making the market for stolen data larger, and more complex.

All of these findings indicate that focusing on preventing breaches from happening in the first place is likely to be more economically efficient than attempting to disrupt distribution channels, or recovering from a data breach once it has happened.

Finally, measurements of "dark web" marketplaces focus on the retail end of the stolen data ecosystem, and are thus an imperfect signal, particularly when it comes to tracing stolen data back to a specific breach. Nevertheless, such measurements give us important information on the health and evolution of the market for illicitly acquired data, and on the monetization techniques in use.

It is thus important to continue supporting these documentation efforts, so that we can decipher the evolving economic and business models that support stolen data markets. Indeed, understanding the criminals' economic incentives is critical to determine which combination of defensive strategies (technical, legal, economic) are most likely to disrupt criminal business models, thereby creating adverse incentives for miscreants to engage in these activities in the first place.

---

[22] Clayton, Richard, Tyler Moore, and Nicolas Christin. Concentrating Correctly on Cybercrime Concentration. In the *Proceedings (online) of the 14th Workshop on Economics of Information Security (WEIS 2015)*. Delft, Netherlands. June 2015.
[23] Leontiadis, Nektarios. *Structuring Disincentives for Online Criminals*. PhD Thesis. Carnegie Mellon University. August 2014.

CSIS | CENTER FOR STRATEGIC & INTERNATIONAL STUDIES

Statement Before the

House Committee on Financial Services
Subcommittee on Terrorism and Illicit Finance

"After the Breach: The Monetization and Illicit Use of Stolen Data"

A Testimony by:

James Andrew Lewis

Senior Vice President

Center for Strategic and International Studies (CSIS)

March 15, 2018

2128 Rayburn House Office Building

I thank the Committee for the opportunity to testify on this subject.

In the 1990s, when the Internet was commercialized, there was a strong millennial belief that this was part of a new age of peace and harmony, with the end of the Cold War and with what some went so far as to call the end of history. All countries would be market democracies, Russia and China would be friends, the role of government would shrink and be replaced by a new multi-stakeholder governance model, the boundaries between countries would fade and the Internet would be the glue that held this new world together.

Millennial optimism has proven to be badly mistaken, but it still undergirds some of our thinking about the Internet, such as the benefits of anonymity, often justified as essential for dissent, but which remains an immense benefit for criminals. The last few years have shown that the Internet has a dark underside that is deeply troubling. The Internet has brought tremendous economic benefit, but this comes with the costs created by espionage and crime. The loss per victim from cybercrime can be low, but there are many victims and the costs and risks of engaging in cybercrime are even lower, making this an irresistible criminal activity. The task for policymakers and legislators is to find a way to reduce that cost without sacrificing the Internet's benefits.

Cybercrime is big business. How big a business is a subject for dispute and, like so many things connected to information technology and the Internet, also a subject of imprecision, hype and exaggeration. CSIS has conducted three studies, with the support of McAfee, to estimate the losses from cybercrime. In interviews for our studies, one senior official called it "the greatest transfer of wealth in human history," while another, a member of the Council of Economic Advisors at that time, called it a "rounding error in a fourteen trillion-dollar economy." Through our work, we hoped to narrow the gap between these two extremes.

Our first study of cybercrime, done in partnership with Stewart Baker, attempted to establish upper and lower bounds for cybercrime by looking at other categories of crime for which there was available data, such as narcotics, maritime piracy, pilferage and (for an estimate to social cost) automobile crashes. This comparison let us estimate how much crime a society can tolerate as part of everyday life, and suggested a range of 0.5% to 1.5% of national income. For perspective, when you hear that cybercrime costs the U.S. a trillion dollars a year, this would be roughly 6% of national income, an immense sum unmatched by any other category of crime. We assessed this as unlikely.

Estimation of the cost of cybercrime is challenging because data collection is woefully inadequate. Even major economies do not collect statistics on cybercrime. This is somewhat understandable because many victims prefer not to report their losses. We tried to account for this in our estimate. There are also valuation problems in deciding how much stolen intellectual property is actually worth or what the market prices are for stolen personal information, a price that can fluctuate with supply. Additionally, there is no common definition on what should count as cybercrime. Some countries count anything where a computer is used. Others do not count intellectual property theft. Developing a common global standard of what should be counted and making the collection of data a priority would help us asses the scope of the

problem. Until then, estimates must do. This may change over time, as insurance companies collect actuarial data on cybercrime, or it may require government intervention in the same way we estimate the cost of narcotics-related crime.

One major difficulty for estimating the cost of cybercrime and cyber espionage is the problems that criminals face in monetizing the results of their theft. Even if we know the value of what was taken, in many cases criminals cannot gain the full value, particularly for personally identifiable information (PII) or intellectual property (IP). It is harder (in some cases, much harder) to monetize the result of a successful hack than it is to hack itself. One reason we believe that cybercrime continues to increase is that criminals have become better at monetization, in part because of the availability of cryptocurrencies like Bitcoin.

Monetization is easiest is when a criminal can transfer funds directly from the victim to a bank account. In the past, this was done by using "mules" or "cashers" to launder money extracted from breached accounts. Cybercriminals transferred stolen funds to the mules' accounts; the mules will take a "commission" (often between 5-10% of the total) and forward the rest to overseas accounts. These older processes were both risky and inefficient. The development of cryptocurrencies reduced risk and increased returns, by increasing the anonymity and ease of criminal transactions. The cybercrime monetization process is increasingly digitized, with criminals moving stolen funds rapidly among accounts with the goal of using it to buy cryptocurrencies in untraceable ways.

Business confidential information can also be monetized easily, by providing the criminal acquirer an advantage in business negotiations or an ability to conduct a transaction at a lower cost than would otherwise be the case. Accounting firms and law offices have become favorite targets for this category of cybercrime since many are small and not well protected. Advance information on quarterly results or mergers and acquisitions could allow a sophisticated criminal to take advantage of the market in ways that could be difficult to trace, making the manipulation of stock prices and other financial assets one of the more difficult aspects of cybercrime. This kind of financial manipulation avoids many of the problems related to monetization.

Monetization of stolen data, whether IP or PII, has always been a problem for cybercriminals. Digital currencies have helped to change that, but they have not solved the fact that there can be a broad gap between what cybercriminals steal and what they able to exploit. Criminals cannot monetize everything they take. Millions of individuals can lose their credit card data in a single incident, but only a fraction of those affected will experience monetary loss. Similarly, thieves and spies may take intellectual property that cost billions to develop, but they face real challenges in their ability to turn this IP into competing products. The gain from the crime to the criminals will vary from product to product depending on how easy it is to turn the stolen IP into a product that can be sold on the market. The theft of a formula for some product like house paint or furniture, for example, allows a competitor to begin production almost immediately. The theft of IP for high tech products like semiconductors, however, might not be useful at all without a modern industrial base that can manufacture products based on the stolen IP.

As an aside, this is part of the explanation as to why China has tried in the last few years to acquire semiconductor companies in the West. China's economic espionage actions before 2015

included the acquisition of IP related to semiconductors, but the Chinese, despite massive investment, lacked the "know-how" to turn the stolen IP in products. While the purchase of western companies has been blocked by regulatory tools, such as the CFIUS process, China's immense government investments and use of joint ventures will eventually allow them to them to overcome the "know-how" obstacle.

Our second report developed a model to estimate of the global cost of cybercrime, based on data from interviews with government officials in a number of countries as well as published and private data on nations' aggregate cybercrime losses. We found information on thirty-two countries that account for a significant portion of global income, and used this data to construct a global estimate for cybercrime. We looked at a broad range of costs, including recovery costs, damage to brand and liability, and opportunity costs– the value of opportunities or benefits that cannot be realized because resources have been expended to protect or recover from cybercrime. We estimated that in 2014 the global loss was between $375 and $425 billion a year.

Our third and most recent study used the same methods and refined this approach by looking at countries by income group (high, medium or low income, using World Bank categories). The study showed an increase in cost, and estimated that cybercrime cost the world between $450 and $600 billion a year, roughly a twenty percent increase. This increase can be explained by the increasing sophistication of cybercriminals, by the larger number of Internet users and volume of Internet transactions which increases the pool of potential victims, and by improvements in the ability of cybercriminals to monetize stolen data.

This improved ability to monetize stolen data is in good measure the result of availability of cryptocurrencies and the continued growth of cybercrime black markets in what some call the "Dark Web." The dark web, websites accessible only through special programs or networks like Tor, has created a space for sophisticated criminal markets and transactions to operate outside the reach of law enforcement and has made the Internet a central hub for global criminal activity in drugs, child pornography, arms, and malware. Cryptocurrencies are an essential part of these marketplaces, allowing transactions to occur with far less visibility than ever before. The development of the dark web and cryptocurrencies support the growth of a sophisticated cybercrime ecosystem, and have eased the challenges of monetizing the spoils of cybercrime.

Digital currencies are cumbersome to use for many transactions, fluctuate in value, and are not widely accepted by mainstream commercial vendors. In 2017, the largest daily amount of bitcoin transactions was around $5 billion. For context, data from the Bank of International Settlements suggests that $5 trillion is traded every day in currency trades. Bitcoin is a rounding error in global financial transactions. Perhaps someday this will change, but for now, cryptocurrencies are primarily a vehicle for currency speculation, online gaming, and for cybercrime.

The preferred currency for anonymous transactions remains the U.S. $100-dollar bill, with more than twelve billion bills in circulation according to the Treasury Department.[1] Cash is still preferred for crime and tax evasion, but for cybercrime, cryptocurrencies have an advantage by avoiding the need for cumbersome and detectable physical transfers or bank transfers subject to

[1] https://www.federalreserve.gov/paymentsystems/coin_currcircvolume.htm

regulation. The failure to counter the proliferation of unregulated digital currency exchanges, and the availability of strong encryption has created opportunities for cybercriminals, state-sponsored cybercrime, sanctioned governments, and terrorists as they effectively evade money laundering controls and find it easier than ever to move large sums quickly and anonymously.

Cryptocurrencies are the digital equivalent of cash, and can allow for untraceable financial transactions. Bitcoin has long been the favored currency for darknet marketplaces, with cybercriminals taking advantage of its pseudonymous nature and decentralized organization to conduct illicit transactions, demand payments from victims, and launder the proceeds from their crimes. Bitcoin's oft-cited anonymity is not perfect, however, which has led to the emergence of a new generation of privacy-enhanced cryptocurrencies offering far greater protection to help cybercriminals conceal the details of their transactions and evade law enforcement. There are dozens of different cryptocurrencies on offer world-wide. Transactions using cryptocurrencies are difficult to trace and once the cryptocurrency is obtained in the commission of a crime, it is relatively easy to use the Internet to transfer it to a bank and exchange it for fiat currency.

Monetization opportunities have also increased due to the flourishing black markets found in cyberspace. Encryption, the dark web and cryptocurrencies have created a safe haven for cybercrime. These black markets are not accessible from the visible Internet, nor can they be discovered by widely used search engines. Access to these markets is usually restricted. On them, you can buy the latest hacking tools or recently stolen PII, learn of recently discovered vulnerabilities, and rent "botnets" --tens of thousands of computers remotely controlled for criminal purposes, usually used for conducting denial of service attacks or engaging in cryptocurrency mining. These black markets can be highly specialized. Some sellers offer guarantees, product ratings, and customer service. Personal information - credit card numbers, social security numbers, and bank accounts – can be bought in lots of thousands or even millions, and buyers have the choices of 'raw' information or personal information that that has been tested for accuracy. These markets are one reason why cybercriminals are adaptive and dynamic in developing new tools and techniques that challenge cyber defenses.

The tools available for crime on the Dark Web continue to improve. Cybercrime attracts innovators and is a dynamic technological environment. There used to be a lag of somewhere between three and five years between the use of hacking tools developed by advanced intelligence agencies and their spread to cybercrime markets for purchase or rental, but this lag is shrinking. The evidence for this is anecdotal, but the trend has been consistent for several years. Recent events, such as the leak of advanced hacking tools on WikiLeaks or through the "Shadow Brokers," has accelerated the improvement in capabilities in both criminal groups as well as nations. Both of these recent leaks are probably the result of Russian intelligence activities, and the Russian state and cybercrime groups are deeply intertwined.

Russia is a haven for the most advanced cybercrime groups and no clear line delineates the criminal world from the government. The Kremlin sees Russian cybercriminals as a strategic asset, and one of the most difficult problems for reducing cybercrime is that Russia, along with North Korea, will not cooperate with Western law enforcement. High-end cybercriminal groups in Russia have hacking capabilities that are better than most nations. A Russian hacker was responsible for the Yahoo breach, compromising more than a billion credentials which were used

for both criminal and intelligence purposes. NoPetya was Russian malware designed to collect both intelligence and commercial information. Russian cyber criminals have likely hacked law firms, accountants, and investment companies to gain information that will let them manipulate financial markets.

The other state that supports cybercrime is North Korea (DPRK). North Korean government agencies have long used criminal activities to gain hard currency for the regime. The North has always relied on criminal activities - smuggling, counterfeiting, to gain hard currency, and in recent years, it has used the hacking skills of its principal intelligence agency, the North Korean Reconnaissance General Bureau (RGB), for cybercrime. The most famous examples of North Korean state cybercrime are the hack of the Bangladeshi Central Bank and Wannacry ransomware event.

These attacks provide a lucrative means to supplement the North Korean government's limited access to foreign currency and to evade sanctions. The DPRK uses variants of malware available on the cybercrime black market and has been successful mainly against poorly protected targets. North Korean cyber capabilities have not yet reached the level that would allow them to go after the most advanced targets (such as American banks), duplicate Stuxnet or the Russian attack on Ukrainian power facilities.

North Korea has also turned to cryptocurrency theft to help fund its regime. North Korean hackers have targeted at least three South Korean cryptocurrency exchanges in 2017.[2] Cryptocurrencies are a particularly valuable target for North Korea, who is able to use Bitcoin's anonymity to circumvent international sanctions. There is some speculation that North Korea has also installed bitcoin mining software on hacked computers to mine for cryptocurrencies. The Pyongyang University of Science and Technology now offers its students classes in bitcoin and blockchain.

Protected spaces on the dark web, an innovative cybercrime ecosystem, cryptocurrency and countries that engage in and support cybercrime – this is a daunting list of problems, but there are solutions. Each of these ideas deserves longer discussion, but in brief,

- The U.S. and its allies must develop an appropriate and effective strategy for punishing states that support cybercrime. This may need to go beyond traditional law enforcement activities to disrupt cybercriminal networks, software programs, and financial resources, much as the Navy had to take action against the Barbary Pirates. In general, the U.S. needs to develop retaliatory strategy, since as long as there are no penalties for malicious cyber action, our opponents see no reason to behave in cyberspace, and this applies to cybercriminals as well as states.

- Many countries are moving to regulate or even block cryptocurrencies. This is a draconian solution to the problem. Cryptocurrencies whose use can be done in ways consistent with anti-money laundering and other financial regulations should be allowed to operate. Those cryptocurrencies and related "mixing services" designed to evade

[2] Luke McNamara, "Why Is North Korea So Interested in Bitcoin?," FireEye, September 11, 2017

money laundering requirements should be banned.

- Widespread adoption of effective cybercrime laws by all countries remains essential, as countries with weak cybercrime laws tend to experience a higher rate of crime. The best vehicle at this time is the Budapest Convention.

- We are unlikely to ever be able to suppress the Dark Web, so efforts to disrupt and dismantle criminal networks should be expanded through increased resources and technology for law enforcement agencies and increased international cooperation.

- Expanded international law enforcement cooperation and the modernization of important tools like Mutual Legal Assistance Treaties (MLAT) are essential for countering cybercrime.

- Companies should ensure their cyber defenses are adequate. In the U.S. this has been done on a voluntary basis. Other countries are moving to a more regulatory approach that requires companies to meet higher standards of cybersecurity.

- Encryption remains a vexing problem. Opinion in many other countries is moving slowly toward greater constrains on the use of the kinds of encryption that create problems for law enforcement, but restrictions would face opposition from privacy groups in the U.S. and other countries. There is no consensus on possible solutions to the encryption problem. These solutions fall into two broad categories: restricting access to encryption that does not allow for recovery of plaintext by third parties or, alternatively, increasing law enforcement capabilities and resources to break encryption or use metadata to deal with the evidentiary problems encryption creates.

- Harmonization of international requirements for cybersecurity in important sectors like finance would both improve security and reduce the compliance burden on multinational companies.

- Finally, all nations would benefit from a serious effort at the national and international level to develop common definitions and measurements for cybercrime and collect data on its cost. We do this now for transnational crimes like narcotics or piracy, and cybercrime should be added to this list.

I thank the Committee for the opportunity to testify and welcome any questions.

Sex, drugs, and bitcoin: How much illegal activity is financed through cryptocurrencies? *

Sean Foley [a], Jonathan R. Karlsen [b], Tālis J. Putniņš [b, c]
[a] *University of Sydney*
[b] *University of Technology Sydney*
[c] *Stockholm School of Economics in Riga*

January, 2018

Abstract

Cryptocurrencies are among the largest unregulated markets in the world. We find that approximately one-quarter of bitcoin users and one-half of bitcoin transactions are associated with illegal activity. Around $72 billion of illegal activity per year involves bitcoin, which is close to the scale of the US and European markets for illegal drugs. The illegal share of bitcoin activity declines with mainstream interest in bitcoin and with the emergence of more opaque cryptocurrencies. The techniques developed in this paper have applications in cryptocurrency surveillance. Our findings suggest that cryptocurrencies are transforming the way black markets operate by enabling "black e-commerce".

JEL classification: G18, O31, O32, O33

Keywords: blockchain, bitcoin, detection controlled estimation, illegal trade

* We thank an anonymous referee, Tristan Blakers, Andrew Karolyi, Maureen O'Hara, Paolo Tasca, Michael Weber, as well as the conference/seminar participants of the RFS FinTech Workshop of Registered Reports, the Behavioral Finance and Capital Markets Conference, the UBS Equity Markets Conference, and the University of Technology Sydney for comments and suggestions. We also thank Tristan Blakers, Adrian Manning, Luke Anderson, Yaseen Kadir, Evans Gomes, and Joseph Van Buskirk for assistance relating to data. Jonathan Karlsen acknowledges financial support from the Capital Markets Co-operative Research Centre. Tālis Putniņš acknowledges financial support from the Australian Research Council (ARC) under grant number DE150101889. The Online Appendix that accompanies this paper can be found at goo.gl/GvsERL.
Send correspondence to Tālis Putniņš, UTS Business School, University of Technology Sydney, PO Box 123 Broadway, NSW 2007, Australia; telephone: +61 2 95143088. Email: talis.putnins@uts.edu.au.

1. Introduction

Cryptocurrencies have grown rapidly in price, popularity, and mainstream adoption. The total market capitalization of bitcoin alone exceeds $250 billion as at January 2018, with a further $400 billion in over 1,000 other cryptocurrencies. The numerous online cryptocurrency exchanges and markets have daily dollar volume of around $50 billion.[2] Over 170 "cryptofunds" have emerged (hedge funds that invest solely in cryptocurrencies), attracting around $2.3 billion in assets under management.[3] Recently, bitcoin futures have commenced trading on the CME and CBOE, catering to institutional demand for trading and hedging bitcoin.[4] What was once a fringe asset is quickly maturing.

The rapid growth in cryptocurrencies and the anonymity that they provide users has created considerable regulatory challenges. An application for a $100 million cryptocurrency Exchange Traded Fund (ETF) was rejected by the US SEC in March 2017 (and again in 2018) amid concerns including the lack of regulation. China has banned residents from trading cryptocurrencies and made initial coin offerings (ICOs) illegal. Central bank heads have publically expressed concerns about cryptocurrencies. While cryptocurrencies have many potential benefits including faster and more efficient settlement, regulatory concerns center around their use in illegal trade (drugs, hacks and thefts, illegal pornography, even murder-for-hire), potential to fund terrorism, launder money, and avoid capital controls. There is little doubt that by providing a digital and anonymous payment mechanism, cryptocurrencies such as bitcoin have facilitated the growth of "darknet" online marketplaces in which illegal goods and services are traded. The recent FBI seizure of over $4 million of bitcoin from one such marketplace, the "Silk Road", provides some idea of the scale of the problem faced by regulators.

This paper seeks to quantify and characterize the illegal trade facilitated by bitcoin. In doing so, we hope to better understand the nature and scale of the "problem" facing this nascent technology. We develop methods for identifying illegal activity in bitcoin. These methods can also be used in analyzing many other blockchains.

Several recent seizures of bitcoin by law enforcement agencies (including the US FBI's seizure of the "Silk Road" marketplace), combined with the public nature of the blockchain, provide us with a unique laboratory in which to analyze the illegal ecosystem that has evolved in the bitcoin network. Although individual identities are masked by the pseudo-anonymity of a 26-35 character alpha-numeric address, the public nature of the blockchain allows us to link bitcoin transactions to individual "users" (market participants) and then further identify the users that had bitcoin seized by authorities. Bitcoin

[2] SEC Release No. 34-79103, March 10, 2017; and https://coinmarketcap.com.
[3] Source: financial research firm Autonomous Next and cnbc.com.
[4] Bitcoin futures commenced trading on the CME (Chicago Mercantile Exchange) on December 18, 2017 and on the Chicago Board Options Exchange (CBOE) on December 10, 2017. A bitcoin futures contract on CBOE is for one bitcoin, whereas on CBOE it is five bitcoins. At a price of approximately $20,000 per bitcoin at the time the CME bitcoin futures launched, one CME bitcoin futures contract has a notional value of around $100,000.

seizures (combined with a few other sources) provide us with a sample of users known to be involved in illegal activity. This is the starting point for our analysis, from which we apply two different empirical approaches to go from the sample to the estimated population of illegal activity.

Our first approach exploits the trade networks of users known to be involved in illegal activity ("illegal users"). We use the bitcoin blockchain to reconstruct the complete network of transactions between market participants. We then apply a type of network cluster analysis to identify two distinct communities in the data—the legal and illegal communities. Our second approach exploits certain characteristics that distinguish between legal and illegal bitcoin users, applying detection controlled estimation models (simultaneous equation models with latent variables). For example, we measure the extent to which individual bitcoin users take actions to conceal their identity and trading records, which is a predictor of involvement in illegal activity.

We find that illegal activity accounts for a substantial proportion of the users and trading activity in bitcoin. For example, approximately one-quarter of all users (25%) and close to one-half of bitcoin transactions (44%) are associated with illegal activity. Furthermore, approximately one-fifth (20%) of the total dollar value of transactions and approximately one-half of bitcoin holdings (51%) through time are associated with illegal activity. Our estimates suggest that in the most recent part of our sample (April 2017), there are an estimated 24 million bitcoin market participants that use bitcoin primarily for illegal purposes. These users annually conduct around 36 million transactions, with a value of around $72 billion, and collectively hold around $8 billion worth of bitcoin.

To give these numbers some context, a report to the US White House Office of National Drug Control Policy estimates that drug users in the United States in 2010 spend in the order of $100 billion annually on illicit drugs.[5] Using different methods, the size of the European market for illegal drugs is estimated to be at least €24 billion per year.[6] While comparisons between such estimates and ours are imprecise for a number of reasons (and the illegal activity captured by our estimates is broader than just illegal drugs), they do provide a sense that the scale of the illegal activity involving bitcoin is not only meaningful as a proportion of bitcoin activity, but also in absolute dollar terms.

The use of bitcoin in illegal trade has interesting time-series patterns. In recent years (since 2015), the proportion of bitcoin activity associated with illegal trade has declined. We attribute this trend to two main factors. The first is an increase in mainstream and speculative interest in bitcoin. For example, we find that the proportion of illegal activity in bitcoin is inversely related to the Google search intensity for

[5] The report, prepared by the RAND Corporation, estimates the user of cocaine, crack, heroin, marijuana, and methamphetamine, and is available at (www.rand.org/t/RR534). A significant share of the illegal activity involving bitcoin is likely associated with buying/selling illegal drugs online (e.g., Soska and Christin, 2015), which is what motivates the comparison with the size of the market for illegal drugs.

[6] The estimate is from the European Monitoring Centre for Drugs and Drug Addiction / Europol "EU Drug Markets Report" for the year 2013 (http://www.emcdda.europa.eu/attachements.cfm/att_194336_EN_TD3112366ENC.pdf).

the keyword "bitcoin". Furthermore, while the *proportion* of illegal bitcoin activity has declined, the *absolute amount* of such activity has continued to increase, indicating that the declining proportion is due to rapid growth in legal bitcoin use. The second factor is the emergence of alternative cryptocurrencies that are more opaque and better at concealing a user's activity (e.g., Dash, Monero, and ZCash). We find that the emergence of such alternative cryptocurrencies is also associated with a decrease in the proportion of illegal activity in bitcoin. Despite these two factors affecting the use of bitcoin in illegal activity, as well as numerous darknet marketplace seizures by law enforcement agencies, the *amount* of illegal activity involving bitcoin at the end of our sample in April 2017 remains close to its all-time high.

Bitcoin users that are involved in illegal activity differ from other users in several characteristics. Differences in transactional characteristics are generally consistent with the notion that while illegal users predominantly (or solely) use bitcoin as a payment system to facilitate trade in illegal goods/services, some legal users treat bitcoin as an investment or speculative asset. Specifically, illegal users tend to transact more, but in smaller transactions. They are also more likely to repeatedly transact with a given counterparty. Despite transacting more, illegal users tend to hold less bitcoin, consistent with them facing risks of having bitcoin holdings seized by authorities.

We find several other robust predictors of involvement in illegal activity. A user is more likely to be involved in illegal activity if they trade when there are many darknet marketplaces in operation, few shadow coins in existence, little bitcoin hype or mainstream interest, and immediately following darknet marketplaces seizures or scams. A user is also more likely to be involved in illegal activity if they use "tumbling" and/or "wash trades"—trading techniques that help conceal one's activity.

The network of bitcoin transactions between illegal users is three to four times denser, with users much more connected with one another through transactions. The higher density is consistent with illegal users transacting more and using bitcoin primarily as a payment system in buying/selling goods.

It is important to consider the differences between cryptocurrencies and cash. After all, cash is also largely anonymous (traceable only through serial numbers) and has therefore traditionally played an important role in facilitating crime and illegal trade (e.g., Rogoff, 2016). The key difference is that cryptocurrencies (similar to PayPal and credit cards) enable digital transactions and thus e-commerce. Arguably, the ability to make digital payments revolutionized retail and wholesale trade. Online shopping substantially impacted the structure of retailing, consumption patterns, choice and hence welfare, marketing, competition, and ultimately supply and demand. Until cryptocurrencies, such impacts were largely limited to legal goods and services due to the traceability of digital payments. Cryptocurrencies have changed this, by combining the anonymity of cash with digitization, which enables efficient anonymous online and cross-border commerce. Cryptocurrencies therefore have the potential to cause an important structural shift in how the black market operates.

While the emergence of illegal darknet marketplaces illustrates that this shift has commenced, it is not obvious to what extent the black market will adopt the opportunities for e-commerce and digital payments via cryptocurrencies—this is an important empirical question. Our findings illustrate the dynamics of this adoption process and suggest that eight years after the introduction of the first cryptocurrency, the black market has indeed adopted this form of electronic payment on a meaningful scale. Thus, our results suggest that cryptocurrencies are having a material impact on the way the black market for illegal goods and services operates.

Our findings have a number of further implications, which we discuss in Section 6. Blockchain technology and the systems/protocols that can be implemented on a blockchain have the potential for revolutionizing numerous industries. In shedding light on the dark side of cryptocurrencies, we hope this research will reduce some of the regulatory uncertainty about the negative consequences and risks of this innovation, facilitating more informed policy decisions that assess both the costs and benefits. In turn, we hope this contributes to these technologies reaching their potential. Second, our paper contributes to understanding the intrinsic value of bitcoin, highlighting that a significant component of its value as a payment system derives from its use in facilitating illegal trade. This has ethical implications for bitcoin as an investment, as well as valuation implications. Third, our paper moves the literature closer to understanding the welfare consequences of the growth in illegal online trade. A crucial piece of this puzzle is understanding the extent to which illegal online trade simply reflects a migration of activity that would have otherwise occurred on the street, versus the alternative that by making illegal goods more accessible, convenient to buy, and less risky to buy due to anonymity, "black e-commerce" could lead to growth in the aggregate black market. Our estimates contribute to understanding this issue, but further research is required to relate these estimates to trends in the offline black market to further our understanding of the welfare consequences.

This paper also makes a methodological contribution. The techniques developed in this paper can be used in cryptocurrency surveillance in a number of ways, including monitoring trends in illegal activity, its response to regulatory interventions, and how its characteristics change through time. The methods can also be used to identify key bitcoin users (e.g., "hubs" in the illegal trade network) which, when combined with other sources of information, can be linked to specific individuals. The techniques in this paper can also be used to study other types of activity in bitcoin or other cryptocurrencies / blockchains.

Our paper contributes to a few areas of recent literature, which we discuss in more detail in Section 6. We add to the literature on the economics of cryptocurrencies and applications of blockchain technology to securities markets by showing that one of the major uses of cryptocurrencies as a payment

system is in settings where anonymity is valued (e.g., illegal trade).[7] Our paper also contributes to the computer science literature that analyzes the degree of anonymity in bitcoin by developing algorithms that identify entities/users/activities in bitcoin's blockchain.[8] We exploit algorithms from this literature to identify individual users in the data, and we add new methods to the literature that go beyond observing individuals, to identification of communities and estimation of populations of users. Finally, our paper is also related to studies of darknet marketplaces and the online drug trade, including papers from computer science and drug policy.[9] We contribute to this literature by quantifying the amount of illegal activity that involves bitcoin, rather than studying a single market (e.g., Silk Road) or indirect lower-bound measures of darknet activity such as the feedback left by buyers. Empirically, we confirm that the estimated population of illegal activity is several times larger than what can be "observed" through studying observable darknet marketplaces and their customers.

The next section provides institutional details about bitcoin and the blockchain, darknet marketplaces in which illegal goods and services are bought/sold using bitcoin, and law enforcement efforts to monitor and disrupt illegal online activity. Section 3 describes the blockchain data used in this paper. Section 4 explains three approaches that we use to construct a sample of illegal activity and characterizes that sample. The sample forms the input to our empirical methods in Section 5 that quantify the total amount of illegal activity, its trends, and its characteristics. A discussion of the implications of the results and how they relate to existing studies is in Section 6, while Section 7 concludes.

2. Institutional details

2.1. The structure of the bitcoin blockchain

Bitcoin is an international currency, not associated with any country or central bank, backed only by its limited total supply and the willingness of bitcoin users to recognize its value.[10] Bitcoins are "mined" (created) by solving cryptographic puzzles that deterministically increase in difficulty and once solved can be easily verified. Each solution results in a new "block" and provides the miner with the "block reward" (currently 12.5 bitcoins), which incentivizes the miner. The difficulty of the cryptographic puzzles is adjusted after every 2,016 blocks (approximately 14 days) by an amount that ensures the average time between blocks remains ten minutes.

[7] See: Malinova and Park, 2016; Khapko and Zoican, 2016; Yermack, 2017; Huberman et al., 2017; Easley et al., 2017.
[8] See: Meiklejohn et al., 2013; Ron and Shamir, 2013; Androulaki et al., 2013; Tasca et al., 2016.
[9] See: Soska and Christin, 2015; Barratt et al., 2016a; Aldridge and Décary-Hétu, 2016; Van Buskirk et al., 2016.
[10] As of January 2017, over 16 million bitcoins had been mined out of a maximum of 21 million. This maximum limit is built into the protocol (Nakamoto, 2008).

Each block, as well as expanding the supply of bitcoin, confirms a collection of recent transactions (transactions since the last block). Each block also contains a reference to the last block, thereby forming a "chain", giving rise to the term "blockchain". The blockchain thus forms a complete and sequential record of all transactions and is publically available to any participant in the network.

Bitcoins are divisible to the "Satoshi", being one hundred millionth of one bitcoin (currently worth less than two hundredths of a cent). Each bitcoin holding (or parcel) is identified by an address, analogous to the serial number of a banknote. Unlike banknotes, bitcoin does not have to be held in round units (e.g., 5, 10, 50). Unless a holding of bitcoin with a given address is exactly spent in a transaction, the "change" from the transaction is returned to a new address forming a new parcel of bitcoin.

A bitcoin "user" (a participant in the network) stores the addresses associated with each parcel of bitcoin that they own in a "wallet". Similar to a conventional cash wallet, a bitcoin wallet balance is the sum of the balances of all the addresses inside the wallet. While individual bitcoin addresses are designed to be anonymous, it is possible to link addresses belonging to the same wallet when more than one address is used to make a purchase.

2.2. Darknet marketplaces and their microstructure

The "darknet" is a network like the internet, but that can only be accessed through particular communications protocols that provide greater anonymity than the internet. The darknet contains online marketplaces, much like EBay, but with anonymous communications, which also makes these marketplaces less accessible than online stores on the internet. Darknet marketplaces are particularly popular for trading illegal goods and services because the identities of buyers and sellers are concealed. The darknet is estimated to contain approximately 30,000 domains (Lewman, 2016).

To access a darknet marketplace, a user is generally required to establish an account (usually free) at the marketplace in order to browse vendor products (Martin, 2014a; Van Slobbe, 2016). Similar to the way PayPal propelled EBay, the secure, decentralized, and anonymous nature of cryptocurrencies has played an important role in the success of darknet marketplaces. While bitcoin is the most widespread cryptocurrency used in such marketplaces, other currencies have occasionally been adopted, either due to their popularity (such as *Ethereum*) or improved anonymity (such as *Monero*). Despite the availability of alternate currencies on some marketplaces, the vast majority of transactions on the darknet are still undertaken in bitcoin.[11]

A user that wants to buy goods or services on a darknet marketplace must first acquire cryptocurrency (typically from an online exchange or broker) and then deposit this in an address

[11] A recent estimate from a darknet marketplace operator identified bitcoin as accounting for 98% of transactions: https://www.wired.com/2017/01/monero-drug-dealers-cryptocurrency-choice-fire/.

belonging to the darknet marketplace (often termed a "hot wallet"). These funds are held in "escrow" by the marketplace. Vendor prices on darknet markets are often quoted inclusive of a marketplace fee. The escrow system also assists marketplace administrators in mediating disputes between buyers and sellers and minimizing scams in which money is collected without the intention of ever shipping any goods (Aldridge and Décary-Hétu, 2014; Christin, 2013). Funds are released when the vendor indicates the goods have been sent. In some marketplaces, the funds are held until the buyer indicates that the goods have been received. The escrow function of the darknet marketplaces sometimes leads to "exit scams", whereby a marketplace ceases operations but does not return bitcoin held in escrow. Many such scams have been perpetrated by marketplaces in the last five years, including *Sheep Marketplace* (2013), *Pirate Market* (2014), *Evolution* (2015), and *Nucleus* (2016).

The evolution of dark marketplaces allows sellers of illegal goods and services to reach global audiences (Van Buskirk et al., 2016). This internationalization of illegal trade necessitates more complex methods of communications and logistics to avoid detection. To this end, buyers placing an order with an online seller typically communicate using PGP (Pretty Good Privacy) encryption, which encodes and decodes messages using a pair of public and private keys (Cox, 2016). On some (typically more recent) marketplaces, this functionality is built into the site. Logistically, items are typically delivered by mail and the process by which this occurs has been widely documented (Christin, 2013; Van Hout and Bingham, 2013; Lavorgna, 2016; Van Slobbe, 2016). Many methods are used to minimize the chance of such deliveries being intercepted by law enforcement, including professional logos, vacuum sealed bags, posting small quantities of product, and including a (fake) return address (Christin, 2013; Basu, 2014; Tzanetakis et al., 2016). Customers are advised by marketplaces to avoid using their real name or address to minimize the risk of being caught by law enforcement agencies (Martin, 2014b).

After receiving their goods, buyers are encouraged to leave feedback about the seller, commenting on the arrival (or otherwise) of the goods, their quality, and overall service (Van Slobbe, 2016). Such feedback is paramount for developing a reputation in a marketplace that is primarily based on trust between participants, with few ramifications for "scamming" purchasers (Aldridge and Décary-Hétu, 2014; Tzanetakis et al., 2016).

To get a sense of how a buyer navigates a darknet marketplace, Figure 1 provides screenshots from one of the first darknet marketplaces, "Silk Road". Panel A provides an example of the "Drugs" page illustrating that a wide variety of illegal drugs, weapons, and forgeries can be purchased using bitcoin. Panel B provides an example of information about individual items and sellers. Clicking on the appropriate headings, one can obtain further information about the items (detailed description, insurance/refund policies, available postage methods and locations, security and encryption, and so on) and about the seller (their rating from buyers, detailed feedback from buyers, history of sales, and so on).

Panel C shows the interface for depositing bitcoin to Silk Road's escrow account, how to transfer bitcoins to a given seller, and how to withdraw bitcoins from escrow.

< Figure 1 >

By providing an anonymous, digital method of payment, bitcoin did for darknet marketplaces what PayPal did for EBay—provide a reliable, scalable, and convenient payment mechanism. What was also required was an anonymous way of hosting and accessing those illegal marketplaces. This issue is solved through the use of The Onion Router (TOR), originally developed by the US Navy. By routing the message through several nodes in the TOR network, TOR obfuscates the path (and hence the IP address) of a message sent between two clients.

The combination of TOR for covert communications and bitcoin for covert payments has led to the proliferation of darknet marketplaces. The most well-known marketplace was the "Silk Road" started in 2011. Since its shutdown by the FBI in 2013, numerous other marketplaces have sprung up (see Table A2 in Appendix A for a list). Despite frequent shutdowns, seizures and scams, measures of darknet marketplace activity indicate steady growth in the number of market participants and products (Matthews et al., 2017). For example, one of the largest marketplaces in 2017, "AlphaBay", had over 350,000 items available for sale in categories such as drugs, weapons, malware, and illegal pornography.

2.3. Surveillance and cryptocurrency seizures from darknet marketplaces

Cryptocurrencies have proven effective not only in facilitating illegal trade, but also in the detection of illegal activity due to the public nature of the blockchain. Even though bitcoin has been used extensively in illegal activity, some argue that the blockchain actually makes it easier for law enforcement to detect illegal activity, despite the currency's anonymity. Koshy, Koshy, and McDaniel (2014) show that by monitoring transactions transmitted from computers to the blockchain, they are able to link individual transactions to the IP address of the sender. Meiklejohn et al. (2013) describe how tracing a bitcoin theft on the blockchain to bitcoin exchanges could be used by authorities with subpoena powers to potentially identify perpetrators. Yermack (2017) hypothesizes that the growing popularity of bitcoin will inevitably lead to a growing market for de-anonymizing technologies, leading to increased transparency of the users making transactions on the blockchain. In response to these pressures, supporters of the anonymity provided by cryptocurrencies are actively developing new currencies that challenge law enforcement's detection methods. Such currencies include *Monero*, which hides user's public keys among a group of public keys that contain the same amount (known as "Ring Signatures"), and *ZCash* (launched

in 2016), which uses zero-knowledge proofs that hide sender, recipient, and transaction amount (Noether, 2015; Ben-Sasson et al., 2014).

Recently, law enforcement agencies have been successful in seizing bitcoin from a number of darknet marketplaces. For example, the Silk Road marketplace was raided by the FBI on October 2, 2013, seizing bitcoin from customer and supplier escrow accounts (hot wallets) and from the owner/operator, Ross William Ulbricht. After the closure of the Silk Road, law enforcement agencies successfully seized bitcoin from several other illegal sites/individuals (see Table A1 of Appendix A). Numerous darknet sites were raided and shut down in "Operation Onymous"; an international collaboration between US and European law enforcement agencies that targeted illegal darknet sites. Despite the seizures, illegal darknet marketplaces continue to operate, with many new ones being created since the seizures.

The seized bitcoin from these operations allows us to identify bitcoin users (customers, suppliers, and marketplace operators) involved in illegal activity. These observations provide a starting point from which to estimate the extent of illegal activity involving bitcoin.

Law enforcement agencies use a number of strategies to detect illegal activity on the darknet, ranging from cyber-surveillance to forensic analysis. Given that detected illegal activity feeds into our identification techniques, it is important to understand law enforcement strategies. Christin (2013) and Kruithof et al. (2016) describe a number of such strategies, including: infiltrating the TOR network to determine individual IP addresses, decoding the financial infrastructure of bitcoin to identify individuals, and using traditional forensic and investigative techniques on seized packages. Law enforcement agencies monitor suspicious packages passing through the postal service. Agencies also order drugs on darknet marketplaces to investigate the return address on the package. For example, an unusual amount of outgoing mail from a large Australian drug dealer led authorities to seize over 24,000 in bitcoin, along with a wide array of drugs and cash. Investigators also sometimes pose as suppliers to gather addresses of customers and reveal their identities. Finally, by conducting major seizures, agencies can create distrust in the online trade of illegal drugs among participants (Van Slobbe, 2016; Christin, 2013). Large-scale initiatives such as "Operation Onymous", in which law enforcement agencies shut down several illegal marketplaces and made 17 arrests across 17 countries, can discourage illegal online activity by increasing the risk of detection (Franklin, Paxson, Perrig, and Savage, 2007).

3. Data and descriptive statistics

We extract the complete record of bitcoin transactions from the public bitcoin blockchain, from the first block on January 3, 2009, to the end of April 2017. For each transaction, we collect the transaction ID, sender and recipient address, timestamp, block ID, transaction fee, and transaction amount.

3.1 Identifying users in transaction-level bitcoin data

The data that make up the bitcoin blockchain reveal "addresses" (identifiers for parcels of bitcoin) but not the "users" (individuals) that control those addresses. A user typically controls several addresses. This one-to-many mapping occurs partly as a result of various activities that users employ to preserve their anonymity and partly due to transaction mechanics (e.g., when a user receives "change" in a transaction, the change is given a new address).[12] We find addresses connected to a single user with the Union-Find algorithm, which is developed by Cormen, Leiserson, Rivest, and Stein (2001) and Ron and Shamir (2013) and used in several related papers such as Meiklejohn et al. (2013). This algorithm transforms the transaction-level data into user-level data, linking each transaction to the associated users.

The following illustrates how the Union-Find algorithm works. A transaction usually involves several addresses from one user. For example, the payer ("sender") of bitcoin might send bitcoin from multiple addresses and also receive change to a new address. Because a user must control the private key of each address from which bitcoin is sent in a given transaction, all of the sender's addresses in one transaction are almost certainly associated with one user. Transitivity is then used to link the addresses of a user across multiple transactions. For example, suppose two separate transactions are observed; one in which bitcoin is sent from addresses A and B and another in which bitcoin is sent from addresses B and C. The first transaction identifies addresses A and B as belonging to the one user, while the second identifies B and C as belonging to the same user. By transitivity, all three addresses (A, B, and C) belong to the same user.

None of the existing algorithms that cluster bitcoin addresses by user has perfect accuracy.[13] The Union-Find algorithm is the most widely used approach, primarily because the errors it makes (too little clustering of addresses rather than too much clustering) are conservative in most applications (Meiklejohn et al, 2013). The Union-Find algorithm might fail to cluster together two sets of addresses controlled by the one user if the user never makes a transaction that uses an address from each set. In such instances, two or more address clusters might in fact correspond to one user.[14] In contrast, the Union-Find algorithm (unlike other approaches such as those that exploit the change from transactions) is very unlikely to make the opposite and more severe error of incorrectly clustering together sets of addresses that involve *more* than one user. In our application, too little clustering (and thus having instances where two or more clusters correspond to one actual user) is unlikely to have severe consequences for our empirical methods,

[12] For example, individuals can send bitcoin to a "tumbling" service which then returns the bitcoin (minus a fee) to a new address, or by sending bitcoin to oneself using a newly generated address as the recipient of the transaction (Ron and Shamir, 2013).

[13] For example, Androulaki et al. (2013) examine two approaches using simulations and find that many, but not all, of the users can be correctly identified by clustering algorithms even when users try to enhance their privacy by creating new addresses.

[14] Meiklejohn et al. (2013) empirically find that this error is "not too common" in bitcoin blockchain analysis.

whereas incorrectly joining multiple users into a single cluster would be far more problematic.[15] Therefore, the Union-Find algorithm is a suitable choice given our requirements.

3.2 Filters

In this study, we are primarily interested in quantifying the amount of illegal trade that uses bitcoin. Currency conversion transactions (between bitcoin and fiat currency or other cryptocurrencies), which are mainly done via bitcoin exchanges, are also recorded on the bitcoin blockchain but do not involve trade in the sense of buying or selling goods or services. In our baseline analysis, we therefore remove bitcoin exchanges (and their transactions) from the data to avoid inflating activity with currency conversion transactions. We also remove the major known bitcoin "miners" and their transactions as their role in the network is one of providing transaction confirmations, i.e., the infrastructure of the bitcoin network. They receive block creation rewards and fees in the process of providing transaction confirmation services.[16] The exchanges and miners are identified via "Wallet Explorer".[17]

We also exclude transactions that have a value of less than $1 on the day of the transaction.[18] Such transactions reflect negligible transfers of value and are therefore used for purposes such as messages, test transactions, and tips. Failure to exclude these transactions could significantly skew our data, particularly measures of the proportion of transactions. Other than these exclusions, we include all other bitcoin users and transaction activity on the bitcoin blockchain.

3.3 Descriptive statistics of user-level variables

Our sample has a total of approximately 106 million bitcoin users, who collectively conduct approximately 606 million transactions, transferring around $1.9 trillion.[19] For each user, we calculate a collection of variables that characterize features of their bitcoin transaction activity (e.g., transaction count, transaction size, transaction frequency, and number of counterparties). We also calculate a range of user-level variables that are more specific indicators of the nature of the activity in which a user is likely to be engaged, such as the number of illegal darknet marketplaces that operate at the time the user transacts, the extent to which the user engages in transactions designed to conceal their activity, and the

[15] For example, if a single actual user appears in the data as two or more clusters, all of those clusters could be correctly classified with the user's actual type (illegal or legal), whereas if a legal and illegal user are incorrectly clustered together, there is no way to assign a correct classification to the cluster.
[16] We remove 83 exchanges and 28 miners, collectively accounting for 15.3% of the total number of transactions.
[17] Wallet Explorer joins transactions into "wallets" (the equivalent of our "users") using a similar procedure to the one described above and then classifies a large number of wallets by type either on the basis of (i) having observed an address being advertised as part of a given entity (e.g., a known address from a bitcoin exchange), or (ii) having identified an entity's wallet by sending a small amount of bitcoin to the entity, where that address is linked to the larger wallet of the entity (similar to Meiklejohn et al., 2013). Data available from https://www.walletexplorer.com.
[18] These small transactions represent 23.9% of all transactions, but less than 0.0001% of total bitcoin volume.
[19] Exact numbers are in Table 3.

degree of interest in bitcoin at the time the user transacts (using Google search intensity). The detailed definitions of these variables are reported in Table 1.

< Table 1 >

Table 2 reports descriptive statistics about the user-level variables. Focusing on the variables that characterize a user's bitcoin transaction activity (Panel A), we see that a typical (median) user engages in three bitcoin transactions (mean *Transaction Count* is 5.7 transactions) with three different counterparties (mean of *Counterparties* is 4.2). Thus, a typical user has a low degree of concentration in counterparties, in that they do not repeatedly transact with the same counterparty (our measure of *Concentration*, which is a normalized Herfindahl–Hirschman Index, has a median of zero). There are a small number of highly active entities, with the most active having 11.4 million transactions and 4.4 million counterparties.

The average transaction size is around $5,000, but a typical transaction (the median *Transaction Size*) is much smaller at $112. Some transactions are very large, with the largest exceeding $90 million. For most users, their first and last bitcoin transaction occurs within the same month (the median *Existence Time* is one month), although some users are present for many years (the maximum *Existence Time* is 101 months, or just over eight years).

The other variables (Panel B) are more specific indicators of the nature of the activity in which a user is likely to be engaged and are thus important in our empirical models. We therefore define and discuss these variables when we turn to the empirical models.

< Table 2 >

4. Identifying a sample of illegal users

We identify a sample of addresses (and therefore users) involved in illegal activity using three approaches described below.

4.1. First approach: Bitcoin seizures by law enforcement agencies

Our first approach exploits bitcoin seizures by law enforcement agencies such as the US FBI. We manually identify bitcoin seizures from news articles (via searches using Factiva) and US court records (via searches of the digital PACER records). Table A1 in Appendix A reports the list of seizures that we use. For each seizure, we extract information from court records and law enforcement agency disclosures about any identified bitcoin addresses or transactions (amounts and dates). From these details we uniquely

identify the users involved in the illegal activity, by matching up the bitcoin address or transaction identifier with our user-level data constructed from the bitcoin blockchain.

In some cases (e.g., the US FBI's seizure of Silk Road and Ross Ulbricht's holdings, and the Australian law enforcement's seizure of Richard Pollard's holdings) the law enforcement agency auctioned the seized bitcoin to the public. Given the public nature of the auctions, we are able to identify the auction transactions on the bitcoin blockchain and work backwards to identify the seized bitcoin addresses, which in turn identify those individuals that were involved in illegal activity and had some or all of their bitcoin holdings seized by law enforcement agencies. Using this approach we are able to identify 1,016 known illegal users, which we refer to as "*Seized Users*".

4.2. Second approach: Illegal darknet marketplaces and their users

Our second approach exploits the known "hot wallets" of major illegal darknet marketplaces. These are central accounts, many of which operate like escrow accounts, into which users of darknet marketplaces deposit or withdraw funds. We are able to identify 17 such marketplaces using data from the Wallet Explorer service, which in turn identifies these marketplaces using an approach similar to Meiklejohn et al. (2013), i.e., on the basis of small "probing" transactions undertaken with a given entity.

From these hot wallets, we identify slightly over 6 million darknet marketplace users as individuals that send to and/or receive bitcoin from a known darknet marketplace. We refer to the darknet marketplace hot wallets and their contributors/recipients as "*Black Market Users*".

An underlying assumption is that the trade that occurs in darknet marketplaces is illegal. This assumption is supported by ample anecdotal evidence, objective empirical evidence in the form of darknet market scrapes that show the goods and services traded there (e.g., Christin, 2013; Aldridge Décary-Hétu, 2014; Van Buskirk et al., 2014; Soska and Christin, 2015), as well as actions by law enforcement agencies, including indiscriminate seizures of *all* bitcoin from such markets.

4.3. Third approach: Users identified in darknet forums

Our third approach exploits information contained in the darknet, in particular the bitcoin addresses of users identified in darknet forums as selling goods/services. We use systematic scrapes of darknet forums from 2013 to 2017.[20] This allows us to identify users that might never have been caught by authorities and might not be otherwise identified in the data through transactions with known darknet marketplaces. Users often post bitcoin addresses in cases such as fraud (they did not receive their goods), quality checking, and for the purposes of advertising the address to which funds should be sent, including

[20] A list of known darknet markets is in Table A2 of Appendix. An archive of darknet forums during 2013-2015 is available at https://www.gwern.net/index. We scrape information from active darknet sites during 2016-2017.

in privately negotiated trade. While other studies have also scraped darknet marketplaces for certain types of information (e.g., Soska and Christin, 2015; Van Buskirk et al., 2016), as far as we know no other study has used scrapes to identify the bitcoin addresses of illegal users.

Using this approach, we identify an additional 448 users that were not already identified in either of the previous two approaches. We refer to these as "*Forum Users*".

4.4. The sample of illegal users

Table 3 shows the number of illegal users identified using the three approaches above and various measures of their activity.[21] Together, there are 6,223,337 "observed" illegal users, representing 5.86% of all bitcoin participants. They account for an even larger share of transactions—a total of 196 million transactions, or around one-third of all transactions (32.38%). They also account for an even larger share of bitcoin holdings—throughout the sample period, the average dollar value of the bitcoin holdings of observed illegal users is around $1.3 billion, which is close to half (45.28%) of the average dollar value of holdings for all users.[22] Observed illegal users control around one-quarter (26.33%) of all bitcoin addresses, and the dollar value of their transactions is approximately 12.96% of the total dollar value of bitcoin transactions.

Within the three subgroups of illegal users, the largest group in terms of number of users is the "*Black market users*", followed by "*Seized users*" and then "*Forum users*". *Seized users* and *Forum users* are nevertheless meaningful subgroups, for example, they account for 3.93% and 2.47% of all transactions, respectively.

< Table 3 >

The results in Table 3 indicate that the sample of "observed" illegal users is already a substantial proportion of users and bitcoin transaction activity, without yet having applied methods to estimate the population of illegal users/activity. Capturing a relatively large sample of illegal activity is important because it provides rich information to our empirical methods that estimate the totality of illegal activity. The fact that the sample of illegal activity is drawn from three different approaches is also likely to help the subsequent empirical models by providing a more diverse sample.

[21] Given a transaction has two sides (a sender and a receiver) and it is possible for the different sides to be users from different groups, throughout the paper we (double) count the number of transactions and volume by considering each transaction from the perspective of the sender and receiver.

[22] The average holdings numbers are considerably lower than current holdings because for the first few years of bitcoin's existence, its market capitalization was much lower than it is currently.

Finally, given the nature of illegal activity could change through time, it is also important that our sample of observed illegal users spans different time periods and is not completely concentrated at one point in time. Figure 2 indicates that this is the case for our sample of observed illegal users and their activity. Figure 2 plots the time-series of the observed illegal users and their activity as a percentage of: total users (Panel A), total number of transaction (Panel B), the dollar value of all transactions (Panel C), and the dollar value of all bitcoin holdings (Panel D).

These time-series show that the observed illegal users are present during all points in time throughout our sample period. Their share of activity is highest at the start of the sample in 2009, and then again during a period from 2012 to the end of 2015. The first of these periods (the year 2009) is not particularly economically meaningful as the first year or two of bitcoin's existence involves a very small number of users and transactions compared to subsequent years. In contrast, the activity in the second period, 2012-2015, is meaningful. This period corresponds to the time when illegal darknet marketplaces grew rapidly in number and popularity. Silk Road 1 was established in January 2011 and soon became a popular venue in which to buy and sell illegal goods and services (e.g., Soska and Christin, 2015). After Silk Road 1 was shut down by the US FBI in October of 2013, a large number of other illegal darknet marketplaces commenced operating throughout 2013-2015 (see Table A2 of Appendix A). Thus, perhaps somewhat unsurprisingly, the peak activity of our sample of observed illegal users coincides with substantial darknet marketplace activity. However, we also observe a reasonable number of illegal users and illegal activity outside of this peak window.

< Figure 2 >

5. Quantifying and characterizing all illegal activity

Having identified a substantial sample of bitcoin users that are involved in illegal activity, our next step is to use the information in this sample to estimate the totality of illegal activity that uses bitcoin. We use two different methods to classify users into those that are primarily involved in illegal activity ("illegal users") and those that are primarily involved in legal activity ("legal users"). Subsequently, we measure the size and activity of the two groups.

At an intuitive level, the first method exploits the network topology—the information about who trades with whom. Trade networks reveal "communities" of users and can thereby identify other illegal users that were not part of our initial sample. In contrast, the second method exploits characteristics that distinguish illegal users from legal users (controlling for non-random detection). Both methods allow a user that was initially classified as an "observed" illegal user to be reclassified as a user that is predominantly engaged in legal activity (a "legal user"). This feature of the methods allows for the

possibility that some of the users identified in the previous stage as having engaged in illegal activity actually engaged in more legal activity than illegal activity.

The two methods provide independent estimates of the illegal activity and its characteristics. Given that the methods rely on completely different assumptions and exploit different information, their concurrent use provides robustness and the ability to cross-validate results. The methods are described below in separate subsections. We then report the results of how many users and how much trade is estimated to be associated with illegal activity, after which we characterize the nature of the illegal users and their trading activity compared to legal users.

5.1. Method 1: Network cluster analysis

The first method exploits network topology to identify "communities" of users based on the transactions between users. In simple terms, the method works as follows. If users A, B, and C are known to be involved in illegal activity (e.g., their bitcoin was seized by law enforcement agencies), a user X that trades exclusively or predominantly with users A, B, or C is likely to also be involved in illegal activity. Similarly, a user Y that trades predominantly with users that are not identified as illegal is likely to be a legal user. This intuition drives the classification of users into legal and illegal on the basis of their transaction partners.

More formally, the method we apply is a network cluster analysis algorithm that takes as inputs the set of users ("nodes" in network terminology) and the trades between users ("edges" or "links" in network terminology). The output of the algorithm is an assignment of users to communities such that the "modularity" of the communities (density of links within communities and sparsity of links between communities) is maximized. The method labels a user as illegal (legal) if the disproportionate share of their transactions is with members of the illegal (legal) community. The method does not assume that users only engage in either legal or illegal activity—users can do both. Therefore, there will be some trades between the legal and illegal communities.

We apply a variant of the Smart Local Moving (SLM) algorithm developed by Waltman and van Eck (2013), adapted to our specific application. The algorithm's name ("smart moving") comes from the fact that the algorithm finds the underlying community structure in the network by moving nodes from one community to another, if such a move improves the model fit. The SLM algorithm is among the leading network cluster analysis algorithms.[23]

Applied to our data, the algorithm is as follows.

[23] For example, Emmons et al. (2016) in their comparison of multiple methods find that the SLM algorithm performs the best in terms of maximizing cluster quality metrics.

- Step 1: Assign all the observed illegal users to the illegal community and all of the remaining users to the legal community.
- Step 2: Loop through each user, performing the following action on each:
    - If the user disproportionately transacts with members of the user's currently assigned community, then leave the user in that community[24];
    - Otherwise, move the user to the other community (if the user is assigned to the illegal community, move the user to legal community, and vice versa).
- Step 3: Repeat Step 2 until, in a complete loop through all users, no user switches between communities. At that point the assignment to communities is stable and ensures that each member trades disproportionately with other members of the same community.

Note that due to the iterative moving in the algorithm, not all of the "observed" illegal users will necessarily remain in the illegal community. For example, it is possible that some of the users that had bitcoin seized by authorities were involved in some illegal activity (hence getting bitcoin seized) but were mainly using bitcoin for legal purposes. This will be recognized by the algorithm in Step 2 and the user will be moved to the legal community.

5.2. Method 2: Detection controlled estimation (DCE)

The second method we use to estimate the population of users involved in illegal activity ("illegal users") is detection controlled estimation (DCE). Intuitively, this method exploits the differences in the characteristics of legal and illegal users of bitcoin to probabilistically identify the population of illegal users. If we had a random sample of illegal users and a set of characteristics that differ between legal and illegal users (e.g., measures of the extent to which a user has employed tools to conceal their activity), this task would be relatively simple and could be achieved with standard techniques (regression, discriminant analysis, and so on). A complication is that detection (as in most settings where violators attempt to conceal their illegal activity from authorities) is not random, and this non-randomness must be accounted for to obtain unbiased estimators.[25] We use "detection" in the broad sense of an illegal user having been identified by any of the three approaches described in the previous section (had bitcoin seized by a law enforcement agency, was identified in darknet forums, or was observed in the blockchain data as having

[24] "Disproportionately" is if the proportion of transactions the user makes with other members of the same community is greater than or equal to the community's proportion of total transactions. In robustness tests we consider the proportion of volume rather than transactions and find consistent results.

[25] A further complication is that the determinants of this non-randomness are not separately observed (unlike, for example, non-respondents in a survey, or people that choose not to participate in the labor force) and therefore the classic tools to deal with sample selection bias (e.g., Heckman models) cannot be applied.

transacted with a known illegal darknet marketplace). Thus, "detected" illegal users are the observed illegal users described in Section 4.

Fortunately this econometric challenge is not unique to illegal activity in bitcoin and methods to overcome it exist. The same challenge occurs in quantifying other forms of misconduct such as tax evasion, fraud, insider trading, and market manipulation, as well as contexts such as nuclear power plant safety regulation breaches, cancer detection by mammograms, and so on. The standard tool for these settings is DCE. Since its development by Feinstein (1989, 1990), DCE models have been applied to various financial misconduct settings including tax evasion (Feinstein, 1991), corporate fraud (Wang et al., 2010), and market manipulation (Comerton-Forde and Putniņš, 2014). By explicitly modelling both underlying processes (violation and detection) simultaneously, one can obtain unbiased estimates of the illegal activity, which is otherwise only partially observed.

< Figure 3 here >

Figure 3 illustrates the two-stage DCE model that we estimate. On the left is the starting point, the data, which in our case is the set of all bitcoin users. In the middle we have the two processes, violation (undertaking illegal activity) and detection (e.g., bitcoin seizures). On the right-hand side are the joint outcomes of those processes: the observable classifications of users into detected illegal users (the set A) and other users (the complement set A^C, comprising legal users and undetected illegal users).

The first branch models whether a bitcoin user, i, is predominantly involved in illegal or legal activity. This branch is modelled as an unobservable binary process (L_{1i}) driven by a continuous latent function (Y_{1i}) of a vector of characteristics, x_{1i}, that can distinguish between legal and illegal users:

$$Y_{1i} = \beta_1 x_{1i} + \epsilon_{1i} \tag{1}$$

$$L_{1i} = \begin{cases} 1 & (Illegal\ user) \\ 0 & (Legal\ user) \end{cases} \text{ if } \begin{matrix} Y_{1i} > 0 \\ Y_{1i} \leq 0 \end{matrix} \tag{2}$$

The second branch models whether or not an illegal user is "detected" (they enter our sample of observed illegal users). This detection process is modelled as another unobservable binary process (L_{2i}) driven by a different continuous latent function (Y_{2i}) of a vector of characteristics, x_{2i}, that affect the probability that an illegal user is detected:

$$Y_{2i} = x_{2i}\beta_2 + \epsilon_{2i} \tag{3}$$

$$L_{2i} = \begin{cases} 1 & (Detected) \\ 0 & (Not\ detected) \end{cases} if \begin{matrix} Y_{2i} > 0 \\ Y_{2i} \leq 0 \end{matrix} \tag{4}$$

Both stages of the model are estimated simultaneously using maximum likelihood. The likelihood function for the model is derived in Appendix B. Intuitively, this process finds estimates for the vectors of

model parameters, β_1 and β_2, that maximize the likelihood of the observed data (the classification of users into sets A and A^C). From the estimates of β_1 and β_2, we compute each user's probability of being involved in illegal activity and construct a binary classification of legal and illegal users.

Similar to the SLM approach, the DCE model does not assume that detected illegal users were engaged solely or predominantly in illegal activity. Once the DCE model is estimated, the classification of users into legal and illegal categories can result in some detected illegal users being re-classified as predominantly legal users.[26]

Similar to Heckman models, identification in a DCE model without instruments is possible, relying on functional form and distributional assumptions. However, more robust identification is achieved through instrumental variables that affect one process but not the other. We take the more robust route of using instrumental variables. The next subsection describes the instrumental variables and their descriptive statistics.

5.3. Variables used in the DCE model and their descriptive statistics

One of the instrumental variables associated with illegal activity is the extent to which the user employs methods to conceal their identity or obfuscate their transaction history. For example, to partially conceal their identities from an observer of the bitcoin blockchain, users can use "tumbling" and "wash trades" to alter the addresses of their bitcoin holdings, increasing the difficulty of tracing their activity. Tumbling, in its simplest form, involves a user sending bitcoin to a tumbling provider who (in return for a small fee) returns the balance to a different address controlled by the user. Wash trades involve a user sending bitcoin from one address to another (new) address that they also control. Legal users have little reason to take such actions to conceal their actions (and incur associated costs). In contrast, users involved in illegal activity are likely to use these concealment techniques. As such, the use of tumbling services and wash trades is likely to be a predictor of whether a user is involved in illegal activity. Importantly (for this to be an instrumental variable), using wash trades and tumbling does not alter the probability of "detection" by law enforcement agencies via the seizures of bitcoin from darknet sites. The seizures confiscated all bitcoin held in darknet marketplace escrow accounts ("hot wallets") irrespective of whether the user employed tumbling or wash trades. For each user, we measure the percentage of their transactions that are tumbling or wash trades and call this variable *Tumbling*.

[26] For example, suppose a user was involved in some illegal activity and had bitcoin seized by authorities but was mainly using bitcoin for legal purposes. Such a user will have characteristics that are similar to those of legal users and not very similar to illegal users, which would lead to a classification by the DCE model into the legal user category. In contrast, a predominantly illegal user, even if not detected or observed, is likely to have characteristics similar to other illegal users and therefore (after controlling for the differences in characteristics due to non-random detection) the user is likely to be classified as illegal by the DCE model.

Another set of instruments for the likelihood that a user is involved in illegal activity involves time-series variables that are likely to correlate with the type of activity in which bitcoin users are engaged. For example, for each user we construct a measure of the average number of operational illegal darknet marketplaces at the time the user transacts (we label the variable *Darknet Sites*). All else equal, illegal transactions (and thus users involved in illegal activity) are more likely when there is a lot of illegal darknet marketplace activity than when there is little or no illegal darknet activity.

In a similar spirit, we construct a measure of the average number of opaque cryptocurrencies in existence (Dash, Monero, and ZCash) at the time the user participates in bitcoin (labelled *Shadow Coins*). These major alternative "shadow coins" were developed, at least in part, to provide more privacy than bitcoin. If some of the online black market starts using these shadow coins instead of bitcoin, the number of such coins in existence at the time a user transacts in bitcoin is likely to inversely correlate with the user's likelihood of being involved in illegal activity.

For each user, we also construct a measure of the amount of mainstream interest and hype associated with bitcoin at the time of their participation in bitcoin (we label the variable *Bitcoin Hype*). We take the average Google Trends search intensity for the keyword "bitcoin" at the time of the user's bitcoin transactions. If Google search intensity for "bitcoin" correlates with speculative trading in bitcoin and mainstream (legal) use, this variable will have an inverse association with the likelihood of the user being involved in illegal activity.

Our final instrument for involvement in illegal activity exploits the anecdotal evidence that significant darknet marketplace shocks such as seizures of darknet marketplaces by law enforcement agencies or closures of such marketplaces for scams or hacks result in a brief spike of transaction activity by illegal users as they turn to alternative marketplaces or relocate their holdings in response to the shock. At the same time, shocks to darknet marketplaces are unlikely to materially affect the activity of legal users. Therefore, for each user, we measure the fraction of the user's transaction value that occurs in the one week period after each major darknet marketplace shock (marketplace "raids", "scams", and "hacks" in Table A2 of Appendix A). We label this variable *Darknet Shock Volume*.

As determinants of the probability of detection, we include a binary variable for whether the user started using bitcoin (date of first bitcoin transaction) before the first seizure of bitcoin by law enforcement agencies from Silk Road 1 (we label the variable *Pre-Silk-Road User*). Because users that enter the bitcoin network after the first seizure can only be detected in subsequent seizures, post-Silk-Road-seizure users are likely to have a lower detection probability.

A few things are worth noting about the variables used in the DCE model. First, while the instrumental variables help identify the model, they are not the only characteristics that help separate legal and illegal users—the full set of characteristics used in the model serve that purpose, including variables

common to both detection and violation equations (they have different coefficients in each equation). The full list of variables is presented in Table 1. Second, identification of the model requires only one variable that is associated with either the probability of being involved in illegal activity or the probability of detection, but not both. We have more candidate instrumental variables than this minimum of one, and in robustness tests we examine how sensitive the results are to the assumptions about these instruments. We do so by relaxing the assumed exclusion restrictions on a subset of the instruments one at a time, from which we conclude that the results are not particularly sensitive to any individual instrumental variable's exclusion restriction.

Table 2 Panel B reports descriptive statistics about the variables that serve as instruments. *Darknet Sites* indicates that for the average bitcoin participant, there are on average 17 operational darknet marketplaces around the time of their transactions. This number ranges from a minimum of zero to a maximum of 27. *Tumbling* indicates that only a relatively small proportion of users (less than 25%) engage in "tumbling" and/or "wash trades", which are used to obscure the user's holdings. Thus, while techniques exist to help a bitcoin user conceal their activity, it appears that few bitcoin users adopt such techniques.

The variable *Shadow Coins* indicates that for the average bitcoin participant, there are around two opaque alternative cryptocurrencies in existence at the time of their transactions. The variable *Darknet Shock Volume* indicates that while most users do not trade in the period immediately following darknet shocks (median of zero), some users conduct a large fraction of their trading during these periods, with the average bitcoin user undertaking 17% of their trading following darknet shocks.

The variable *Bitcoin Hype* indicates that for the average user, the intensity of Google searches for "bitcoin" is around 28% of its maximum of 100%. The *Pre-Silk-Road User* dummy indicates that only around 7% of all bitcoin participants started transacting before October 2013, when the first darknet marketplace seizure by law enforcement agencies occurred (the seizure of Silk Road 1 by the FBI).

5.4. How much illegal activity involves bitcoin?

Both methods—network cluster analysis (SLM) and detection controlled estimation (DCE)—arrive at probabilistic classifications of bitcoin users into those primarily involved in legal activity and those primarily involved in illegal activity. Once the users have been partitioned into the legal and illegal "communities", we use those categorizations to quantify the size and activity of the two groups.

Table 4 presents the main results at the aggregate level, for the whole sample period. Panel A reports the estimated size of the groups and their level of activity, while Panel B re-expresses these values as percentages for each group. First, the percentage of bitcoin users estimated to be predominantly involved in illegal activity is 29.12% using the SLM and 21.37% using the DCE, giving a midpoint

estimate of about one-quarter of bitcoin users (25.24%, the average of the estimates from the two models). The 99% confidence interval around this estimate is 21.73% to 28.76%.[27] The midpoint estimate suggests around 26.82 million bitcoin users are predominantly involved in illegal activity, versus 79.42 million legal users.

The estimated number of illegal users is around four times larger than our sample of observed illegal users. Given our sample of observed illegal users is based on a comprehensive approach and includes all users that can be observed transacting with one of the known darknet marketplaces, the results suggest that without empirical methods such as the SLM or DCE, illegal activity that can be inferred from involvement with known darknet marketplaces represents only a small (and likely non-random) fraction of all illegal activity. Thus, our results suggest that studies of known/identifiable darknet markets (e.g., Soska and Christin, 2015; Meiklejohn et al., 2013) only scratch the surface of all illegal activity involving bitcoin.

< Table 4 >

Table 4 also indicates that illegal users account for an even larger share of all transactions—around 44.33% (45.67% using the SLM and 42.99% using the DCE) or approximately 269 million transactions. Thus, the average illegal user is involved in more transactions than a legal user. This result is consistent with the notion that illegal users are likely to use bitcoin as a payment system (which involves actively transacting), whereas legal users may hold bitcoin for reasons such as speculation. A similar proportion is observed for holding values—illegal users on average hold around one-half (51.28%) of the outstanding bitcoin. One reason for the large share of illegal user holdings (relative to their share of the number of users) is related to the calculation of this variable as a time-series average. A high fraction of illegal users in the early parts of the sample (when there are fewer bitcoin users) can generate such a result even if the holdings *per user* are lower among illegal users compared to legal users.

Illegal users are estimated to control around 38.21% of bitcoin addresses and account for about one-fifth (20.30%) of the dollar volume of bitcoin transactions. In dollar terms, illegal users conduct approximately $378 billion worth of bitcoin transactions. Because illegal users account for a larger share of transactions than their share of dollar volume, they tend to make smaller value transactions than legal

[27] We use a form of bootstrapped standard errors to form the confidence interval. First we obtain standard errors from the DCE model using a bootstrap of 200 samples in which, for computational reasons, we are forced to reduce the sample size by taking a random sample (this is a conservative step as it inflates the estimated standard errors relative to the standard errors for the full sample size). We then apply the conservative bootstrapped DCE standard errors to approximate the error in the midpoint estimate. This step assumes the SLM standard errors (which we cannot compute as a bootstrap would not be appropriate when one needs to use the transaction network in the model) are similar in magnitude to the DCE standard errors.

users. This result is consistent with illegal users primarily using bitcoin as a payment system rather than holding it as an investment or speculative asset.

Three general conclusions can be drawn from the results in Table 4. First, illegal users account for a sizeable proportion of both users and trading activity in bitcoin, with the exact proportion varying across different measures of activity and the two estimation models. Second, the estimates from both the SLM and DCE are fairly similar across the various activity measures, despite relying on completely different assumptions and information. Third, even a fairly comprehensive approach to identifying illegal activity directly (such as the approach used in the previous section and that used in other darknet market studies) only captures a small fraction of the total illegal activity, highlighting the importance of extrapolation beyond a directly observed sample.

5.5. How does the illegal activity vary through time?

There is interesting time-series variation in the amount of illegal activity and its share of all bitcoin activity. Figures 4 to 7 plot the estimated amount of illegal activity that uses bitcoin through time from the first block in 2009 to 2017. The figures show the estimated number of illegal users, the number and dollar value of their transactions, and the value of their bitcoin holdings. Panel B of each of the figures shows these activity measures as a percentage of the total across all bitcoin participants.[28]

< Figure 4 here >
< Figure 5 here >
< Figure 6 here >
< Figure 7 here >

A pattern that is observed across all activity measures is that illegal activity, as a percentage of total bitcoin activity, tends to be high at the start of the sample in 2009, and then again from 2012 to the end of 2015, after which it steadily declines through to 2017. The activity levels indicate that there is only a very small (negligible) level of activity in bitcoin until about the middle of 2011, so the activity at the start of the sample is not economically meaningful. In contrast, the high relative level of illegal activity between 2012 and 2015 is noteworthy and coincides with the growth in the number of illegal darknet marketplaces, starting with the Silk Road in 2011. After the Silk Road was shut down in October of 2013, a large number of other illegal darknet marketplaces commenced operating between 2013 and 2015 (Table A2 of Appendix A).

[28] Figures 4-7 use the average of the SLM and DCE model estimates. The SLM and DCE time-series estimates are separately reported in Figures A1-A8 of the Online Appendix.

What could drive the decline in the relative level of illegal activity from the end of 2015 onwards? The first thing to note is that the decline is observed in relative terms (that is, illegal activity as a fraction of total bitcoin activity), but *not* in absolute terms. Thus, it is not the case that the level of illegal activity in bitcoin has declined in recent years, rather, there has been a disproportionate increase in the legal use of bitcoin since the end of 2015. For example, from the end of 2015 to April 2017, the estimated number of illegal bitcoin users increases from around 16 million to around 24 million, reflecting growth of around 50%, whereas the estimated number of legal bitcoin users increases from around 15 million to around 58 million, reflecting growth of around 290%. The rapid growth of legal use is likely driven by factors such as increased interest from investors and speculators (e.g., the emergence of "cryptofunds", and more recently bitcoin futures) and increased mainstream adoption as a payment system (e.g., cafes and internet merchants accepting bitcoin).

The time-series of legal and illegal activity levels show strong growth in both illegal and legal activity throughout the sample period, in particular since 2012. Interestingly, the strong growth in illegal activity precedes the strong growth in legal activity—by about three or four years. Thus it seems illegal users were relatively early adopters of bitcoin as a payment system.

Finally, because of the rapid growth in the legal use of bitcoin in the final two years of the sample, the aggregate measures of the illegal proportion of bitcoin activity reported in the previous subsection understate the proportion seen throughout most of the sample period. For example, for most of the sample period, the estimated proportion of illegal users is closer to one-half than one-quarter (the aggregate estimate). The aggregate estimate is heavily influenced by the large number of legal users that enter in the last two years of the sample. Similarly, for much of the sample period, the estimated proportion of bitcoin transactions involved in illegal activity is between 60% and 80%, contrasting with the aggregate estimate of around 44%.

The most recent estimates of illegal activity (at the end of our sample in April 2017) suggest there are around 24 million illegal users of bitcoin. These users conduct around 36 million bitcoin transactions annually, valued at around $72 billion, and collectively hold around $8 billion in bitcoin.[29]

5.6. What are the characteristics of illegal users?

We assess the differences between legal and illegal user characteristics in two ways: univariate statistics that compare observed or estimated illegal users with their legal counterparts, and multivariate tests exploiting the coefficients of the estimated DCE model.

[29] For these estimates, we have halved the double-counted volumes so that the estimates can be interpreted as the volume/value of goods/services bought/sold by the illegal users.

< Table 5 >

Starting with a univariate difference in means, Table 5 compares the characteristics of the sample of "observed" illegal users with the characteristics of other users. Note that the "other users" are not all legal users—they contain a mix of legal users and undetected illegal users. Therefore, the table also compares the characteristics of users classified by the SLM and DCE models as being involved in illegal activity with those of users classified as legal. Interestingly, despite being based on completely different assumptions, the SLM and DCE models generally agree on how the characteristics of legal users differ from illegal users. This is true for the signs of the mean differences for all but one characteristic (*Transaction Frequency*).

The SLM and DCE models agree that illegal users tend to transact more (have a two to three times higher *Transaction Count*), but use smaller sized transactions (about half the average size of legal transactions). This result could be a reflection of illegal users predominantly using bitcoin to buy and sell goods and services, whereas some legal users also use bitcoin for investment and speculation.[30]

The models also agree that illegal users tend to hold less bitcoin (measured in dollar value) than legal users; their average *Holding Value* is about half that of legal users. This characteristic is consistent with the previous conjecture—legal users might tend to hold larger bitcoin balances because some use bitcoin for investment/speculation purposes, whereas for an illegal user that buys/sells illegal goods and services using bitcoin, holding a large balance is costly due to (i) opportunity costs of capital, and (ii) risks associated with having holdings seized by authorities. For these reasons, illegal users are likely to prefer holding less bitcoin and this tendency is supported by the data.

Illegal users tend to have more counterparties in total, reflecting their larger number of transactions, but tend to have a higher counterparty concentration. This suggests that illegal users are more likely to repeatedly transact with a given counterparty. This characteristic might be a reflection of illegal users repeatedly transacting with a given illegal darknet marketplace or other illegal user once trust is established from a successful initial exchange. Illegal users have a longer *Existence Time* (time between their first and last transactions in bitcoin), consistent with our observations from the time-series that illegal users tend to become involved in bitcoin earlier than legal users. Similarly, the differences in means also show that there is a higher proportion of Pre-Silk-Road users among the illegal users than the legal users (as indicated by the variable *Pre-Silk-Road User*).

[30] While the result could also reflect illegal users engaging in techniques to conceal their trading, this is less likely to be an explanation because a similar result holds in multivariate (DCE) tests that control for tumbling and wash trades.

The more specific indicators of illegal activity also show significant differences between the two groups. Illegal users tend to be more active during periods in which there are many illegal darknet marketplaces operating (a higher mean for the variable *Darknet Sites*). They make greater use of tumbling and wash trades to conceal their activity (two to four times more *Tumbling*). On average, a larger proportion of illegal volume, compared to legal volume, is transacted immediately following shocks to darknet marketplaces (*Darknet Shock Volume*). This finding is consistent with anecdotal evidence that illegal users turn to alternative marketplaces in response to darknet marketplace seizures or scams.

Interestingly, illegal users are more likely to transact with bitcoin when there are fewer opaque "shadow coins" in existence, suggesting such coins do get used as alternatives to bitcoin in illegal transactions. This result (for the variable *Shadow Coins*) is consistent with anecdotal accounts of shadow coins becoming recognized by the illegal community for their increased privacy, as well as recent examples of hackers demanding ransom payments in shadow coins rather than bitcoin.

Another interesting result is that legal users tend to be more active in bitcoin when there is less *Bitcoin Hype*, measured by the Google search intensity for "bitcoin". It therefore appears that Google searches for "bitcoin" are associated with mainstream (legal) adoption of bitcoin for payments, and/or speculative/investment interest in bitcoin.

In summary, the comparison of transactional characteristics (number and size of transactions, holdings, and counterparties) is consistent with the notion that illegal users predominantly use bitcoin for payments, whereas legal users are more likely to treat bitcoin as an investment asset. Furthermore, legal and illegal users differ with respect to when they are most active in bitcoin, with illegal users being most active when there are more darknet marketplaces, fewer shadow coins, less bitcoin hype, and immediately following shocks to darknet marketplaces. The differences in characteristics for the instrumental variables are consistent with the hypothesized differences, lending support to their use as instruments.

< Table 6 >

The DCE model coefficients reported in Table 6 provide multivariate tests of how the characteristics relate to the likelihood that a user is involved in illegal activity. The results confirm most of the observations made in the simple comparison of means. The effects of all of the instrumental variables are consistent with their hypothesized effects. A user is more likely to be involved in illegal activity if they trade when: (i) there are many darknet marketplaces operating, (ii) there are fewer shadow coins in existence, (iii) there is little bitcoin hype, and (iv) darknet marketplaces experience seizures or scams. A user is also more likely to be involved in illegal activity if they use tumbling and/or wash trades, transact frequently in small sized transactions, and tend to repeatedly transact with a given counterparty.

The marginal effects in Table 6, reported in parentheses below the coefficient estimates, provide a sense of the magnitudes of the effects and their relative importance.[31] For example, the marginal effects indicate that a one standard deviation increase in the number of illegal darknet marketplaces at the time a user transacts in bitcoin increases the probability of that user being involved in illegal activity by a factor of 0.435, or 43.5% of what their probability would otherwise be.[32] The magnitudes generally show that most of the determinants of involvement in illegal activity and determinants of the detection probability are economically meaningful. In particular, the instrumental variables *Darknet Sites*, *Shadow Coins*, *Bitcoin Hype*, and *Darknet Shock Volume* all have strong relations with the probability that a user is involved in illegal activity.

The DCE model also sheds light on the determinants of the likelihood that an illegal user is "detected" by either of our three approaches. The main instrument, *Pre-Silk-Road User* has a strong relation with detection, indicating that illegal users that commence transacting in bitcoin prior to the first darknet marketplace seizure in October 2013 have a higher probability of being detected. Similarly, those users that transact in bitcoin for a longer period of time (higher *Existence Time*), trade more frequently (higher *Transaction Frequency*), or tend to trade repeatedly with a given counterparty such as a darknet marketplace (higher *Concentration*) have a significantly higher detection probability.

Model 2 in Table 6 adds further control variables to the models, including *Holding Value* and *Transaction Count*, and finds that the main results do not change much in response to additional control variables. A risk of adding too many transactional control variables is co-linearity between such variables.

5.7. What are the characteristics of the illegal user network?

Exploiting the fact that the bitcoin blockchain provides us with a complete record of every transaction between every pair of counterparties, we briefly explore how the trade network of illegal users differs from that of legal users. Our approach is to compute a few descriptive network metrics that capture different aspects of network topology and structure for each of the two groups or "communities" separately and then compare the values between the two communities. In mapping the networks, users form the "nodes", and transactions between users form the "edges" or "links" between nodes.

< Table 7 >

[31] To make the comparisons and interpretation easier, before estimating the DCE models, we standardize all variables to have mean zero and standard deviation of one. We also log transform the right skewed variables (*Transaction Frequency*, *Size*, and *Count*, and *Holding Value*) and winsorize the variables at +/- three standard deviations to reduce the influence of extreme values.

[32] As an example of how to interpret the marginal effect of 0.435, if a user's illegal probability is say 20%, the predicted effect of a one standard deviation increase in *Darknet Sites*, holding all else constant, is to increase the user's probability to 20% × 1.435 = 28.7%, an increase of 43.5% of what their probability would otherwise be.

Table 7 reports the results. The first metric, *Density*, takes the range [0,1] and indicates how highly connected the users are within a community (versus how sparse the connections are between users); it is the actual number of links between users within the given community (a "link" between two users means that they have transacted with one another) divided by the total potential number of links. It shows that the illegal trade network is three to four times denser in the sense that users are much more connected to one another through transactions. This observation is consistent with the fact that illegal users tend to transact more than legal users. It is also consistent with the notion that in the illegal community, bitcoin's dominant role is likely that of a payment system in buying/selling goods, whereas in the legal community, bitcoin is also used as an investment or for speculation.

Reciprocity takes the range [0,1] and indicates the tendency for users to engage in two-way interactions; it is the number of two-way links between users within the given community (a two-way link is when two users send and receive bitcoin to and from one another) divided by the total number of links within the given community (two-way and one-way). While *Reciprocity* is higher among illegal users than it is among legal users, it is generally very low in both communities (1% among legal users and 3% among illegal users). Thus, interactions between bitcoin users are generally only one-way interactions with one counterparty receiving bitcoin from the other but not vice versa.

Entropy measures the amount of heterogeneity among users in their number of links to other members of the community. It takes its minimum value of zero when all users have the same number of links (same degree).[33] The results suggest that illegal users are a more heterogeneous group in terms of the number of links each user has with other members of the community. A driver of that heterogeneity could be that the illegal community at one end of the spectrum has darknet marketplaces that have hundreds of thousands of links to vendors and buyers, and at the other end has individual customers of a single marketplace, potential with only the one link.

A concluding observation is that both the SLM and DCE models provide a consistent picture of how legal and illegal users differ, this time in the context of their trade networks. Again, this suggests that the two different models tend to agree about the nature of the illegal activity in bitcoin.

5.8. Robustness tests

We conduct a number of different robustness tests. Perhaps the most rigorous robustness test of an empirical model is to compare its results with results from a completely different model/approach that makes different assumptions and draws on different information. Throughout the paper we put our two

[33] Formally, $Entropy = -\sum_d P(d)\log[P(d)]$, where $P(d)$ is the degree distribution (probability density of the degree for each user, where a user's degree, d, is the number of links the user has with other members of the same community).

empirical models through this test. The two models, one based on a network cluster analysis algorithm and the other on a structural latent variables model drawing on observable characteristics, provide highly consistent results. The two models tend to agree, within a reasonable margin of error, on the overall levels of illegal activity, as well as the differences between legal and illegal users in terms of characteristics and network structure.

We also subject each of the models to specific tests that vary key assumptions or modelling choices. Table 8 reports the estimated amount of illegal activity for the most notable of these tests. For the SLM, we re-estimate the model using transaction volumes as the measure of interaction between users rather than transaction counts (*SLM Alternative 1*). We also consider a modification of the SLM algorithm in which we impose a constraint that does not allow the sample of "observed" illegal users to be moved to the legal community (*SLM Alternative 2*). For the DCE model, one set of robustness tests involves examining the sensitivity to relaxing key exclusion restrictions. For example, in the baseline model, *Darknet Sites* (the number of operational darknet marketplaces at the time a user transacts) is included only as a determinant of illegal activity. As a robustness test (*DCE Alternative 1*), we include it in both equations, allowing it to also affect the probability of detection. Of all the determinants of illegal activity, *Darknet Sites* has the most plausible reasons for possibly also affecting detection—the existence of many darknet marketplaces might be a catalyst for increased surveillance and enforcement by law enforcement authorities. We also test sensitivity to the key exclusion restriction in the detection equation by including *Pre-Silk-Road User* in both equations (*DCE Alternative 1*), thereby allowing it to also affect the probability of illegal activity.

< Table 8 >

Table 8 shows that the estimated overall levels of illegal activity across the various activity measures are not overly sensitive to modifications of the baseline model. Similarly, the estimated characteristics of illegal users are not overly sensitive to these modifications (results not reported for conciseness). The Online Appendix Table A1 reports the coefficient estimates for the two DCE models described above in which we relax key exclusion restrictions, showing that the coefficients are also not particularly sensitive to these modifications.

We also examine the robustness of the DCE model to the initial parameter values used in estimating the model. We initialize the model with different starting values (-1, 0, +1, and randomly drawn starting values), and find that our results are not sensitive to the choice of starting values, suggesting convergence to a global rather than local maximum of the likelihood function.

Finally, we re-estimate the standard errors used in confidence bounds around the estimated illegal activity and significance tests. Instead of the bootstrapped standard errors that we use in the main results, we instead compute standard errors using analytic expressions. We find that the analytic standard errors are considerably smaller than the bootstrapped standard errors. This finding suggests that using bootstrapped standard errors in the main results is the more conservative of the two approaches.

Finally, the characteristics of illegal users could change through time (for example, in response to seizures by law enforcement agencies), which could lead to model mis-specification. To examine this possibility, we repeat the difference-in-means comparison of legal and illegal users, partitioning the sample into a pre-Silk-Road seizure period and a post-Silk-Road seizure period (pre/post October 2013). Tables A2 and A3 in the Online Appendix report these results for both SLM and DCE classifications of illegal users.[34] For most characteristics, the differences between legal and illegal users take the same sign in both the pre/post periods, typically with similar magnitudes. In cases where the pre and post periods are different, the difference is often driven by a change in the characteristics of legal rather than illegal users (the change in the legal user mean is larger than the change in the illegal user mean). This tendency suggests while some characteristics do change through time, the changes are more likely to reflect general trends rather than a response of illegal users to events such as darknet marketplaces seizures.

6. Discussion

6.1. Implications

Blockchain technology and the systems/protocols that can be implemented on a blockchain have the potential to revolutionize numerous industries. Possible benefits to securities markets include reducing equities settlement times and costs (Malinova and Park, 2016; Khapko and Zoican, 2016), increasing ownership transparency leading to improved governance (Yermack, 2017), and providing a payments system with the network externality benefits of a monopoly but the cost discipline imposed by free market competition (Huberman et al., 2017). The technology has even broader applications beyond securities markets, from national land registries, to tracking the provenance of diamonds, decentralized decision making, peer-to-peer insurance, prediction markets, online voting, distributed cloud storage, internet domain name management, conveyancing, medical record management, and many more.

This technology, however, is encountering considerable resistance, especially from regulators. Regulators are cautious due to their limited ability to regulate cryptocurrencies and the many potential but poorly understood risks associated with these innovations. The negative exposure generated by anecdotal accounts and salient examples of illegal activity no doubt contributes to regulatory concerns and risks

[34] The comparison excludes characteristics that have little or no variation with the pre or post periods, such as *Pre-Silk Road User* dummy variable, *Shadow Coins*, and *Darknet Sites*.

stunting the adoption of blockchain technology, limiting its realized benefits. In quantifying and characterizing this area of concern, we hope to reduce the uncertainty about the negative consequences of cryptocurrencies, allowing for more informed decisions by policymakers that assess both the costs and benefits. Hopefully, by shedding light on the dark side of cryptocurrencies, this research will help blockchain technologies reach their potential.

A second contribution of this paper is the development of new approaches to identifying illegal activity in bitcoin, drawing on network cluster analysis and detection controlled estimation techniques. These methods can be used by law enforcement authorities in surveillance activities. For example, our methods can be applied to blockchain data going forward as new blocks are created, allowing authorities to keep their finger on the pulse of illegal activity in bitcoin. Applied in this way, one could monitor trends in illegal activity such as its growth or decline, its response to various regulatory interventions such as seizures, and how its characteristics change through time. Such information could help make more effective use of scarce regulatory and enforcement resources.

Another surveillance application is in identifying individuals/entities of strategic importance, for example, major suppliers of illegal goods. Combining these empirical methods with other sources of information can "de-anonymize" the nameless entities identified in the data. This might be done, for example, by tracing the activity of particular individuals to the interface of bitcoin with either fiat currency or the regulated financial sector (many exchanges and brokers that convert cryptocurrencies to fiat currencies require personal identification of clients). The methods that we develop can also be used in analyzing many other blockchains.

Third, our finding that a substantial amount of illegal activity is facilitated with bitcoin suggests that bitcoin has contributed to the emergence of an online black market, which raises several welfare considerations. Should policymakers be concerned that people are buying and selling illegal goods such as drugs online and using the anonymity of cryptocurrencies to make the payments? This is an important question and the answer is not obvious. If the online market for illegal goods and services merely reflects a migration of activity that would have otherwise occurred "on the street" to the digital world of e-commerce, the illegal online activity facilitated via bitcoin might not be bad from a welfare perspective. In fact, there are many potential benefits to having illegal drugs and other goods bought and sold online rather than on the street. For example, it might be safer and lead to reduced violence (e.g., Barratt et al., 2016a). It could also increase the quality and safety of the drugs because darknet marketplaces rely heavily on user feedback and vendor online reputation, which can give a buyer access to more information about a seller's track record and product quality than when buying drugs on the street (e.g., Soska et al., 2015). There is also more choice in the goods offered, which has the potential to increase consumer welfare.

However, by making illegal goods more accessible, convenient, and reducing risk (due to anonymity), the darknet might encourage *more* consumption of illegal goods and increase reach, rather than simply migrating existing activity from the street to the online environment (Barratt et al., 2016b). Presuming illegal goods and services have negative net welfare consequences, then bitcoin and other cryptocurrencies could decrease welfare by enabling the online black market. Such negative consequences would have to be weighed up against welfare gains that also accompany cryptocurrencies.

Therefore, while our paper does not provide a definitive answer to the question of welfare effects, it does get us closer to an answer by having estimated both the trends and scale of illegal activity involving bitcoin (the most widely used cryptocurrency in darknet marketplaces). Future research might quantify the relation between drug trafficking on the street vs online (drawing on our methods or estimates) to understand to what extent we are experiencing a simple migration vs an expansion in the overall market. It might also quantify the benefits of moving to an online market and contrast them with the negative consequences of any expansion in the market as a result of it being more accessible / convenient / safe.

Our results also have implications for the intrinsic value of bitcoin. The rapid increase in the price of bitcoin in recent times has prompted much debate and divided opinions among market participants and even policymakers / central banks about whether cryptocurrency valuations are disconnected from fundamentals and whether their prices reflect a bubble. In part, the debate reflects the uncertainty about how to value cryptocurrencies and how to estimate a fundamental or intrinsic value. While we do not propose a valuation model, our results provide an input to an assessment of fundamental value in the following sense. One of the intrinsic uses of cryptocurrencies, giving them some fundamental value, is as a payment system. To make payments with bitcoin, one has to hold some bitcoin; the more widespread its use as a payment system, the greater the aggregate demand for holding bitcoin to make payments, which, given the fixed supply, implies a higher price. Our results suggest that currently, as a payment system, bitcoin is relatively widely used to facilitate trade in illegal goods and services and thus the illegal use of bitcoin is likely to be a meaningful contributor to bitcoin's fundamental value.

This observation—that a component of bitcoin's fundamental value derives from its use in illegal trade—raises a few issues. First, an ethical investor might not be comfortable investing in a security for which a meaningful component of the fundamental value derives from illegal use. Second, changes in the demand to use bitcoin in illegal trade are likely to impact its fundamental value. For example, increased attention from law enforcement agencies or increased adoption/substitution to more opaque alternative cryptocurrencies could materially decrease the fundamental value of bitcoin. Conversely, continued migration of the black market to online with continued use of bitcoin, could further increase bitcoin's fundamental value. Third, recent price appreciation of bitcoin greatly exceeds the growth in its use in

illegal activity, suggesting either a substantial change in other components of bitcoin's fundamental value or a dislocation of the bitcoin price from its fundamental value.

6.2. Relation to other literature

This paper contributes to three branches of literature. First, several recent papers analyze the economics of cryptocurrencies and applications of blockchain technology to securities markets (e.g., Malinova and Park, 2016; Khapko and Zoican, 2016; Yermack, 2017; Huberman et al., 2017; Easley et al., 2017). Our paper contributes to this literature by showing that one of the major uses of cryptocurrencies as a payment system is in settings in which anonymity is valued (e.g., illegal trade).

Another related, although small, branch of literature examines the degree of anonymity in bitcoin by quantifying the extent to which various algorithms can identify entities/users in bitcoin blockchain data and track their activity (e.g., Meiklejohn et al., 2013; Ron and Shamir, 2013; Androulaki et al., 2013; Tasca et al., 2016). In doing so, some of these papers also provide insights about the different types of activities that use bitcoin. Of these papers, one of the closest to ours is Meiklejohn et al. (2013), who explore the bitcoin blockchain up to April 2013, clustering addresses into entities/users and manually identifying some of those entities by physically transacting with them. They are able to identify the addresses of some miners, exchanges, gambling services, and vendors/marketplaces (including one darknet marketplace), suggesting bitcoin entities are not completely anonymous. Tasca et al. (2016) use a similar approach to explore the different types of activity in bitcoin, focusing only on the largest entities, so-called "super clusters", and within that set, only those with a known identity.

None of these papers attempt to categorize all of the activity in bitcoin, nor do they try and quantify or characterize the population of illegal bitcoin users, which is the focus of our paper. We exploit the lack of perfect anonymity that is documented in these studies and draw on some of the techniques from this literature to construct an initial sample of known illegal users. We add new methods to this literature, extending the empirical toolkit from making direct observations about individuals, to identification of communities and estimation of populations of users.

Finally, another related branch of literature is the recent studies of darknet marketplaces and the online drug trade, including papers from computer science and drug policy. For example, Soska and Christin (2015), use a web-crawler to scrape information from darknet marketplaces during 2013-2015, collecting a variety of data. Their paper provides valuable insights into these markets, including information about the types of goods and services traded (largely drugs), the number of goods listed, a lower bound on darknet turnover using posted feedback as a proxy (they do not have data on actual transactions/sales), the number of vendors, and the qualitative aspects of how these marketplaces operate (reputation, trust, feedback). The related drug policy studies often draw on other sources of information

such as surveys of drug users and contribute insights such as: (i) darknet marketplaces like the Silk Road facilitate initiation into drug use or a return to drug use after cessation (Barratt et al., 2016b) and can encourage drug use through the provision of drug samples (Ladegaard, 2017); (ii) darknet forums can promote harm minimization by providing inexperienced users with support and knowledge from vendors and more experienced users (Bancroft, 2017); (iii) darknet marketplaces tend to reduce systemic violence compared with in-person drug trading because no face-to-face contact is required (Barratt et al., 2016a; Martin, 2017; Morselli et al., 2017); (iv) about one-quarter of the drugs traded on the Silk Road occur at a wholesale scale, suggesting that such markets might also indirectly serve drug users "on the street" by impacting dealers (Aldridge and Décary-Hétu, 2016); and (v) there are interesting cross-country differences in the use of the darknet marketplace "Agora" (Van Buskirk et al., 2016).

We contribute to this literature by quantifying the amount of illegal activity undertaken using bitcoin. All of the illegal activity captured by the existing studies of one or several darknet marketplaces is also in our measures because one of the approaches we use to construct a sample of observed illegal activity involves measuring transactions with known darknet marketplaces. However, our estimates include much more than this activity—we use direct measures of transactions rather than a lower-bound measure such as feedback, consider all known darknet marketplaces (rather than one or a few), include two other methods of obtaining a sample of illegal activity, and most importantly, we estimate models that extrapolate from the *sample* of observed illegal activity to the estimated *population*. This yields vastly different and more comprehensive estimates. Empirically, we confirm that studies of darknet marketplaces only scratch the surface of the illegal activity involving bitcoin—the estimated population of illegal activity is several times larger than what can be "observed" through studying known darknet marketplaces. Furthermore, the studies of darknet marketplaces do not analyze how the characteristics of illegal and legal bitcoin users differ, or how recent developments such as increased mainstream interest in bitcoin and the emergence of new, more opaque cryptocurrencies impacts the use of bitcoin in illegal activity. These are further contributions of our paper.

7. Conclusion

As an emerging FinTech innovation, cryptocurrencies and the blockchain technology on which they are based could revolutionize many aspects of the financial system, ranging from smart contracts to settlement, interbank transfers to venture capital funds, as well as applications beyond the financial system. Like many innovations, cryptocurrencies also have their dark side. We shed light on that dark side by quantifying and characterizing their use in illegal activity.

We find that illegal activity accounts for a sizable proportion of the users and trading activity in bitcoin, as well as an economically meaningful amount in dollar terms. For example, approximately one-

quarter of all users and close to one-half of transactions are associated with illegal activity, equating to around 24 million market participants with illegal turnover of around $72 billion per year in recent times.

Illegal users of bitcoin tend to transact more, in smaller sized transactions, often repeatedly transacting with a given counterparty, and they tend hold less bitcoin. These features are consistent with their use of bitcoin as a payment system rather than for investment or speculation. Illegal users also make greater use of transaction techniques that obscure their activity, and their activity spikes following shocks to darknet marketplaces. The proportion of bitcoin activity associated with illegal trade declines with increasing mainstream interest and hype (Google search intensity), with the emergence of more opaque alternative cryptocurrencies, and with fewer darknet marketplaces in operation.

Our results have a number of implications. First, by shedding light on the dark side of cryptocurrencies, we hope this research will reduce some of the regulatory uncertainty about the negative consequences and risks of this innovation, thereby allowing more informed policy decisions that weigh up the benefits and costs. In turn, we hope this contributes to these technologies reaching their potential.

Second, the techniques developed in this paper can be used in cryptocurrency surveillance in a number of ways. The methods can be applied going forward as new blocks are added to the blockchain, allowing authorities to keep their finger on the pulse of illegal activity and monitor its trends, its responses to regulatory interventions, and how its characteristics change through time. Such information could help make more effective use of scarce regulatory and enforcement resources. The methods can also be used to identify individuals of strategic importance in illegal networks.

Third, our paper suggests that a significant component of the intrinsic value of bitcoin as a payment system derives from its use in facilitating illegal trade. This has ethical implications for those that view bitcoin as an investment, as well as valuation implications. For example, changes in the demand to use bitcoin in illegal trade (e.g., due to law enforcement crackdowns or increased adoption of more opaque cryptocurrencies in illegal trade) are likely to impact its fundamental value.

Finally, our paper moves the literature closer to answering the important question of the welfare consequences of the growth in illegal online trade. A crucial piece of this puzzle is understanding the extent to which the online illegal trade simply reflects migration of activity that would have otherwise occurred on the street, versus the alternative that by making illegal goods more accessible, convenient to buy, and less risky due to anonymity, the move online could lead to growth in the aggregate black market. Our estimates of the amount of illegal trade facilitated with bitcoin through time contribute to understanding this issue, but further research is required to relate these estimates to trends in the offline black market.

Appendix A: Bitcoin seizures and darknet sites

Table A1: Bitcoin seizures

This table reports major bitcoin seizures, the seizing authority, the owner of the seized bitcoin, the date of the seizure, and the amount (in bitcoin) seized.

Seizing authority	Seized entity	Owner of seized bitcoins	Date of seizure	Bitcoin seized
Australian Government	Individual	Richard Pollard	December 1, 2012	24,518
US government	Individual	Matthew Luke Gillum	July 23, 2013	1,294
ICE and HSI	Individual	Cornelius Jan Slomp	August 27, 2013	385,000
FBI	Individual	Ross William Ulbricht	October 1, 2013	144,000
FBI	Site	Silk Road escrow accounts (many users)	October 2, 2013	29,655

Table A2: Darknet sites accepting bitcoin, current and past

This table reports the 30 known darknet marketplaces with the longest operational history. For sites that remain operational (as at May 2017), the *End date* column states "Operational" and thus there is no *Closure reason*. *Days operational* is the number of days the site was operational before closure. Data are sourced from www.gwern.net.

Market	Launch date	End date	Closure reason	Days operational
Dream	November 15, 2013	Operational		>1,207
Outlaw	December 29, 2013	Operational		>1,163
Silk Road 1	January 31, 2011	October 2, 2013	Raided	975
Black Market Reloaded	June 30, 2011	December 2, 2013	Hacked	886
AlphaBay	December 22, 2014	Operational		>805
Tochka	January 30, 2015	Operational		>766
Crypto Market / Diabolus	February 14, 2015	Operational		>751
Real Deal	April 9, 2015	Operational		>697
Darknet Heroes	May 27, 2015	Operational		>649
Agora	December 3, 2013	September 6, 2015	Voluntary	642
Nucleus	October 24, 2014	April 13, 2016	Scam	537
Middle Earth	June 22, 2014	November 4, 2015	Scam	500
BlackBank	February 5, 2014	May 18, 2015	Scam	467
Evolution	January 14, 2014	March 14, 2015	Scam	424
Silk Road Reloaded	January 13, 2015	February 27, 2016	Unknown	410
Anarchia	May 7, 2015	May 9, 2016	Unknown	368
Silk Road 2	November 6, 2013	November 5, 2014	Raided	364
The Marketplace	November 28, 2013	November 9, 2014	Voluntary	346
Blue Sky Market	December 3, 2013	November 5, 2014	Raided	337
Abraxas	December 13, 2014	November 5, 2015	Scam	327
Pandora	October 21, 2013	August 19, 2014	Scam	302
BuyItNow	April 30, 2013	February 17, 2014	Voluntary	293
TorBazaar	January 26, 2014	November 5, 2014	Raided	283
Sheep	February 28, 2013	November 29, 2013	Scam	274
Cloud-Nine	February 11, 2014	November 5, 2014	Raided	267
Pirate Market	November 29, 2013	August 15, 2014	Scam	259
East India Company	April 28, 2015	January 1, 2016	Scam	248
Mr Nice Guy 2	February 21, 2015	October 14, 2015	Scam	235
Andromeda	April 5, 2014	November 18, 2014	Scam	227
Topix 2	March 25, 2014	November 5, 2014	Voluntary	225

Appendix B: Derivations for the DCE model

We define $I(.)$ and $D(.)$ to be monotonic link functions that map $x_{1i}\beta_1$ and $x_{2i}\beta_2$ to latent probabilities of a bitcoin user being involved predominantly in illegal activity, and detection of an illegal user, respectively.[35] That is,

$$I(x_{1i}\beta_1) = \Pr(L_{1i} = 1) \tag{B.1}$$

$$D(x_{2i}\beta_2) = \Pr(L_{2i} = 1 \mid L_{1i} = 1) \tag{B.2}$$

Table B1 reports the probabilities of various joint outcomes (represented by cells in the table). The joint outcomes are mutually exclusive and exhaustive, so the probabilities in Table B1 sum to one.

Table B1: Two-stage DCE model probability matrix

	Illegal user	Legal user
Detected	$I(x_{1i}\beta_1)D(x_{2i}\beta_2)$	0
Not detected	$I(x_{1i}\beta_1)[1 - D(x_{2i}\beta_2)]$	$1 - I(x_{1i}\beta_1)$

The log likelihood of the users that end up in the detected (seized) illegal users category (A) is the log of the sum (over users in A) of the probabilities of that joint outcome:

$$log\ L_A = \sum_{i \epsilon A} \log[I(x_{1i}\beta_1)D(x_{2i}\beta_2)] \tag{B.3}$$

Similarly, the log likelihood of the users that end up in the other category (A^C) is the log of the sum (over users in A^C) of the probabilities of that joint outcome (the probability that the user is a legal one plus the probability that an illegal user is not detected):

$$log\ L_A c = \sum_{i \epsilon A^C} \log[I(x_{1i}\beta_1)[1 - D(x_{2i}\beta_2)] + 1 - I(x_{1i}\beta_1)] \tag{B.4}$$

$$log\ L_A c = \sum_{i \epsilon A^C} \log[1 - I(x_{1i}\beta_1)D(x_{2i}\beta_2)] \tag{B.5}$$

Sets A and A^C constitute the universe of all bitcoin users. Therefore the full-sample log likelihood is:

$$log\ L = \sum_{i \epsilon A} \log[I(x_{1i}\beta_1)D(x_{2i}\beta_2)] + \sum_{i \epsilon A^C} \log[1 - I(x_{1i}\beta_1)D(x_{2i}\beta_2)] \tag{B.6}$$

Maximum likelihood estimation involves selecting parameter vectors β_1 and β_2 such that the function $log\ L$ is maximized.

[35] In our implementation, the link functions are cumulative logistic distribution functions, that is, $I(x_{1i}\beta_1) = \frac{1}{1+e^{-x_{1i}\beta_1}}$, $D(x_{2i}\beta_2) = \frac{1}{1+e^{-x_{2i}\beta_2}}$.

References

Aldridge, J., and D. Décary-Hétu, 2014, 'Not an ebay for drugs': The cryptomarket 'Silk Road' as a paradigm shifting criminal innovation, *Unpublished manuscript.*

Aldridge, J., and D. Décary-Hétu, 2016, Hidden wholesale: The drug diffusing capacity of online drug cryptomarkets, *International Journal of Drug Policy* 35, 7–15.

Androulaki, E., G. Karame, M. Roeschlin, T. Scherer, and S. Capkun, 2013, Evaluating user privacy in bitcoin, In *Proceedings of Financial Cryptography 2013.*

Bancroft, A., 2017, Responsible use to responsible harm: Illicit drug use and peer harm reduction in a darknet cryptomarket, *Health, Risk and Society* 19, 336–350.

Barratt, M.J., J.A. Ferris, and A.R. Winstock, 2016a, Safer scoring? Cryptomarkets, social supply and drug market violence, *International Journal of Drug Policy* 35, 24–31.

Barratt, M.J., S. Lenton, A. Maddox, and M. Allen, 2016b, 'What if you live on top of a bakery and you like cakes?'—Drug use and harm trajectories before, during and after the emergence of Silk Road, *International Journal of Drug Policy* 35, 50–57.

Basu, G., 2014, The strategic attributes of transnational smuggling: Logistics flexibility and operational stealth in the facilitation of illicit trade, *Journal of Transportation Security* 7, 99–113.

Ben-Sasson, E., A. Chiesa, C. Garman, M. Green, I. Miers, E. Tromer, and M. Virza, 2014, Zerocash: Decentralized anonymous payments from bitcoin, In *2014 IEEE Symposium on Security and Privacy.*

Christin, N., 2013, Traveling the silk road: A measurement analysis of a large anonymous online marketplace, In *Proceedings of the 22nd International Conference on World Wide Web.*

Comerton-Forde, C., and T.J. Putniņš, 2014, Stock price manipulation: Prevalence and determinants, *Review of Finance* 18, 23–66.

Cormen, T.H., C.E. Leiserson, R.L. Rivest, and C. Stein, 2001, *Introduction to Algorithms*, Cambridge: MIT Press.

Cox, J., 2016, Staying in the shadows: The use of bitcoin and encryption in cryptomarkets, In *The Internet and Drug Markets, Edited by EMCDDA*, 41–47, *Lisbon: EMCDDA.*

Emmons, S., S. Kobourov, M. Gallant, and K. Börner, 2016, Analysis of network clustering algorithms and cluster quality metrics at scale, *PLOS ONE* 11, 1–18.

Feinstein, J.S., 1989, The safety regulation of US nuclear power plants: Violations, inspections, and abnormal occurrences, *Journal of Political Economy* 97, 115–154.

Feinstein, J.S., 1990, Detection controlled estimation, *Journal of Law and Economics* 33, 233–276.

Feinstein, J.S., 1991, An econometric analysis of income tax evasion and its detection, *RAND Journal of Economics* 22, 14–35.

Franklin, J., A. Perrig, V. Paxson, and S. Savage, 2007, An inquiry into the nature and causes of the wealth of internet miscreants, In *Proceedings of the 14th ACM Conference on Computer and Communications Security.*

Huberman, G., J.D. Leshno, and C. Moallemi, 2017, Monopoly without a monopolist: An economic analysis of the bitcoin payment system, *Unpublished manuscript.*

Easley, D., M. O'Hara, and S. Basu, 2017, From mining to markets, *Unpublished manuscript.*

Khapko, M., and M.A. Zoican, 2016, "Smart" settlement, *Unpublished manuscript.*

Koshy, P., D. Koshy, and P. McDaniel, 2014, An analysis of anonymity in bitcoin using p2p network traffic, In *18th International Conference on Financial Cryptography and Data Security.*

Kruithof, K., J. Aldridge, D. Décary-Hétu, M. Sim, E. Dujso, and S. Hoorens, 2016, Internet-facilitated drugs trade, *Unpublished manuscript.*

Ladegaard, I., 2017, Instantly hooked? Freebies and samples of opioids, cannabis, MDMA, and other drugs in an illicit e-commerce market, *Journal of Drug Issues* (forthcoming).

Lavorgna, A., 2016, How the use of the internet is affecting drug trafficking practices, In *Internet and Drug Markets, EMCDDA Insights.*

Lewman, A., 2016, Tor and links with cryptomarkets, In *Internet and Drug Markets, EMCDDA Insights.*

Malinova, K., and A. Park, 2016, Market design for trading with blockchain technology, *Unpublished manuscript.*

Martin, J., 2014a, *Drugs on the Dark Net: How Cryptomarkets are Transforming the Global Trade in Illicit Drugs,* Berlin: Springer.

Martin, J., 2014b, Lost on the silk road: Online drug distribution and the "cryptomarket", *Criminology and Criminal Justice* 14, 351–367.

Martin, J., 2017, Cryptomarkets, systemic violence and the "gentrification hypothesis", *Addiction* (forthcoming).

Matthews, A., R. Sutherland, A. Peacock, J. Van Buskirk, E. Whittaker, L. Burns, and R. Bruno, 2017, I like the old stuff better than the new stuff? Subjective experiences of new psychoactive substances, *International Journal of Drug Policy* 40, 44–49.

Meiklejohn, S., M. Pomarole, G. Jordan, K. Levchenko, D. McCoy, G.M. Voelker, and S. Savage, 2013, A fistful of bitcoins: Characterizing payments among men with no names, In *13th ACM Internet Measurement Conference.*

Morselli, C., D. Décary-Hétu, M. Paquet-Clouston, and J. Aldridge, 2017, Conflict management in illicit drug cryptomarkets, *International Criminal Justice Review* 27, 237–254.

Nakamoto, S., 2008, Bitcoin: A peer-to-peer electronic cash system, *Unpublished manuscript.*

Noether, S., 2015, Ring signature confidential transactions for monero, In *IACR Cryptology ePrint Archive*, 1098.

Rogoff, K., 2016, *The Curse of Cash*, (Princeton, NJ: Princeton University Press).

Ron, D., and A. Shamir, 2013, Quantitative analysis of the full bitcoin transaction graph, In *17th Financial Cryptography and Data Security International Conference.*

Soska, K., and N. Christin, 2015, Measuring the longitudinal evolution of the online anonymous marketplace ecosystem, In *Proceedings of the 24th USENIX Conference on Security Symposium.*

Tasca, P., S. Liu, and A. Hayes, 2016, The evolution of the bitcoin economy: Extracting and analyzing the network of payment relationships, *Unpublished manuscript.*

Tzanetakis, M., G. Kamphausen, B. Werse, and R. Von Laufenberg, 2016, The transparency paradox. Building trust, resolving disputes and optimising logistics on conventional and online drugs markets, *International Journal of Drug Policy* 35, 58–68.

Van Buskirk, J., A. Roxburgh, M. Farrell, and L. Burns, 2014, The closure of the silk road: What has this meant for online drug trading?, *Addiction* 109, 517–518.

Van Buskirk, J., S. Naicker, A. Roxburgh, R. Bruno, and L. Burns, 2016, Who sells what? Country specific differences in substance availability on the agora cryptomarket, *International Journal of Drug Policy* 35, 16–23.

Van Hout, M.C., and T. Bingham, 2013, 'Surfing the silk road': A study of users' experiences, *International Journal of Drug Policy* 24, 524–529.

Van Slobbe, J., 2016, The drug trade on the deep web: A law enforcement perspective, In *Internet and Drug Markets, EMCDDA Insights.*

Waltman, L., and N. Jan Van Eck, 2013, A smart local moving algorithm for large-scale modularity-based community detection, *The European Physical Journal B* 86, 1–14.

Wang, T.Y., A. Winton, and X. Yu, 2010, Corporate fraud and business conditions: Evidence from IPOs, *Journal of Finance* 65, 2255–2292.

Yermack, D., 2017, Corporate governance and blockchains, *Review of Finance* 21, 7–31.

House Committee on Financial Services

Hearing: "After the Breach: the Monetization and Illicit Use of Stolen Data"

March 15, 2018

Questions for the Record from U.S. Representative Ted Budd (R-NC.)

New threats to penetrate networks are being drawn up by hackers daily. This has resulted in an excess of "endpoint security solutions" whose complexity makes it difficult to manage a company's network defense. A pathway is needed for cybersecurity vendors to collaborate in this space.

Mr. Bernik, can you tell me what McAfee, and others, throughout industry are doing to confront this challenge?

Obviously there have been some big-time data breaches in the news as of late. And it seems now, more than ever, companies that hold large amounts of their customer's personal data are scrambling to ensure they are protected against the threat of data-breach.

Mr. Bernik, can you describe in greater detail how these companies and corporations should deploy security measures that are strong enough to earn them cyber-liability protections that industry has asked for?

To earn liability protections, companies should deploy end to end cyber security solutions and strategies. Best of breed companies, those truly dedicated to improving their cyber security capabilities, invest in cyber security solutions, processes and people with the necessary skills to align to NIST's cyber security frame work. The NIST cyber security frame work pushes companies to think holistically about cyber security and to invest in protecting every aspect of their organization. To date, the NIST framework has achieved a high degree of adoption from the vast majority of critical infrastructure companies.

Despite the success of the NIST framework, we believe that the private sector, working with NIST, can still do more. We need a different approach where technology – enabled by strong collaboration – can be deployed rapidly to security platforms so they can communicate with each other over open communication protocols. Organizations in both the public and private sector need security tools that are interoperable and interchangeable to protect against existing and prospective threats. As cybersecurity solutions become interoperable, they become more efficient and cost-effective. They also become easier to maintain than an IT environment of disparate systems. Over time, more interoperable cybersecurity systems also will contribute to closing the skills gap as these systems become more widely deployed and require less manual intervention.

McAfee has called on the cybersecurity industry to design technology to an open standard, on an open platform so customers are not locked into proprietary technologies that don't work with each other or allow for change. Customers deserve the ability to deploy best-of-breed security solutions, but if they need to install a complete infrastructure just to do so, then customers lose. By having interoperable standards for interface and exchange formats, the industry could move to a more plug-and-play

capability for security products. This has been successful in the past with efforts such as the Security Content Automation Protocol (SCAP), currently in use in the HBSS and CDM programs. SCAP provides a wide variety of vendors the ability to exchange compliance and patch validation content.

McAfee has long believed in breaking down the walls that separate vendors' security products, and we have taken a major step toward doing just that by opening our Data Exchange Layer (DXL) – a communications fabric that enables unprecedented collaboration in an open-source, real-time system – to other developers and vendors to use at no expense. OpenDXL™ is at the core of our mission to enable security devices to share intelligence and orchestrate security operations at rapid speed. As of today, a growing set of security companies, including IBM and Cisco, are actively making connections to the DXL ecosystem. Open DXL is a big part of what we mean by Together Is Power. No single industry partner can cover the vast spectrum of security and privacy problems, just as no single industry partner will catch every issue every time. Only by working collaboratively in the private and public sectors can we build the tools and infrastructure needed to defeat cyber attackers.

We encourage the government to work with the private sector to make the vision of a truly open and interoperable cybersecurity ecosystem become a reality. Such an ecosystem, based on open, international standards, promotes a great deal of competition and innovation. At the same time, it also promotes collaboration – making sure that systems work together. The real benefit is an environment that promotes enough competition to deliver innovative solutions, but enough collaboration to ensure that these new and innovative solutions can work together. Much like the railroad industry that agreed on basic rules of the road – e.g., size and gauge of the tracks and right of ways – the security industry needs rules of the road to allow cooperation, so that firms can compete on implementations to allow for as much innovation as possible.